

Chapter 2

Hedonic Housing Price Theory Review

2.1 Introduction

The most commonly applied methods of housing price evaluation can be broadly divided into two groups: traditional and advanced methods. There are five traditional mainstream standard recognized valuation methods in the field of property valuation: comparative method (comparison), contractor's method (cost method), residual method (development method), profits method (accounts method), and investment method (capitalization/income method). Advanced methods include techniques such as hedonic price modeling, artificial neural networks (ANN), case-based reasoning, and spatial analysis methods.

Hedonic price modeling is the most commonly applied of these. Many scholars (e.g., Griliches 1961) have referred to the work of Court (1939) as an early pioneer in applying this technique. He used the term hedonic to analyze price and demand for the individual sources of pleasure, which could be considered as attributes combined to form heterogeneous commodities. It was an important early application of multivariate statistical techniques to economics.

In this chapter, several aspects of hedonic modeling will be investigated in-depth, including the theoretical basis, theoretical criticism, estimation criticism, and its use in pricing housing attributes, including accessibility (the subject of this book). Accordingly, the conclusion will mainly focus on the theoretical aspects of hedonic price modeling that are relevant to the question of which function form to choose in this study.

2.2 Hedonic Model

In regards to the theoretical foundations, the hedonic model is based on Lancaster's (1966) theory of consumer's demand. He recognized composite goods whose units are homogeneous, such that the utilities are not based on the goods themselves but instead the individual "characteristics" of goods—its composite attributes. Thus, the consumers make their purchasing decision based on the number of good characteristics as well as per unit cost of each characteristic. For example, when people choose a car, they would consider the quantity of characteristics of the car, such as fast acceleration, enhanced safety, attractive styling, and increased prestige, and so on.

Although Lancaster was the first to discuss hedonic utility, he says nothing about pricing models. Rosen (1974) was the first to present a theory of hedonic pricing. Rosen argued that an item can be valued by its characteristics; in that case, an item's total price can be considered as a sum of price of each homogeneous attributes, and each attribute has a unique implicit price in an equilibrium market. This implies that an item's price can be regressed on the characteristics to determine the way in which each characteristic uniquely contributes to the overall composite unit price.

As Rothenberg et al. (1991) describes, the hedonic approach has two significant advantages over alternative methods of measuring quality and defining commodities in housing markets. First, compressing many characteristics of housing into one dimension allows the use of a homogenous commodity assumption, and thus, the hedonic construction avoids the complications and intractability of multicommodity models. Furthermore, the hedonic approach reflects the marginal trade-offs that both supplier and demanders make among characteristics in the markets, so that differences in amounts of particular components will be given the weights implicitly prevailing in the marketplace.

2.2.1 Theoretical Basis

Housing constitutes a product class differentiated by characteristics such as number of rooms and size of lot. Freeman III (1979b) argued that the housing value can be considered a function of its characteristics, such as structure, neighborhood, and environmental characteristics. Therefore, the price function of house hi can be demonstrated as

$$P_{hi} = P_h(S_{i1}, \dots, S_{ij}, \dots, N_{i1}, \dots, N_{ik}, \dots, Q_{i1}, \dots, Q_{im}) \quad (2.1)$$

where

The S_j , N_k , and Q_m indicate the vectors of site, neighborhood, and environmental characteristics respectively.

Empirical estimation of Eq. (2.1) involves applying one of a number of statistical modeling techniques to explain the variation in sales price as a function of

property characteristics. Let X represents the full set of property characteristics (S_j , N_k and Q_m) included in the empirical model. The empirical representation of the i th housing price is as follows:

$$p_i = p(X_i, \beta, \varepsilon) \quad (2.2)$$

where

β is a vector of parameters to be estimate;

ε is a stochastic residual term;

p_i is the implicit price respected to that characteristics

Such as hedonic price models aim at estimating implicit price for each attributes of a good, and a property could be considered as a bunch of attributes or services, which are mainly divided into structural, neighborhood, accessibility attributes, etc. Individual buyers and renters, for instance, try to maximize their expected utility, which are subject to various constraints, such as their money and time.

Freeman (1979) explains that a household maximizes its utility by simultaneously moving along each marginal price schedule, where the marginal price of a household's willingness to pay for a unit of each characteristic should be equal to the marginal implicit price of that housing attribute. This clearly locates the technique within a neo-classical economics framework—a framework that analytically computes prices on the assumption that markets equilibrate under an “invisible hand” with perfect information and no transaction costs. It is noted that although the theory of hedonics has been developed with this limiting theoretical context discussed above, the technique is typically applied as an econometric empirical model and does not rely on the utility maximization underlying theory.

To understand if a household is in equilibrium, the marginal implicit price associated with the chosen housing bundle is assumed equal to the corresponding marginal willingness to pay for those attributes. To unpack this, I begin with considering how a market for heterogeneous goods can be expected to function, and what type of equilibrium we can expect to observe (Fig. 2.1).

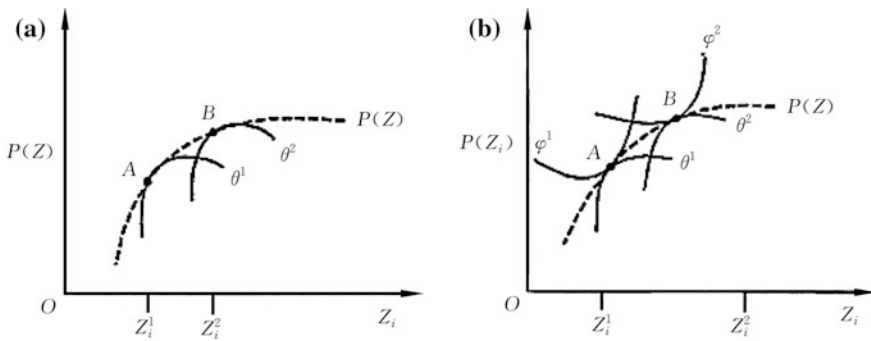


Fig. 2.1 Demand and offer curves of hedonic price function. *Source* Follain and Jimenez 1985, pp. 79

Following Follain and Jimenez's works (1985), a utility function can interpret a household decision, $u(x, z)$, where x is a composite commodity whose price is unity, and z is the vector of housing attributes. Assume that households want to maximize utility subject but with the budget constraint $y = p(z) + x$, where y is the annual household income. The partial derivative of the utility function with respect to a housing attribute is the household's marginal willingness to pay function for that attribute. A first order solution requires $\frac{u_{z_i}}{u_x} = p_i = \frac{\partial p(z)}{\partial (z_i)}$, $i = 1, 2, \dots, n$, under the usual properties of u .

An important part of the Rosen model is the bid rent function:

$$\theta(z_i, u, y, a) \quad (2.3)$$

where a is a parameter that differs from household to household.

This can be characterized as the trade-off a household is willing to make between alternative quantities of a particular attribute at a given income and utility level, while remaining indifferent to the overall composition of consumption.

$$u = u(y - \theta, z, a) \quad (2.4)$$

θ^1 pictured in the upper panel of Fig. 2.2 shows that when solving the schedule for θ . θ^1 represented by households is everywhere indifferent along θ^1 and θ schedules that are lower, which depend on its higher utility levels. It can be shown that

$$\theta_i = \frac{u_z}{u_x} \quad (2.5)$$

which is the additional expenditure a consumer's willingness to pay for another unit of z_i and be equally well off (i.e., the demand curve). Figure 2.2 denotes two such equilibria: a for household θ^1 and b for household θ^2 .

