Housing Choice, Residential Mobility, and Hedonic Approaches

9

David M. Brasington

Contents

Abstract

This chapter explores the literature on residential mobility and house price hedonics. Residential mobility studies the decision of economic agents to move or not and, if they move, their choice of new residence. Topics covered in this chapter include the theory behind the move-or-stay decision, modeling intra- and interregional moves, empirically validated determinants of moving, and macro- and microlevel studies on mobility. Next, house price hedonics explain the price of a house as the sum of all the things that give a house value, from structural characteristics like the number of full bathrooms to public services and neighborhood characteristics that the house experiences. The chapter discusses the theory behind hedonics, applications of the technique, and empirical approaches to identify hedonic house price studies and second-stage hedonic regressions of the demand and supply of characteristics that give a house its value.

D.M. Brasington
Department of Economics, University of Cincinnati, Cincinnati, OH, USA
e-mail: David.Brasington@UC.Edu

9.1   Introduction

Everyone moves at some point in their life, and if they are lucky, they move into a dwelling. This chapter takes up the topics of residential mobility and house price hedonics: the movement of people from one dwelling to another and the factors that influence the price of the dwelling they move into.

The first part of this chapter covers residential mobility. First, the theoretical underpinnings of residential mobility are covered: why do people move? Life cycle models, cost-benefit models, and neighborhood change models are discussed. The chapter then discusses attempts to model moves within a region and between different regions. The factors influencing moves are explored, including the role of job changes, social capital, public policy, climate, and differences in the mobility of men and women. It includes a look at macrolevel and microlevel studies. The chapter finishes with a discussion of whether jobs follow people or if people follow jobs.

The second part of the chapter covers house price hedonics. First, the theoretical framework of the hedonic method is discussed. The issue of estimating a second-stage demand or supply regression from the first-stage hedonic is then explored, with particular attention to identifying the second stage from the first. The various applications of the hedonic method are discussed: (i) capitalization of taxes and public services into house price, (ii) quantifying the relative importance of non-market goods to house prices, (iii) constructing hedonic price indexes, (iv) evaluating policy alternatives, (v) applications to real estate finance, and (vi) testing for market segmentation. The discussion of hedonics concludes with an exploration of a handful of ways researchers have tried to identify hedonic parameters, including instrumental variables, spatial econometrics, the borders approach, and the mixed index approach. The final part of the chapter offers concluding thoughts and directions for future research.

9.2   Residential Mobility

Residential mobility studies the decision of economic agents to move or not and, if they move, their choice of new residence. Residential mobility is a key driver of economic development. By means of agglomeration economies, residents make a city more productive the larger a city becomes. The forced relocation of peasants to large urban areas in the early years of the Soviet Union was an attempt to create the conditions for a revolution of the proletariat described by Marx but was also an attempt to kick-start economic development. The migration of rural poor to urban areas of China has been a driver and consequence of the economic development of China's cities in our day.

Most of the literature in regional science on residential mobility takes an existing urban setting as its starting point. Winstanley, Thorns, and Perkins (2002) provide

a valuable background for the literature, noting that mobility of households within an urban area was traditionally considered a negative phenomenon because it interfered with social networks and moves were often provoked by deteriorating housing stock or neighborhoods or by adverse health events in a family. When Rossi (1955) challenged this view, a new literature was born on the premise that it is perfectly normal for a household to change locations over time. As people marry, have children, age, and experience other *household-formation events* like divorce and retirement, their housing needs change and they move to meet these new needs. These needs include house-specific needs like extra bedrooms but also encompass the need for schools, shopping, leisure and employment opportunities, and proximity to family and friends.

Winstanley, Thorns, and Perkins (2002) divide the literature on residential mobility into three strands: *life cycle models*, cost-benefit models, and *neighborhood change models*. From a regional science point of view, all models can be reduced to cost-benefit models if the costs and benefits are defined appropriately. Recent work by Rabe and Taylor (2010) illustrates the linkage between life cycle models and neighborhood change models, finding that life cycle events may be associated with a move to a better neighborhood (having a baby) or a worse one (a single person moving out of a parent's basement or a husband's unemployment).

Brueckner (2011) provides a useful overview of residential mobility using Tiebout (1956) sorting within a single urban area as a starting point. Tiebout sorting says that under certain conditions, households vote with their feet to end up in a municipality whose levels of public services exactly match consumer preferences, without the community changing its public service level.

In theory, with a head tax, an urban area consists of a collection of municipalities with internally homogeneous public service level demand. If income levels drive public service demand, municipalities are stratified by income. In the real world, *Tiebout sorting* is imperfect: there are a limited number of municipalities to choose from, so sorting is restricted to the "closest" match, leaving internally heterogeneous communities. Urban areas characterized by city-county governments are especially restricted in opportunities for Tiebout sorting, while there are better opportunities for sorting in large urban areas with many municipalities.

But even in a theoretical world with unlimited potential for Tiebout sorting, *economies of scale* provide a limit to the number of municipalities that will form. In the presence of important economies of scale, a minimum number of residents are required to lower the average cost of public service provision to a level competitive with its neighbors. That is, even if community A provides exactly the level of public service desired by a household, the household may choose community B if the cost of public service provision, and therefore taxes, is significantly lower than A, while the level of public service provision is sufficiently close to the most-preferred level. Scale economies may occur at different population levels for different services, which may result in a municipality

providing its own police services but cooperating with other municipalities for schooling services.

If higher income levels equate to higher preference for public service levels, a rich municipality chooses a high public service provision with a high head tax. But if models allow preferences to deviate from income levels, lower-income individuals may wish to sort into a high-service municipality. A sufficiently high head tax may be enough to discourage high-preference poor people from entering the rich municipality. Whether because of income levels, a prohibitive head tax, or both, a poor community forms with a low head tax and a low level of public service.

