ESTIMATING THE DEMAND FOR THE CHARACTERISTICS OF HOUSING

Raymond B. Palmquist*

Abstract—Estimates of the demand for the characteristics contained in a differentiated product such as housing can be based on hedonic regressions. However, previous studies have not dealt with several potential problems in such estimation. A more appropriate framework for the estimation is provided in this paper. Also data from a number of cities are used to eliminate the identification problem, and the endogeneity of the marginal prices derived from non-linear hedonic equations is considered in the estimation. The results on own and cross price elasticities, expenditure elasticities, and the effects of socio-economic variables accord well with expectations.

THERE has been considerable interest in understanding the markets for differentiated products. There are many goods that are not homogeneous and yet are traded on well-integrated markets. The utility provided by such goods is based on the utility yielded by the various characteristics of the differentiated good. For many years hedonic regressions have been used to study the contributions of the various characteristics to the price of the composite good. Rosen (1974) provided a widely-used model to explain the theoretical origins of the hedonic results. In that model there are demands and supplies for the various characteristics of a differentiated product.

Most of the research on estimating the underlying demands and supplies for characteristics has been done in housing markets. Residential housing is obviously a differentiated product, and obtaining the necessary data has been comparatively easy. Estimating the demands for housing characteristics has been desirable for a number of reasons. Just as information on the parameters of the demand functions in explicit markets is useful, this same information is important in implicit markets. This is particularly true in evaluating the numerous government programs that influence the quantity and form of housing in the United States. Knowledge of the demands for the characteristics of housing would also allow the formation of more reasonable aggregates measuring "housing services." Finally, some of the characteristics of housing have the nature of public goods, so information on the demands for these goods is not readily available except in these implicit markets.

Recently various problems with the estimation techniques used have been pointed out. The purpose of this paper is to implement an estimation technique that avoids these potential problems so that more reliable estimates of the parameters of the demand functions can be derived. The estimation insures identification of the demands without arbitrary assumptions by using data from a number of cities. The endogeneity of the implicit marginal price derived from a non-linear hedonic equation is considered in the choice of estimation technique. Finally, the adjustments that are necessary to make the estimated equations consistent with consumer behavior are undertaken.

The next section discusses the model and estimation techniques that have commonly been used in this area. An alternative view of the problem being estimated is then provided, and the techniques to be used here are discussed. The third section describes the data sources used. The fourth section provides the hedonic and demand results from the estimation and compares the elasticities derived here with those obtained in other studies.

Current Techniques and Potential Improvements

An early hedonic study of property values by Ridker and Henning (1967) generated both interest in the possibilities of hedonic housing studies and a lively debate over the correct interpretation of their results.[1] The main shortcoming with Ridker and Henning was that it was not based on a well-defined theoretical model that implied the correct interpretation of the estimates. Such a model was provided in Rosen (1974) and Freeman

Received for publication July 14, 1983. Revision accepted for publication December 20, 1983.

*North Carolina State University.

The author is grateful to Dennis Epple, Paul Johnson, Thomas Johnson, V. Kerry Smith, and an anonymous referee for helpful comments on an earlier version of this paper. The support provided by the U.S. Department of Housing and Urban Development under contract HC-5035 is also gratefully acknowledged.

[1] The main contributions to the debate can be found in Freeman (1971, 1974a), Polinsky and Shavell (1975), Small (1975), and Harrison and Rubinfeld (1978b).

(1974b). Following Rosen, differentiated products are assumed to be made up of various characteristics. These characteristics are not explicitly traded on markets, but the implicit marginal prices of these characteristics can be revealed by hedonic regressions. However, in general the estimation of such a regression reveals little information about the underlying demands and supplies of characteristics unless all consumers or all suppliers are identical. Rather, the equation represents the locus of equilibrium points. The hedonic equation is determined by the bids that consumers are willing to make for different bundles of characteristics and the offers of those bundles by suppliers.

The bid and offer functions are of more significance to the economist since they do not depend specifically on the temporal or spatial location of the market. Rosen proposed the following procedure for estimating the bid and offer functions. The hedonic price equation need not be linear in the absence of costless repackaging of the characteristics. In this case, the marginal prices of the characteristics will differ depending on the quantities of the characteristics purchased. Maximization of utility results in the marginal bid being equated to the marginal price, while maximization of profits results in the marginal offer being equated to the marginal price. Rosen proposed estimating the marginal bid and offer functions as with any simultaneous supply and demand system. Many economists have applied his suggestion empirically in the housing market. Witte, Sumka, and Erekson (1979) and Nelson (1978) have estimated supply and demand simultaneously. Harrison and Rubinfeld (1978a, 1978b); Bender, Gronberg, and Hwang (1980); and Blomquist and Worley (1981), on the other hand, have contended that the supply of characteristics is perfectly inelastic at any location and used ordinary least squares (OLS) to estimate the demand equation.

Despite this almost universal use of Rosen's model explaining equilibria to attempt to estimate individual demands in a demand and supply framework, in many cases there is a more appropriate model based on optimization. The typical study today is based on a cross-section of individual house purchasers. In this case, each consumer is a price-schedule-taker and cannot influence equilibrium prices.[2] This is somewhat different than a price-taker in the typical market for a homogeneous product because the consumer can influence the marginal price paid by varying the quantity of the characteristics purchased, but he cannot influence the overall price schedule. The problem reduces to the consumer maximizing utility subject to an exogenous non-linear budget constraint. The supply of characteristics was important in determining the hedonic price schedule but is exogenous for a given consumer.