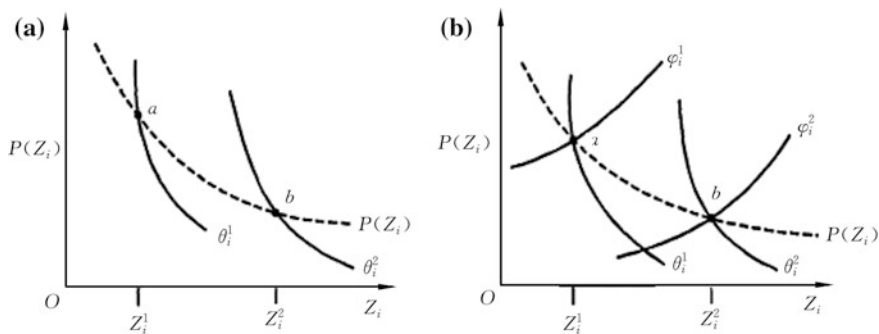


Fig. 2.2 Marginal implicit price of an attribute as a function of supply and demand. *Source* Follain and Jimenez 1985; pp. 79

The supply side could also be considered as $P(Z)$ is determined by the market. When $P(Z)$ is shown as given, and constant returns to scale are assumed, each firm's costs per unit are assumed to be convex and can be denoted as $c(z, \beta)$, where the β denotes factor price and production-function parameters. The firm then maximizes profits per unit $\pi = p(z) - c(z, \beta)$, which would yield the condition that the additional cost of providing that i th characteristics, C_i , is equal to the revenue that can be gained, so that $p_i = c_i$.

Rosen (1974) emphasized that in fact the function P_h is determined by a market in a clearing condition, where the amount of commodities offered by sellers at every point must be equal to amounts demanded by consumers choosing. Both consumers and producers base their locational and quantity decisions on maximizing behavior and equilibrium prices are determined so that buyers and sellers can be perfectly schedulers. Generally, a market-clearing price is determined by the distributions of consumer tastes as well as producer costs.

However, Rosen did not formally present a functional form for the hedonic price function. His model clearly implies a nonlinear pricing structure.

2.2.2 Hedonic Price Criticism

One of the most important assumptions to come under attack is the one relating to perfect equilibrium. For this assumption to hold, it requires perfect information and zero transaction costs (Maddison 2001). If the equilibrium condition does not hold, the implicit prices derived from hedonic analysis are biased, because there is no a priori reason to suppose that the extent of disequilibrium in any area is correlated with the levels of particular amenities contributing to the hedonic house price. The consequence of disequilibrium is likely to be in increased variance in results rather than systematic bias (Freeman III 1993). Furthermore, Bartik (1987) and Epple (1987) also pointed out that the hedonic estimation is not the result of demand–supply interaction, as in the hedonic model, the decision of an individual consumer does not affect the hedonic price function, which implies that an individual consumer's decision cannot affect the suppliers.

Follain and Jimenez (1985) argued that the marginal price derived from the hedonic function does not actually measure a particular household who is willing to pay for a unit of a certain characteristic. Rather, it is a valuation that is the result of demand and supply interactions in the entire market. Under the restrictive condition of homogeneous preferences—another limitation of the neo-classical model—the hedonic equation can reveal the underlying demand parameters for the representative household. When all households are similar with homogeneous characteristics of income and socioeconomic and supplies are different, the hedonic coefficient will be the marginal willingness to pay. Only in extreme cases, when all consumers have identical incomes and utility functions will the marginal implicit price curve be identical to the inverse demand function for an attribute. With identical incomes and utility functions, these points all fall on the same marginal willingness to pay curve

(Freeman 1979). Hence, the implicit price of an attribute is not strictly equal to the marginal willingness to pay, and hence demand for that attribute.

Another issue raised by Freeman (1979) is the speed of adjustment of the market to changing condition of supply and demand. If adjustment is not complete, observed marginal implicit price will not accurately measure household marginal willingness to pay. When the demand for an attribute is increasing, marginal implicit prices will underestimate true marginal willingness to pay. This is because marginal willingness to pay will not be translated into market transactions that affect marginal implicit price until the potential utility gains pass the threshold of transactions and moving cost.

Finally, the market for housing can be viewed as a stock-flow model where the flow is a function form, but the price at any point in time is determined only by the stock at that point in time. This raises a concern about the accuracy of the price data itself. Given that the data is based on assessments, appraisals, or self-reporting, it may not correspond to actual market price. The errors in measuring the dependent variable will tend to obscure any underlying relationship between true property value measures and environmental amenities. But the estimation of the relationship will not be biased unless the errors themselves are correlated with other variables in the model.

2.2.3 *Estimation Criticism*

The hedonic price model relies on regression technology, which is criticized by some authors for a series of econometric problems that can lead to the bias of estimation, such as function specification, spatial heterogeneity, spatial autocorrelation, housing quality change, multicollinearity, and heteroscedasticity.

2.2.3.1 *Function Specification*

Hedonic models are sensitive to the choice of functional form, as economic theory gives no clear guidelines on how to select the functional form. Rosen (1974) demonstrated that the hedonic price functional form is a reduced form equation which reflects mechanisms of both supply and demand. A further important task researchers facing is how to function the relationships of dependent variable and the explanatory variables naturally, which imposes an incorrect functional form on the regression equation, and that will lead to misspecification bias. The simple approach is the ordinary linear approach, but if the true functional form of the hedonic equation is not linear, there will occur inconsistent estimation in the resulting coefficients (Linneman 1980). Freeman (1979) specified the Box–Cox transformation, which allows choice of the proper function form based on the structure of a particular dataset. Typically, hedonic price regression models can be classified into four simple parametric functional forms.

- (a) Linear specification: both the dependent and explanatory variables enter the regression with linear form.

$$p = \beta_0 + \sum_{k=1}^K \beta_k x_k + \varepsilon \quad (2.6)$$

where

p denotes the property value;
 ε is a vector of random error term;
 β_k ($k = 1, \dots, K$) indicates the marginal change of the unit price of the k th characteristic x_k of the good

- (b) Semilog specification: in a regression function, dependent variable is log form and explanatory variable is linear, or dependent variable is linear and explanatory variable is log form.

$$\text{Lnp} = \text{Ln}\beta_0 + \sum_{k=1}^K \beta_k x_k + \varepsilon \quad (2.7)$$

where

p denotes the property value;
 ε is a vector of random error term;
 β_k ($k = 1, \dots, K$) indicates the rate at which the price increases at a certain level, given the characteristics x

- (c) Log-log specification: in a regression function, both the dependent and explanatory variables are in their log form.

$$\text{Lnp} = \text{Ln}\beta_0 + \sum_{k=1}^K \beta_k \text{Ln}x_k + \varepsilon \quad (2.8)$$

where

p denotes the property value;
 ε is a vector of random error term;
 β_k ($k = 1, \dots, K$) indicates how many percent the price p increases at a certain level, if the k th characteristic x_k changes by 1 %

- (d) Box–Cox transform: determine the specific transformation from the data itself then enter the regression in individual transformed form.

$$p(\theta) = \beta_0 + \sum_{k=1}^K \beta_k x_k^{(\lambda_k)} + \varepsilon \quad (2.9)$$

where

$$\begin{aligned} p^{(\theta)} &= \frac{p^{(\theta)} - 1}{\theta}, & \theta &\neq 0 \\ &= \ln p, & \theta &= 0 \\ x^{(\lambda_k)} &= \frac{x^{(\lambda_k)}}{\lambda_k}, & \lambda_k &\neq 0 \\ &= \ln x_k, & \lambda_k &= 0 \end{aligned}$$

From the Box–Cox transform equation, we can see if the θ and λ_k are equal to 1, the model will transform to the basic linear form. If the θ and λ_k are equal to 0, the model will transform to the log-linear form. If the value θ is equal to 0 and λ_k are equal to 1, then the model can be the semilog form.

2.2.3.2 Debate About the Hedonic Function

Unfortunately, economic theory provides little guidance, and there is no specific function form for the hedonic price models suggested by Rosen (1974), Freeman (1979), Halvorsen and Pollakowski (1981), and Cassel and Mendelsohn (1985), so it is reasonable to try several functional forms to find the best performance. Among the four types of function forms in hedonic literatures, the semilogarithmic form is much more prevalent, as it is easy to interpret its coefficients as the proportionate change in price arising from a unit change in the value of the characteristic. Furthermore, unlike log–log models, the semilog model can deal with dummy variables for characteristics that are either present or absent (0 or 1). Diewert (2003) argued that the errors from a semilog hedonic function are homoskedastic (have a constant variance).

