Chapter 8
Hedonic Analysis of Housing Markets

Stephen Sheppard*
Oberlin College

1997

1. Introduction

Imagine, for a moment, that you are a private investigator or market researcher studying the demand for food. You have a particular disadvantage, however, in that you have been banned from entering the local grocer. You have found a place outside where you can sit and photograph shoppers as they approach the checkout counter, and from these photographs you can pretty much tell what foods each customer has purchased (although some items may be obscured in the shopping basket) and the total cost of all items combined. By bribing a contact at the local bank, you are able to find out each shopper's income. That is all the information you have. From this, can you infer the demand for eggs? Can you determine how much households would be willing to pay to remove sugar import quotas?

Such a difficult assignment is essentially what is undertaken in the hedonic analysis of housing markets. We have no direct observations of attribute prices. We have somewhat imprecise observations of what attributes are purchased. We have reasonably good observations of what is spent for the entire bundle, and how much income the household has. From this we attempt to answer some quite important questions.

Housing and residential construction are of central importance for determination of both the level of welfare in society and the level of aggregate economic activity. In many economies a residence represents the most valuable single asset owned by most individuals, and a very large share of total household wealth. In all economies the share of income spent on housing represents a very large fraction of total expenditure. It is thus to be expected that economists would devote considerable effort towards understanding the structure of the demand for housing and equilibria in these markets.

The first difficulty which arises in this effort stems from the obvious heterogeneity of the product. To attempt a comparison between the price of houses in China and Chile, or between housing prices in 1996 and 1946 requires us to address the similarity and appropriateness of making such a comparison. We observe a low-income central city resident who pays $500 per month for his housing, while the more affluent suburban family in the same city has a monthly expenditure of $1500. Can we conclude that the central city resident is lucky and faces a price of housing which is lower than in the suburbs? More subtly, suppose we observe that these residences are really very different. Are the different choices made by the two households attributable to different incomes, different preferences, non-market constraints on choice, or differences in the effective prices of the attributes which characterize each house? To answer such questions, we have no choice but to undertake hedonic analysis of the market.

The important work of Griliches [1961] and Griliches [1971] did much to introduce hedonic analysis and techniques for dealing with commodity heterogeneity to a wider audience of economists. Griliches and many others have rightly referred to the work of Court [1939] as an early pioneer in application of these techniques, as well as the first to apply the term 'hedonic' to analysis of prices and demand for the individual sources of pleasure – the attributes which combine to characterize heterogeneous commodities. At least a decade before the appearance of Court's

*I would like to thank Joseph Gyourko, Paul Cheshire, and Ed Mills for helpful comments on an earlier draft of this chapter.

work, however, the study of Waugh [1929] appears to be the first to provide a systematic analysis of the impact of 'quality' on the price of a commodity. Waugh characterizes quality using a variety of observable attributes, and estimates the implicit price of each of these attributes. Not only is the study the first to estimate what we now call a 'hedonic price function', it is an important early application of multivariate statistical techniques to economics. It nicely illustrates the way in which hedonic analysis was, and is, an important source of econometric innovation for the profession.

This chapter reviews several aspects of the hedonic analysis of housing markets, focusing broadly on the theoretical foundation of such analysis followed by the practical implementation. Within each of these areas, we consider the issues which affect the determination and estimation of implicit prices for housing attributes, as well as the difficult problems which confront the use of these prices to estimate the demand for characteristics.

This reflects a basic dichotomy which is also present in the literature. There is a large body of work which focuses more or less exclusively on inference of the implicit prices of housing and environmental characteristics. Such information is of considerable value in the construction of price indices which take proper account of changes in the quality of the goods produced, and also in estimating or forecasting values of real estate assets. A second branch of the literature uses these implicit prices to estimate the structure of household demand for housing attributes, and to evaluate the welfare consequences of changes in attribute prices or environmental quality.

There have been several excellent surveys of all or important parts of this literature. The papers by Mayo [1981], Follain and Jimenez [1985], Palmquist [1991] and others provide particularly useful accounts of the progress made in this area, and one might 'review' this area of urban economics by simply referring the reader to these. The discussion below, in addition to highlighting some of the more recent contributions in this area, provides an alternative view of some of these developments.

2. Theoretical foundation

Before discussing the actual estimation and use of hedonic models of housing markets, we begin with a discussion of some of the theoretical constructs which are at the foundation of this approach. We begin with some preliminary remarks concerning 'implicit markets' before turning to discussion of hedonic price functions and the structure of demand for housing attributes.

2.1. Implicit Markets

The notion of implicit markets denotes the process of production, exchange, and consumption of commodities which are primarily (perhaps exclusively) traded in 'bundles'. The explicit market, with observed prices and transactions, is for the bundles themselves. Such a market, however, might be thought of as constituting several implicit markets for the components of the bundles themselves. This is of particular importance when the bundles are not homogeneous, but vary due to the varying amounts of different components which they contain.

There are at least two possible perspectives to take on implicit markets, which differ more in their emphasis and orientation than in the final theory whose development is thus motivated. On the one hand, we might regard the demand for all goods – even those which appear to be homogeneous – as being based not on the goods themselves but upon the characteristics which they embody. The household purchases these goods and uses them as a type of 'input', transforming them into utility, the level of which depends upon the quantity of characteristics embodied in the goods purchased. This approach, motivated at least in part by the work of Lancaster [1966], tends to place particular emphasis on household production and the properties of household demand for the (sometimes unobservable) characteristics.

An alternative view emphasizes the idea that some goods are usefully combined and thought of as being traded in a single 'market', but they are quite heterogeneous – automobiles, workers, and houses are all reasonable examples. Such markets are not capable of being analyzed with the usual economic models because they are not characterized or even approximated by a single price, but rather by a range of prices which depend upon the quality of the commodity or the characteristics it contains. The hedonic approach attacks this difficulty by asserting that these goods, while globally heterogeneous, are composed of aggregates of (more or less) homogeneous parts, and while the aggregate bundle may not have a common price, the component attributes do (or at least a common price structure). The hedonic approach provides a methodology for identifying the structure of prices of the component attributes (estimation of the hedonic price function). Analysis of demand can then proceed using these prices, estimating a demand system in which the attributes are treated as goods. This involves an implicit assumption that a variety of aggregate bundles are available in the market so that consumers can choose any bundle of attributes they wish, being constrained only by their incomes and the price of the resulting bundle.

If we let Z represent a vector of quantities of these attributes which differentiate the commodities, the hedonic approach starts by recognizing that first, each consumer may consume a different commodity, in the sense that each consumer may consume a unit of the good with a different amount of Z embodied. Second, each consumer may pay a different price for the good, so that a range of marginal prices may exist which, in general, depend upon the quantity of Z.

2.2. Hedonic Price Functions

Hedonic price functions are estimated for two primary reasons: for use in construction of overall price indices which account for changes in the quality of goods produced, and as an input in the analysis of consumer demand for attributes of heterogeneous goods. To understand the appropriate estimation techniques and problems, and to interpret the results, we must begin with an understanding of how a market for heterogeneous goods can be expected to function, and what types of equilibria we can expect to observe.

We begin with a recapitulation of the basic theory of hedonic markets. Consumers are assumed to derive utility from consumption of a commodity which embodies a vector Z of J different characteristics, plus consumption of a composite good Y. They have fixed income M and face a price function $P(Z)$ which gives the price of the heterogeneous good (which we refer to as housing) as a function of the embodied characteristics Z. The preferences of the household are represented by the utility function

$$u = u(Z, Y, \alpha) \tag{2.1}$$

where α is a vector of observed and unobserved parameters which characterize the preferences of the household. Households are therefore assumed to be fully characterized by an income M and parameter vector α, with distribution over possible values described by the joint probability $f(\alpha, M)$.

From the utility function 2.1 we can derive the amount which a household would be willing to pay for a house as a function of the embodied characteristics, given the income and achieved utility level for the household. This household 'bid rent' function $\beta(Z, M, u, \alpha)$ is defined implicitly by:

$$u = u(Z, M - \beta, \alpha) \tag{2.2}$$

The derivative of the bid rent function $\frac{\partial \beta}{\partial Z_i}$ gives the rate at which the household would be willing to change expenditure on a house as characteristic i increases, while holding the utility level constant. This is the inverse of the compensated demand curve.

The household chooses a house with characteristics Z, and consumption of composite commodity Y, to solve:

$$\max_{Z,Y} \quad u(Z, Y, \alpha) \text{ subject to } M \geq P(Z) + Y \tag{2.3}$$

First order conditions for this problem require that

$$\frac{u_i}{u_Y} = P_i \quad \forall\, i \tag{2.4}$$

where subscripts indicate partial derivatives so that $u_i = \frac{\partial u}{\partial Z_i}$ and $P_i = \frac{\partial P}{\partial Z_i}$. The derivative P_i is usually referred to as the *hedonic price* of characteristic i, and the function $P(Z)$ as the *hedonic price function*.

Combining the first order conditions in 2.4 with implicit differentiation of 2.2 yields the familiar result that optimal choice of a house is characterized by equality between the slope of the bid rent and the hedonic price for each characteristic:

$$\frac{\partial \beta}{\partial Z_i} = \frac{u_i}{u_Y} = P_i \tag{2.5}$$

This observation is part of the justification of the hedonic approach to analysis of markets, for it indicates that if we can 'observe' (or estimate) the hedonic price for a characteristic and the choice made by the consumer, then under the assumption of optimizing behavior the observation provides local information about the consumer's preferences or willingness to pay for attributes in the neighborhood of the observed choice. In this sense the problem seems similar to the standard analysis of consumer behavior, in which observed choices and prices provide local information about consumer preferences, and given sufficient data we can hope to make accurate inferences about consumer behavior and price determination. As we shall see below, the hedonic market problem is substantially more complex.

To complete the model, we must provide an explanation for the determination of the hedonic price function P. For this we introduce producers of the heterogeneous good who are characterized by cost functions $C(Z, N, \gamma)$ which depend upon the characteristics Z of the houses supplied, the number N of such houses built, and a vector γ of parameters which characterize each producer. Thus the profit of a producer is given by:

$$\pi = P(Z) \cdot N - C(Z, N, \gamma) \tag{2.6}$$

There are a variety of producers, and their distribution is described by the probability density $g(\gamma)$. Each of these producers is assumed to take the price function $P(Z)$ as given and solve:

$$\max_{Z,N} \quad P(Z) \cdot N - C(Z, N, \gamma) \tag{2.7}$$

The first order conditions for solution of this require that:

$$\begin{aligned} P_i &= C_i \quad \forall\, i \\ P(Z) &= C_N \end{aligned} \tag{2.8}$$

Thus each producer equates the marginal cost of each characteristic to its hedonic price, and builds houses until the marginal cost of building another house (of type Z) is equal to the value of the house $P(Z)$. There will typically be a large number of housing producers active in the market, with sellers of existing housing consisting of a special type of producer with $N = 1$ and a cost function $C(.)$ which is determined by the costs and technology of house repair and remodeling.

Equilibrium in this market for heterogeneous goods requires a hedonic price function $P(Z)$ that equates supply and demand for every type of house Z. This equilibrium is conventionally represented as a locus of tangencies between a series of marginal cost curves, C_i, and the derivatives β_i of the bid rent curves. In general, as pointed out in Epple [1987] *inter alia*, the equilibrium price function depends upon the distributions $f(\alpha, M)$ and $g(\gamma)$. This dependence has also been discussed in the context of labor markets in Tinbergen [1959] and Sattinger [1980].