Property taxes finance a large share of local public services in the USA, and to the extent that they do, it allows more internal heterogeneity within a municipality. All else equal, a poor person would like to consume the larger amount of public service. But the taxes required are too high under a head tax. If a property tax replaces the head tax and the poor person can find a small house in the rich municipality, the poor person can pay a small property tax and still enjoy the large amount of public service.

The influx of poor residents into a rich municipality will have fiscal consequences for the rich residents. Because the poor residents pay less for the service, the municipality may run a fiscal deficit on these residents, a deficit that must be made up for by rich residents if the most-preferred level of public services is to be maintained. In response, rich residents may choose to engage in fiscal zoning, so that new construction pays for the cost of the public services provided to the new residence.

The literature long held that job changes primarily affect moves between urban areas but that households do not generally move within an urban area in response to a change in jobs. However, Clark and Davies-Withers (1999) show that, in fact, a person who takes a new job in the same urban area as the old job is 2.4 times more likely to find a new residence than people who experience no change in job. Such *intra-urban moves* are more likely for young households, single-earner households, and renters; in contrast, older households, dual-income households, and homeowners are more likely to commute to their new job from their current residence.

Apart from mobility within an urban area, households also may move from one urban area to another. Such long-distance moves to new labor markets are often spurred on by changes in employment opportunities, including households moving after retirement (Clark and van Lierop 1986). Changes in human capital are often cited as a cause of *interregional migration*.

While changes in employment opportunities are important to mobility, recent research explores reasons why such moves do not always happen. For instance, Kan (2007) explores the role of social capital in the decision to move, finding evidence that the more social ties a household has in an area, the less likely it is to move to a new location within the same labor market and that social ties make a move to a new labor market even more unlikely.

The issue of household mobility, be it intraurban or interregional, involves three decisions: (1) whether to move, (2) where to move, and (3) what to move into.

These decisions may be viewed as sequential as in a nested framework (e.g., Fischer and Aufhauser, 1988) or as a simultaneous set of decisions. The decision of whether to move has been modeled as *housing dissatisfaction* (e.g., Hanushek and Quigley, 1978). Before Hanushek and Quigley, most research on residential mobility was descriptive. Hanushek and Quigley motivate their work by noting that moving is a dynamic response to a household's changed circumstances. The decision to move depends not only on housing demand but also on search costs, the transaction costs of moving, and the distribution of house prices available to a prospective mover. Hanushek and Quigley model the decision to move as a probit function where the dependent variable is the difference between a household's housing consumption and its desired housing consumption in the next period, divided by its desired housing consumption in the next period. They use a data set of renters from Phoenix and Pittsburgh who were interviewed initially, then one year later, then a year later again. Housing demand is estimated by an ordinary least squares model that regresses monthly rent on income, wealth, age, household size, household education level, race, and whether the apartment has a refrigerator, air conditioner, or stove. The percentage difference between housing consumption expenditures and predicted housing consumption expenditures is Hanushek and Quigley's measure of disequilibrium. The decision to move is modeled as a function of demographic characteristics like age, income, race, and household size. The authors extend their model to a multinomial probit case of whether a household searches and moves, searches for new housing without moving, and does not search for new housing. Hanushek and Quigley's model is readily extended from housing dissatisfaction to neighborhood dissatisfaction, providing a link between housing dissatisfaction and neighborhood change models.

More recent research on household mobility has focused on the role of public policy. California's Proposition 13 prohibits reassessment of the base-year value of a property except in the case of new construction or a change in ownership. The law provides incentive for residents to stay in their current dwelling. Recent work by Ferreira (2010) estimates the impact on household mobility by analyzing recent amendments to Proposition 13 that allow residents over 55 years of age to sell their house and still transfer the tax savings to a new house. Ferreira finds that 55-year-old homeowners are 25 % more likely to move than comparable 54-year-olds.

A wide variety of approaches to modeling mobility have been used in the literature, as discussed by Clark and Van Lierop (1986). These include linear programming models, gravity models, discrete choice, and behavioral models. The methods are described below, and readers may refer to Clark and Van Lierop (1986) for specific examples of research using each method.

Perhaps the oldest approach to modeling residential location is the linear programming model, which is the counterpart to the Alonso location model. In the linear programming model, households choose a market basket of non-housing goods to maximize the ability of the household to spend on housing. The outcome of choosing the optimal basket results in the household choosing the optimal housing location, optimal in the Pareto sense that no household can relocate to

increase its expenditures on rent without displacing another household, thus reducing the other household's expenditures on rent.

Gravity models were sometimes used to model residential location instead of linear programming models. *Gravity models* are derived from Newton's law of gravity and are also known as distance decay models, *spatial interaction models*, and *origin-destination flow models*. These models describe the attraction between two objects, in this case people in one location and their attraction to another location. It is common in the residential mobility literature to parameterize spatial interaction models with a log-linear functional form to look at the aggregate movement of households between areas. These log-linear models regress the number of moves between two areas as a function of explanatory variables such as the net migration from each area, the control variables, the distance between each area, and the distribution of flows between areas. Shortcomings of these models, particularly their failure to consider spatial dependence, are discussed in LeSage and Fischer (2010).

The micro counterpart to the log-linear spatial interaction models is the use of discrete choice models, pioneered by Hanushek and Quigley (1978). Even today, most research follows this methodology. Households receive utility from choosing a dwelling, and the utility they receive affects the probability that they will select the dwelling. The discrete choice framework is used to model the move-or-stay decision, the rent or own decision, as well as the decision to select a new place of residence from a variety of locations. The discrete choice framework is also used in more complex procedures like sequential choices of move versus stay, then choice of a specific neighborhood from a finite option set, then a choice of a specific dwelling within the neighborhood.

