# 4

# Specifying the Demand for Housing Characteristics: The Exogeneity Issue

GLENN BLOMQUIST    LAWRENCE WORLEY

## I. HEDONIC PRICES AND DEMAND FOR HOUSING CHARACTERISTICS

The rapidly growing literature on the demand for housing characteristics has begun to take into account the complex and subtle nature of the implicit market for these characteristics.[1] Although early research attempted to infer demand directly from the hedonic housing equation, it is now more common to interpret the equation as a market-clearing price function. As such, it is influenced by the suppliers as well as the demanders of housing. Thus, the standard practice is to estimate the demands employing a two-step procedure.

Following Rosen (1974), housing is viewed as a bundle of characteristics and estimating the demand for characteristics involves (*a*) estimating a hedonic housing equation and determining the implicit prices of the characteristics; and (*b*) using these implicit prices along with information about the consumers and producers of housing to estimate and, if necessary, identify the demand equations for these characteristics. Freeman (1979) observes that in only two special cases is there agreement concerning demand estimation. One case is where the hedonic is linear. There is no variation in the prices of characteristics, and demand (at least as we usually define it) cannot be estimated. The other case is where the hedonic is nonlinear and where all consumers have the same utility func-

---

[1] For a critical review of this research on housing markets see Smith (1980).

ISBN 0-12-214840-1

tion and have equal values of all arguments in the utility function. In this second case, the demand can be obtained directly from the hedonic equation. The demand is the marginal function obtained from the total willingness-to-pay function—the hedonic housing equation.

The treatment of housing supply particularly lacks harmony. Harrison and Rubinfeld (1978) and Bender, Gronberg, and Hwang (1980)—investigating the demand for the amenity, clean air in the neighborhood—assume that the supply of houses with some specified air quality is completely inelastic and that it is completely inelastic at any specified air quality. Using this assumption, they estimate the inverse demand for clean air. Linneman (1977)—investigating the demands for rooms, access, and neighborhood homogeneity—and Blomquist and Worley (1981)—investigating the demands for several neighborhood amenities and rooms—assume that the supply of houses with some specified quantity of a characteristic is perfectly elastic and that this holds for any quantity of the characteristic. Accordingly, they estimate the ordinary (quantity is the dependent variable) demand for various housing characteristics. Nelson (1978) and Witte, Sumka, and Erekson (1979) assume that the supply is neither completely inelastic nor perfectly elastic and estimate demand and supply simultaneously.[2] Which specification is appropriate depends upon the housing characteristic whose demand is being estimated and the supply mechanism for the characteristic. Misspecification will lead to biased estimates since the regressor that is assumed to be independent is, in fact, correlated with the error term. Unnecessary estimation of a simultaneous system may introduce problems associated with the choice of instrumental variables.

Freeman (1979, p. 166) argues that the correctness of each assumption depends on the speed of adjustment in supplying housing characteristics relative to the speed at which housing prices adjust to supply. He emphasizes that additional housing amenities, such as clean air, can be supplied through either additional amenities for existing houses or additional houses in areas with a higher level of amenities. If demand adjusts relatively more quickly, then the quantities of the particular bundle of housing characteristics can be assumed to be exogenous and demand can be estimated inversely. If supply adjusts relatively more quickly, then the

---

[2] To some extent, most of the studies that estimate the value of housing amenities do consider the importance of the supply assumption. Harrison and Rubinfeld do check for simultaneous equation bias by estimating the demand for clean air using 2SLS. Linneman uses a Wu test to check for endogeneity and, where appropriate, employs instrumental variables. Blomquist and Worley estimate benefits of increased housing characteristics under several assumptions. An exception is Brown and Pollakowski (1977), who assume fixed supplies and estimate the value of shoreline from the hedonic housing equation.

implicit prices can be assumed to be exogenous and demand can be estimated in the usual way (i.e., with quantity consumed as a function of price). If the speeds of adjustment are approximately equal, then demand must be estimated in a way that accounts for the simultaneity. Freeman's conclusion, which is the point of departure for this chapter, is that the question of appropriate specification is essentially an empirical one.

In Section II we estimate the hedonic housing equation for a particular housing market using a maximum likelihood search procedure to determine the best functional form of the hedonic and whether or not it is linear. After calculating the implicit prices for the housing characteristics, we proceed in Section III to estimate the demand for each characteristic, again determining the best functional form using a maximum likelihood search procedure. In Section IV we test for the relative exogeneity of implicit prices and quantities of the housing characteristics. We present our conclusions about the supply conditions for each housing characteristic, the nature of the implicit market, and implications for measuring the value of housing amenities in Section V.