This dependency is clear when one considers extreme examples. Suppose that there are a variety of consumers but only one type of producer, so that $g(\gamma)$ is degenerate. In this case the equilibrium will have only a single C_i for each characteristic, and all β_i will be tangent to it. An equilibrium hedonic price function would be the cost function C. In general, then, the equilibrium hedonic price function depends upon the distribution of consumers and producers. It is possible to endogenize the distribution g using a zero profit condition to generate 'entry' and 'exit' of house producers who behave as described above, increasing or decreasing the density of each producer type γ until a type of 'long run' equilibrium condition is met.

This provides a theoretically coherent foundation for explaining the relationship between the price of a house (or other heterogeneous good) and the characteristics it possesses. It describes the actions of market participants, and provides equilibrium conditions in which these actions combine to determine the hedonic price relationship. Furthermore, it does so in a way which underscores the potential usefulness of observing this relationship. Apart from the difficulties in actually carrying out the estimation required for this ‘observation’, there are still some gaps which have been addressed by researchers or provide opportunities for future research.

The gaps which seem of greatest relevance in the present context relate to aspects which set the market for houses apart from the markets for other heterogeneous commodities such as breakfast cereals or automobiles. Some of these aspects are shared with a few other markets, but in any event are not incorporated into the standard formulations of hedonic models of housing markets. Consider the following features of housing markets:

- **Housing markets involve search** – it is costly to collect information concerning the characteristics embodied in a particular structure, and as a result there is uncertainty about the exact nature of the hedonic price function. A consumer purchases a house by examining a sequence of structures, eventually beginning the process of making offers on properties (possibly continuing to sample) and continuing until the expected increase in utility from continued search is less than the cost. In this sense the housing market is similar to other ‘matching’ type markets such as the labor market.

- **Housing markets are intrinsically spatial** – houses involve varying quantities of land, and possess particular locations. Many empirical studies do not even include location as a characteristic of the house, let alone take advantage of the special nature of this characteristic. The special nature arises when all (or most) relevant local amenities have been accounted for, so that the only factor which distinguishes one location from another is the transport costs. In this case the bid rent curve for locations with greater accessibility is determined by the transport cost function.

- **Housing markets involve both new and existing homes** – perhaps the most significant difference between housing and other heterogeneous goods markets is the importance of the sale of existing (previously produced) goods. The share of house sales accounted for by new construction is relatively small in almost all housing markets. Consumers can substitute between them, and seek out the type, whether new or old, which maximizes utility. While this poses no problem for the consumer side of the market, how should this factor be incorporated into the supply side of the market?

In addition to the observations made above, we can also question whether the ‘hedonic approach’ itself has any testable implications. Are there observations which could be made concerning the implicit price of housing attributes which would refute some or all of the theory outlined above? One issue concerns the convexity of the hedonic price function. This in turn translates into the convexity of the consumer’s budget set, which if violated would make problematic the first-order characterization of consumer choice outlined in equation 2.4 above.

In a restrictive setting (all attributes and commodities perfectly divisible) Jones [1988] shows that in equilibrium, hedonic price functions must be convex. He also shows that in this setting they must be linearly homogeneous. In practice, most estimated hedonic price functions satisfy convexity (see Anderson [1985]) but a great many fail to satisfy the homogeneity restriction.

2.3. Structure of demand and welfare evaluation

A primary reason for undertaking hedonic analysis of housing markets is to understand the structure of demand for housing attributes and environmental amenities. Such understanding is essential for predicting the response to changes in the housing market and for providing welfare estimates of the costs and benefits associated with such changes. In this section we explore some of the properties which can be expected of such demand, and the way in which these properties inform and constrain the application of hedonic analysis. In this context we begin with a consideration of the extent to which ‘demand functions’ can be derived within the hedonic framework, and which properties these relationships might be expected to exhibit. We then discuss the way in which these relationships might be applied to evaluate the welfare consequences of changes which affect the prices of attributes or the levels of environmental amenities.

2.3.1. Properties of attribute demand

Because the hedonic price of an attribute is typically not constant, an ambiguity arises in presenting housing 'demand' within the context of a hedonic model. If we know the structure of preferences for housing and environmental characteristics, we can present the demand functions that characterize a household with those preferences if they faced constant prices. Alternatively, we can present behavioral functions which represent the choices which would be made by a utility maximizing household, as a function not of hedonic prices, but rather of the parameters of the hedonic price function. The first would present quantities chosen by the household as a function of 'prices' and income. The second presents choices as a function of the parameters which determine the hedonic price function and income.

The first has the advantage of presenting household choice as a function of income and prices in a structure which is reassuringly (but deceptively) familiar. By presenting a representation of household choice in a conventional structure, demand functions which express household choice as a function of prices and income allow calculation of price and income elasticities which facilitate comparison of demand for housing with the demand for other goods. Unfortunately, such comparisons can be misleading since households do not, in fact, face linear budget constraints.

The second technique, which presents household choice as a function of income and the parameters of the hedonic price function has the advantage of providing (subject to the limits of estimation) a quantitative description of actual household behavior. Unlike the 'linearized' demand functions, these would provide accurate descriptions of household responses to changes in the hedonic price function. There are two difficulties with these 'hedonic demands' which perhaps account for the infrequency with which they appear in applied research. First, they can be somewhat difficult to estimate, since they will almost always necessitate the use of nonlinear models. Second, they are difficult to use in developing what might be called a 'housing market heuristic'. For example, we often confront changes in housing markets such as an increase in household incomes or a change in availability of land for residential development, which are expected (and typically do) have clear qualitative consequences for hedonic prices. It can be difficult to translate these into qualitative predictions of housing attribute demand using the 'hedonic demands', and quite easy to do so with the 'linearized demands'.

The ease with which such a heuristic is developed, however, is of little benefit if it is fundamentally misleading. Our first question, therefore, is to inquire as to the usefulness of estimating these linearized demands. We shall see below that estimation of this type of 'demand' provides a technique for estimating the structure of preferences, which is a principal objective of hedonic analysis of housing markets, and from which the more complete 'hedonic demands' can be derived analytically. There are a variety of properties which demand functions possess in the conventional setting, and we inquire as to those that might be expected to carry over to the linearized demands derived for a consumer facing a nonlinear budget. In the subsection that follows, we consider problems that arise in using the linearized demands for welfare analysis in housing markets.

Suppose that we proceed along the lines suggested in Rosen [1974], and in fact implemented in a number of studies, taking estimated hedonic prices as if they were actual prices, combining them with household income and observed household choices to estimate the structure of demand. We might attempt this by postulating a parametric form either for the utility function or, equivalently, for the expenditure function, and then derive a demand function for each attribute of the form

$$q_i\left(P_1, ..., P_J, M, \alpha\right) \tag{2.9}$$

We could then proceed to estimate this in the form of a share equation in an additive error term:

$$w_i = \frac{P_i \cdot q_i\left(P_1, ..., P_J, M, \alpha\right)}{M} + \varepsilon \tag{2.10}$$

There are two important problems with this strategy: the first concerns the estimation technique and lack of stochastic independence between the 'variables' P_i and the error term ε. This issue has received a great deal of attention and is discussed more completely in section 3.2 below. The second problem concerns the appropriateness of this approach when the budget set faced by the household is potentially non linear due to the nonlinearities which may be present in $P\left(Z\right)$.

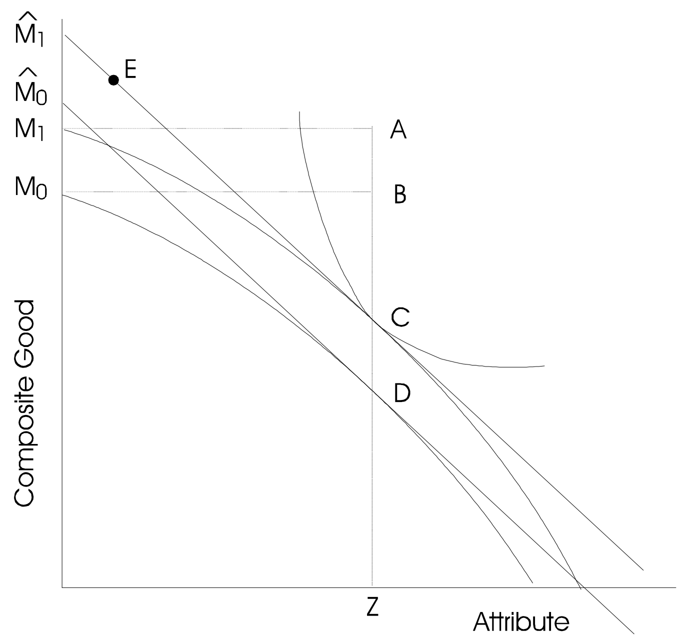

Figure 2.1: Hedonic budget lines and consumer choice

To understand the importance of the nonlinearity, consider figure 2.1. Here we see two 'budget lines'. The budget line is defined by the equality

$$Y = M - P(Z) \tag{2.11}$$

where we shall assume that the hedonic price function is convex so that the household's budget set will be convex. As long as the hedonic price function is independent of household income (which will in general be true if sellers of houses are price-takers) the budget lines themselves are 'quasi linear'– the slope of the budget line is $-P'(Z)$ and independent of consumption of the composite good. Thus the distance $(A - C) = (B - D) = P(Z)$, and $\widehat{M_1} - \widehat{M_0} = M_1 - M_0 = (C - D)$.

A household with income M_1, facing hedonic price $P(Z)$ and making choice C exhibits the same behavior as a household with income $\widehat{M_1}$facing an attribute price which is constant and equal to $\frac{\partial P}{\partial Z}$. As discussed in section 2.2 the first order conditions for household choice imply a tangency between the indifference curve, the true budget line M_1, and the linearized budget line $\widehat{M_1}$ at the point C. Provided that preferences are strictly convex, demand will be a function, and the linearized demand will be the same as the hedonic demand.[1]

Thus if we want to estimate the structure of household preferences by estimating household demand for attributes, we can take as data the hedonic prices $P_i = \frac{\partial P(Z)}{\partial Z_i}$, and the actual household choices, but we must use income level $\widehat{M} = M - P(Z) + \sum_{i=1}^{J} P_i \cdot Z_i$. Based on this data, the structure of demand can be estimated.

Such demand is what Murray [1983] refers to as 'mythical' demand, stressing that the resulting estimated demands are those that would obtain if the consumer faced linearized prices. While the demand functions might be 'mythical', there is nothing mythical about the household preferences which can be inferred from the estimated demands. If a parametric form for demand functions is chosen to facilitate integrability, then we can determine the parameters of a utility or expenditure function directly from the estimated demand systems. For example if a Generalized CES (as in Quigley [1982] or Follain and Jimenez [1983]) or Almost Ideal Demand System (as

[1] A discrepency clearly arises if preferences are not *strictly* convex. While a point such as E in figure 2.1 would not be an element of hedonic demand, it could be an element along with point C in linearised demand if the indifference curve had a flat portion which was colinear with the linearised budget constraint $\widehat{M_1}$. While of minor theoretical interest, this is of minimal relevance in applications since estimation of demand functions already requires the assumption of strictly convex preferences.

in Cheshire and Sheppard [1998] or Parsons [1986]) is applied, then the parameters of the utility or expenditure function are directly obtainable from estimates of the demand.