Although more and more researchers prefer to use the Box–Cox transformation function, letting the dataset drive the function form, Cassel and Mendelsohn (1985) pointed out four inconsistencies of the Box–Cox transformation. Firstly, the large number of coefficients estimated with Box–Cox reduces the accuracy of every single coefficient, which could lead to poorer estimates of price. Secondly, the traditional Box–Cox functional form is not suited to any dataset containing negative numbers. Furthermore, the Box–Cox function may be invalid for prediction, as the mean predicted value of the untransformed dependent variable need not equal to the mean of the sample upon which is estimated. The predicted untransformed variables

will be biased, and the predicted untransformed variables may also be imaginary. Fourthly, the nonlinear transformation results in complex estimate of slopes and elasticities, which are often too cumbersome to use properly.

Taking the least errors as the choice criterion, Cropper et al. (1988) compared six function forms: linear, semilog, double-log, Box–Cox linear, quadratic, and quadratic Box–Cox, testing the best goodness of fit using data for Baltimore. He found that no function produced the lowest β_i for all the attributes, though the quadratic Box–Cox function had the lowest normalized errors. However, the linear Box–Cox function had the lowest error variance, and based on the criterion, the linear Box–Cox performed the best and quadratic and double-log functions the worst. On the other hand, when variables are replaced or omitted, the Box–Cox linear function was the best of the six.

Having said all this, Halvorsen and Pollakowski (1981) rightfully pointed out that the true hedonic function form is unknown: we can only estimate it in some particular dataset, though as I have shown, we do have methods to help choose the most appropriate parametric hedonic function form.

2.2.3.3 Housing Submarkets

Housing property cannot be regarded as a homogeneous commodity. A unitary metropolitan housing market is unlikely even to exist. Instead, it is likely to be composed of interrelated submarkets (Adair et al. 1996; Tu 1997; Goodman and Thibodeau 1998; Whitehead 1999; Watkins 2001). Straszheim (1974) suggested that the housing market is a series of single markets, which requires different hedonic functions. According to Schnare and Stuyk (1976), housing submarkets arise when competition in a housing market is insufficient to ensure spatial equalization of physical housing attributes. Thus, the submarkets existence is the results of inelasticity (or high inelasticity) of demand and supply of housing at least in a short term. Basu and Thibodeau (1998) define housing submarket as follows:

Housing submarket are typically defined as geographic area where the prices per unit of housing quantity (defined using some index of housing characteristic) are constant. (Basu and Thibodeau 1998).

Goodman and Thibodeau (1998) argued that the existence of submarket questions the validity of the traditional assumption that urban housing markets can be modeled on the basis of a single market-wide house price equation. Adair et al. (1996) also argued that the failure to accommodate the existence of housing submarket will introduce bias and error into regression-based property valuation. Orford (2000) demonstrated that submarkets could be considered as relatively homogeneous subgroups of the metropolitan housing market, people's preference for each housing attribute may vary in different submarkets while remain the same within each submarket. However, the theory assumes that the implicit price for per unit of each housing attribute is stationary over space, and this assumption ignores that different geographical demand and supply characterized by different classes of

people can lead to the spatial disequilibrium of housing market in a metropolitan area. Thus, parameters estimated by a simple hedonic function for the whole market sometimes seem misleading.

Goodman and Thibodeau (2007) emphasized that housing submarkets are important in house price modeling for several reasons. Firstly, the assigning of properties to housing submarkets is likely to increase the accuracy of the prediction of the statistical models, which are used to estimate house prices. Secondly, identifying housing submarket boundaries within metropolitan areas will increase the chance of researchers deriving better spatial and temporal variations in their models of prices. Thirdly, the accurate allocation of properties to submarkets will improve the abilities of lenders and investors to price the risk related to the financing of homeownership. Finally, the provision of submarket boundary information to housing consumers will decrease their search costs.

In terms of the specification of housing submarkets, Goodman and Thibodeau (1998) stated that a metropolitan housing market might be segmented into groups of submarkets according to the factor of demand and/or supply. Watkins (2001) also suggested that housing submarkets exist as dwelling can generate different price due to the interaction between segmented demand characterized by consumer groups, and segmented supply characterized by product groups. As such, housing submarkets may be defined by dwelling type (e.g., town house, flat, and detached house); by structural characteristics (numbers of bedroom and building style); and by neighborhood characteristics (e.g., school quality). Alternatively, housing markets may be segmented by age, income, and race of households (Schnare and Struyk 1976; Gabriel and Wolch 1984a; Munro 1986; Allen et al. 1995). In that case, higher income households tend to pay more for housing (per unit of housing services) and the attributes of other home-owners—to protect the homogeneity of their neighborhood, life chances of children, and so on. Finally, racial discrimination may produce separate housing submarkets for majority and minority households (King and Mieszkowski 1973). Several empirical studies of submarkets have found that spatial characteristics are more important than structural characteristics. Ball and Kirwan (1977) found housing affordability and the availability of mortgage finance are important shapers of sub markets, despite spatial constraints. Historical characteristics can also contribute to housing market segmentation. More recently, scholars have been more aware of the importance of both spatial and structural factors as the specification criterions of housing submarket (Adair et al. 1996; MacLennan and Tu 1996).

Although many researchers agree on a submarket definition based on structural and locational features, there is little consensus as to how a submarket should be identified in practice. The most welcomed procedure for testing submarket existence was introduced by Schnare and Struyk (1976) and has been employed subsequently (for example, Dale-Johnson 1982; Munro 1986). The test procedure involves three stages.

First, the functions of hedonic house price are estimated for each potential market segment in order to compare the submarket price for a *standard* dwelling. Secondly, a chow test is computed in order to show whether there are significant

differences between the submarket specific prices. Thirdly, a weighted standard error is calculated for the submarket model, which acts as a further *common-sense* test of the significance of price differences for standard dwellings in different submarkets. This procedure also enables us to compare the effects on the accuracy of the house price models when different submarket definitions and stratification schemes are being compared.

Bourassa et al. (1999) stressed the need to test whether boundaries of submarkets are stable over time or not. Adding a dynamic part to the analysis makes it even more difficult to specify submarket models since markets are constantly changing.

2.2.3.4 Spatial Autocorrelation

A further discussion in terms of the application of hedonic price modeling is spatial dependency, also known as spatial autocorrelation. One of the basic assumptions underlying the regression model is that observations should be independent of each other. However, from the first law of geography, attributed to Tobler (1970), “everything is related to everything else, but near things are more related than distant things,” the independence of observations assumption is clearly a fallacy. Spatial autocorrelation is concerned with the degree to which objects or activities at some place in the earth’s surface are similar to other objects or activities located nearby (Goodchild 1986). This is important in the sense that it is a special feature of spatial data (Can 1990); for example, houses that are close in geographic space are likely to have similar attributes. Generally, if the spatial effect is ignored, it is more likely that the real variance of the data is underestimated and thus leads to bias of the results (Ward and Gleditsch 2008). According to the works of Dunse and Jones (1998), Bowen et al. (2001), Gillen et al. (2001), and Orford (1999), there are at least three sources of spatial autocorrelation, including property characteristics, the evaluation process, and misspecification in the OLS model.

Firstly, spatial dependency exists because nearby properties have similar property characteristics, in particular structural features, as the properties are developed at the same time and also share the same locational conditions (Gillen et al. 2001; Bourassa et al. 2005). Secondly, spatial autocorrelation also arises from the valuation process, as the transaction price agreed between buyers and sellers will affect the price of the surrounding area (Bowen et al. 2001), especially where valuers use the comparison method, which is very common in the residential real estate industry. Thirdly, misspecification of a model can result in spatial autocorrelation (Orford 1999), when the model is missing important variables, has unimportant extra variables, and/or an unsuitable functional form. Anselin (1988) also stated that spatial autocorrelation is associated with spatial aggregation, the presence of uncontrolled-for non-linear relationships, and the omission of relevant variables.