A consumer with a vector of socio-economic characteristics α derives utility from the various characteristics of the house, $z_1, \dots, z_n$, and from the numeraire non-housing good, x. This utility function,

$$U = U(z_1, \dots, z_n, x, \alpha) \tag{1}$$

is maximized subject to the budget constraint

$$y = x + P(z_1, \dots, z_n). \tag{2}$$

This would be a standard consumer optimization problem except that the budget constraint may be non-linear.[3] Fortunately, just as the budget constraint is tangent to the indifference surface at the optimum point, a budget constraint that has been linearized around the optimum point is also tangent to the indifference surface at that point. Thus, converting the constraint to a linear form allows the use of conventional theory and estimation techniques (see Diewert, 1982, pp. 586–587 or Deaton and Muellbauer, 1980, pp. 247–248). Since consumers are price-schedule-takers, the uncompensated demand functions will have the typical form with the quantity demanded of a characteristic being a function of the marginal price of the characteristic itself, prices of related characteristics, expenditure, and socio-economic factors. This differs from the form utilized in studies following Rosen (1974) where the quantities of other characteristics rather than their prices are held fixed.

Linearizing the budget constraint avoids the extreme non-linearity that would be likely if one solved the maximization problem in equations (1) and (2) directly. However, the income or expenditure variable in the estimation must be adjusted so

[2] This point of supply exogeneity has been raised theoretically by Diamond and Smith (1982) and empirically by Bartik (1983). However, in both of these studies the inverse form for the demand equation was maintained as in Rosen (1974).

[3] Labor economists have been interested in estimating labor supply decisions with a piece-wise linear budget constraint. See Hausman (1983) and the works cited therein.

that the linearized budget constraint would still be tangent at the observed consumption point. This adjustment can be seen by adding $\sum_{i=1}^{n} P_{zi} z_i$ to both sides of equation (2) where P_{zi} is the marginal price of characteristic z_i. If $P(z)$ is then subtracted from each side of the equation the result is

$$y_a \equiv y - P(z) + \sum P_{zi} z_i = x + \sum P_{zi} z_i \qquad (3)$$

where y_a is the adjusted income. Failure to make this adjustment will result in biased estimates of the coefficients of the price variables in the demand regressions.

Even with the linearized budget constraint, the marginal prices of the characteristics are still endogenous. The marginal price of a characteristic depends on the quantities of the characteristics consumed. In this situation ordinary least squares will not provide consistent estimates. It is necessary to find instruments for the characteristics price variables and the adjusted income variable to obtain consistent estimates. All of the characteristics and their prices may be correlated with the error terms, so the obvious choices for forming instruments are income and the socio-economic variables.

A problem that has been previously ignored in empirical work on the demand for housing characteristics is the identification issue raised by Brown and Rosen (1982) and discussed in a non-housing context by Brown and Mendelsohn (1980). The two relevant sets of equations in the current context are the demand functions and the marginal price functions derived from the hedonic regressions. To insure the desired estimation, the demand functions must be identified. This identification could be accomplished in a variety of ways. If the correct functional form for the demand equations were known, identification would be obtained. Alternatively, if some of the characteristics did not enter the demand functions, identification might be possible. These two methods of identification have been used without comment in many of the empirical studies (e.g., Nelson, 1978, and Harrison and Rubinfeld, 1978a). The recent introduction of the use of flexible functional forms at both stages of the estimation (e.g., Bender et al., 1980) makes these methods of identification more difficult. Another method of identification has been used by Quigley (1982). He assumed a functional form for the utility function that is homothetic. A non-linear hedonic equation provided varying marginal prices. The homotheticity allowed scaling the purchases of individuals with differing income to yield many observations on a given indifference curve. Each of these techniques would be acceptable if the necessary assumptions corresponded to the "truth" about the demand or utility functions. Unfortunately, there is no "a priori" way to evaluate these assumptions.

An alternative method that avoids these assumptions has been advocated by Brown and Mendelsohn and by Brown and Rosen. This technique uses data from spatially or temporally distinct markets so that separate hedonic equations can be estimated for each market. This variation in the marginal price schedule between locations allows identification of the demand functions. An additional advantage of using data from multiple markets is that the hedonic regressions need not be arbitrarily restricted to be non-linear for demand estimation.

Identical Data for Multiple Markets

In this research it is assumed that there is no market segmentation within an urban area, since there is mobility among housing types and locations and little evidence of price discrimination. Even racial barriers were breaking down before 1977. However, the costs of moving between cities prevent the complete integration of real estate markets between cities. It is further assumed that differences in consumers within and between cities are measurable and may be controlled for. Given these assumptions it is possible to estimate separate hedonic equations for each of several cities.[4] This yields distinct price surfaces in each of the cities which makes it possible to identify the demand curves without imposing assumptions on the structure of the model.