It is important to qualify the interpretation of the estimated demands and preferences by recognizing that households subject to nonlinear budget constraints do not exhibit the same behavior as households subject to linear constraints, even if (locally) the linear and nonlinear budgets are equivalent. In this sense, the 'linear approximation' demands are not so much mythical as erroneous, with the error depending upon the nonlinearity of the hedonic prices and the preferences of the household.

Conventional demands derived from linear budget constraints exhibit a variety of standard properties which have proven useful in applied demand analysis. These can be used either as the basis of tests of the model, or potentially as a source of 'non-sample' information to improve the quality of estimates when, for example, sample data exhibit colinearity. Which of these properties are preserved in the linearized demands?

The study by Turnbull [1994] derives some properties of overall housing demand in a context which explicitly considers the endogenous location choices which determine the spatial structure of housing consumption. This paper (along with related studies such as Blackley and Follain [1987], DeSalvo [1985], or Turnbull [1993]) establish clearly that demand for housing is affected in important ways by endogenous location choice, and that such demand does not possess the same properties as conventional demand. In particular, the 'law of demand' may be violated for housing even if housing is not an inferior good. Turnbull derives sufficient conditions to rule this out, and shows that with convex equilibrium rent functions this is less likely to occur near the urban periphery. The fact that it can occur at all might make one despair of using these properties to test or inform housing demand estimates. While the analyst might seek comfort in the fact that the analysis focuses on location choice, the difficulty is much more general.

The problem arises because of the endogeneity of prices, and the potential for a household to respond to a pattern of price increases by not only altering the amount of space they consume, but also their location (and hence the price of space and their effective income). In a non-spatial setting, an almost identical problem arises, in which an increase in the general price of an attribute can generate not only a change in the consumption of that attribute, but also a reconfiguration of the entire dwelling and neighborhood. In general this will change the marginal price of the attribute under consideration, and the final adjustment in consumption of the attribute may be quite unexpected.

The most complete treatment of this problem (surprisingly rarely cited in the hedonic literature) has been provided by Blomquist [1989], who investigates the general properties exhibited by demands which are obtained by maximizing a strictly quasi-concave differentiable utility function subject to a general nonlinear budget constraint $M \geq P(Z,\theta)$, where θ is a vector of parameters of the hedonic price function[2]. Blomquist assumes that P is strictly increasing in Z and differentiable, and is such that a unique differentiable solution exists to the household optimization problem. Clearly it is sufficient for this if P is convex. Let $Z_i(M,\theta,\alpha)$ represent the hedonic demand for housing attribute i, that is the i^{th} component of the solution to the problem:

$$\max_{Z,Y} \quad u(Z,Y,\alpha) \text{ subject to } M \geq P(Z,\theta) + Y \tag{2.12}$$

For a fixed (M^*,θ^*) ,we have well defined household choice given by $Z_i(M^*,\theta^*,\alpha)$, and can define the linearized demand $Z_i^L\left(\widehat{M},p\right)$as the i^{th} component of the solution to the problem:

$$\max_{Z,Y} \quad u(Z,Y,\alpha) \text{ subject } \quad \text{to } \widehat{M} \geq p'Z + Y \tag{2.13}$$

$$\text{where } p_i = \frac{\partial P(Z_i(M^*,\theta^*,\alpha),\theta^*)}{\partial Z_i} \qquad \text{and } \widehat{M} = \sum_{i=1}^{J} p_i \cdot Z_i(M^*,\theta^*,\alpha)$$

Let S represent the Slutsky matrix of substitution terms derived from the linearized demands Z^L, and let H^P

[2]Blomquist [1989] does not refer to P as a 'hedonic price' function. The focus of his paper is oriented towards complex tax schedules or wage functions.

represent the matrix of second derivatives of the hedonic price function with typical element given by:[3]

$$H^P_{i,j} = \frac{\partial^2 P\left(Z_i\left(M^*, \theta^*, \alpha\right), \theta^*\right)}{\partial Z_i \partial Z_J} \tag{2.14}$$

Blomquist shows that there is a well defined link between the income effects of the true hedonic demand and the income effect of the associated linearized demand. The relation is given by:

$$\frac{\partial Z}{\partial M} = \left(I - S \cdot H^P\right)^{-1} \frac{\partial Z^L}{\partial \widehat{M}} \tag{2.15}$$

This elegant formulation illustrates how the curvature of the true budget constraint (captured via H^P) and the curvature of preferences (captured via S) interact to produce a complex relation between the estimated income effects of the linearized demand and the income effects which can be expected of a utility maximizing household. A similar expression allows evaluation of the impacts of a change in the parameters θ which determine the hedonic price function. Define a vector $H^P_\theta = \frac{\partial^2 P}{\partial Z_i \partial \theta_j}$, whose i^{th} element gives the impact on the hedonic price of attribute i of a change in parameter θ_j. Taking

$$A = S \cdot H^P_\theta - \frac{\partial P}{\partial \theta_j} \frac{\partial Z^L}{\partial \widehat{M}} \tag{2.16}$$

we can write the impact of the parameter θ_j

$$\frac{\partial Z}{\partial \theta_j} = \left(I - S \cdot H^P\right)^{-1} A \tag{2.17}$$

Again, we see the possibility of expressing the actual change in household attribute demand as a function of the structure of linearized demand (via S and the 'income effect' component of A), the curvature of the budget constraint, and the way in which the parameter in question affects the budget constraint.

This analysis demonstrates that there is nothing problematic or incorrect in estimating linearized demands using the derivatives of the hedonic price function as prices and using the modified income $\widehat{M}$ defined in equation 2.13. Not only is such a procedure capable of providing estimates of the parameters of the utility or expenditure function, but the true comparative statics of hedonic demand in the neighborhood of the equilibrium choice can be obtained by applying equations 2.15 and 2.17.

This analysis is not, however, always helpful in constraining or informing the estimation of the linearized demands. Some writers (such as McConnell and Phipps [1987]) prefer to avoid the term 'demand' altogether, since the 'correct' hedonic demands don't depend upon prices (but rather on the parameters of the hedonic price function) and the linearized demands which do depend on prices are 'mythical'. Whatever terminology is preferred, what is clear is that it is at least in principle possible to infer the parameters of attribute demand directly from observations of household choices and hedonic prices.

The primary difficulty which confronts such inferences is the nonlinearity of the household budget set. Though it increases the complexity of analysis, this difficulty is not insurmountable. For example, the analysis of Ohsfeldt and Smith [1990] derives exact expressions for calculating elasticities taking into account the non-linearity of household budgets. Their research makes an insight which will be important in the discussion of welfare analysis below: that 'linearized' elasticities underestimate the substitution which occurs between housing attributes and non-housing goods.

2.3.2. Welfare analysis

Consider two examples of analysis which might constitute plausible goals for application of hedonic analysis of housing markets: evaluation of the welfare effects of housing regulation, and anticipation of the consequences of an increase in income on housing demand. We continue to assume that the hedonic price function is convex, so that household choice sets are also convex.

[3] Clearly this matrix comprises the first J rows and columns of the hessian matrix of the hedonic price function.

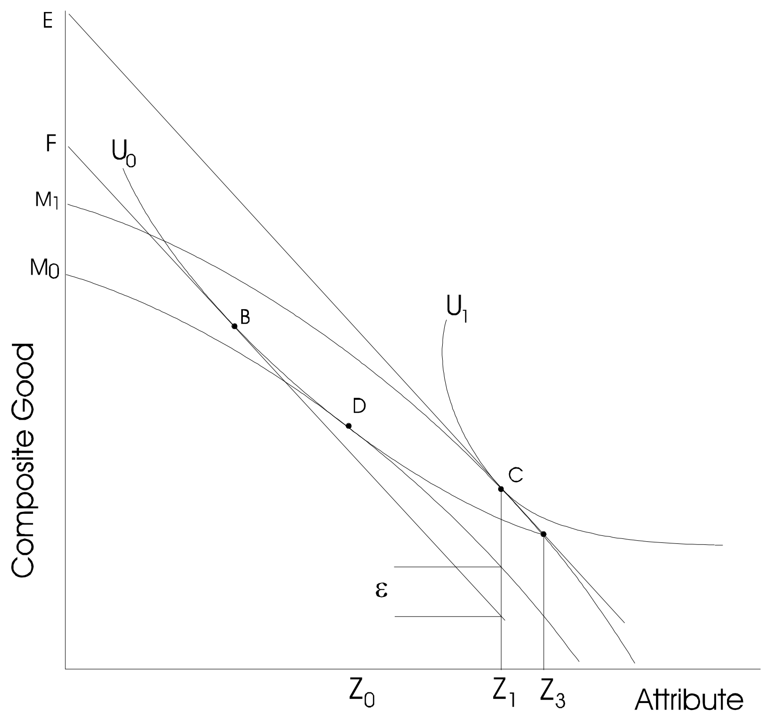

Figure 2.2: Estimated attribute demand can overstate welfare costs

Figure 2.2 illustrates the situation of a household having income M_1 who achieves the maximum utility level U_1 by purchasing a house with quantity Z_1 of the attribute. Suppose that a constraint on household behavior requires them to instead purchase quantity Z_3 of the attribute[4]. This forces them to utility level U_0 and could be reasonably said to generate a welfare cost whose income equivalent is $M_1 - M_0$.

Suppose we utilize the standard expenditure function derived from the preferences we have estimated using 'linearized' budgets and adjusted income as outlined above. From the estimated expenditure function we could parametrize utility to determine utility levels U_1 and U_0. We might then try to estimate the welfare cost using this expenditure function, and the estimate would be $E(P_Z, U_1) - E(P_Z, U_0) = E - F$ in figure 2.2, where the hedonic price of the attribute P_Z is held constant. What relation does this estimate have to the correct magnitude $M_1 - M_0$?

As indicated in the diagram, the welfare loss estimated from the conventional expenditure function is usually **greater** than the true welfare cost, and the error is indicated by ε. This error results from the convexity of the hedonic budget constraint and the failure of the usual technique using traditional expenditure functions to account for the endogeneity of price. This is related to the observation noted by Ohsfeldt and Smith [1990] that non-linearity of the budget generates greater substitution between housing attributes and non-housing goods (the composite commodity).

One circumstance in which the traditional approach gives an exact evaluation is if there is 'zero income effect' in demand for the attribute. In such an unlikely case the utility maximizing choice for a household facing the reduced income level M_0 would involve no change in the quantity of the attribute, and the change in total expenditures would equal the equivalent change in income.

The nonlinearity of budgets also has implications for forecasting the response of house attribute demand to a change in household income. Again referring to figure 2.2, the estimated demand structure would lead us to expect an increase in attribute demand from Z_0 to Z_1 to be generated by an increase in income of $E - F$. In fact, demand will

[4] We might call this example 'large lot zoning', except that strictly speaking such a system wide regulation would also change the equilibrium hedonic price function $P(Z)$. Perhaps the extra required painting of a home with 'historic' designation would be an appropriate interpretation.

be somewhat more income elastic, with this change in attribute demand being generated by the smaller increase in income of $M_1 - M_0$. Clearly, the accuracy of approximation obtained using the conventional 'linear budget' constructs to analyze household choice in a hedonic framework depends upon the structure of preferences and the curvature of the hedonic price function.