Generally, spatial autocorrelation analysis is applied for testing whether the observed value of a variable is independent of the values of the variable with neighbors. The function of a spatial autocorrelation index is to measure the degree of interdependence among variables, the strength, and nature of that

interdependence. It may be categorized as positive and negative, respectively. Positive autocorrelation occurs when high or low values of the random variable tend to cluster in space, whereas negative autocorrelation takes place when locations tend to be surrounded by neighbors with very dissimilar characteristics. Commonly, Moran's I test measures spatial dependency in the residuals of a regression model, and it checks the similarities among the housing price and attribute data in relation to the spatial relationships (Bowen et al. 2001). If there are N observations on a variable x at locations i, j , then the formula for Moran's I is:

$$I = \frac{(N/S_0) \sum_i \sum_j w_{ij} (x_i - \mu)(x_j - \mu)}{\sum_i (x_i - \mu)^2} \quad (2.10)$$

where

μ is the mean of the x variable;

w_{ij} are the elements of the spatial weights matrix;

S_0 is the sum of the elements of the weights matrix

The range of Moran's I is from -1 to $+1$, and an expected value (zero) shows absence of autocorrelation in samples. Moran's I compares the relation between the deviations from the mean across all neighbors of I , adjusted for the variation in y and the number of neighbors for each observation. Higher value of Moran's I indicates stronger positive clustering, which means the values from neighboring units are similar to one another.

2.2.3.5 Housing Quality Change

Although the hedonic model has been used for housing market analysis for more than 30 years, and most studies rely on one-shot studies of one place at one point in time (Richardson et al. 1990), the modeling approach has been criticized for its instability of the coefficients over time (Case and Quigley 1991; Quigley 1995; Case et al. 2006). According to the Dhrymes (1971), most studies consider the time variables as essential explanatory variables, which implies that they are a proxy for change in quality over time. Griliches (1996) constructed a hedonic price index for automobiles in an attempt to measure the change in quality over time, and he noticed the coefficient of the characteristics were unstable and changed over time. McMillen (2003) took an alternative repeat-sales model to identify changes in house price distance gradients in Chicago. He employed the transaction data of Chicago from January 1, 1983 to December 31, 1998, and he found that the distance from the CBD did not affect house prices in the City of Chicago significantly in the early to mid-1980s. However, the situation reversed in the 1990s as a significant CBD house price gradient was reestablished, and by the end of the 1990s, house values fell by more than 8 % with each additional mile of distance from the CBD. He explained that this change in gradients was caused by a rapid

appreciation of values near the city center as new housing was built, which increased demand for housing near Chicago's center.

Similarly, Hulten (2005) criticized the hedonic price method as failing to capture dynamics since price inflation and quality both change over time. Inflation could lead to an upward shift in the hedonic function because some or all factors become more expensive, whereas quality change can be caused by changes in compositions of housing attributes and by product innovation. Changes in compositions of varieties can occur with changes in people's income, tastes, demographics, and environmental preferences. For example, opening a new train station will affect people's choice of location in respect to living near job centers. On the other hand, product innovation occurs when the cost of acquiring a number of characteristics is reduced. For example, in the past whether the property has a fireplace can affect the property value, but nowadays the fireplace is less significant as the construction cost for installing a fireplace is lower.

Some scholars attempt to explore the relationship between property value and certain attributes in temporal, such as the question of how soon the housing price will increase due to the kind of attribute adding the value. For example, Gatzlaff and Smith (1993) examined the impact of the development of the rail system on the residential property values in Miami. Using hedonic model, they found that the rail developmental announcement's impact on residential property value is weak. However, comparing a house price index for properties located near rail stations with the housing price index for the Miami MSA indicates a weak relative increase in house value closing to the station. Based on the results, they concluded that the Metro rail system has little effect on accessibility improvement. Noonan (2007) examined the effect of historic landmarks on property value in Chicago through a hybrid hedonic model and repeat-sales model. The results suggested that housing near landmark buildings sold at a small premium during the 1990s.

2.2.3.6 Multicollinearity

Another issue that researchers often encounter when they attempt to estimate the hedonic function is multicollinearity (Lake et al. 2000; Orford 2002). Multicollinearity is a statistical phenomenon when two above exploratory variables in a multiple regression model are highly correlated. For example, it is well known that both traffic noise and air pollution have a negative impact on housing properties. However, traffic noise and air pollution are also highly correlated, as high level traffic flows could result in poor air quality. In this case, regression analysis finds it difficult to tease apart the separate influence on property price; consequently, the estimation for each parameter is no longer reliable. There is no easy solution to the problem of multicollinearity and the parameters estimated maybe implausibly large or have the wrong sign (e.g., the opposite relationship). Sometimes, it is

possible to overcome multicollinearity by measuring the variables more accurately, or applying principal components analysis to combine the highly correlated variables into one index.

However, multicollinearity does not reduce the predictive power and reliability of whole model, at least within the sample size. It only affects calculation regarding individual predictors. There are several methods to detect multicollinearity such as VIF, Condition Number test, and Farrar–Glauber test. The studies in this book use the variance inflation factor (VIF) to measure the multicollinearity, and its formula is as following:

$$\text{VIF} = \frac{1}{1 - R_j^2} \quad (2.11)$$

where R_j^2 is the coefficient of determination of a regression of explanatory j on all the other explicators. If VIF is more than 10, then it indicates there is a serious multicollinearity problem.

2.2.3.7 Heteroscedasticity

Heteroscedasticity is a phenomenon where the variance of the disturbance or error term of the hedonic model is unequal (Fletcher et al. 2000) or changes across the sample (Hendry 1995). One assumption of OLS is that the variance of the error should not be correlated with the dependent variables. Although it can happen in both time series and cross-sectional data, it is more commonly present in studies employing cross-sectional data (Nghiep and Al 2001). According to Gujarati (2003), heteroscedasticity (constant variance) does not cause ordinary least squares coefficient estimates to be biased. The estimates are consistent but inefficient, because their variances are no longer minimized, even if the sample size is increased. However, it could lead to unreliable confidence intervals and t and F tests.

There are two main sources of heteroscedasticity: statistical methods of modeling and economic behavior of the sample (Hendry 1995). A common statistical source of error of measurement arises in the situation where data for a high level of geographical aggregation, such as census area data, is used to represent individual properties. An example of a behavioral source is the effect of different bargaining power of housing consumers. This can induce heteroscedasticity as the more the consumers bargain, the lower the price possibly be (Harding et al. 2003).

Thus, heteroscedastic data could bias the estimations for the relationship between the predictor variable and the outcome. However, heteroscedasticity can be detected by series of diagnostic tests, such as Park test, Glejser test, White test,

Breusch–Pagan test, Goldfeld–Quandt test, and Cook–Weisberg test. In this study, the White test would be applied to detect whether there is heteroscedasticity in the error terms;

$$Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + u_i \quad (2.12)$$

$$u_i^2 = \alpha_1 + \alpha_2 X_{i2} + \alpha_3 X_{i3} + \alpha_4 X_{i2}^2 + \alpha_5 X_{i2}^2 + \alpha_6 X_{i2} X_{i3} \quad (2.13)$$

Step 1: use OLS procedure, obtain $\hat{\beta}_1, \hat{\beta}_2$ and $\hat{\beta}_3$.

Step 2: Square the residual $\hat{u}_i^2 = \left(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_{i2} - \hat{\beta}_3 X_{i3} \right)^2$

Step 3: Regress the squared residual $\hat{u}_i^2$ against a constant $X_{i2}, X_{i3}, X_{i2}^2, X_{i2}^2$ and $X_{i2} X_{i3}$,

Step 4: Compute the statistics nR^2 where the n is the size of the sample and R^2 is the unadjusted R-square from the step 3

Step 5: Reject the null hypothesis that $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 = 0$ if $nR^2 > X_5^2(\alpha)$, the upper 5 % point on the chi-square distribution with 5 d.f.