Behavioral models of residential mobility examine a household's level stress, disequilibrium, or, as in Hanushek and Quigley (1978), dissatisfaction with housing consumption (discussed earlier in this chapter). Much attention has been given to the search process involved in relocation. Often, such models contain a stopping parameter that guides when a household finalizes its decision, such as when the expected gap between the expected utility resulting from additional search and the highest utility available is small enough. *Search models* have incorporated preferences over housing consumption and neighborhood consumption, as well as incorporating the preferences of single-earner households, dual-earner households, and the preferences of households with children that allow the children's preferences to enter into the decision calculus.

Perhaps due to data constraints, many of the early studies on residential mobility were macro-oriented, dealing with the moves of an aggregated number of people between geographic areas like urban areas, census tracts, or municipalities. The increased availability of data has led to the micro approach – the examination of the residential mobility decisions of a single household – gathering the lion's share of attention in the literature. Nevertheless, recent research reviewed by Dieleman (2001) has rediscovered the macro side of the equation, uncovering new factors important for mobility between geographic areas. *Turnover rates of the housing stock* and housing prices vary across regions and across time. The renewal of the

housing stock allows for more potential choice of housing by area residents, while low housing prices make home ownership more affordable to young people and household formation by young people into renter-occupied dwellings more feasible. Household mobility is twice as prevalent in the American South as in the American Northwest. Research suggests that moves are more common where the local economy is stronger: such areas are marked by increased construction of new dwellings and an influx of young workers. Endogenous growth in the number of young workers also plays a role, as Southerners are more likely to raise children and Northwesterners are more likely to raise dogs. Differences in public policy might also help explain variations between regions in mobility, with the urban growth boundaries of Portland providing a marked contrast with the urban sprawl of Atlanta and the regulatory difficulty of constructing new dwellings in California contrasted with the relative lack of zoning laws in Texas. Internationally, such differences help explain the relative lack of housing stock turnover and household formation in Europe. There is some truth to the joke of the Italian man who proudly proclaims, "Before I was 35, I lived with my mother; now that I'm over 35, my *mother* lives with *me*." Variation in *housing prices* tends to persist over time, although there are places where the variation in house prices over time is greater, such as the coastal regions of the USA, which experience boom and bust cycles to a greater degree than inland cities. Higher house prices are found in larger urban areas, more rapidly growing urban areas, and areas where construction costs – including regulatory approval costs – are higher. Direct government intervention in the housing market also affects the price of housing, such as rent controls, government construction of low-income dwellings, urban growth boundaries and development taxes, and government treatment of renting versus owning.

Mobility is a subject with many aspects to be investigated: renting versus owning, within versus interurban mobility, and mobility by education, marital status, child status, racial composition, and more. Perhaps surprisingly, little research has investigated the *mobility of women* relative to men. The field commonly assumed that women were less mobile than men without engaging in much empirical research to confirm or refute the supposition. Faggian, McCann, and Sheppard (2007) find evidence that women in the UK are more mobile than men. They point out that previous research on the issue almost entirely ignores the role of human capital in the relative migration of men and women. Faggian, McCann, and Sheppard study a sample of UK university graduates. They follow these individuals from their homes to university and from university to place of first employment. After controlling for human capital and regional economic conditions, they find women are more interregionally mobile than men, arguing that this increased mobility represents attempts by women to overcome gender bias in the labor market. Faggian, McCann, and Sheppard also find that racial minorities, older individuals, unemployment in the individual's hometown, and distance of the individual's hometown to London reduce mobility of the four types of migrants studied, relative to nonmigrants. They use multinomial logit, dichotomous logit, and conditional logit regression approaches.

In recent decades, the role of climate in attracting new residents has drawn a lot of research attention. The population shift from New England and the Midwest in the USA to the "Sunbelt" areas of the South and West is attributed in large part to climate. Rappaport (2007) challenges the claim that the invention of air conditioning is responsible for the shift, finding increased population not only in areas with warmer winters but also in areas with cooler, less humid summers. Rappaport also notes that the shift to nice weather is driven not just by elderly retirees but also is happening at almost the same rate by working-age people. Rappaport details the literature on *climate and residential mobility*, pointing out that the traditional approach is the *compensating differential approach*: estimating the wages that workers forego to live in places with a better climate. But Rappaport notes that there are several important drawbacks to the compensating wage differential approach used by previous studies. One such problem is that the level of detailed observations required is only available for places with over 100,000 people. This population requirement introduces sample selection bias because data is only available from places where a lot of people have already chosen to live. A second problem is the inability to properly control for individual-specific and house-specific characteristics. This is particularly a problem because high-income individuals may choose to live in areas with a high quality of life, so that the unobserved characteristics of these individuals are likely positively correlated with quality of life variables, including climate. Given the limitations of the compensating differential approach, Rappaport instead studies climate's role in changes in quality of life and productivity. Weather's contribution to changes in quality of life and changes in productivity has occurred for four reasons: (i) the decline of agricultural employment has made weather less important for agricultural productivity, (ii) air conditioning and improved heating technology have made extreme temperatures more bearable, (iii) rising incomes have increased the consumption of good weather, and (iv) the utility value of weather has increased because of the rise in the number of affluent retirees. His approach is unable to distinguish between the quality of life and the productivity effects of climate, and the approach cannot quantify the size of the effects, but it avoids the problems of the compensating differential approach and opens up a new angle of attack with which to research a long-standing issue. Rappaport uses county-level data, which provides a large number of observations and full coverage of the USA, avoiding the sample selection problem of using urban area-level data. The annual growth rate of population density is regressed as a function of average January high temperature, average July high temperature, precipitation, and other controls. Population growth is found to be statistically significantly related to high temperatures in January and July in most regressions.