Existing studies have relied on census data (e.g., Nelson, 1978; Harrison and Rubinfeld, 1978a) or local data sets (e.g., Witte et al., 1979) for information on housing characteristics. For a study that considers a number of urban areas, one would have difficulty obtaining local data that were comparable in all aspects. The census data are available throughout the nation, but only census tract

[4] This procedure was originally proposed by Freeman (1974b), but he later stated that multiple cities were not necessary for the estimation (Freeman, 1979).

averages are reported, and few of the necessary characteristics are represented.

For these reasons a new data base had to be developed for this research. Adequately specifying structure and lot characteristics requires information on individual dwellings and the locations of those dwellings. Because there is some debate about the accuracy of owner or assessor estimates of house values, data on actual sales were desired. In addition it is necessary to have information on the purchaser of the house. The only national data set that was available with these attributes was the Federal Housing Administration records on mortgages insured by FHA. These FHA tapes contain information on the selling price, a wide variety of the characteristics of the house, the income and wealth of the purchaser, and a variety of other attributes of the purchasing individual or family. In more recent years the tapes also include information on the census tract in which the house is located.

The next type of data necessary was information on neighborhood characteristics. This was obtained from the "Special Area Profiles" issued by the Bureau of the Census and based on the Census of Population and Housing. It makes available, by census tract, data on characteristics of the population such as ethnic background and education as well as the average quality of the houses in the census tract.

Since the primary emphasis of the current research is on residential location decisions, an important component of the data set concerned the location of the houses. The FHA data give the census tract in which the house is located. Converting this to a spatial location required the use of the Census Bureau's Urban Atlas Files. For each census tract these files give the latitudes and longitudes of every corner of the census tract boundary including "islands" and "holes" in census tracts. This boundary information was used to calculate geographical centroids and areas for each census tract.

The locations of air quality monitoring stations were obtained from the Environmental Protection Agency's Directory of Air Quality Monitoring Sites. Air quality data were available from the EPA's Storage and Retrieval of Aerometric Data (SAROAD) tapes.

The final information necessary was the location of employment and commercial activity within the urban areas. A source of such data that covers a number of cities is the portion of the Annual Housing Survey concerning "Travel-to-Work." This supplement to the AHS is sponsored by the Department of Transportation and provides, by census tract, the origin and destination of all work trips taken by the interviewed households.

Empirical Results

The estimation results are for seven Standard Metropolitan Statistical Areas for which data were available from all sources. The cities used in this study are Atlanta (2863 sales), Denver (4775), Houston (3185), Louisville (2050), Miami (2390), Oklahoma City (3058), and Seattle (1976). The sample is restricted to single-family, owner-occupied houses. The data are for mortgages recorded in 1977. Since some mortgages recorded then were for sales in 1976, a dummy variable was included to indicate the earlier sales.

The variables used for the hedonic regressions are defined in table 1. The selection of most of the variables was based on previous experience with hedonic regressions and the experiences of appraisers rather than experimenting with the current data set. This is desirable so that the usual statistical tests can be correctly interpreted. It was also desirable to have the specifications comparable between cities which would have been impossible if the specifications had been tailored to the individual cities. The only exception to using identical variables in all cities was where a characteristic was not present or present in less than 2% of the houses in the sample in a city. In this case the variable for that characteristic was omitted for those SMSAs.

One characteristic requires special attention. It would be anticipated that the number of square feet of living space would not simply have a linear effect on price. As the number of square feet increases, construction costs do not increase proportionally since such items as wall area do not typically increase proportionally. Appraisers have long known that price per square foot varies with the size of the house. To allow for these nonlinearities the square of the living area was also included in the regressions.

Information on the choice of functional form for the hedonic equations was obtained for each city. To reduce the costs of estimation, the search

TABLE 1.—VARIABLE DEFINITIONS IN THE HEDONIC AND DEMAND REGRESSIONS

Name	Description	Source[a]
LOTAREA	Area of lot in square feet	
IMPAREA	Finished interior area in square feet	
SQSQFT	Finished interior area squared	
BATHS	Number of bathrooms	
YRBUILT	Year of construction	
GARCAP1	Number of stalls in garage	
GARCAP2	Number of stalls in carport	
DETGAR	Dummy variable, 1 if garage is detached from house	
UNDG	Dummy variable, 1 if there is underground wiring	
DW	Dummy variable, 1 if there is a dishwasher	
DSIP	Dummy variable, 1 if there is a garbage disposal	
CAC	Dummy variable, 1 if there is central air conditioning	1
WAC	Dummy variable, 1 if there are wall air conditioning units	
FAN	Dummy variable, 1 if there is a ceiling fan	
D76	Dummy variable, 1 if the date of sale was 1976	
EXC	Excellent condition	
FAIR	Fair condition	
POOR	Poor condition	
BRICK	Brick or stone exterior finish	
BSMTFL	Dummy variable, 1 if there is a full basement	
BSMTPT	Dummy variable, 1 if there is a partial basement	
FRPL	Dummy variable, 1 if there are one or more fireplaces	
SWIM	Dummy variable, 1 if there is a swimming pool	
TSP	The annual arithmetic mean of the particulate air pollution level	2
MEDAGE	The median age of the residents of the census tract	
MEDINC	The median family income of residents of the census tract	
BLUECOL	The percentage of workers in the census tract that has a blue collar job	
MOVENEW	The percentage of houses in the census tract that has changed ownership within the last five years	3
NONWHITE	The percentage of the population of the census tract that is classified as non-white	
HIGRAD	The percentage of the population of the census tract over 24 years old that has graduated from high school	
LESS1	The percentage of the structures in the census tract with 1.00 or less persons per room	
FREQSQMI	The number of work destinations within the census tract divided by the area of the census tract	4
TEXPADJ	Adjusted monthly housing expenditure	1
PSQFT	Hedonic price of square feet of living space	
PBATHS	Hedonic price of bathrooms	
PHIGR	Hedonic price of the percentage of the census tract with high school degrees	5
PWH	Hedonic price of racial homogeneity	
PLOT	Hedonic price of lot area	
PYRBLT	Hedonic price of reduction in age of house	
AGE	Age of the purchaser	
UNMAR	Dummy variable, 1 if the purchaser is single	
DEPEND	Number of dependents in the family making the purchase	1
BLACK	Dummy variable, 1 if the purchaser is black	