2.3 Housing Attributes

The basic hypothesis of hedonic housing models is that housing price can be considered as willingness to pay for a bundle of characteristics. Empirical studies have generally grouped determining variables into four subsets:

- (a) Structural or internal attributes describing the physical characteristics of housing (e.g., numbers of bedroom, swimming pool, and garage).
- (b) Locational attributes including the distance to major places of employment, to major amenities (e.g., shopping mall and public facilities), and to road infrastructure and transport access points (e.g., train station, subway station, major streets, highways, and airports).
- (c) Neighborhood attributes depicting the quality of the economic and social characteristics of the neighborhood (e.g., income status and racial composition).
- (d) Environmental attributes describing environmental quality and environmental amenities, such as air pollution, water pollution, noise, aesthetic views, and proximity to recreational sites or public service.

These are discussed in the following Table 2.1.

Table 2.1 Selected previous studies on hedonic price model

Type of housing attributes	Characteristic	References	Sample size	Other variables used	Impact on housing price
Structure characteristic	Square footage	Sirmans et al. (2006)	58	The geographical location, time trend, real median household income, and the data source	+
	Lot size		41		+
	Age		71		—
	Bedrooms		33		+
	Bathrooms		50		+
	Garage		29		+
	Swimming pool		34		+
	Fireplace		29		+
	Air conditioning		28		+
	Basic residential quality	Kain and Quigley (1970)	579/275	Six factors derived from 39 variable	+
	Dwelling unit quality				+
	Age of structure				—
	Number of rooms				+
	Number of bathrooms				+
	Parcel size				+
Locational characteristic	Distance to ocean	Richardson et al. (1990)	9078/10,928	19 variables (e.g., number of bathrooms, condition, living area, age, month of sale, and income)	—
	Distance to CBD				—

(continued)

Table 2.1 (continued)

Type of housing attributes	Characteristic	References	Sample size	Other variables used	Impact on housing price
	Distance to Santa Monica				—
	Distance to South Bay				+
	Distance to CBD	Heikkila et al. (1989)	10,928	19 variables (e.g., number of bathrooms, condition, living area, age, month of sale, and income)	+
	Labor market accessibility	Osland and Thorsen (2008)	2788	11 variables (e.g., age, living area, lot size, garage, number of toilets, and rebuild)	+
Neighborhood characteristic	Crime	Dubin and Goodman (1982)	589/1178	50 variables (e.g., criminal homicide, rape, pupil-to-staff ratio, and average teacher experience)	—
	Education				+
	Race				—
	Primary school performance	Gibbons and Machin (2003)	2900/2998/1544	7 variables (e.g., dwelling type, mean rooms, and portion in social housing)	+
Environmental characteristic	Distance to hazards	Brasington and Hite (2005)	5051	25 variable (e.g., house size, age, porch, school quality, portion of white, and portion of graduate degree)	—

(continued)

Table 2.1 (continued)

Type of housing attributes	Characteristic	References	Sample size	Other variables used	Impact on housing price
	Distance to parks	Poudyal et al. (2009)	11,125	27 variables (e.g., brick exterior, stories, population density college degree, and park size)	—
Other characteristic	Cemetery view	Tse and Love (2000)	1,550,000	8 variables (e.g., area, age, car park, shopping center, and estate type)	—

2.3.1 *Structural Characteristics*

Structural attributes describe the physical structure of property goods and land parcel. Compared with locational attributes, the structural attributes are easier to account for and accurately perceived. Follain and Jimenez (1985) summarized the most welcomed structural attributes from previous researches and noted that measures of living space have been reduced to lot size, floor area, and the number of rooms while structural quality is measured by age, style, and interior and exterior quality scores.

Sirmans et al. (2005) summarized the top twenty characteristics that have been used to specify hedonic pricing equations. He described the total number of times a characteristic has been used and the number of times its estimated coefficient has been positive, negative, or not significant. Age shows up the most frequently in hedonic models and typically has the expected negative sign though it is seen to be positive or not significant in some studies. The age effect will depend on the period studied and the age of the city. In historic cities, age may have a positive influence on price, but only in particular housing markets. Age in a modern part of a historic city may have a negative influence since the quality of modern era houses is typically inferior to those in historic quarters. Square footage is the next most used characteristics and typically has the expected positive effect in selling price. Other characteristics that appear frequently are garage, fireplace, and lot size. Each typically has the expected positive effect. Garage never has a negative sign, but it has been insignificant in a number of studies. Fireplace shows negative in only a few studies and lot size never shows up negative. Other characteristics that show up frequently are number of bedrooms, bathrooms, swimming pool, and basement.

Bedrooms show up negative in some studies but bathrooms almost never do. Research shows that a swimming pool never has a negative impact on selling price though it has been insignificant in certain studies, possibly because of the liability of maintenance. Basement is usually positive but sometimes negative or even insignificant in some studies, possibly related to endemic dampness problems in some property markets.

In addition, Morris et al. (1972), in their pilot study in San Juan, Puerto Rico, examined structural quality by using the dimension of availability of plumbing facilities and other service facilities such as cooking equipment, refrigeration, and lighting. They differentiated plumbing facilities into “inside, for exclusive use,” “inside, shared,” and “other.” These measures, which reflected the quality of the dwelling without associating them with the locational or neighborhood attributes, were found to be effective proxies for measuring quality features. Kain and Quigley (1970) investigated the impact of housing quality on housing prices. They used measures such as condition of exterior structure, walls, condition of floors, drives and walks, windows, and levels of housekeeping. These quality features (e.g., number of bathrooms, the number of rooms, and lot size) were found to have as much effect on the price of housing.

Sirmans et al. (2006) examined the effect of nine housing characteristics on housing price that appeared in most hedonic pricing models, including square footage, bedrooms, lot size, age, bathrooms, fireplace, swimming pool, garage, and air condition. They found that the coefficient of the square footage, lot size, age, bathroom, swimming pool, and air condition are sensitive to some geographical locations and to the number of variables in the hedonic model but not to time, household income, or to source of data. However, garage, fireplace, and bedrooms coefficients are not affected by some geographical location, time, income, or the type of data. In contrast, Kohlhase (1991) found that the significance of structural attributes can change over time and may vary between nations. While attributes relating to the number of rooms and floor area are relatively important across nations, other attributes change with the tradition of building style or the climate.

Theoretically, a property’s structural attributes and its location within the city are related, since they reflect the growth of the urban structure (Muth 1969), which implies that an element of location will be inherent within the physical structure of the property. This is indicated in studies such as Cubbin (1970) and Kain and Quigley (1970), which revealed a high degree of multicollinearity between structural attributes and the results suffered from spatial autocorrelation.

2.3.2 Locational Characteristics

A property represents not only an amount of structural characteristics, but also a set of specific locational characteristics, which has long been regarded as the fundamental influence in the modeling of residential location. Von Thunen’s classical

land use model was the first to formally correlate value with systematic locational characteristics—distance to a central marketplace.