Like Rappaport (2007), Partridge and Rickman (2003) also investigate a long-standing issue with a new approach. For decades, regional scientists have investigated whether *jobs follow people or people follow jobs*. The lack of finality in the issue stems from the fact that both job growth and population growth are endogenous, leading multiple researchers to call this a "chicken and egg" problem. Job growth draws new residents, while an influx of new residents creates new jobs.

Early attempts to identify the direction of causality employed simultaneously estimating employment and migration equations using instrumental variables, but such an approach suffers from the difficulty of finding appropriate instruments. It also fails to estimate short-run versus long-run responses. *Vector autoregression (VAR) models* have been used more recently to try to disentangle the job-people question, but previous attempts have fallen short in various ways. Some VAR studies assumed that all contemporaneous employment innovations were labor-demand shocks. Other VAR studies fail to include relevant equations in their models, such as migration equations or wage equations, making it impossible to disentangle labor demand and supply. Instead, Partridge and Rickman use a structural VAR model that incorporates a labor market model. The labor market model is used to create long-run identifying restrictions for the structural VAR model. Partridge and Rickman also distinguish between *labor-supply and labor-demand shocks* noting that wages respond in opposite directions to each type of shock. Finally, Partridge and Rickman decompose labor-supply shocks into those due to new residents and original residents. They use data from the 48 contiguous United States from 1969 to 1998, finding labor-demand shocks account for about 46 % of the variance in employment forecasts. Migration (labor-supply) shocks account for about 33 % of the variance, and internal labor-supply shocks the remaining 21 %. So while people seem to follow jobs more than jobs follow people (46 % vs. 33 %), when you sum the 33 % from migration and 21 % from internal labor-supply shocks, it is clear that the role of supply is as important as that of labor demand. Partridge and Rickman's model suggests that Sunbelt states are more influenced by migration shocks, and Rustbelt, Farm Belt, and Energy states are more influenced by labor-demand shocks.

9.3 House Price Hedonics

House price hedonic regressions attempt to break the price of a house into each aspect of the house that gives it value, from structural characteristics like house size to neighborhood characteristics like crime rates and to public finance characteristics like tax rates. Adding up the price of all the things that give a house its value will tell you the market value of the house. But a house is a bundle of services, each providing value but few being easily added onto or subtracted from a house. That is, while it may be possible to find the market value of adding another full bathroom to a house by getting a quote from a contractor, it is not as easy to take the same house and purchase an extra unit of air quality or public safety for it. The house price hedonic estimation is a way to use statistics to estimate the value that an extra unit of public safety would provide toward a house.

Residential mobility is linked to house price hedonics through many channels. As households *Tiebout sort* into an urban area, their demand for specific housing characteristics will be reflected in house price through the relative supply of that attribute in the area. If houses with high-quality public schools are in short supply relative to demand, households will have to pay a premium to live in such houses.

Housing turnover affects residential mobility, but the lack of new housing will tend to push the price of newer houses up, especially in the face of supply restrictions like development taxes and urban growth boundaries. And while residential mobility is likely to be less pronounced in urban areas with low *population growth rates*, the level of house prices is likely to be lower in such areas as well.

Although its use goes back to the early 1900s, the modern era of house price hedonics began when Sherwin Rosen (1974) detailed the theoretical underpinnings of the method. Rosen notes that the number of rooms in a house provides value to the house, but a house is a bundle of attributes, including the number of rooms. If we let z be the amount of housing consumed, then $z = (z_1, z_2, z_3, \ldots, z_n)$ lets the amount of housing consumed be a function of the n characteristics of the house, so that summing z_1 through z_n yields the amount of housing consumption. These quantities can be converted to prices p so that $p(z) = p(z_1, z_2, z_3, \ldots, z_n)$: the price of a house $p(z)$ is a function of the price of its attributes.

Buyers and sellers come together in the market to determine the price of the attributes of a house, the end result of which determines the selling price of the house. Consumers have a budget constraint in which their incomes are exhausted on purchases of housing attributes and a numeraire good. Consumers maximize utility over housing attributes and the numeraire good subject to a budget constraint. The first-order conditions of the constrained maximization equate the marginal rate of substitution for house characteristic z_i to p_i, the marginal price of characteristic z_i. And because the house is a bundle of characteristics that cannot be easily bought and sold separately on the market, each p_i is an *implicit price*, not an explicit price.

Consumers bid for housing attributes given a fixed level of income and utility. A bid is the marginal rate of substitution between a housing attribute z_i and money, which shows the consumer's reservation demand price for an additional unit of z_i. Next, $p(z)$, the minimum price a consumer must pay in the market for various quantities of z_i, is graphed along with consumers' bid functions, and tangencies show the optimal consumption of an attribute by a consumer.

Rosen next considers *producers of housing*, which are assumed to operate in a perfectly competitive environment. Producers maximize profits by producing additional housing attributes until marginal cost equals marginal revenue. A producer's offer function shows the unit prices a producer will accept for various house designs, given a fixed level of profits. Producers maximize their offer prices subject to the price they can get in the market, $p(z)$. Rosen graphs a producer's offer function for characteristic z_i along with the price $p(z)$ obtainable in the market, and tangencies between these curves reveal the optimal amount of characteristic z_i the firm should offer, given the optimal quantity of other housing characteristics.