[a] See appendix.

was restricted to the four functional forms most frequently used: linear, semi-logarithmic, log-linear and inverse semi-logarithmic. The choice can be based on maximizing the likelihood function in Box and Cox (1964). For the four forms considered, such selection is identical to minimizing the residual sum of squares for regressions where the dependent variable has been standardized by multiplying by the inverse of the geometric mean of that variable. While the selection can be based on this rule, statistical tests are not readily available because the alternative hypotheses are non-nested.

The linear form was preferable in all urban areas except Miami where the semi-logarithmic form was used.

The results of the hedonic regressions are given in table 2. Almost all of the coefficients have the expected signs and magnitudes. With approximately 200 coefficients estimated, there are only 17 with incorrect signs and none of these are for the most important variables.

The linear term for improved living area consistently has a coefficient that is positive, strongly significant, and of a reasonable magnitude. The

TABLE 2.—RESULTS OF HEDONIC REGRESSIONS

Variable	Atlanta	Denver	Houston	Louisville	Miami[a]	Oklahoma City	Seattle
LOTAREA	0.0813	0.1474	0.0998	0.0745	3.20×10^{-6}	0.1423	0.6542
	(6.019)	(7.908)	(4.825)	(2.938)	(2.841)	(7.542)	(3.113)
IMPAREA	15.0576	12.7203	12.7237	8.4252	5.51×10^{-4}	8.6116	17.2921
	(8.662)	(11.474)	(8.834)	(3.202)	(9.539)	(8.646)	(8.653)
SQSQFT	−0.0022	−0.0019	−0.0002	−0.0012	-9.69×10^{-8}	0.0007	−0.0032
	(−3.853)	(−4.552)	(−0.355)	(−1.045)	(−4.447)	(1.986)	(−4.027)
BATHS	1821.3200	1881.8610	477.7357	3611.4540	3.73×10^{-2}	1169.3990	2527.3200
	(7.478)	(16.122)	(2.131)	(11.104)	(6.448)	(6.353)	(12.086)
YRBUILT	134.4473	79.3402	111.4022	71.2724	2.15×10^{-3}	106.0038	101.9034
	(11.071)	(12.968)	(10.088)	(8.113)	(6.183)	(14.182)	(13.643)
GARCAP1	1451.0940	2198.9280	1838.5800	1602.4330	5.91×10^{-2}	1694.6060	1319.1420
	(8.780)	(25.204)	(17.240)	(8.490)	(9.812)	(18.477)	(9.712)
GARCAP2	1198.0810	601.4742	682.5717	999.5843	2.33×10^{-2}	1097.1160	483.6459
	(8.497)	(2.952)	(3.796)	(2.691)	(4.753)	(4.854)	(2.520)
DETGAR	−1006.9100	−820.9986	−739.4174	−409.3972	-2.83×10^{-2}	−1277.0800	−479.6224
	(−2.613)	(−4.251)	(−3.571)	(−1.332)	(−1.466)	(−6.732)	(−1.776)
UNDG	710.0944	510.1540	1239.9450	2156.1050	-5.80×10^{-3}	449.7995	672.2081
	(1.048)	(3.720)	(7.725)	(4.536)	(−0.422)	(2.008)	(2.960)
DW	1710.1180	984.5379	1153.7380	2027.1380	1.80×10^{-2}	1028.8940	1006.5220
	(8.225)	(7.904)	(5.982)	(5.420)	(1.911)	(6.972)	(4.721)
DISP	292.5529	473.8454	783.4335	1214.1630	4.81×10^{-2}	866.3541	696.6563
	(0.995)	(2.991)	(3.864)	(2.610)	(4.006)	(6.043)	(3.102)
CAC	1937.3910		1998.0340	2113.5660	5.78×10^{-2}	1606.4410	
	(9.648)		(9.196)	(9.125)	(7.205)	(11.945)	
WAC	604.6657		984.1632	642.5249	2.80×10^{-2}	285.4088	
	(1.662)		(4.828)	(2.012)	(5.446)	(1.314)	
FAN	344.7140	570.0075	−165.0138	977.5475	1.75×10^{-2}	560.0915	300.8057
	(1.979)	(4.361)	(−1.083)	(4.548)	(1.470)	(2.348)	(1.567)
D76	−1114.5000	−2432.4900	−1758.7300	−1179.6900	-2.17×10^{-2}	−1616.0800	−2207.6000
	(−5.285)	(−16.173)	(−13.891)	(−5.535)	(−4.294)	(−13.076)	(−14.148)
EXC	1007.5020	1434.4560	759.2975	384.2958	2.75×10^{-2}	1084.7870	1243.1500
	(2.936)	(10.138)	(3.497)	(0.803)	(2.316)	(7.494)	(3.937)
FAIR	−2227.3700	−2095.1300	−1352.8500	−2538.3400	2.85×10^{-2}	−1042.0700	−1626.3600
	(−4.792)	(−7.796)	(−5.864)	(−3.892)	(4.372)	(−5.005)	(−4.606)
POOR		−4316.1300		−3390.7400	-9.19×10^{-2}	−8880.1200	−153.1606
		(−3.691)		(−1.106)	(−4.748)	(−3.637)	(−0.051)
BRICK	622.3300	979.6494	1568.2720	2390.8420		1241.0530	3981.1320
	(3.760)	(7.5045)	(7.895)	(10.714)		(9.277)	(4.386)
BSMTFL	1859.1940	2229.4430		3219.7330			3712.4100
	(7.320)	(17.509)		(14.699)			(19.726)
BSMTPT	1108.2920	2218.8050		2201.4890			2748.4830
	(4.746)	(12.789)		(6.450)			(9.829)
FRPL	1114.5690	2118.6430	2418.9860	1604.1510		2416.3650	1334.0850
	(5.150)	(16.161)	(12.852)	(5.050)		(15.070)	(7.903)
SWIM	3274.7250				0.1080	3436.9250	
	(3.029)				(7.946)	(3.448)	
MEDAGE	−58.1812	49.0941	119.2348	47.7201	6.19×10^{-3}	2.8702	−108.9760
	(−1.897)	(4.109)	(4.105)	(1.610)	(11.395)	(0.215)	(−3.468)
MEDINC	0.0788	0.1655	−0.0044	0.0249	2.42×10^{-5}	−0.0976	0.3584
	(0.747)	(6.547)	(−0.063)	(0.299)	(9.558)	(−2.080)	(3.488)
BLUECOL	52.0316	−15.0316	27.0144	−44.1273	1.76×10^{-3}	−51.5020	−76.8352
	(2.387)	(−1.486)	(2.839)	(−2.055)	(3.412)	(−5.036)	(−4.727)