The traditional neo-classical micro-economic theory (Alonso 1964; Muth 1969; Mills 1972) developed Von Thunen's model (and the underlying Ricardian value model) into emphasisean "access-space" trade-off model that describes transportation costs as a trade-off against land rents. The trade-off model was developed under the assumption of a monocentric city on an isotropic plane with a housing market in perfect competition. The key idea behind the monocentric model is several restrictive assumptions such as spatially centralized workplaces, which makes the CBD the major determinant of location-specific land values and site rents. Beckmann (1973) developed models of urban housing markets based on the central assumption that housing and employment accessibility were jointly purchased in the residential choice decision. Most early economic studies of housing price found there is a downward sloping housing price curve with distance from the central business district (CBD). However, the monocentric model has inherent limitations and has increasingly been criticized by researchers (e.g., Boarnet 1994). The criticisms can be summarized into three types. Firstly, many authors have noted that employment is not concentrated in a central business district (e.g., McDonald 1987; Heikkila et al. 1989). Secondly, it has been questioned whether urban commuters engage in "wasteful commuting" in their journey to work, which could also be interpreted as the question of whether the distribution of jobs and residence is the primary determinant of the journey to work (Hamilton and Röell 1982; Hamilton 1989). For example, if persons do not choose their residential location to minimize their commute to work, then there are non-transportation factors, which are also influential in residential location (White 1988b; Small and Song 1992). Thirdly, some researchers have questioned the monocentric assumption of exogenous employment location (Steinnes 1977, 1982). The monocentric idea that the residential location is endogenous to employment location, but whether the employment location is largely exogenous to residential location is questionable. If employment location is endogenous to residential location, the partial equilibrium approach of most monocentric models is inappropriate.

Many scholars conclude that workplace accessibility has been over emphasized in the urban economics empirical literature. For example, as stated by Heikkila et al. (1989):

with multiple-worker households, multiple workplaces are common; given a high degree of residential mobility, sites offering accessibility to many employment nodes are more valuable because it is not very likely that successive owners will work in the same workplace.

Richardson et al. (1990) found a significantly negative value of the coefficient related to distance from the LA CBD in 1970, and this variable was found not to influence house prices in 1980. McMillen (2003) found that in many cities the CBD no longer appears to exert a significant influence on house value. Take Chicago as an example, it was long viewed as a monocentric city but one in which the centrality has declined steadily in importance over time.

In fact, cities rarely have a simple monocentric structure, and the monocentric city is a special case of the standard urban model (Bender and Hwang 1985). Employment and amenity centers are often located outside of the city center, which may cause the house price gradient to be complex (Orford 1999). Button and Taylor (2000) noted in the 1990s, when metropolitan areas were in a state of reformulation that the CBD is no longer the only place one may find gainful employment, as many suburban employment centers have arisen to combat its draw. On the other hand, there is little consensus as to the appropriate method for identifying the subemployment center, such as Giuliano and Small's clustering methods (1991) and McMillen's nonparametric methods (2001). Clustering methods rely on ad hoc definitions of density and total employment cutoffs and parametric models which make strong assumptions regarding parametric form, leading to misspecification (Redfearn 2007). Empirical research on the nature of property price with a polycentric urban context has been scarce. For example, in the case of Baltimore, Dubin and Sung (1987) concluded that "the CBD appears to behave like the other secondary center: it has an impact, but this effect is limited to a relatively small area." Similar results were also found by Jackson (1979) in Milwaukee and Bender and Hwang (1985) in Chicago.

2.3.2.1 Accessibility

Accessibility has been discussed in geographic contexts from numbers of perspective (Kwan 1998). There is little consensus on the precise definition of accessibility. Stewart described accessibility as the population—over-distance relationship or "population potential," while Hansen (1959) defined accessibility as: "the potential of opportunities for interaction." Accessibility is a measurement of "the spatial distribution of activities about a point, adjusted for the ability and the desire of people or firms to overcome spatial separation" (Hansen 1959). Accessibility also can be defined as "the ability of individuals to travel and to participate in activities at different locations in an environment" (Des Rosiers et al. 1999). According to Des Rosiers et al. (2000), accessibility relates to the ability of the individual to travel and to participate in activities at different locations. In a transportation model, accessibility is defined as "...the distribution of some defined activity measure versus the travel impedance (time, cost and distance) to reach that activity from the selected zones..." (Adair et al. 2000).

Ball (1973) found that while most studies showed distance variables are significant, not all agree on the measure of distance. Recently, more sophisticated measures of accessibility have been proved to perform better than purely Euclidean distance in many studies on property value (Niederhorn and Ammari 1987; Hoch and Waddell 1993). Heikkilä et al. (1989) suggested considering the possibility that accessibility to nodes other than the CBD might be important. Bowes and Ihlanfeldt (2001) proposed that railway stations raise the value of nearby properties, as that reduce people's commuting costs, and station area should therefore be better able to attract retail activity. Various researchers have explored the relationship between

specific measures of accessibility and property value, and as showed in Fig. 2.3, accessibility has been measured at aggregate level and individual level. At individual level, accessibility could account for the distance or time cost from a location to certain facilities, such as school, transit station, employment center, and shopping center (e.g., Landau et al. 1981; Henneberry 1998).

Debrezion et al. (2006) explored the impact of the railway network on house price in the Netherlands by hedonic price approach. The railway access variables the authors used include the distance to railway station, the frequency of train services, and the distance to railway tracks. They found that housing in close proximity to a railway station commands a market price that is about 25 % higher than equivalent housing at a distance of 15 km or more from a station. Munoz-Raskin (2010) paid attention on the relationship of bus rapid transit and property values within walking area. He found housing marketplaces value premiums on properties in immediate walking proximity of BRT feeder lines.

In contrast, Andersson et al. (2010) examined the effect of High-speed rail station (a long-distance rail) accessibility on real estate price in Taiwan. The estimated results show that HSR accessibility has at most minor effects on house price. Rolon (2005) also found a new station does not bring substantial accessibility improvement and the marginal effect on land and property values is negligible. However, there was a negative impact of proximity to a transit station due to noise, vibration, pollution, visual impacts, and safety issues (Bowes and Ihlanfeldt 2001).

However, at the aggregate level, accessibility can be considered as a point of attractiveness or proximity to an opportunity. For example, Hwang and Thill (2010) examined the impact of job accessibility on housing price in the Buffalo and Seattle metropolitan areas. They compute a travel-time-based job-accessibility measure at the employment level of census tracts, according to Hansen (1959)'s formulation. The results suggest that suburbanites are more willing to pay for additional increase in job accessibility in housing consumption than urban residents in the Buffalo-Niagara Falls MSA, whereas the situation is opposite in Seattle.

Song and Sohn (2007) also criticized accessibility measures based on distance from a housing unit to the CBD, regional and commercial center, arguing that it

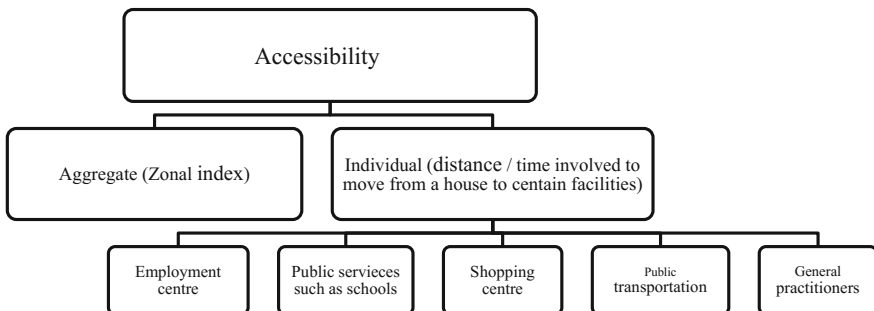


Fig. 2.3 Accessibility measurement types

cannot capture the overall level of accessibility for retail service. They applied a spatial accessibility index to evaluate the effect of enhanced access to retailing in the single family housing market in the city of Hillsboro, Oregon. This spatial accessibility index considered the numbers of neighborhood retail store, the size of retail store, as well as the distance to the retail store by units of census block (e.g., Weibull 1976; Shen 1998). The results showed that spatial accessibility to retailing as a service is capitalized into residential price.

Adair et al. (2000) focused on the relationship between accessibility and housing price in the Belfast urban area. Instead of traditional studies using the CBD as a reference for accessibility indicator, they calculated the accessibility index by considering all trip attractors and generators in the area of 182 traffic zones with sample size of 2648. For the whole housing market, they found accessibility is of little significance in explaining variation of house prices. However, the authors found that accessibility can be an important factor at submarket level, particular in low-income areas. Johnson and Ragas (1987) found that the biggest obstacle to finding declining rent gradients within a small area is that there are many other confounding factors that affect the land rent other than just spatial locational characteristics.