Rosen next puts consumers' bids together with producers' offers, and tangencies denote a joint envelope all along $p(z)$, the implicit price function of characteristic z_i. Rosen notes that $p(z)$ may be nonlinear: the price of a characteristic may change as more of it is consumed.

Rosen's work would have been influential enough if he had stopped there, but Rosen goes on to describe how the supply and demand for *implicit house price*

hedonic characteristics may be estimated. Step 1: regress p(z) as a function of all z_i. This is the house price hedonic. Step 2: compute the marginal implicit prices $\partial p(z)/\partial z_i = p_i(z)$ for each buyer and seller, evaluated at the actual quantities consumed. Step 3: use the marginal implicit prices endogenously in a system of supply and demand for characteristics, being careful to include some exogenous shift characteristics in each equation. Step 4: estimate supply and demand simultaneously, making sure both supply and demand equations are identified from each other.

Later, James Brown and Harvey Rosen (1982) claimed the need to modify Sherwin Rosen's method of estimating a second-stage supply and demand system for hedonic attributes. They point out that while Sherwin Rosen was careful to identify the supply curve from the demand curve and vice versa, neither the supply nor the demand is identified from the initial house price hedonic in Step 1 above. All the information from the house price hedonic is embedded in the marginal implicit prices included in the second-stage supply and demand regressions, so that attribute prices are not independent of errors in the second stage.

Because of the difficulties in *identifying the demand for a hedonic attribute* like air quality from the underlying house price hedonic, most researchers stop at estimating a first-stage hedonic regression. However, there have been suggestions for how to identify second-stage supply and demand equations from the first-stage hedonic. The first, and the most frequently applied, is Brown and Rosen's (1982) suggestion to segment the housing market.

Market segmentation is the identification strategy taken by Brasington and Hite (2005). Brasington and Hite regress the natural log of house prices on twenty-two explanatory variables, including the focus variable, DISTANCE TO HAZARD, representing the distance from each house to the nearest environmental hazard. They run separate first-stage hedonic regressions for the six largest urban areas in Ohio, using a *spatial Durbin model* to help control for the influence of omitted variables. They use the parameter estimate of DISTANCE TO HAZARD to calculate the marginal implicit price of *environmental quality* for each of the 44,255 houses in the sample, although values are averaged at the census block group level (5,051 census block groups). The implicit prices for the six urban areas are based on six different parameter estimates of DISTANCE TO HAZARD, but all implicit prices in the sample are pooled to estimate a single second-stage demand curve for environmental quality. The implicit prices are treated endogenously using instrumental variables. Brasington and Hite find a small, statistically significant relationship between house prices and environmental quality in five of the six urban areas. They estimate the demand for environmental quality using two-stage least squares, limited information maximum likelihood, a fixed effect model, and a spatial Durbin model. There seems to be no relationship between environmental quality and lot size in the spatial regression. However, house size seems to be a substitute for environmental quality, and school quality and environmental quality seem to be complements.

Other than Brown and Rosen's (1982) market segmentation by geography, researchers have used or promoted identification of demand and supply from

first-stage hedonics by segmentation by time, functional form assumptions, and instrumental variables approaches. Of particular note is a paper by Cheshire and Sheppard (1998), who use a spatial lag of the two nearest houses as an instrument for implicit prices. Cheshire and Sheppard obtain similar results for their instrumented and so-called underidentified demand system, suggesting that while identification of demand from initial hedonic regressions is theoretically important, it does not affect regression results much. Also of note are Ekeland, Heckman, and Nesheim (2002), which claim that Rosen (1974) was right. Specifically, they claim that when all the information in the hedonic model is exploited, second-stage demand equations are identified from the first-stage hedonic even in single markets and without having to impose an arbitrary functional form assumption. The problem with second-stage models is that the hedonic model is generically nonlinear, and it is only when people use a linearization that identification problems arise. They show theoretically that a normal linear quadratic model achieves identification, but they do not actually estimate such a model.

Far more common than two-stage demand models are models that simply apply the first-stage hedonic house price regression to various problems of interest to regional scientists. These applications fall into at least six categories: (i) capitalization of taxes and public services, (ii) quantifying the relative importance of non-market goods to house prices, (iii) constructing hedonic price indexes, (iv) evaluating policy alternatives, (v) applications to real estate finance, and (vi) testing for market segmentation.

One of the first *applications of the hedonic method* to regional science was to investigate the capitalization of taxes and public services into constant-quality house prices. The idea of capitalization of taxes, for example, is that, holding public services and everything else constant, low taxes provide value to a house. Low taxes provide value to a house in the current period, the next period, and all successive periods for the life of the house (or land) subject to a discount rate. The sum of this discounted stream of value boosts house price above what it otherwise would be, so that low taxes are "capitalized" into constant-quality house price. Oates (1969) is the most-cited paper in the field of public finance capitalization, although Orr (1968) preceded that article or was at least contemporaneous. Oates found capitalization of both property taxes and school spending and claimed that it proved Tiebout (1956) was right. But the interpretation of *capitalization* has a large literature of its own. Several papers claim that capitalization should not occur in theory and that Tiebout's paper itself would not predict capitalization. Other papers say that capitalization does exist and that its existence is a result either of a scarcity of desirable public service/tax combinations, local governments using more resources than necessary to produce public services, or local governments spending too much or too little on public services. Another paper says that both major views of capitalization are right: it should occur and should not occur simultaneously, depending on whether the house is in an area of high or low housing supply elasticity. See Brasington (2002) for a detailed discussion and empirical exploration of this issue.