TABLE 2. —(*continued*)

Variable	Atlanta	Denver	Houston	Louisville	Miami[a]	Oklahoma City	Seattle
MOVENEW	−32.4515	−61.5007	−30.7874	5.9894	-7.65×10^{-4}	−2.7746	−30.8822
	(−4.524)	(−16.808)	(−5.924)	(2.620)	(−3.487)	(−0.588)	(−3.728)
NONWHITE	−1516.5000	−4465.6700	−2455.1100	−5561.1600	4.06×10^{-2}	−3412.1300	−6155.9100
	(−3.335)	(−9.981)	(−6.422)	(−10.650)	(1.907)	(−8.405)	(−8.862)
HIGRAD	1.2341	0.3271	0.4575	0.9166	1.71×10^{-5}	0.2563	−0.3741
	(6.434)	(3.633)	(5.179)	(4.735)	(3.287)	(2.824)	(−1.837)
*LESS*1	35.9097	64.3941	80.8062	−16.2552	-1.36×10^{-3}	−39.3587	199.9203
	(4.165)	(14.615)	(4.203)	(−3.834)	(−2.180)	(−1.833)	(4.223)
FREQSQMI	16.6915	13.9396	26.8967	3.7791	1.09×10^{-3}	−7.0332	8.7971
	(1.816)	(1.501)	(2.201)	(0.513)	(2.669)	(−1.119)	(1.458)
TSP	−45.4710	−26.0403	−11.8616	−1.5987	-3.05×10^{-3}	−0.2232	−8.5865
	(−5.055)	(−7.192)	(−3.215)	(−0.482)	(−11.347)	(−0.118)	(−1.477)
INTERCEPT	−9337.3200	4398.5110	−12156.8000	1116.2100	9.3257	2901.1820	−9526.0500
	(−3.519)	(3.407)	(−5.666)	(0.433)	(102.618)	(1.217)	(−2.069)
R^2	0.7993	0.7330	0.8708	0.8003	0.6484	0.9001	0.8238

Note: t-ratios are given in parentheses below parameter estimates.
[a]Semi-logarithmic form by Box-Cox test.

coefficient for the squared term for that variable is negative in six of the seven cities and is frequently statistically significant. Combining these two terms indicates that increasing living space significantly increases the price of the house but at a decreasing rate as hypothesized. The coefficients of the number of bathrooms and the year built are always positive, significant, and generally of reasonable magnitude.

The relationships among the variables also accord with expectations. For example, the number of carport stalls is always significantly positive, but, more importantly that coefficient is always smaller than the coefficient for the number of garage stalls. In addition, if a given garage is detached, its value is reduced. Similarly, central air conditioning is consistently positive and of a reasonable magnitude, and the coefficient is always larger than the coefficient of wall air conditioning units. The same relationship holds for full basements and partial basements. Overall, the structural variables perform remarkably well.