Another issue emphasized by Song and Sohn (2007) is the inaccuracy of the accessibility measure due to spatial aggregation. Some spatial information is lost and become insignificant as “households in each zone or area are typically represented by a single point (i.e., centroid or weighted center) in calculating distance to and from the zone.” Furthermore, they argued that the arbitrariness of spatial units could distort the real accessibility level of individual household. Cheshire and Sheppard (1997) also argued that much of the data used in hedonic analyses still lacks land and location information. Different accessibility indices can interpret the different locational information but not all the indices have proved to affect housing value. Besides that, these aggregate empirical studies generally have not found accessibility to be as major a factor in controlling residential mobility as traditional trade-off models imply (Adair et al. 2000).

2.3.3 Neighborhood

Neighborhood attributes are also typically included in the estimation of housing price models. Important among these are income level (which is a surrogate for other things, neighbor externalities, the “snob” factor, quality of housing, level of expenditure on housing investment and maintenance, school quality, and so on). Generally, higher income neighborhoods are assumed to be of higher quality (e.g., higher quality education and lower criminal rate). This leads to the idea that all households prefer to live in higher income neighborhoods. Set against this is the observation and theoretically plausible idea that households prefer to live in neighborhoods dominated by households similar to themselves (Gans 1963). Hedonic models of housing values that use a set of social and economic status

variables (in terms of age, income, ethnicity, and lifestyle), controlling for neighborhood and amenity quality (e.g., school) can help tease out these effects.

2.3.3.1 School Quality

There is a wide recognition that school quality is the most essential determinant of housing price, in particular within the US and the UK contexts, as it is close related with local property tax bands. Generally, in the field of education, some authors choose an indicator to control school quality, such as pupil–teacher ratio and standardized test scores (e.g., Oates 1969; Haurin and Brasington 1996).

Dubin and Goodman (1982) estimated the impact of school characteristics and crime measure on 1765 house prices in Baltimore in 1978. They measured school characteristics by pupil-to-staff ratio, average teacher experience, the percent of staff with master degrees or above, and a battery of third and fifth grade test scores. As the school variables were highly correlated, they used principle component analysis to reduce the data. They confirmed the school characteristics had a significant effect on house price, but it is still difficult to determine which school characteristics contribute more effect.

Hefner (1998) examined the impact of school characteristics on house price by conducting two measurements for school quality. The first measurement focused on management, including teacher’s salary, teacher/pupil ratios, teacher tenure, and percentage of teachers with advanced education degrees. The second measurement is considered by support for and participation in gifted and talented programs. He found that administrative and leadership choices made by school and parents can increase the prices of surrounding properties.

Gibbons and Machin (2003) investigated the impact of primary school performance on housing prices in England with a pooling dataset of 7444 “postcode area” during the years from 1996 to 1999. They considered school type as the instrument for school quality and found a positive effect of local school quality on house price. The results show that 1 % increase in the proportion of children meeting raises property values by 0.67 %.

However, the question of how change in school characteristics relate to changes in real estate value remains open in part because, as Mieszkowski and Zodrow (1989) noted, many existing efforts to determine the extent of capitalization have been flawed due to inadequate data.

2.3.3.2 Social Economics Status and Ethnicity

The socioeconomic characteristics of the neighborhood, such as the social status and population characteristics (in terms of age, income, ethnicity, and lifestyle) of a neighborhood, also play a role in the choice behavior of house buyer, and therefore have an effect on house price (Visser et al. 2008). Dubin and Sung (1990) showed that the socioeconomic status and racial composition of the neighborhood affect

housing price more than the quality of public services. Racial segregation behavior studies in some USA cities (Harris 1999) may influence housing price, depending on a community's willingness to pay to keep its identity.

Baumont and Legros (2009) examined the impact of neighborhood on the housing value in the metropolitan area of Paris. Social capital, social status, social externalities, and urban renewal policies all have impacts on housing prices, positively or negatively.

Schafer (1979) looked at Boston 1970 census data. When the data are divided into submarkets defined as the central city ghetto, the central city transition area, the central city white area, and suburban white area, he reports that price differentials paid by the blacks vary greatly, depending on the submarket. He found that house prices are higher in the ghetto and transition area, relative to the white area.

Dougherty et al. (2009) measured the effect of both elementary school test scores and racial composition on household's purchase choice over a 10-year period. Overall, while both test scores and race help explaining the variation in housing price significantly, they found that the influence of school performance declined, and racial composition became nearly seven times more influential during the study period.

Differences in house prices across racial groups have been carefully analyzed in the past. As Yinger (1979) and Chambers (1992) pointed out, it is crucial to include indicators for the household, the neighborhood, and the submarket, while controlling for the characteristics of the house when testing for price differentials. If relevant characteristics are excluded, the estimated coefficients will be biased.

2.3.4 *Environmental*

Since Ridker and Henning's (1967) pioneering study, there has been growing interest in using property value as a source of information on the benefits to be expected from controlling environment disamenities. To be more specific, property price models have become one of the common ways of valuing environmental externalities. The method is often applied to variations in housing prices which reflect the value of local environmental indicators, and various empirical studies have used a single environment indicator in a hedonic price model (Anderson and Crocker 1971; Wilman 1981; Murdoch and Thayer 1988). Generally, environmental characteristics can be subdivided into two categories: environmental quality and environmental amenities. Environmental quality includes air pollution, water pollution, and noise, while environmental amenities can be interpreted as aesthetic views and proximity to recreational sites.

2.3.4.1 Environmental Quality

Some studies have examined the impact of environmental quality of air, water, and traffic on house price by hedonic models. Generally, these studies show that urban disamenities have a negative effect on house price, which means people have a low demand and do not have the willingness to pay more for these characteristics—on the contrary, they are willing to pay less for them. Day et al. (2007) and Bateman et al. (2001) considered aircraft noise, while Schipper et al. (1998) and Nelson (1982) assessed the impact of multiple resources of noise from transportation. Air quality has been proved to have a negative relationship with property value (Graves et al. 1988; Smith and Huang 1995). Water quality, such as pH level, clarity or visibility has been found to be positively and significantly related to sale price (Steinnes 1992; Michael et al. 1996). Proximity to hazardous waste sites unsurprisingly has a negative impact on property value (Kiel 1995; Farber 1998).

Brasington and Hite (2005) used the pollution site data of Ohio in USA to examine the relationship between house price and environmental disamenities. They used spatial autoregressive method to confirm that nearby point-source pollutants depress house price. Epp and Al-Ani (1979) studied waterfront residential properties located along small rivers and streams in Pennsylvania and found that pH levels too low to limit recreational use affect housing price. They found that acidity from minerals and carbon dioxide, which affects pH levels, significantly influenced housing price.

A distinction can be made between studies principally aimed at deconstructing house price and those undertaken to value an environmental “bad.”. It is likely that the latter kind of impact studies may not be so comprehensive in identifying a full range of independently predictive variables. Such evaluations have been criticized for difficulties in capturing imperfect knowledge on the attributes of each location and measuring intangible influences and individuals’ perceptions (Wardman and Bristow 2004).

2.3.4.2 Environmental Amenities

As mentioned above, proximity to a certain environmental amenity, such as a river and open space, could be considered as a dimension of accessibility adding value to property values. Stegman (1969), Richardson (1977), and Pollakowski (1982) show that house prices are also determined by the environmental attributes of the location, which connotes lower accessibility of areas peripheral to the city center tends to outweigh increased access to open space. Gillard (1981) argues: “even when a park may not be used for recreation because of crime problems, it may still be valued for aesthetic reasons by residents with a view of the park.” McLeod (1984) discovered that the river views were particularly important, and had a greater influence than a view of park. In particular, industrial, business, and transportation land uses can have a negative effect upon property with respect to aesthetic qualities (Powe et al. 1995).