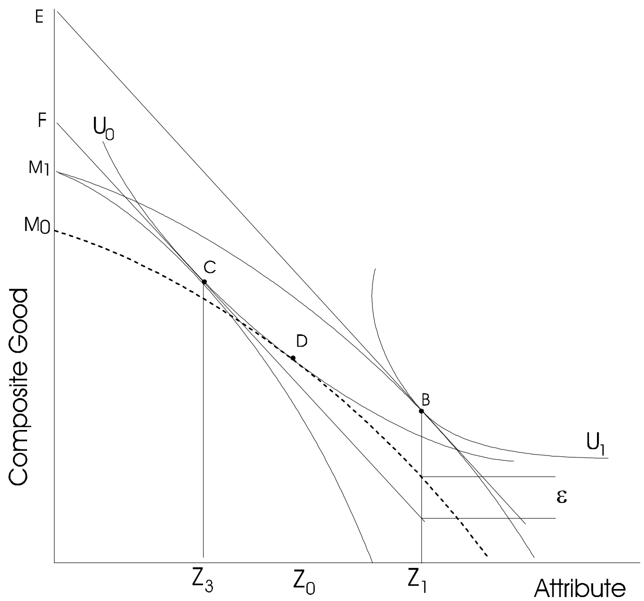

Figure 2.3: Measuring welfare effects of a price change

Similar problems arise when we seek to evaluate the impact of a change in the hedonic price function. Of course, there are a wide variety of possible changes, but suppose we consider an 'increase' in the hedonic price in which at every level of attribute consumption, the hedonic price increases so that the budget line is 'steeper'. Such a situation is illustrated in figure 2.3, where the budget line produced by the hedonic price function is rotated downwards by the price increase. The household responds by changing from optimal choice B to choice C. It is reasonable to seek to measure the variation in income which is equivalent to this price change.

Figure 2.3 indicates that if the hedonic attribute prices were kept constant but income were reduced by $M_1 - M_0$, then the effect on welfare would be the same as the increase in attribute price. Using a conventional measure of the welfare effect based on the expenditure function would give $E - F$ as the change in income which is equivalent to the price increase. As indicated this overestimates the actual welfare effect by ε. The accuracy of the approximation will depend on the nonlinearity of the hedonic price function and the structure of preferences.

Thus a basis exists for using estimated hedonic prices and modified income levels to estimate the structure of demand. A variety of econometric difficulties confront such estimation, and these are discussed below in section 3.2. To be most useful a parametric form would be chosen which is explicitly derived from an underlying expenditure or utility function, but in this respect hedonic analysis of house attribute demand is no different from other types of demand analysis.

Bartik [1988] has provided a comprehensive guide to accurate measurement of benefits from changes in house attributes or environmental amenities using hedonic price models. The methods discussed implicitly provide a way to estimate the error ε in the linearized welfare measure noted above. Although difficult to implement, it is possible to obtain an exact welfare measure. In practice, many studies communicate the approximate welfare measure, but it is important to note that this is at best an approximation.

2.4. Hedonic price indices

Before proceeding to discuss implementation of hedonic analysis of demand, we note that a great deal of interest focuses purely on the use of estimated hedonic price functions for the construction of price and quantity indices. Indeed, such indices have been the primary motivation for much of the hedonic literature, particularly that which begins with Griliches [1961]. There have been a variety of interesting contributions in this area, including a relatively comprehensive method for constructing price indices presented in Feenstra [1995]. Comparisons of some of the alternative methods used for estimation have been presented by Meese and Wallace [1997].

3. Implementation

3.1. Estimation of hedonic price functions

As mentioned in the introduction, the pioneering study by Waugh [1929] represented not only the first attempt to estimate a hedonic price function, but also a very early application of multivariate techniques to the analysis of economic data. It is perhaps a surprise that one of the first applications of multiple regression should have been to a problem which is in fact very difficult. This may account for the continued attention which the problem draws from econometricians and applied economists. Estimation of hedonic prices confronts the economist with a rich sampling of the standard difficulties that arise in estimation using cross-section data. These include choices of the proper parametric specification – both of functional form and of variables to be included – coping with colinearity and ill-conditioned data, potential heteroscedastic and non-normal errors, regressors subject to measurement error, and maximum likelihood estimation of relationships which are non-linear.

Confronted with these difficulties, a variety of approaches are possible. We review here some which appear to be most useful or promising, along with the traditional. Choosing an estimation methodology cannot be done independent of consideration of the particular data sources available and the objectives of the analysis. For example, the 'best' approach to use will generally depend on whether the estimated hedonic price function is to be used to infer the implicit prices of attributes, or simply to forecast or appraise the value of individual properties.

We first discuss the traditional parametric approaches in which a specific functional form is chosen, and the parameters which define it are estimated. This is followed by noting the alternative non-parametric or semi-parametric approaches which make inferences about the implicit prices of attributes without imposing *a priori* a functional relation between the total price of the house and the quantities of attributes which characterize it.

3.1.1. Parametric approaches

From the early work of Waugh [1929], Court [1939], and Griliches [1961] to the most recent studies, the standard approach to estimation of the hedonic price function has been to adopt some form of parametric approach. This means that the analyst must choose a functional form whose actual values are determined by a finite number of parameters. The estimation then proceeds by selecting those parameter values resulting in a hedonic price function which gives the 'best fit' to the data.

This procedure has evolved in several ways, as allowed by the constraints of computation technology, the availability of data, and understanding of the nature of statistical problems involved. Studies such as Ridker and Henning [1967], Kain and Quigley [1970], Harrison and Rubinfeld [1978], Goodman [1978], and Linneman [1981] have tended to rely on linear or logarithmic parametric forms which performed reasonably well and were computationally feasible to estimate. In the 1980's, beginning with the work of Linneman [1980], hedonic studies began to make use of more flexible functional forms obtained by applying the Box-Cox transformation either to housing prices or non-dichotomous attribute quantities. This transformation uses a single parameter λ to transform the variable x as follows:

$$x^{(\lambda)} = \begin{array}{l} \frac{x^{\lambda}-1}{\lambda} \text{ if } \lambda \neq 0 \\ \ln x \text{ if } \lambda = 0 \end{array} \tag{3.1}$$

The model proceeds by expressing transformed price as a linear or quadratic function of transformed attribute quantities, with the transformation parameters being possibly different between the variables, and estimated along with the other parameters.

Although more complicated to estimate, particularly using computationally simple techniques (as noted by Spitzer [1982] and further investigated by Blackley, Follain, and Ondrich [1984]) increased computing power eventually made maximum likelihood estimation of such models relatively straightforward, even with a large number of attributes. Combined with the analysis of Halvorsen and Pollakowski [1981] which identified some particularly strong features of models based on the Box-Cox transformation, many studies began, and still do, rely upon this sort of parametric approach.

This approach is not, however, free of all difficulties. A problem observed by Cassel and Mendelsohn [1985] is that focus on traditional model specification criteria which emphasize the predictive capability of the model is not always appropriate in estimating a hedonic model. If the model is estimated for the sole purpose of predicting total house values, then choosing the parametric form which gives the best possible 'fit' is quite appropriate. Often, however, the objective of estimating a hedonic price function is to determine the implicit prices of attributes. For this objective, minimization of the squared prediction error may be quite inappropriate, and a model which 'fits the data' well in this sense may be less satisfactory than another with less predictive power but more stable parameter estimates.

The interesting paper of Cropper, Deck, and McConnell [1988] investigates precisely this issue. They simulate housing market equilibria using housing stock data from an actual urban area and varying the parameter of the household utility functions. They can then determine analytically the true equilibrium price functions, and the prices of structures can be used to estimate hedonic price functions and marginal bids, comparing the estimates obtained from different functional forms with the 'true' values. They evaluated linear, log-linear, quadratic, linear Box-Cox, and quadratic Box-Cox models, and obtained several surprising results. Models were evaluated not by how well they fit the data, but by how accurately they estimated true marginal bids. When all valuable attributes were observed, and without measurement error, the model which was linear in Box-Cox variables performed the best while the quadratic Box-Cox performed the worst.

When some important attributes were omitted or observed imprecisely, the simpler forms such as linear Box-Cox and logarithmic forms were the most accurate. They attribute this at least in part to the fact that in quadratic forms, each hedonic price depends on more coefficients than in the more nearly linear cases.

Related to the issue of the parametric form of the hedonic price function are issues of which variables to include and the way in which they should enter. Consider, for example, the role of lot size or land area. Examination of any standard presentation of urban economic theory reveals the central role played by models of land value determination. One might reasonably assert that a central principal of urban economics is that the price of land will vary with location, and this varying land price is what produces the varying types and intensity of land use in cities.

Given this observation, it is surprising how many hedonic models lack either a variable for land area or a variable which explicitly identifies the location of the structure[5]. Very few models explicitly incorporate a land value function which depends upon location. Two exceptions to this are Jackson and Kaserman [1984] and Cheshire and Sheppard [1995]. Both of these studies are able to obtain reasonable estimates of the value of land in residential use, and to reject the hypothesis that this value is constant over locations. Recent research by Colwell and Munneke [1997], who studied prices of vacant land, has also shown that land prices are most definitely not constant over locations.

While it is commonplace to justify adoption of a flexible form for estimation of a hedonic price function by observing that 'theory places few restrictions' on the form, there are at least **some** restrictions. As observed above,urban economic theory suggests that the form should include land values which depend on location. Furthermore, Jones [1988] has shown that in equilibrium the hedonic price functions should be convex. This guarantees convexity of consumer budget sets and, with convexity of preferences, establishes continuity properties of household attribute demand. Although Anderson [1985] develops and applies a test of hedonic function convexity, it is again

[5] Even worse, some studies lack both.

surprising how few subsequent studies have applied this simple test. The study of Colwell and Munneke [1997] is worth mentioning again at this point because in examination of prices of vacant land they obtain concavity of the price function for land. Whether this holds more generally for land as an attribute in housing markets has not been widely investigated.

A third way in which theory might suggest appropriate restrictions on the parametric structure of the hedonic price function estimated concerns public goods or environmental amenities. Parsons [1990] argues for weighting local public goods and neighborhood amenities by lot size, based on land market equilibrium and the fact that consumption of these amenities is limited by available land for residential consumption. While his argument depends on an assumption that consumption of these attributes is non-exclusive (and therefore may not be appropriate for some local attributes such as schools or parks), for a great many environmental amenities his analysis suggests a reasonable theoretical constraint on the parametric specification for hedonic models.

Specification Many of the issues raised above might be viewed under the general heading of 'model specification' for the hedonic price function. Over the past decade, this topic has generated considerable interest amongst econometricians. Several studies have considered techniques for testing or evaluating alternative specifications for the hedonic price function. Among these are Butler [1982], Milon, Gressel, and Mulkey [1984], Dubin and Sung [1990], Burgess and Harmon [1991], and Craig, Kohlhase, and Papell [1991].

This last paper is interesting because it introduces a new test which is particularly appropriate in the setting of hedonic estimation. They apply the test to housing market data to test for the systematic nonlinearities in the relationship between structure price and attribute quantity. In addition to testing for the appropriateness of nonlinearities in the functional form, the test can also provide an indication of important omitted variables. Interestingly (and *a propos* the comments above about the need to include land area which varies with location) their analysis identified an inability to 'discover a specification for lot size which ... passes' the test they derive. Whether a specification that did not impose spatially invariant land value would pass was not tested.

Colinearity and error structure Two econometric problems which seem intrinsic to estimation of hedonic price functions are collinearity or ill-conditioned data and lack of stochastic independence between observations. The first problem has received considerable attention in the literature. Only recently has the second begun to be addressed.