Because a number of cities were studied it is also interesting to note consistencies and inconsistencies in the non-structural characteristics coefficients across cities. Several different forms of the accessibility variable were tried, with some surprising results. The traditional theoretical result in urban economics is that land values and property values should be negatively related to the distance from the central business district. These models assume, however, that all employment is in the CBD. When the distance to the CBD was used as a measure of accessibility, the signs and significance of the coefficients fluctuated between cities. This lack of a stable relationship existed whether the variable was entered in linear or nonlinear form. This is not surprising because of the significant dispersion of work places today. The DOT-AHS data set allows the determination of the locations of these diverse work centers. The measure of accessibility that performed most successfully was the number of work destinations per square mile in the census tract of residence.

Several of the other neighborhood characteristics are worth noting. The percentage of the population of the surrounding census tract that was non-white proved to have a strongly negative effect on property values in all but one of the urban areas. Generally, neighborhood quality as indicated by the percentage of the population of the census tract with a high school degree increases property values, as do less crowded structures in the neighborhood (*LESS*1). Neighborhoods with more transient populations typically have lower property values (*MOVENEW*). Air quality in this data set was measured only by annual average concentrations of suspended particulate matter. The most consistent explanatory power was obtained when the annual arithmetic mean at the nearest monitoring station was used. However, even in this case the coefficients were statistically insignificant in three of the seven cities.[5]

[5] The purpose of this paper is to implement a new estimation technique for determining the demand for the characteristics of

The partial derivatives of the hedonic equations with respect to the characteristics were then used to provide marginal prices for the characteristic demand equations. Because of the large size of the data set in each city random samples were taken, and the data set for the demand estimation was formed by merging these samples. The variables used in these second-step regressions are also defined in table 1.

Because of the squared term for living space in the hedonic regressions, as well as the non-linear equation for Miami, the marginal prices of the characteristics often depend on the quantities consumed. Thus, all prices that are non-constant within an urban area are endogenous in the demand regressions. When marginal prices are non-constant it is also necessary to adjust the expenditure variable if only marginal prices are used in estimating demand. This adjusted expenditure variable is also endogenous. It was assumed that expenditures on housing were weakly separable from other consumer expenditures, so a measure of total housing expenditure as calculated by FHA was adjusted as in equation (3) to form *TEXPADJ*. Instruments for the nonconstant marginal characteristics prices and *TEXPADJ* were developed by regressing the endogenous variables on all linear and quadratic terms in the exogenous socioeconomic variables and a set of dummy variables for the urban areas. These equations were then used to calculate predicted values for the endogenous variables. These predicted values were used as instrumental variables.

The second-stage regression of greatest interest concerns the demand for square feet of living space. The results are provided in table 3, column 1. The results conform well with expectations. The price per square foot of living space has a signifi-

housing. For this reason the paper focuses on some of the more important characteristics that determine housing prices. It would have been desirable also to estimate the demand for air quality to allow comparison with the considerable research that has been done on this topic. However, the data set used here was not well adapted to such a study, and pursuing this line of research did not look promising. The variation in air quality within several of the cities was limited. In addition, two of the cities were in compliance with the EPA standards for total suspended particulates and yet had large and statistically significant negative coefficients on TSP. This was probably because TSP was the only pollutant in the data set, yet several other pollutants also have important and perceptible effects on residents. These other pollutants are correlated with TSP but not perfectly. For these reasons the estimation of the demand for air quality with the current data set was not fruitful. A study has subsequently been undertaken using data on four major air pollutants and cities that are better adapted to an air pollution study (Palmquist, 1983).

TABLE 3.—LOG-LINEAR DEMANDS FOR HOUSING CHARACTERISTICS

Variable	*IMPAREA*	*BATHS*	*HIGRAD*	*WH*
TEXPADJ	0.491	0.612	0.305	0.257
	(23.641)	(15.290)	(7.753)	(4.116)
*AGE*1	4.158×10^{-4}	2.421×10^{-4}	-4.813×10^{-4}	-1.060×10^{-3}
	(0.905)	(0.273)	(−0.553)	(−0.766)
UNMAR	-1.024×10^{-2}	-8.432×10^{-3}	3.576×10^{-2}	1.928×10^{-2}
	(−0.966)	(−0.413)	(1.780)	(0.604)
DEPEND	1.173×10^{-2}	7.396×10^{-3}	-5.004×10^{-3}	-8.171×10^{-3}
	(3.550)	(1.162)	(−0.800)	(−0.822)
BLACK	4.300×10^{-2}	6.291×10^{-2}	-2.988×10^{-2}	−0.448
	(3.463)	(2.629)	(−1.271)	(−11.992)
PSQFT	−0.962	−1.329	-9.458×10^{-2}	−0.355
	(−9.925)	(−7.116)	(−0.515)	(−1.215)
PBATHS	−0.172	−0.228	0.259	3.198×10^{-2}
	(−9.164)	(−6.316)	(7.300)	(0.567)
PLOT	-5.561×10^{-2}	6.391×10^{-2}	0.374	-1.677×10^{-2}
	(−2.971)	(1.772)	(10.541)	(−0.298)
PYRBLT	0.440	0.644	0.393	0.146
	(13.736)	(10.433)	(6.475)	(1.509)
PHIGR	−0.196	−0.313	−0.291	−0.104
	(−5.984)	(−4.975)	(−4.710)	(−1.060)
PWH	−0.198	−0.217	−0.124	−0.142
	(−8.223)	(−4.668)	(−2.722)	(−1.954)
INTERCEPT	7.139	0.492	5.400	−0.444
	(14.742)	(0.528)	(5.887)	(−0.305)