Public open space and urban parks could enhance the value of environment, as well as quality of life, by improving air quality, providing recreational opportunities, and enhancing aesthetic value, among many other benefits (Nowak and McPherson 1993). Previous research has revealed that the price of house increases with proximity to nearby parks (Tyrväinen 1997; Thorsnes 2002); while, other studies reveal that the housing values increase when a nearby urban park expands its size. (Tyrväinen 1997). This lends weight to the idea that it is size of parks rather than number per se, that matters in meeting the open space needs of a city.

Poudyal et al. (2009) examine how the demand for green parks captured by property value. They employed a traditional hedonic price model to confirm that urban recreation parks increase nearby property values. Expanding the average size of parks by 20 % from the current level would in areas the per household consumer surplus by 160 dollars.

Netusil (2005) investigated how far open space can affect house price, taking the empirical study of Portland, Oregon as example. At the radii of 30 m within the open space, the study found that impact of locational advantage on the sale price of houses is insignificant. However, at distances farther than 30 m and up to 450 m from open space, houses were found to sell for statistically greater price than those located over 450 m from open space.

Lutzenhiser and Netusil (2001) used the same data, to explore the open space effect more deeply. They classified open space into urban park, natural park, and specialty park and found homes within 450 m from natural parks were of great significance. Other types of open space have a statistically significant influence, including golf courses (13.3 %), specialty parks/facilities (8.5 %), and urban parks (1.8 %). They found being proximity to open space does have positive impact on property values, but that is still dependent on the type of open space and distance from the open space.

Anderson and West (2006) explored the effect of neighborhood parks, regional, state, federal parks, natural areas, and cemeteries on property value in the Minneapolis–St. Paul metropolitan area. They found the value of properties which are proximal to neighborhood parks and special parks falls, because it increases the distance to the CBD. The benefits of proximity to neighborhood parks on housing price become higher when neighborhood has more children.

2.3.5 Others

In addition, there are some local context attributes, which influence property values. For example, Tse and Love (2000) found that a cemetery view has a negative impact on a property's price in Hong Kong. Generally, dwellings that have a cemetery view are not accepted, as that is bad *fengshui* (geomancy).

Interestingly, there have been some studies that demonstrated the influence of *fengshui* beliefs in the power of “lucky” and “unlucky” properties. For example, Bourassa and Peng (1999), who used sales transactions from 1989 to 1996, found

that lucky house numbers (e.g., 3, 6, 8, and 9) have significant positive hedonic prices and are capitalized into the sale prices of houses in Auckland, New Zealand. Chau et al. (2001) also found similar results in the predominantly Cantonese society in Hong Kong. Their results, however, showed that lucky floor numbers (e.g., 8, 18, or 28) are sold at significantly higher premiums during periods of property boom than the time of property slumps.

2.3.6 *Summary*

This chapter presents a wide-ranging literature review of hedonic price models, which can be summarized from numbers of aspects.

Firstly, based on Rosen's work (1974), it is possible to state that under a perfectly competitive market, and when demand equals supply, the implicit price of each attribute in the hedonic price model is the price people are willing to pay for. However, there are many scholars who have criticized Rosen's hedonic price estimation for the assumption of market equilibrium, in which case the implicit price of an attribute estimated is not strictly equal to the marginal willingness to pay.

Secondly, it has been noted that economic theory gives no clear guidelines on how to select hedonic price functional form. Because the study area of the first two empirical studies of this book is a part of metropolitan area of Cardiff, UK, and Orford (1999) explored the impact of locational externalities on housing price in Cardiff, choosing the semilog function form, this study will follow Orford (1999)'s approach. This is further backed up by the consensus in the literature that semilog equations have a more meaningful and intuitive interpretation.

In addition, the literature review also indicates that the hedonic price model suffers from a series of potential econometric issues, which could lead to estimation bias, such as spatial autocorrelation, multicollinearity, heteroscedasticity, and non-measurement of housing quality change. Thus, several types of econometrics tests will be applied in the chapters that follow. To make sure the models correctly or as accurately as possible, they will estimate the implicit prices of housing attributes, including accessibility. Among all these econometric problems, multicollinearity is most fatal for the estimation, as it can result in coefficients with opposite signs to their real relationship. Therefore, it is necessary to specify the function form, and choose the variables carefully.

Thirdly, despite the importance of housing submarkets, there is little consensus on how to specify submarkets, or identify their boundaries. Thus, this study attempts to contribute to the theory and practice of identifying housing submarket by using hedonic price models. The traditional method mainly emphasizes that within geographic area the price per unit of housing characteristics is constant, which means within a certain space, people have identical choice preferences. In that case, they could use school districts, postcode areas, and so on. Researchers could use building types to specify the housing submarket. However, there arises a

question of how to specify housing submarkets where the building type is homogeneous, and the price per unit of housing characteristics is not constant within some geographic unit. Indeed, some scholars have approached this matter by using cluster analysis for non-spatial information by minimizing estimated error for social economics indices. However, these alternative methods are criticized since the results are unstable over time, and they require high quality database and are difficult for policy-analysts to interpret. The innovative method the author explores in this book is to see if detailed morphological metrics can be used in defining housing submarkets. This takes the discussion of accessibility in hedonics a step further, to see if fine-grained systemic accessibility measures, taken from an urban grid, can not only help refine the prediction of individual house prices, but also help define geographical areas with stable and homogeneous valuations of housing attributes.

Fourthly, it is found that housing attributes can be divided into structural attributes, locational attributes, neighborhood attributes, and environmental attributes. Based on the discussions above, it is found that there is a debate saying that the locational attributes influence house price. However, empirical evidence can be contradictory. The New Urban Economics theory suggests that in a monocentric model, emphasis on the location (the distance to CBD) is the most essential factor to property value, because it accounts for the minimum travel cost, time, and distance behavior. The result is a negative gradient curve demonstrating the relationship between distance to a center and housing price in markets. However, due to urban expansion and polycentricity and the growing complexity of urban systems, many empirical studies have found that the power of location attributes have become weak. Current studies measure accessibility at both an aggregated and disaggregated level. However, aggregated measurements apply advanced algorithms calculating potential opportunities or attractiveness, which have been criticized for the loss of spatial information when aggregated. In contrast, disaggregated measurement requires *a priori* specification for terminals (such as bus station and train station), and they mainly rely on Euclidean distance measurement.

Hence, the focus of this book attempts to contribute the hedonic price theory by following aspects:

Firstly, previous studies need *a priori* specification within a predefined area, identifying the impact of local attractions on housing property, and most of them ignore the spatial information contains in street layout, which is the most essential element in the field of urban planning. Furthermore, compared with traditional measurement for locational attributes, the accessibility derived from urban configuration does not require *a priori* knowledge and can be easily employed at disaggregated level.

Secondly, despite of previous specifications for housing submarket in developed countries, there is less knowledge about how to delineate submarkets in property markets where the building type is simple (apartments) and social neighborhood characteristics are not long established and changing quickly over time, such as most cities in China. Thus, the author attempts to establish a new framework for delineating submarket in that situation, by clustering urban configuration features,

as urban configuration features are assumed to be associated with both spatial information and people's preference.

Thirdly, there has been little evidence of the micro-level determinants of house price volatility in urban land use planning. It is known that China is undergoing a process of rapid urbanization, whose scale is perhaps the largest the world has ever experienced. Although some researchers have investigated the impact of accessibility value, they have failed to confirm the dynamic relationship. Thus, the author posits that transformations in urban configuration accompany the urbanization process and that the corresponding continuous changes in street network configuration associated with urban growth and the attendant changes in accessibility are one of the key determinants of micro-level house price volatility in a city.



<http://www.springer.com/978-981-10-2761-1>

Urban Morphology and Housing Market

Xiao, Y.

2017, XIII, 189 p. 27 illus., Hardcover

ISBN: 978-981-10-2761-1