It seems natural to expect collinearity to pose a problem for the estimation of hedonic prices. Because of similarity in the preferences of households and limits on the technology of house construction, there are intrinsic limits to the extent of variance of attributes which we are likely to observe. The more limited is this variance (and the greater the extent to which the variables tend to move together) the less will be the precision with which model parameters are estimated. This is a particular problem for hedonic estimation where the precision of parameter estimation is important for obtaining accurate estimates of attribute prices.

As is reasonably widely known (see for example Belsley and Welsch [1980]) there is really only one way to 'solve' the problem of collinearity: get more information. This information might come from larger or richer data sources. Alternatively, we might turn to formal ways of incorporating non-sample information into our estimates using some sort of Bayesian technique. Recent studies by Knight, Hill, and Sirmans [1993] and Gilley and Pace [1995] have applied and evaluated this approach. The first study uses Monte Carlo techniques to evaluate Bayesian versions of Stein-like estimators and shows that according to several criteria the approaches based on non-sample data empirically dominate least squares approaches.

The study by Gilley and Pace is particularly important because it illustrates an innovative use of information which should be available in almost all applied settings. As noted in section 2.2 above, equilibrium in an implicit market involves a tangency between consumer bid functions and producer cost functions. Thus the hedonic price should be equal to both the household marginal willingness-to-pay (equation 2.4) and also the marginal cost of making the attribute available (equation 2.8). In many markets there are standard construction cost data sources which give some estimate of the marginal cost of attribute supply. While these would apply primarily to new construction, they may constitute a valuable source of non-sample information which can be used to inform a prior distribution in a Bayesian estimation procedure.

The spatial structure of housing markets has been noted above, and is obvious to anyone engaged in analysis of the housing market. This might lead one to be concerned about spatial linkages between errors in hedonic price models: that is a lack of stochastic independence between observations, in which the error of an observation is correlated with those observations that are located nearby. In analogy with the common time series difficulty, this problem is termed spatial autocorrelation, and if present it leads to similar difficulties as arise in time series models.

The presence of this problem and possible corrections have been investigated by Dubin [1988], Dubin [1992], and Can [1992]. Alternative estimation strategies which account for the spatial autocorrelation are seen to be warranted, and the results suggest that models which account for these types of spatial structure are likely to produce more reliable results.

3.1.2. Non-parametric approaches

An alternative to specification of a parametric functional form for the hedonic price function is to adopt a non-parametric or 'semi-parametric' approach to estimation which attempts to infer attribute prices directly from the data without benefit of an assumed functional relationship. The difficulty which arises in application of these techniques is the extremely large amounts of data they require (formally, the slow rate at which these estimation techniques converge to the true value as sample size increases).

Feasible estimates can, however, be obtained by considering approaches which consider truly non-parametric combinations of simple parametric forms. Papers which have investigated this approach include Knight, Hill, and Sirmans [1993], Pace [1993], Pace [1995], and Anglin and Gencay [1995]. These approaches turn out to be much more robust to specification and measurement error than many parametric estimates, although they can be considerably more computationally complex to estimate.

The analysis of Pace [1995] suggested in particular that semi-parametric estimates of hedonic price functions seem to suffer fewer incorrectly signed parameters, a common problem with collinear data. Even more interesting, the study of Anglin and Gencay [1995] compared a parametric estimator with a semiparametric one. The parametric model was easily rejected in tests which compare it to the semiparametric. Furthermore, this was not for a poorly specified parametric model, but one that passed a variety of standard tests of model specification. This suggests that considerable gains in accuracy may be available by utilizing semiparametric techniques for estimating hedonic price functions.

3.2. Estimation of demand for housing attributes

Estimation of the demand for any heterogeneous good such as housing presents a variety of difficult problems, some of which are rarely encountered in other economic contexts. It is a perhaps peculiar attribute of the literature concerning estimation of demand for housing attributes that during the past decade much of it has been obsessed with clarifying the problem of endogeneity of price.

Partly, this stems from a peculiarity in the development of the literature, in which the problem of estimating the demand for housing attributes was originally thought to represent a 'garden variety' simultaneous equations problem[6]. That this was not so was realized relatively quickly, but it took nearly a decade to produce a clear exposition of exactly how the price endogeneity problem differed from the variety one finds in the usual economist's garden.

What is peculiar about this is that while there have been a variety of papers which discussed and clarified the nature of the problem, there have been few which used actual housing market data to illustrate how much difference a proper accounting of the problem made to final estimates. In part this was no doubt due to the difficulty of producing estimates which do take proper account, but in part it was also due to the somewhat limited intersection between the set of economists interested in elucidating proper econometric practice and the set of economists seeking to use hedonic analysis as a tool for understanding markets or conducting policy analysis.

[6] As characterized in Rosen [1974].

A second peculiar aspect has been the enthusiasm with which the 'lack of consistency' due to failure to use a proper instrumental variables technique has been criticized, while other sources of bias receive little in the way of repeat attention. Most obvious of these other sources has been the failure to include almost certainly relevant variables in hedonic models, particularly the structure location. Again, there is a pragmatic explanation: an individual economist might be expected to be able to adopt different estimation techniques which could improve estimation and account for price endogeneity. An individual researcher is unlikely to be able to compel release of actual (rather than pseudo) census tract numbers in the American Housing Survey.

Economists have become accustomed to accepting data as they find it, warts and all. They then agonize about how optimally to analyze this data, even if the reduction in error obtained from this optimization is trivial compared to the error caused by the structure of (or omissions from) the data. This occupational hazard applies with particular force to practitioners in the hedonic modeling area. A reader of the literature could be forgiven for getting the impression that studies which neglected location, land area, and all local public goods would be welcomed if only they incorporated a maximum-likelihood instrumental variables technique which accounted for price endogeneity. In fact, of course, there is little to recommend adoption of an estimation technique which achieves consistent estimates if it is applied to an obviously misspecified model.

3.2.1. Identification and endogeneity

The most widely perceived difficulty afflicting estimation of demand for housing attributes is the 'identification' problem which arises because of the endogeneity of attribute prices. Although the problem is technically one of identification, it is useful to avoid confusion with the econometric difficulties faced when attempting to estimate the parameters of demand and supply simultaneously. This sort of 'conventional' simultaneity was first noted as a potential problem by Freeman [1979], and alluded to in Rosen [1974]. For studies based on cross sectional data with units of observation large enough to influence market prices, such difficulties may indeed arise. Most hedonic studies, however, are based upon data sets with observations of individual household decisions and incomes.

For estimates based on individual data, the first response might be to argue that there is no problem in estimation (other than the usual ones of specification and dealing with nonlinear structures). The individual demands or 'marginal benefit functions' are determined taking the hedonic prices as exogenous. The situation would seem to be similar to that depicted in figure 3.1. Here we see a three dimensional surface representing the individual demand function which gives the quantity of attribute desired as a function of household income and the hedonic price of the attribute. A few of the many available data points are indicated, and these may be used to determine the demand function using relatively standard maximum likelihood methods.

This perspective (which at least implicitly lies at the foundation of a very large literature) holds that since we have observations of individual behavior, along with the prices and incomes which have produced that behavior, we can estimate the structural equations of household demand as long as we are willing to assume that a common preference ordering determines the decisions of all households. All that is required are data in which there is linearly independent variation in income and hedonic prices, and the estimates will be readily obtained.

The first objection to this perspective is the simple one that the hedonic prices are not known with certainty, but must be estimated. The stochastic nature of the relation between attribute quantities and house price will impart some uncertainty to the estimated hedonic prices. To this one might reply that demand estimation is a classic 'errors in variables' problem, and that as long as the hedonic price function has been correctly specified, the parameters will be consistently estimated and for large samples we need not worry. Of course, for finite samples any hypothesis testing we might wish to do with our demand estimates will be troubled by variance estimates which are biased downwards, but in principle these could be corrected by using the covariance matrix for parameters of the hedonic price function to obtain an estimate of the covariance structure for the hedonic prices. This approach, however, cannot by itself compensate for violations of Gauss-Markov conditions, and produces consistent estimates only if the errors in hedonic prices (and 'linearized income') are independent of the error terms in the demand functions. This is precisely the core of the problem.

The problem is easily seen in figure 3.2. In the upper portion of the figure is shown a hedonic price function

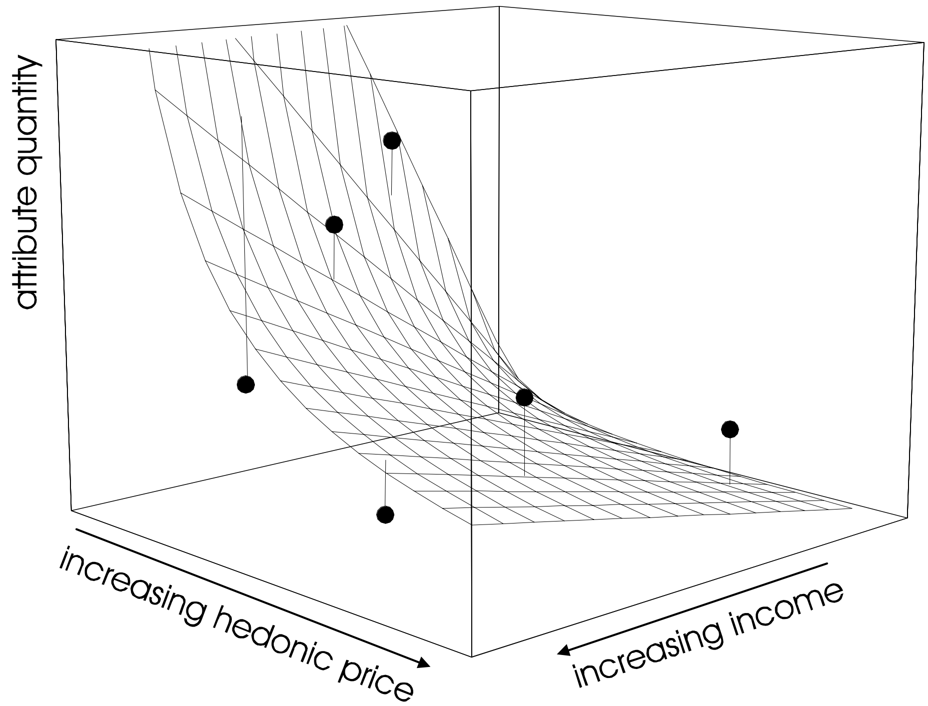

Figure 3.1: Estimation of hedonic demand

$P(Z)$ and a household bid function for the attribute Z which is tangent at point B so that the household would demand Z_1 units of the attribute given this price structure.[7] The object of hedonic analysis is to obtain an estimate of the actual demand or marginal benefit function drawn as a solid line in the lower portion of figure 3.2. For example, we might assume additive error ε and estimate

$$Z = \beta_0 + \beta_1 \cdot P_Z + \varepsilon \tag{3.2}$$

using the observed household consumption and the estimated hedonic price function to determine the hedonic prices P_Z. In the case illustrated, the parameter β_1 is negative, and we try to obtain an estimate $\widehat{\beta_1}$. When the error ε is near zero, we observe the household consuming Z_1 and infer from the hedonic price function that the price which produces this behavior is the slope of $P(Z)$ at point B.

When ε is negative, we observe household consumption of an attribute level such as Z_0. At this consumption we infer a hedonic price which is the slope of $P(Z)$ at point A, which is less than the true slope of the household bid curve and hence less than the true level of the household demand function. When ε is positive, we observe Z_2 and infer a hedonic price which is the slope of $P(Z)$ at point C. The result is an estimated household demand which, as indicated in the figure, is too low for negative ε and too high for positive ε – we are simply tracking the slope of the hedonic price function.