Note: *t*-ratios are in parentheses below parameter estimates. Instruments are used for all non-constant marginal characteristics prices and *TEXPADJ*.

cantly negative effect on the quantity of living space demanded. The coefficient of expenditure was positive and strongly significant as well. The coefficients of the prices of all other characteristics included were significantly different from zero, an important finding since most previous studies have only included the price of the characteristic being considered. In most cases theory does not suggest the sign of the coefficients of these other prices. However, it seems reasonable that living space and bathrooms would be complementary, and this is confirmed. In fact, most of the other characteristics are complementary with living space. The number of dependents has a significant positive effect on demand for living space as expected. Black purchasers have a significantly higher demand for living space, a result that agrees with the findings of Kain and Quigley (1975), p. 241.[6] This result should not be confused with the issue of racial differences in home ownership rates or the consumption of neighborhood quality.

The estimation results for the demand equations for other characteristics are also given in table 3. In addition to living space, the number of bathrooms is an important structural characteristic of housing. The results for this equation are qualitatively similar to the equation for living space. The own-price coefficient is significantly negative while the expenditure coefficient is significantly positive. Blacks have a significantly greater demand for this characteristic (which also accords with the findings of Kain and Quigley). The own-price elasticity of demand for bathrooms is more inelastic than the comparable elasticity for living space, a reasonable result since a smaller fraction of total housing expenditure is devoted to this characteristic.

The variable for the percentage of adults in the neighborhood that had at least completed high school was used as one measure of neighborhood quality. The results for the demand for this quality confirm the earlier findings. The own-price elasticity is again significantly negative and smaller in absolute value than the living space elasticity. The expenditure elasticity is significantly positive, but the other socio-economic characteristics are no longer significant.

[6] Kain and Quigley (1975) do not include characteristics prices in their demand equations (pp. 231–232).

Finally, the results for the measure of racial homogeneity are reported in the last column of table 3. The price coefficient is again negative, but this time it is not quite significant at the 0.05 level (.051). Expenditure elasticity is significantly positive.

When the demand equations are estimated with monthly income substituted for total housing expenditure, the income elasticities are all lower than the expenditure elasticities, but still significant. This is as expected since the best current estimates of the income elasticity of demand for aggregate "housing services" are below one (Polinsky, 1977). This means that an increase in income would result in a less than proportional increase in income allocated to housing expenditure. This new housing expenditure would then be allocated according to a model similar to that described in this paper. Thus income elasticities would be less than expenditure elasticities.

There have been surprisingly few previous empirical studies of the demand for housing characteristics except for studies concentrating on air quality. There appear to be four studies where at least approximate comparisons can be made: King (1976); Witte, Sumka, and Erekson (1979); McMillan (1979); and Blomquist and Worley (1981). With the exception of Witte et al., these studies all used data from a single city so the validity of the demand estimates depends on the appropriateness of the implicit identifying assumptions. All of the studies either are based on aggregates of the characteristics rather than the characteristics themselves or use Census data. Thus direct comparison of the results of this study is not possible. However, rough comparisons of the price and income or expenditure elasticities of demand for living space are enlightening. Three of the studies found very small or insignificant price elasticities, whereas McMillan found the opposite with price elasticities ranging to -7.57. The estimated income elasticities were not significantly different from zero in Witte et al. and Blomquist and Worley. King estimated the expenditure elasticity to be 0.64 which is closer to this study, while McMillan estimated it to be as large as 2.03. It appears that the current study has obtained estimates that accord better with expectations than have studies that ignored the identification and endogeneity issues.

Conclusions

This paper has estimated the demands for the more important characteristics of housing using techniques that avoid the identification and endogeneity problems which have been overlooked in previous empirical studies. By using data from a number of cities it is possible to identify the demand curves without imposing arbitrary assumptions. It is also true that even when studies correctly ignore the supply side of the market, prices are endogenous whenever the hedonic equations are non-linear. Consistent estimates of the parameters in the face of this endogeneity have been made.

The demand equation estimates behave remarkably well. The coefficients have the expected signs and magnitudes and are almost all highly significant. The price elasticity of demand for living space is approximately unitary, while the demands for the other characteristics are more inelastic. The cross-price effects are significant, so studies that have omitted them may have specification bias. The expenditure and income elasticities are found to be somewhat inelastic. These results appear to be more reasonable than those in other studies that have estimated the demands for characteristics.