The house price hedonic has been used extensively to quantify the relative importance of *non-market goods*. The difficulty for economists and regional scientists in studying nonmarket goods is that there is no observable market price. Some researchers estimate the market price by contingent value surveys or estimating travel cost models; others exploit the housing market. If air quality imparts value to a house, it should be possible to see how much value it imparts to a house by regressing house price as a function of house characteristics, neighborhood characteristics, and air quality. The partial derivative of house price with respect to air quality yields its marginal implicit price, from which a marginal willingness to pay can be calculated. Other nonmarket goods whose value has been estimated by house price hedonic regressions include public school quality, crime rates, and proximity to lakes, parks, hospitals, churches, shopping malls, and airports.

Hedonic price indexes are used to track how the price of a typical house changes over time. The house price hedonic is first estimated. A researcher then plugs in the sample means of the explanatory variables which, together with the parameter estimates for each variable, yield the typical house's estimated price in the base year (index = 100). The regression is run again for new time periods, with researchers plugging in the *old* sample means for the explanatory variables and using the new parameter estimates to come up with a new value for the index. There are different ways to construct hedonic price indexes. Gatzlaff and Ling (1994) compare indexes constructed from traditional hedonics based on sale price, hedonics based on assessed value, and the repeat sales method. The repeat sales method follows the sale price of the same houses over time. In theory, it provides an unbiased price index because, unlike the traditional house price hedonic, there are no unobserved characteristics to bias the price – as long as the quality of the house remains constant between sales. The criticisms typically levied against the *repeat sales method* are that (i) it suffers from sample selection bias, as certain starter homes and undesirable homes may sell more often than other houses; (ii) that the same houses may have been remodeled between sales, so that researchers are not really comparing the same house over time; and (iii) that the repeat sales method wastes a tremendous amount of data, because only a small subset of houses sell more than once over a small time period. These and other criticisms of the repeat sales model are detailed by Haurin and Hendershott (1991). Gatzlaff and Ling get around these criticisms by having information about houses that did not sell during the time period, using the entire set of houses for the city of Miami between 1971 and 1991, and controlling for houses that have been substantially remodeled between sales. Gatzlaff and Ling find that all methods provide precise measurements of house price indexes, both in levels and in changes between years, although there was more noise in quarterly price changes. It is especially noteworthy that indexes based on assessed value were similar to those based on actual sale prices, because assessed property value data is much easier and cheaper to obtain than sale price data.

The house price hedonic can be used to *evaluate policy alternatives*. This can be done in a cross-sectional study or a study measuring changes over time. For a cross-sectional study, a researcher pools data from a group of houses that does

not have policy x and a group that does. A dummy variable for policy x is included in the regression, and a statistically significant parameter estimate for the dummy variable tells whether the policy has a positive, negative, or no effect on house prices. Care must be taken to control for every possible influence, so that omitted variables are not driving the dummy variable's parameter estimate. The same procedure can be done across time, with the sample containing house sales before and after the policy change and the dummy variable reflecting sales after the policy change.

There are numerous applications of the house price hedonic to real estate. Researchers have used the technique to estimate the discount or premium for having an unusual house for the neighborhood, a motivated seller discount, differences in sale price for houses transacting with a real estate broker or for-sale-by-owner, and the effect on price of selling by way of a sheriff's auction or bank sale.

As mentioned before, the two-stage hedonic demand technique described by Rosen (1974) is most often conducted by using segmented housing markets in the first-stage regression. Another use of the hedonic technique is to test for *market segmentation*. A different house price hedonic regression is run for every area suspected of being a distinct housing market. If the parameter estimates are significantly different from each other, the houses are deemed to come from segmented markets; that is, each market is subject to a different set of supply and demand conditions.

While hundreds of house price hedonic studies have been conducted, the *identification issue* – whether we believe the parameter estimates are accurate – is an issue that continues to dog the hedonic approach. In response, a variety of estimation techniques have been used. These include ordinary least squares, two-stage least squares, other instrumental variable techniques, spatial regression approaches, the borders approach, and the mixed index approach. Traditionally, ordinary least squares has been used to estimate house price hedonics. One example is Hoehn, Berger, and Blomquist (1987), which deserves special note in this chapter. This paper uses the hedonic method not only to estimate a house price equation but also to estimate a wage equation, an additional application of the hedonic technique. Furthermore, it helps bridge the gap between the hedonic and residential mobility literatures by including a considerable number of quality of life controls including numerous measures of pollution, climate, and other urban amenities.

Subsequent research recognized the drawbacks of ordinary least squares for hedonic regressions. An early adopter of instrumentation is Voith (1991), who simultaneously runs wage and rent regressions using two-stage least squares. Instrumentation is necessary, Voith says, because wages and rents are simultaneously determined. Another attempt to identify house price hedonics using instrumental variables is found in Epple and Sieg (1999).

Although not a statistical method, another approach to identifying parameter estimates from hedonic models is to exploit borders. The technique was pioneered by Gill (1983) and was all but forgotten until the late 1990s, since which time it has become more popular than ever (although Gill's paper is rarely credited).

The idea is to examine houses that are similar in all respects but one: the houses on one side of a border have characteristic x, and those on the other side of the border have characteristic y. The border could be a school district boundary, as in Gill (1983), or it could be a state line, county line, tax abatement zone boundary, or any other boundary of interest to researchers. Gill picks areas with similar housing that have less than 10 % of blacks in neighborhood schools. Some areas are within the Columbus (Ohio) City School District; others are in suburban school districts. Suddenly in 1978, a school busing plan was approved that mandated all city schools to have between 20.9 % and 50.9 % black students. Suburban schools were exempt from busing. Gill finds evidence of increased demand for suburban houses, especially those with four or more bedrooms, when the courts first ordered city schools to desegregate.