This is essentially the problem observed by Brown and Rosen [1982]. If no correction is made for price endogeneity, demand estimates may be determined entirely by the hedonic price function itself. The estimate $\widehat{\beta_1}$ is biased upwards, and is inconsistent since increasing the sample size does nothing to eliminate or change the correlation ρ between the error ε and the hedonic price P_Z. Letting σ_ε and σ_{P_Z} denote the standard deviation of ε and P_Z, respectively, we have:

$$\text{plim } \widehat{\beta_1} = \beta_1 + \rho \frac{\sigma_\varepsilon}{\sigma_{P_Z}} \tag{3.3}$$

[7] Alternatively, the hedonic price function could be used to construct a budget line, and the bid function would be an indifference curve tangent to the budget line at B.

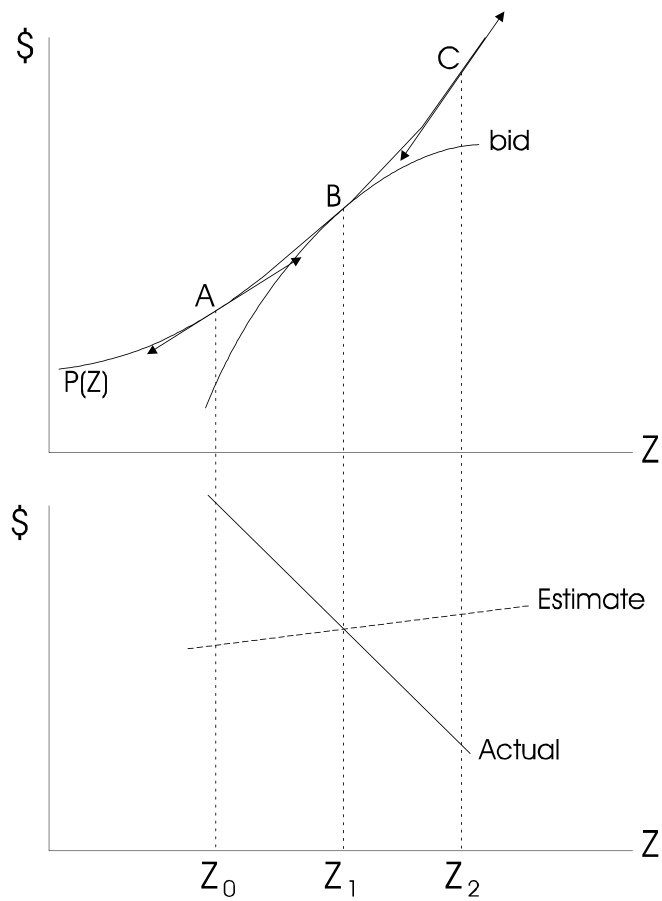

Figure 3.2: Biased estimates from price endogeneity

For strictly convex hedonic price functions, the correlation ρ is positive and we obtain inconsistent demand estimates whose slopes are biased upwards. For more complex (and realistic) demands with several attributes, and linearized income which is also correlated with ε, the formula for the bias is more complex but the basic intuition remains unchanged: **the correlation between estimated hedonic prices and errors in measured demand behavior leads to inconsistent estimates of the structure of demand.**

There are a variety of possible responses to this problem, but most of the recent literature on estimation of housing demand begins with this basic observation. Thus the analyses of Blomquist and Worley [1982], Brown and Rosen [1982], Murray [1983], Diamond and Smith [1985], Ohsfeldt and Smith [1985], Bartik [1987b], Epple [1987] and McConnell and Phipps [1987] all share a common initial theme: with individual household data, there is no structural simultaneity; the difficulty arises because the endogeneity of prices gives rise to correlation between the random error in the model and the 'independent variables'. This results in inconsistent estimates.

The required response to such a problem is to devise consistent, 'instrumental variables' estimates of those variables which appear on the right hand side of the structural equations. Thus, for example, in equation 3.2 we need to identify variables which are uncorrelated with ε which can be used to provide a consistent estimate of P_Z. This estimate is then used in subsequent stages of the procedure to estimate attribute demand. More generally, we identify instruments which permit estimation of all of the hedonic attribute prices and linearized income.

Three basic approaches exist to obtain such instruments. First, we may be able to find or construct other variables which are independent of the errors but sufficiently correlated with hedonic prices to provide admissible instruments. Second, we might take advantage of nonlinearities which exist in actual hedonic price and marginal benefit relations to identify the models (essentially using transformations of the variables as instruments). Third, we might use other variables which occur in the structural equations to obtain a set of instruments for consistently estimating hedonic prices.

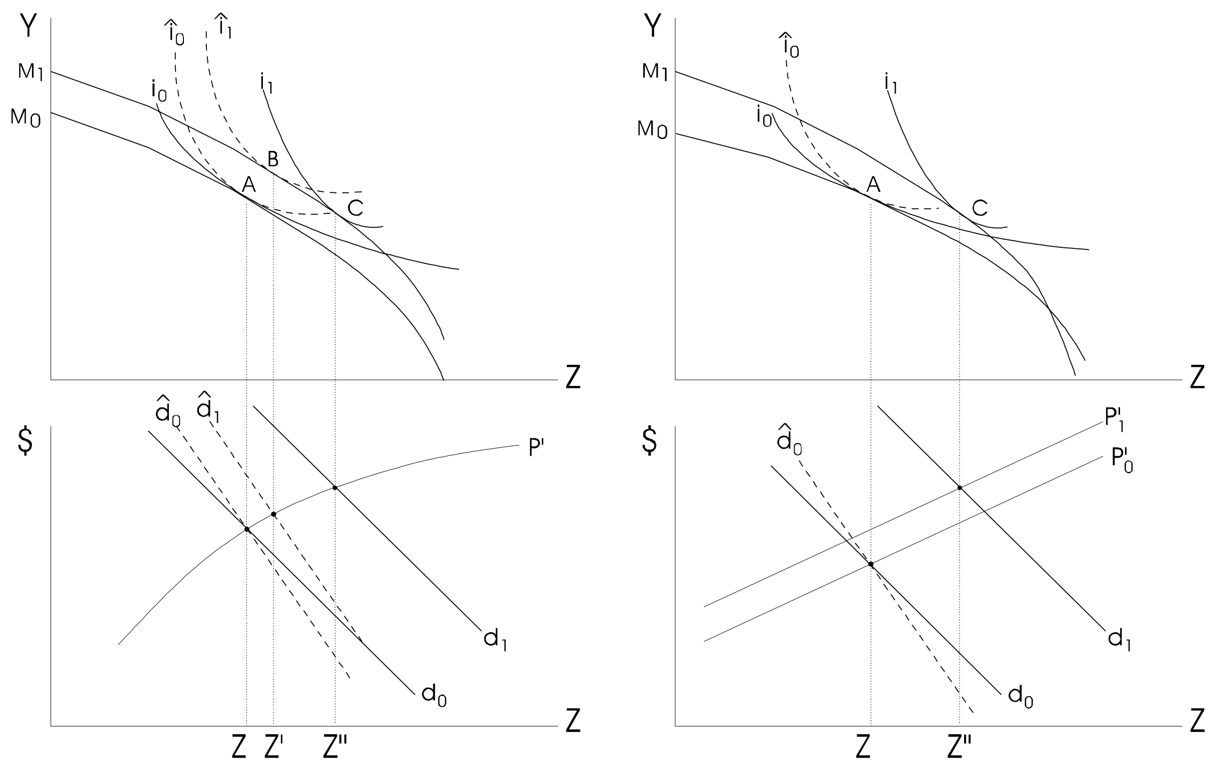

Figure 3.3: Data from single and multiple markets

The third method is familiar from use in solving the identification problem and estimating models in which there is true simultaneity. Its application is stressed as a potential solution in several of the papers cited above, and is illustrated in the right hand portion of figure 3.3. In the upper right hand quadrant of the figure we see two budget curves associated with hedonic price functions from two different markets.[8] In one market, the observed household has income M_0 and optimally chooses attribute level $\underline{Z}$ indicated by tangency A. The problem is to distinguish between the alternative demand structures d_0 and $\widehat{d_0}$ generated respectively by preferences i_0 and $\widehat{i_0}$. This is possible because of the exogenous variation in price schedules. In the second market, the observed household has income M_1 and the choice of attribute level Z'' at tangency B serves to identify the demand structure and preference. This approach, recommended *inter alia* by Brown and Rosen [1982], Diamond and Smith [1985], Ohsfeldt and Smith [1985], Bartik [1987b], and Epple [1987].

There are two primary difficulties with actually applying this 'multiple market' approach. The first is theoretical, the second practical. From a theoretical perspective, estimation of demand from multiple markets requires the assumption of a common demand structure in all markets covered. If these markets are from multiple time periods, we must assume constant preferences over time (as well as cross sectionally). More problematic, if the markets are from widely separated urban areas, we must assume that households have a common preferences structure across all areas. Given the wide variety of environmental amenities, available public goods, and social structures which characterize each urban area this may be regarded as a possible source of specification error. Related to this point is a problem noted by Diamond and Smith [1985]: in a world in which households can choose which of several urban areas in which to live, the urban area itself may well be endogenous and not serve as a valid instrument.

From a practical perspective the difficulties are even larger. There are almost no individual level data sets which are comparable across a large number of urban areas and available to the public. There are none[9] in Europe or Asia. In the United States there are two potential data sources: the American Housing Survey and property market data available from local tax authorities or through real estate listing services. The first of these is available at very low cost and includes a large number of variables for about 30 metropolitan areas. The difficulty is that it does not include location. This makes it impossible to estimate a 'land value' component in the hedonic price function, although there are some accessibility variables which provide information on the journey to work and mode of transport used. The second data source contains much better information concerning location (since property address is included), but typically contains no detail concerning the occupants. In particular, it does not contain household income, nor does it include any data on the age or structure of the occupants. It might be possible (given sufficient resources) to survey a sample of properties to determine values for these other important variables.

Most of the 'multiple market' studies that have been done have used the American Housing Survey.[10] These studies were severely constrained by the nature of the data available. None of these included location, despite the theoretical importance of location in determining the value of land and the relatively large proportion of residential construction costs accounted for by land. None included more than 8 attributes, with such obviously important factors as local school quality or accessibility to parks and open space never appearing in any study. These limitations arise because of the limited data available, and might be corrected in the future. At present, however, it is rather difficult to understand the attraction of seeking to correct one sort of bias (arising from endogenous prices) by introducing bias of another sort (severe specification error by failing to include important attributes). This is particularly true when other approaches exist which might solve the problem.

The second approach mentioned above relies upon imposition of mostly untestable restrictions of nonlinearity on both the hedonic price function and the structure of demand. This approach is illustrated in the left hand side of figure 3.3. In the top left quadrant we see two budget sets drawn from the same hedonic market but generated by the two income levels M_0 and M_1. At the lower income level the household makes optimal choice of attribute level $\underline{Z}$ determined by the tangency A. As before, the problem is to distinguish between the demand structure d_0 and $\widehat{d_0}$ generated respectively by preferences i_0 and $\widehat{i_0}$.