APPENDIX

Data Sources

1. Master Statistical Files, Tape F-31, 1977, Management Information Systems Division, Federal Housing Administration, U.S. Department of Housing and Urban Development.
2. *Directory of Air Quality Monitoring Sites Active in 1977* (EPA-450/2-78-048) and related computer tapes, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency.
   Air Quality data—1977 Annual Statistics (EPA-450/2-78-040) and related computer tapes, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency.
   Urban Atlas Tract Boundary Files, 1970 Census of Population and Housing, Data User Services Division, U.S. Bureau of the Census.[a]
3. Special Area Profiles, 1970 Census of Population and Housing, Data User Services Division, U.S. Bureau of the Census.
4. Travel-to-work Tract Data, Annual Housing Survey Supplement, 1975 and 1976 (C-AHS 76003), U.S. Department of Transportation.
   Urban Atlas Tract Boundary Files, see source 2 above.
5. Hedonic prices estimated in this paper or instruments for those prices.

[a] The use of the data in these files to calculate the geographic centroid and area of each census tract is described in more detail in an appendix available from the author.

REFERENCES

Bartik, Timothy J., "The Estimation of Demand Parameters in Hedonic Price Models," Vanderbilt University (1983).

Bender, Bruce, Timothy J. Gronberg, and Hae-Shin Hwang, "Choice of Functional Form and the Demand for Air Quality," this REVIEW 62 (Nov. 1980), 638–643.

Blomquist, Glenn, and Lawrence Worley, "Hedonic Prices, Demands for Urban Amenities, and Benefit Estimates," *Journal of Urban Economics* 9 (1981), 212–221.

Box, G. E. P., and D. R. Cox, "An Analysis of Transformations," *Journal of the Royal Statistical Society, Series B* 26 (1964), 211–252.

Brown, Gardner, Jr., and Robert Mendelsohn, "The Hedonic-Travel Cost Method," this REVIEW 66 (Aug. 1984).

Brown, James N., and Harvey S. Rosen, "On the Estimation of Structural Hedonic Price Models," *Econometrica* 50 (May 1982), 765–768.

Deaton, Angus, and John Muellbauer, *Economics and Consumer Behavior* (Cambridge: Cambridge University Press, 1980).

Diamond, Douglas B., and Barton A. Smith, "Simultaneity in the Market for Housing Characteristics," Department of Housing and Urban Development (1982) (forthcoming in the *Journal of Urban Economics*).

Diewert, W. Erwin, "Duality Approaches to Microeconomic Theory," in Kenneth J. Arrow and Michael D. Intrilligator (eds.), *Handbook of Mathematical Economics* (Amsterdam: North-Holland Publishing Company, 1982).

Epple, Dennis, "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products," Carnegie-Mellon University (1982).

Freeman, A. Myrick III, "Air Pollution and Property Values: A Methodological Comment," this REVIEW 53 (Nov. 1971), 415–416.

______, "Air Pollution and Property Values: A Further Comment," this REVIEW 56 (Nov. 1974a), 554–556.

______, "On Estimating Air Pollution Control Benefits from Land Value Studies," *Journal of Environmental Economics and Management* 1 (May 1974b), 74–83.

______, *The Benefits of Environmental Improvement* (Baltimore: Johns Hopkins Press, 1979).

Harrison, David, Jr., and Daniel L. Rubinfeld, "Hedonic Housing Prices and the Demand for Clean Air," *Journal of Environmental Economics and Management* 5 (Mar. 1978a), 81–102.

______, "The Air Pollution and Property Value Debate: Some Empirical Evidence," this REVIEW 60 (Nov. 1978b), 635–638.

Hausman, Jerry A., "Taxes and Labor Supply," NBER Working Paper #1102, forthcoming in A. Auerbach and M. Feldstein (eds.), *Handbook of Public Finance* (1983).

Kain, John F., and John M. Quigley, *Housing Markets and Racial Discrimination: A Microeconomic Analysis* (New York: National Bureau of Economic Research, 1975).

King, A. Thomas, "The Demand for Housing: A Lancastrian Approach," *Southern Economic Journal* 43 (Oct. 1976), 1077–1087.

McMillan, Melville L., "Estimates of Households' Preferences for Environmental Quality and Other Housing Characteristics from a System of Demand Equations," *Scandinavian Journal of Economics* 81 (1979), 174–187.

Nelson, Jon P., "Residential Choice, Hedonic Prices, and the Demand for Urban Air Quality," *Journal of Urban Economics* 5 (July 1978), 357–369.

Palmquist, Raymond B., "Estimating the Demand for Air Quality from Property Value Studies," North Carolina State University (1983).

Polinsky, A. Mitchell, "The Demand for Housing: A Study in Specification and Grouping," *Econometrica* 45 (Mar. 1977), 447–461.

Polinsky, A. Mitchell, and Steven Shavell, "The Air Pollution and Property Value Debate," this REVIEW 57 (Feb. 1975), 100–104.

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

Ridker, Ronald G., and John A. Henning, "The Determinants of Residential Property Values with Special Reference to Air Pollution," this REVIEW 49 (May 1967), 246–257.

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

Small, Kenneth A., "Air Pollution and Property Values: Further Comment," this REVIEW 57 (Feb. 1975), 105–107.

Witte, Ann D., Howard J. Sumka, and Homer Erekson, "An Estimation of a Structural Hedonic Prices Model of the Housing Market: An Application of Rosen's Theory of Implicit Markets," *Econometrica* 47 (Sept. 1979), 1151–1173.