## II. THE HEDONIC HOUSING EQUATION

In this section we estimate for residences (bundles of housing described by the vector of characteristics $\mathbf{Z}$), the hedonic price function $P(\mathbf{Z})$ and calculate the implicit prices ($P_{Z_i}$) for each housing characteristic ($Z_i$). The data are composed of observations from the 1970 U.S. Census for Springfield, Illinois. The hedonic equation is estimated using block data (Third Count) for housing prices and housing traits. The demand equations are estimated for block groups (Fifth Count) using an average of the implicit prices for the blocks within the group as well as income and taste variables for the block group.[3] In our sample there are 199 blocks and 38 block groups. Table 4.1 reports the summary statistics for housing prices, housing characteristics, and income and taste variables for the southern section of Springfield.

Given the interest in the functional form of the hedonic equation (with respect to variation in the first partial derivatives), we follow the statistical model of Box and Cox, which permits investigation of the functional specification through search for the best fit as measured by the log likelihood function. The search is limited to power transformations of the

---

[3] Suppression of block data for income and taste variables for reasons of confidentiality prevents estimating the demand equations using block data. We are limited to owner-occupied housing for the same reason.

**TABLE 4.1**
**Summary Statistics for 1970 Housing Data for Springfield, Illinois, by Block**

| Variable | Definition | Mean[a] | Standard deviation[a] |
|---|---|---|---|
| PHOUS | Reported property value (median) | \$19,670 | 9,013 |
| ROOM | Rooms per house | 5.020 rooms | .689 rooms |
| HBAS | Houses with basement | 66.31% | 25.57% |
| HPLB | Houses with adequate plumbing | 98.09% | 4.943% |
| DPP | Distance from electric power plant | 12,140 ft | 5,412 ft |
| NLAK | Proximity to Lake Springfield[b] | 6,182 ft | 4,065 ft |
| NPK | Proximity to the closest park[b] | 26,193 ft | 4,668 ft |
| DSS | Distance from Highway 66 | 4,808 ft | 5,221 ft |
| NFF | Proximity to Interstate 55[b] | 10,680 ft | 3,789 ft |
| WHOS | Houses with nonblack occupants | 92.39% | 12.95% |
| INC | Income per adult | \$6,076 | \$3,538 |
| FAMSZ | Family size | 3.410 persons | .425 persons |
| AGE | Age | 31.35 years | 5.556 years |
| GSPER | Grade school population | 1.805% | .891% |
| HSPER | High school population | .782% | .789% |
| PROF | Professional population | 24.25% | 13.27% |
| SMHOS | In same house 5 years | 29.12% | 26.69% |

[a] These are the summary statistics for 199 census blocks for PHOUS down to WHOS and average values for 38 census block groups for INC down to SMHOS.

[b] These proximity variables are calculated by substracting the distance from the amenity source from the maximum distance in the sample. For Lake Springfield the maximum is 12,710 ft.; for parks, the maximum is 27,010 ft.; and for Interstate 55, the maximum is 17,100 ft.

variables. Thus the hedonic equation is

$$\frac{P^{\lambda} - 1}{\lambda} = b_0 + b_1 \frac{Z_1^{\gamma_1} - 1}{\gamma_1} + \cdots + b_n \frac{Z_n^{\gamma_n} - 1}{\gamma_n} + e, \qquad (4.1)$$

where $\lambda$ is the power transformation of housing price, $\gamma_i$ is the power transformation of each housing site trait, the $b_i$ are constant coefficients and $e$ is an error term that is assumed to be normally distributed with zero mean and constant variance. It should be noted that the linear functional form is the special case where the transformation factor for each variable equals 1 and the natural logarithmic form is the special case where the transformation factor for each variable equals zero.