To get an intuitive feel for the problem, imagine that we restrict attention to *only* these two possible preference structures. While it might be impossible to distinguish between these two demand structures at income level M_0, they may be readily distinguished at other price-income combinations. If households have preferences which generate

[8] We know these are two different markets because the budget lines are not vertically parallel.
[9] None, at least, known to the author.
[10] See *inter alia*, Parsons [1986].

indifference curves $\widehat{i_0}$ and $\widehat{i_1}$, then at income level M_1 the optimum will be at tangency point B and they will choose attribute level Z'. If they have preferences which generate indifference curves i_0 and i_1, their optimum will be defined by point C with attribute level Z''. Subject to the constraint that household demand structures are one of these two types, we only need data with sufficient variation in hedonic prices and household incomes to identify demand.

Of course, it is hardly satisfactory to restrict households to having one of two possible preference orders, but given sufficient variation in incomes and hedonic prices we can impose much less severe restrictions. The analysis of McConnell and Phipps [1987] provides a complete discussion of the restrictions required for identification of preference parameters in this case, and the discussion in Epple [1987] is also usefully comprehensive.

These sources underscore the difficulty, but not the impossibility, of identifying hedonic demand estimates from single market data. If there is sufficient variation in the data to permit unique maximum likelihood estimates of parameters, that alone can be argued to ensure local identifiability. The studies by Quigley [1982] and Kanemoto and Nakamura [1986] pursue variations on this second approach. By restricting attention to a particular class of preference ordering and functional forms for the hedonic price function, they solve for and estimate the parameters of the preference function.

The first of the three approaches to price endogeneity mentioned above was to find or construct other variables which are correlated with hedonic prices faced by the household but not correlated with the error terms of the demand (or marginal benefit) functions. In one sense, this obvious approach is what is being done in the multiple market approach as well: other variables (which index or characterize the particular hedonic markets) are used to construct instruments which are then used to obtain consistent estimates of variables on which the demand functions depend. The usual approach when applying this method to true simultaneous equation models is to rely upon other variables in the structural equations, and most of the discussion on the 'multiple market' approach has followed this tradition.

Restricting the search to other structural variables, however, makes the problem very difficult. Thus McConnell and Phipps [1987] argues that if some attribute is not included in the hedonic price function, it will be impossible to create instruments which are not correlated with the error. The analysis of Bartik [1987a] makes this argument most forcefully, arguing that unless unobserved variability in preferences is assumed not to exist, it will be impossible to construct valid instruments for consistent estimation of household attribute demand. His argument proceeds by examples, however, and is restricted implicitly to an assumption that the source of instruments are other variables which appear in the structural equations of the model.

Are such variables the only source of instruments to which we might turn? Surely that depends upon the stochastic structure of the application, but in principle one need not restrict attention to only those variables. It is helpful to keep in mind that our problem is not one of a truly simultaneous equation system. Consistent estimation of the hedonic demand for J attributes requires that we obtain consistent estimates of the J hedonic prices (and hence also linearized income) which define the budget set of the household. For this we need $J+1$ or more instruments which are uncorrelated with the error in household attribute demand, but which are not so weak as to give extremely imprecise (even if consistent) estimates of the actual hedonic prices.

Murray [1983] makes a variety of interesting suggestions concerning possible instruments, and recently Cheshire and Sheppard [1998], have pursued the idea of using the average attribute prices paid by 'similar' households as instruments. There are numerous dimensions along which one might define similarity, and they consider two: taking those households which occupy the locations which are nearest to the household, and taking those households which have chosen to consume similar houses (including both the attributes and location). For the latter concept of similarity, they construct an index which measures the Euclidean distance in characteristics space weighted by consumption shares. For both geographic distance and characteristics space distance, they experimented using as instruments the prices paid by the two 'nearest' households in their sample.

Whether such an approach is likely to provide a valid set of instruments is, as noted above, dependent on the stochastic specification of the model.[11] If these errors come from simple measurement error on the quantities

[11] This particular importance of stochastic specification is in addition to the important considerations identified by Horowitz [1987].

of attributes, then consideration of prices paid by similar households might provide good instruments as long as the measurement errors were uncorrelated between observations. Whether this is true will naturally depend on the attribute and the nature of the market, but fortunately it is relatively straightforward to test for instrument admissibility. Gourieroux and Monfort [1995] present a test of unknown linear constraints based on Asymptotic Least Squares, and show how to apply it as a test of instrument admissibility.

Suppose we seek to estimate a hedonic demand Z_i which depends on hedonic prices P_k and linearized income $\widehat{M}$ (see equation 2.13 above):

$$Z_i = \beta_0 + \sum_{k=1}^{J} \beta_k \cdot P_k + \beta_M \cdot \widehat{M} + \varepsilon \tag{3.4}$$

Suppose we have a set of $K > J$ instruments Ψ for the T observations of P_k and $\widehat{M}$, and we wish to test the validity of these instruments. That is, we wish to test the hypothesis that

$$\operatorname*{plim}_{T} \frac{1}{T} \sum_{t=1}^{T} \Psi_t \cdot \varepsilon_t = 0 \tag{3.5}$$

where the ε_t error terms have common variance σ_ε. A test statistic for this hypothesis is

$$\xi_T = \frac{1}{\widehat{\sigma_\varepsilon}} \widehat{\varepsilon}' \, \Psi \left(\Psi'\Psi\right)^{-1} \Psi' \widehat{\varepsilon} \tag{3.6}$$

where $\widehat{\varepsilon}$ is the vector of residuals from the two-stage least square estimate obtained using Ψ as instruments. A remarkable feature of this statistic is that it is equal to the product of the sample size T and R^2,the coefficient of determination obtained from regressing the two-stage least squares residuals $\widehat{\varepsilon}$ on the set of instruments Ψ. The statistic ξ_T is distributed χ^2 with $K - J$ degrees of freedom under the hypothesis of admissibility.

This is completely intuitive: we have as an estimate of the (unobserved) error the residuals $\widehat{\varepsilon}$. Validity of the instruments Ψ requires that they be independent of these errors. If the variables Ψ are able to explain the variation in $\widehat{\varepsilon}$, then it seems unlikely that Ψ and ε will be independent.

Using this test, Cheshire and Sheppard have found that the combined characteristics distance instruments are admissible for use in estimating the hedonic demand for land area and for open space amenities. They have found that the geographic distance instruments are admissible for all of the characteristics for which they had data. The geographic distance based instruments were somewhat weaker, although both provided demand system estimates with acceptable (in sample) accuracy.

Thus a potentially reasonable procedure for applied analysis using single market data would be to estimate hedonic demand using an instrumental variables procedure such as two-stage least squares, taking as instruments the hedonic prices which confront the households with the nearest locations. A possible variation on this idea would be to average several nearby households, possibly weighting the average by distance. This approach, of course, requires that the location of the households be part of the available data, but this is almost surely required for proper specification of the model in any event. If it is not available, proximity or similarity based on other household characteristics might serve the same role.

If the data are available, it would be quite reasonable to combine these approaches. Use of nonlinear functional forms, with data from several urban areas, taking as instruments the hedonic prices faced by the nearest households would be defensible on several grounds. Data from multiple markets will always be helpful, and will increase the in-sample variability in income and hedonic prices, and will therefore generally produce more accurate estimates. The point is that multi-market data is not the only approach to estimation of such demand. Tolerating a severely misspecified model in order to access data from several housing markets is unlikely to be a reasonable strategy.

3.2.2. Specification

In estimating household hedonic demand or marginal bid functions, there are three basic approaches which have appeared in the literature. First, and probably least reliable, is to simply estimate some reasonably flexible non-linear demand which depends on (linearized) household income and attribute hedonic prices. The difficulty with this is the lack of connection between the estimated demand and a preference ordering which presumably generates it[12].

A preferable approach is to estimate a demand structure derived from an explicit utility or expenditure function. For example, the studies by Quigley [1982] and Kanemoto and Nakamura [1986] undertake this as a possible solution to the price endogeneity problem discussed above. Without regard to whether their technique is the best solution to price endogeneity, they do certainly obtain estimable demand functions derived from an explicit preference ordering. An alternative is to use a demand system derived from a flexible expenditure function. The Almost Ideal Demand System presented in Deaton and Muellbauer [1980] is used in Parsons [1986] and Cheshire and Sheppard [1998].

While some techniques have been proposed for estimating compensating and equivalent income variations directly from hedonic price functions (see Horowitz [1984]), general welfare analysis will typically require more detailed knowledge of the household expenditure functions. This is most directly obtained by estimation of a completely specified demand system.

3.2.3. Comparison with Discrete choice approaches

Before ending our discussion of attribute demand estimation using hedonic techniques, it is worth noting an alternative approach to valuing house attributes using the 'discrete choice' approach developed by McFadden [1977] and more completely presented in Ellickson [1981] and Lerman and Kern [1983]. This approach uses the basic hedonic model of implicit markets to develop a multinomial discrete choice model of residential choice in which the amount by which increasing an attribute raises the probability that a particular house is chosen conveys information about the value the household attaches to that attribute. While the discrete choice model has fewer problems to deal with in terms of price endogeneity and demand specification, it avoids these problems by imposing considerable structure (implicit in the discrete choice modeling itself) on the nature of household preferences.

This discrete choice approach provides an alternative to hedonic demand estimation, and the study presented in Cropper [1993] compares the two approaches. The study concludes that for small 'marginal' changes in attributes, the traditional hedonic approach using a hedonic price function which is linear in Box-Cox transformed variables provides more accurate evaluation. This advantage is lost, however, for large changes in attribute quantities.

4. Conclusion

Hedonic analysis of housing markets is an important part of the toolbox of applied urban economics. The technique has been evolving over some seventy years of econometric practice and economic understanding, and has become very important in the past 25 years.

The theory can be presented in a formal way based on theories of implicit markets, and it is possible to infer the parameters of household demand from observations of household choices and the implicit prices they face. Implementation of this approach, however, forces confrontation with a variety of difficulties.

Most of the problems associated with estimation of the hedonic price function itself are conventional, even if not easily solved. Models should be specified so that they correspond to the restrictions implicit in the theory of urban housing markets. Estimation must confront inadequate data, and make use of whatever information sources are available.

Use of these hedonic prices to estimate the structure of demand brings more difficulties, many of which have been only poorly understood. In addition to the usual problems of model specification and measurement error, the nonlinearity in household budgets implies endogenous determination of attribute prices. A variety of strategies for meeting the challenge of endogeneity have been presented, ranging from use of multiple market data to construction of alternative 'non-structural' instrumental variables for prices and income. Whatever approach is adopted, it is clear that accurate estimation requires some explicit acknowledgment of the endogeneity, although the actual quantitative significance of the problem may in some cases be modest.

[12] As noted, for example, by McConnell and Phipps [1987] in discussing the example used by Brown and Rosen [1982].

References

Anderson, J. E. (1985): "On Testing the Convexity of Hedonic Price Functions," *Journal of Urban Economics*, 18(3), 334–37.

Anglin, P. M., and R. Gencay (1995): "Semiparametric Estimation of a Hedonic Price Function," Discussion paper, University of Windsor, University of Windsor Working Paper.

Bartik, T. J. (1987a): "Estimating Hedonic Demand Parameters with Single Market Data: The Problems Caused by Unobserved Tastes," *Review of Economics and Statistics*, 69(1), 178–80.

——— (1987b): "The Estimation of Demand Parameters in Hedonic Price Models," *Journal of Political Economy*, 95(1), 81–88.