Despite the current popularity of the *borders approach* to identifying house price hedonics, LaCombe (2004) shows that the use of spatial statistics is better. Note that while spatial statistics is not a technique in itself, it is an approach that incorporates spatial dependence in a variety of regression techniques. LaCombe takes counties that border each side of a state. He studies whether state-level differences in AFDC and food stamp usage is associated with a different prevalence of female-headed households and labor force participation levels by women. LaCombe uses both the borders approach and *spatial statistics*. While the borders approach shows no relationship between AFDC and food stamps on the percentage of female-headed households, the spatial statistical models show a positive, statistically significant relationship, a finding more consistent with theory and previous studies. LaCombe argues that the driver of this result is the presence of spatial dependence: the idea that observations nearby in space are similar to each other but that observations farther away from each other are less similar. The same Brasington and Hite (2005) paper described above for two-stage hedonic demand uses spatial econometrics to address spatial dependence. The specific procedure used in Brasington and Hite is the *spatial Durbin model*. Brasington and Hite describe the intuition behind how their spatial Durbin model incorporates the influence of *omitted variables*, thus helping to identify a house price hedonic (p. 63):

> while a lagged dependent variable in time series regressions relies on observations nearby in time, the spatial lag relies on a linear combination of house values nearby in space. Unmeasured influences help determine the value of neighboring houses and, as explained earlier, the value of neighboring houses is related to the value of our own house. So our own house value is affected by the unmeasured influences of neighboring observations. And the unmeasured influences of neighboring houses are similar to the unmeasured influences for our house because our neighbors are close: the same things that affect our neighbors should affect us, too.

More formal and detailed discussions of spatial models and omitted variable bias are given in Fischer and Getis (2010) and LeSage and Pace (2010; ▶ Chap. 77, "Interpreting Spatial Econometric Models"). There are many models within spatial econometrics and spatial statistics that can incorporate the influence of spatial dependence. These include the spatial autoregressive model, the general

spatial model, the geographically weighted regression, and the spatial Durbin model. Certain models incorporate the influence of omitted variables; others do not. Spatial routines are available in Matlab at www.spatial-econometrics.com and www.spatial-statistics.com and also with S+, Spatial Stata, PySAL, STARS, Open GeoDa, GeoR, and GeoDa Space, most of which are free.

A final technique to identify hedonics comes from an idea of Bowden (1992), operationalized by Brasington and Hite (2008). In their introduction, Brasington and Hite (2008) discuss the shortcomings of the borders approach, instrumental variables approach, panel data approaches, and spatial statistical approaches. They then discuss Bowden's observation that Tiebout sorting leads to endogeneity of house price hedonic variables and that even including buyer characteristics as explanatory variables is insufficient to identify the hedonic. What is needed, Bowden says, is a *mixed index model*. This mixed index model consists of a system of equations estimated simultaneously. This system contains a house price hedonic with explanatory variables for house characteristics and buyer characteristics (Brasington and Hite add neighborhood characteristics to Bowden's list). The system of equations also contains equations with dependent variables representing endogenous explanatory variables from the house price hedonic. Brasington and Hite implement Bowden's mixed index model for the first time. The endogenous variables they choose from the house price hedonic are the homeowner's income, commute time, and air pollution in the homeowner's census block group. These endogenous variables are regressed as a function of additional homeowner characteristics, house characteristics, and neighborhood characteristics. Brasington and Hite find that the mixed index model yields predicted house prices that are statistically significantly different from a traditional ordinary least squares regression and from a traditional regression that adds buyer characteristics. They find the magnitude of the *capitalization of environmental quality* differs a lot when using the mixed index model, and they find the mixed index model has the most accurate policy predictions with the lowest prediction variance and the most favorable skewness and kurtosis performance. The major drawback for implementing the mixed index model is that data on the characteristics of the buyers of each house is rarely available. Furthermore, even the most thoughtful exclusion restrictions in the system of equations are somewhat arbitrary. Still, the favorable statistical properties may make the mixed index model an attractive choice when data on individual *buyer characteristics* becomes more widely available.

9.4   Conclusions

This chapter has examined the state of the literature on residential mobility and house price hedonics. It has discussed the theoretical framework of the literature, looked at the empirical approaches to test theory, and briefly discussed the factors influencing mobility and house prices.

While the theory behind the second-stage demand regressions stemming from house price hedonics continues to receive some attention, the theory behind residential mobility and first-stage house price hedonics seems relatively stable. Rather, it is the empirical side of the story that has received much more exploration in recent decades.

The residential mobility literature has advanced far in testing the determinants of mobility, but less energy has been invested in the empirical techniques used. While traditional probits and the like are appropriate and certain studies like Partridge and Rickman (2003) have applied structural vector autoregression to traditional questions, there is room to apply newer econometric approaches to investigate long-standing issues in the literature. For example, in a discipline in which space plays such a prominent role, there is a notable lack of research applying spatial statistics and econometrics to the residential mobility literature (Fischer and Getis, 2010). There is also room to investigate other determinants of mobility that have been understudied. For instance, what is the impact of having a city-county government on in-migration and out-migration? Is moving more or less common than in urban areas with a large number of highly fragmented political jurisdictions?

Concerning the house price hedonic literature, there has been some advance in regression technique, with several approaches available to help identify parameter estimates in first- and second-stage hedonic regressions. What is needed is more research like Lacombe (2004) to *compare the different techniques* to see which offer the best combination of desirable estimation qualities like unbiasedness and efficiency. The hedonic literature is a major beneficiary of renewed interest in *housing prices* as much of the developed world suffers through a post-bubble *housing bust*. Some of this renewed literature could investigate additional hedonic determinants of rises and falls in housing prices, including the role of climate and views of lakes and oceans and the role of the industries a city specializes in – the hedonic counterpart to Rappaport (2007) and Partridge and Rickman (2003).