——— (1988): "Measuring the Benefits of Amenity Improvements in Hedonic Price Model s," *Land Economics*, 64(2), 72–83.

Belsley, David A., K. E., and R. E. Welsch (1980): *Regression Diagnostics, Identifying Influential Data and Sources of Collinearity*. Wiley, New York.

Blackley, D. M., and J. R. Follain (1987): "Tests of Locational Equilibrium in the Standard Urban Model," *Land Economics*, 63, 46–61.

Blackley, P., J. Follain, James R., and J. Ondrich (1984): "Box-Cox Estimation of Hedonic Models: How Serious Is the Iterative OLS Variance Bias?," *Review of Economics and Statistics*, 66(2), 348–53.

Blomquist, G., and L. Worley (1982): "Specifying the Demand for Housing Characteristics: The Exogeneity Issue," in *The Economics of Urban Amenities*, ed. by D. B. Diamond, and G. S. Tolley, chap. 4, pp. 89–102. Academic Press, New York.

Blomquist, N. S. (1989): "Comparative Statics for Utility Maximization Models With Nonlinear Budget Constraints," *International Economic Review*, 30, 275–296.

Brown, J. N., and H. S. Rosen (1982): "On the Estimation of Structural Hedonic Price Models," *Econometrica*, 50(3), 765–68.

Burgess, James F., J., and O. R. Harmon (1991): "Specification Tests in Hedonic Models," *Journal of Real Estate Finance and Economics*, 4(4), 375–93.

Butler, R. V. (1982): "The Specification of Hedonic Indexes for Urban Housing," *Land Economics*, 58(1), 96–108.

Can, A. (1992): "Specification and Estimation of Hedonic Housing Price Models," *Regional Science and Urban Economics*, 22(3), 453–74.

Cassel, E., and R. Mendelsohn (1985): "The Choice of Functional Forms for Hedonic Price Equations: Comment," *Journal of Urban Economics*, 18(2), 135–42.

Cheshire, P., and S. Sheppard (1995): "On the Price of Land and the Value of Amenities," *Economica*, 62(246), 247–67.

——— (1998): "Estimating the Demand for Housing, Land, and Neighbourhood Characteristics," *Oxford Bulletin of Economics and Statistics*, (forthcoming).

Colwell, P. F., and H. J. Munneke (1997): "The Structure of Urban Land Prices," *Journal of Urban Economics*, 41, 321–36.

Court, A. T. (1939): "Hedonic Price Indexes with Automotive Examples," in *The Dynamics of Automobile Demand*. General Motors, New York.

Craig, S. G., J. E. Kohlhase, and D. H. Papell (1991): "Chaos Theory and Microeconomics: An Application to Model Specification and Hedonic Estimation," *Review of Economics and Statistics*, 73(2), 208–15.

Cropper, M. L., L. B. Deck, and K. E. McConnell (1988): "On the Choice of Functional Form for Hedonic Price Functions," *Review of Economics and Statistics*, 70(4), 668–75.

Cropper, M. L. e. a. (1993): "Valuing Product Attributes Using Single Market Data: A Comparison of Hedonic and Discrete Choice Approaches," *Review of Economics and Statistics*, 75(2), 225–32.

Deaton, A., and J. Muellbauer (1980): "An Almost Ideal Demand System," *American Economic Review*, 70, 312–326.

DeSalvo, J. S. (1985): "A Model of Urban Household Behavior with Leisure Choice," *Journal of Regional Science*, 25, 159–174.

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

Dubin, R. A. (1988): "Estimation of Regression Coefficients in the Presence of Spatially Autocorrelated Error Terms," *Review of Economics and Statistics*, 70(3), 466–74.

——— (1992): "Spatial Autocorrelation and Neighborhood Quality," *Regional Science and Urban Economics*, 22(3), 433–52.

Dubin, R. A., and C. H. Sung (1990): "Specification of Hedonic Regressions: Non-nested Tests on Measures of Neighborhood Quality," *Journal of Urban Economics*, 27(1), 97–110.

Ellickson, B. (1981): "An Alternative Test of the Hedonic Theory of Housing Markets," *Journal of Urban Economics*, 9(1), 56–79.

Epple, D. (1987): "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products," *Journal of Political Economy*, 95(1), 59–80.

Feenstra, R. C. (1995): "Exact Hedonic Price Indexes," *Review of Economics and Statistics*, 77, 634–53.

Follain, J., and E. Jimenez (1983): "The Demand for Housing Characteristics in Developing Countries," Discussion Paper WUDD 43, The World Bank Water Supply and Urban Development Department.

Follain, J. R., and E. Jimenez (1985): "Estimating the Demand for Housing Characteristics: A Survey and Critique," *Regional Science and Urban Economics*, 15, 77–107.

Freeman, A. M. (1979): "Hedonic Prices, Property Values and Measuring Environmental Benefits: A Survey of the Issues," *Scandinavian Journal of Economics*, 81(2), 154–73.

Gilley, O. W., and R. K. Pace (1995): "Improving Hedonic Estimation With an Inequality Restricted Estimator," *The Review of Economics and Statistics*, pp. 609–621.

Goodman, A. C. (1978): "Hedonic Prices, Price Indices and Housing Markets," *Journal of Urban Economics*, 5(4), 471–84.

Gourieroux, C., and A. Monfort (1995): *Statistics and Econometric Models, Volume 2*. Cambridge University Press, Cambridge.

GRILICHES, Z. (1961): "Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change," *The Price Statistics of the Federal Government*, No. 73.

——— (1971): *Price Indexes and Quality Change: Studies in New Methods of Measurement*. Harvard University Press, Cambridge, MA.

HALVORSEN, R., AND H. O. POLLAKOWSKI (1981): "Choice of Functional Form for Hedonic Price Equations," *Journal of Urban Economics*, 10(1), 37–49.

HARRISON, DAVID, J., AND D. L. RUBINFELD (1978): "Hedonic Housing Prices and the Demand for Clean Air," *Journal of Environmental Economics and Management*, 5(1), 81–102.

HOROWITZ, J. L. (1984): "Estimating Compensating and Equivalent Income Variations From Hedonic Price Models," *Economics Letters*, 14, 303–308.

——— (1987): "Identification and Stochastic Specification in Rosen's Hedonic Price Model," *Journal of Urban Economics*, 22(2), 165–73.

JACKSON, JERRY R., J. R. C., AND D. L. KASERMAN (1984): "The Measurement of Land Prices and the Elasticity of Substitution in Housing Production," *Journal of Urban Economics*, 16, 1–12.

JONES, L. E. (1988): "The Characteristics Model, Hedonic Prices, and the Clientele Effect," *Journal of Political Economy*, 96(3), 551–67.

KAIN, J. F., AND J. M. QUIGLEY (1970): "Measuring the Value of Housing Quality," *Journal of the American Statistical Association*, 65, 532–548.

KANEMOTO, Y., AND R. NAKAMURA (1986): "A New Approach to the Estimation of Structural Equations in Hedonic Models," *Journal of Urban Economics*, 19(2), 218–33.

KNIGHT, J. R., R. C. HILL, AND C. F. SIRMANS (1993): "Estimation of Hedonic Housing Price Models Using Nonsample Information: A Monte Carlo Study," *Journal of Urban Economics*, 34(3), 319–46.

LANCASTER, K. (1966): "A New Approach to Consumer Theory," *Journal of Political Economy*, 74, 132–157.

LERMAN, S. R., AND C. R. KERN (1983): "Hedonic Theory, Bid Rents, and Willingness-to-Pay: Some Extensions of Ellickson's Results," *Journal of Urban Economics*, 13(3), 358–63.

LINNEMAN, P. (1980): "Some Empirical Results on the Nature of the Hedonic Price Function for the Urban Housing Market," *Journal of Urban Economics*, 8(1), 47–68.

——— (1981): "The Demand for Residence Site Characteristics," *Journal of Urban Economics*, 9, 129–148.

MAYO, S. K. (1981): "Theory and Estimation in the Economics of Housing Demand," *Journal of Urban Economics*, 10, 95–116.

MCCONNELL, K. E., AND T. T. PHIPPS (1987): "Identification of Preference Parameters in Hedonic Models: Consumer Demands with Nonlinear Budgets," *Journal of Urban Economics*, 22(1), 35–52.

MCFADDEN, D. (1977): "Modeling the Choice of Residential Location," Discussion Paper 477, Cowles Foundation.

MEESE, R. A., AND N. E. WALLACE (1997): "The Construction of Residential Housing Price Indices: A Comparison of Repeat-Sales, Hedonic-Regression and Hybrid Approaches," *Journal of Real Estate Finance and Economics*, 14, 51–73.

Milon, J. W., J. Gressel, and D. Mulkey (1984): "Hedonic Amenity Valuation and Functional Form Specification," *Land Economics*, 60(4), 378–87.

Murray, M. P. (1983): "Mythical Demands and Mythical Supplies for Proper Estimation of Rosen's Hedonic Price Model," *Journal of Urban Economics*, 14(3), 326–37.

Ohsfeldt, R. L., and B. A. Smith (1985): "Estimating the Demand for Heterogeneous Goods," *Review of Economics and Statistics*, 61, 165–171.

——— (1990): "Calculating Elasticities from Structural Parameters in Implicit Markets," *Journal of Urban Economics*, 27, 212–221.

Pace, R. K. (1993): "Nonparametric Methods with Applications to Hedonic Models," *Journal of Real Estate Finance and Economics*, 7(3), 185–204.

——— (1995): "Parametric, Semiparametric, and Nonparametric Estimation of Characteristic Values within Mass Assessment and Hedonic Pricing Models," *Journal of Real Estate Finance and Economics*, 11(3), 195–217.

Palmquist, R. B. (1991): "Hedonic Methods," in *Measuring the Demand for Environmental Quality*, ed. by J. B. Braden, and C. D. Kolstad, pp. 77–120. Elsevier Science Publishers, Amsterdam.

Parsons, G. R. (1986): "An Almost Ideal Demand System for Housing Attributes," *Southern Economic Journal*, 53, 347–363.

——— (1990): "Hedonic Prices and Public Goods: An Argument for Weighting Locational Attributes in Hedonic Regressions by Lot Size," *Journal of Urban Economics*, 27(3), 308–21.

Quigley, J. M. (1982): "Nonlinear Budget Constraints and Consumer Demand: An Application to Public Programs for Residential Housing," *Journal of Urban Economics*, 12, 177–201.

Ridker, R. G., and J. A. Henning (1967): "The Determinants of Residential Property Values With Special Reference to Air Pollution," *Review of Economics and Statistics*, 49, 246–257.

Rosen, S. (1974): "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 82(1), 34–55.

Sattinger, M. (1980): *Capital and the Distribution of Labor Earnings*. North-Holland, Amsterdam.

Spitzer, J. J. (1982): "A Primer on Box-Cox Estimation," *Review of Economics and Statistics*, 64, 307–313.

Tinbergen, J. (1959): "On the Theory of Income Distribution," in *Selected Papers of Jan Tinbergen*, ed. by L. M. K. L. H. Klaasen, and H. J. Witteveen. North-Holland, Amsterdam.

Turnbull, G. K. (1993): "The Substitution Theorem in Urban Consumer Theory," *Journal of Urban Economics*, 33, 331–343.

——— (1994): "Housing Demand Properties in the Monocentric Market Form," *Regional Science and Urban Economics*, 24(2), 253–63.

Waugh, F. V. (1929): *Quality as a Determinant of Vegetable Prices*. Columbia University Press, New York.