Both the residential mobility and the house price hedonic literatures could benefit from a comprehensive data set that would put the Nurses' Health Study and the Panel Study of Income Dynamics to shame. Such a data set would need to cover an entire continent to account for mobility of subjects. It would need to enroll a large number of subjects. All the purchases of these subjects would need to be tracked, as well as the sales of durable goods. The residences and workplaces of these subjects throughout their lifetimes would be recorded, as would health, fertility, and income data. And detailed information about the subjects' residences and neighborhoods would be collected, as well as certain information about the subjects' relatives. The information about relatives would be useful for constructing instrumental variables for endogenous variables. Data on education, religious practice, criminal records, and misbehavior in school would also be useful. Data collected on such persons over the course of a lifetime would prove a research gold mine, if confidentiality could be maintained.

References

Bowden RJ (1992) Competitive selection and market data: the mixed-index problem. Rev Econ Stud 59:625–633

Brasington DM (2002) Edge versus center: finding common ground in the capitalization debate. J Urban Econ 52:524–541

Brasington DM, Hite D (2005) Demand for environmental quality: a spatial hedonic analysis. Reg Sci Urban Econ 35:57–82

Brasington DM, Hite D (2008) A mixed index approach to identifying hedonic price models. Reg Sci Urban Econ 38:271–284

Brown JN, Rosen HS (1982) On the estimation of structural hedonic price models. Econometrica 50:765–768

Brueckner JK (2011) Lectures on urban economics. MIT Press, Cambridge

Cheshire P, Sheppard S (1998) Estimating the demand for housing, land, and neighbourhood characteristics. Oxford Bull Econ Stat 60:357–382

Clark WAV, Davies-Withers S (1999) Changing jobs and changing houses: mobility outcomes of employment transitions. J Reg Sci 39:653–673

Clark WAV, van Lierop WFJ (1986) Residential mobility and household location modelling. In: Nijkamp P (ed) Handbook of regional and urban economics, vol 1, Regional economics. North-Holland, Amsterdam, pp 97–132

Dieleman FM (2001) Modelling residential mobility: a review of recent trends in research. J Hsg Built Environment 16:249–265

Ekeland I, Heckman JJ, Nesheim L (2002) Identifying hedonic models. Am Econ Assoc Papers Proc 92:304–309

Epple D, Sieg H (1999) Estimating equilibrium models of local jurisdictions. J Polit Econ 107:645–681

Faggian A, McCann P, Sheppard S (2007) Some evidence that women are more mobile than men: gender differences in U.K. graduate migration behavior. J Reg Sci 47:517–539

Ferreira F (2010) You can take it with you: proposition 13 tax benefits, residential mobility, and willingness to pay for housing amenities. J Pub Econ 94:661–673

Fischer MM, Aufhauser E (1988) Housing choice in a regulated market. A nested multinomial logit analysis. Geogr Anal 20:47–69

Fischer MM, Getis A (2010) Handbook of applied spatial analysis. Software tools, methods and applications. Springer, Berlin/Heidelberg/New York

Gatzlaff DH, Ling DC (1994) Measuring changes in local house prices: an empirical investigation of alternative methodologies. J Urban Econ 35:221–244

Gill HL (1983) Changes in city and suburban house prices during a period of expected school desegregation. Southern Econ J 50:169–184

Hanushek E, Quigley J (1978) An explicit model of intrametropolitan mobility. Land Econ 54:411–429

Haurin DR, Hendershott PF (1991) House price indexes: issues and results. AREUEA J 19:259–269

Hoehn JP, Berger MC, Blomquist GC (1987) A hedonic model of interregional wages, rents, and amenity values. J Reg Sci 27:605–620

Kan K (2007) Residential mobility and social capital. J Urban Econ 61:436–457

Lacombe D (2004) Does econometric methodology matter? An analysis of public policy using spatial econometric techniques. Geogr Anal 36:105–118

LeSage J, Fischer MM (2010) Spatial econometric methods for modeling origin-destination flows. In: Fischer MM, Getis A (eds) Handbook of applied spatial analysis. Springer, Berlin/Heidelberg/New York, pp 409–433

LeSage J, Pace RK (2010) Introduction to spatial econometrics. CRC Press, Boca Raton

Oates WE (1969) The effects of property taxes and local public spending on property values: an empirical study of tax capitalization and the Tiebout hypothesis. J Polit Econ 77:957–971

Orr LL (1968) The incidence of differential property taxes on urban housing. Nat Tax J 21:253–262

Partridge MD, Rickman DS (2003) The waxing and waning of regional economies: the chicken-egg question of jobs versus people. J Urban Econ 53:76–97

Rabe B, Taylor M (2010) Residential mobility, quality of neighbourhood and life course events. J R Stat Soc Ser A Stat Soc 173:531–555

Rappaport J (2007) Moving to nice weather. Reg Sci Urban Econ 37:375–398

Rosen S (1974) Hedonic prices and implicit markets: product differentiation in pure competition. J Polit Econ 82:34–55

Rossi PH (1955) Why families move: a study in the social psychology of urban residential mobility. Free Press, Glencoe

Tiebout C (1956) A pure theory of local expenditures. J Polit Econ 64:416–424

Voith R (1991) Capitalization of local and regional attributes into wages and rents: differences across residential, commercial, and mixed-use communities. J Reg Sci 31:127–145

Winstanley A, Thorns DC, Perkins HC (2002) Moving house, creating home: exploring residential mobility. Housing Stud 17:813–832