Table 4.2 reports the constrained maximum likelihood estimates for the hedonic equation. To simplify a complex, costly search, the $\gamma_i$ are constrained to equal $\lambda$. By varying $\lambda$ in increments of 0.1, we determined that the preferred functional form is that where $\lambda$ equals 0.1. Construct the statistic $-2\ \ln(L_0/L_a)$, where ln indicates logarithm, $L_0$ is the likelihood under the null hypothesis and $L_a$ is the likelihood under the alternative hy-

**TABLE 4.2**
**Hedonic Equations for Housing Price, PHOUS is the Dependent Variable**

| Housing trait | Coefficient[a] (absolute $t$ value) | | |
|---|---|---|---|
| | $\lambda = \gamma_i = 1.0$ | $\lambda = \gamma_i = .1$ | $\lambda = \gamma_i = 0.0$ |
| ROOM | 86.095 | 2.9833 | 2.147 |
| | (14.09) | (13.04) | (13.62) |
| HBAS | .2747 | .03457 | .02321 |
| | (2.50) | (2.02) | (1.93) |
| HPLB | 1.0774 | .1768 | .2907 |
| | (1.83) | (1.73) | (3.62) |
| DPP | .2476 | .001553 | .01477 |
| | (1.44) | (.30) | (.32) |
| NLAK | .6303 | .06763 | .04745 |
| | (3.70) | (3.04) | (2.78) |
| NPK | .2881 | −.01359 | −.003634 |
| | (1.61) | (.34) | (.13) |
| DSS | .2968 | .08429 | .05963 |
| | (1.86) | (3.33) | (3.00) |
| NFF | .1532 | .02898 | .02670 |
| | (.84) | (.86) | (.98) |
| WHOS | −2.747 | .004009 | .002727 |
| | (1.12) | (.40) | (.44) |
| $R^2$ | .8172 | .7768 | .7704 |
| $F$ | 93.86 | 73.48 | 70.85 |
| ln L | −1008.7 | −988.0 | −988.6 |

[a] Since the coefficients are for different transformations, the values cannot be compared directly, but the signs and $t$ values can.

pothesis. This statistic is approximately distributed chi-square (Mood, Graybill, and Boes, 1974) and can be used to test whether or not the preferred form is significantly different from other forms. The estimate of .1 ($\ln L = -988.0$) is significantly different from that $\lambda$ for the linear form, 1.0 ($\ln L = -1008.7$) at the 1% level. It is not significantly different from that $\lambda$ for the logarithmic form, 0.0 ($\ln L = -988.6$) at the 5% level.

The results based on our data indicate that the linear form of the housing price equation, used in earlier studies such as Ridker and Henning (1967), is inferior to a nonlinear form such as the transformation where $\lambda$ equals .1 or the logarithmic form. Using different data, Linneman (1980) does a constrained search over Box–Cox transformations and finds the best $\lambda$ ranges from .2 to .4. He also does not find them to be significantly different from the log form. For our hedonic equation, .2 and .4 are not significantly different from .1 at the 5% level. Goodman (1978) finds the

**TABLE 4.3**
**Marginal Trait Prices Calculated from the Hedonic Equation, Where $\lambda = \gamma_i = .1$**

| Variable | Average evaluated price[a] | Average price[b] | Standard deviation[c] |
|---|---|---|---|
| PROOM | \$8100/room | \$8264/room | \$2128/room |
| PHBAS | \$9.20/% | \$74.17/% | \$24.39/% |
| PHPLB | \$33.07/% | \$35.42/% | \$12.26/% |
| PDPP | \$.002378/ft | \$.003196/ft | \$.002038/ft |
| PNLAK | \$.1916/ft | \$.5865/ft | \$1.391/ft |
| PNPK | \$−.01244/ft | \$−.06273/ft | \$.3024/ft |
| PDSS | \$.2995/ft | \$.5723/ft | \$.5873/ft |
| PNFF | \$.07939/ft | \$.1346/ft | \$.1103/ft |
| PWHOS | \$.7916/% | \$149.5/% | \$107.1/% |

[a] Trait price evaluated at the mean of PHOUS and $Z_i$ for the 199 census blocks.

[b] Average prices of the 38 census block groups. The averages of the marginal prices exceed the average evaluated prices because of some extreme values of marginal prices among the block groups. The average evaluated price is the marginal price evaluated at the means of the characteristics.

[c] Standard deviation of the price of the 38 census block groups.

best $\lambda$ is .6, which is significantly different from the linear form and for our hedonic equation is significantly different from .1 as well.

The functional form is of special interest because the nonlinear form of the hedonic equation implies that trait prices do indeed vary across the southern section of Springfield. The marginal trait price of any particular trait is given by the partial derivative of the hedonic equation with respect to the quantity of the trait. For the unconstrained Box–Cox hedonic form we have: $\partial P/\partial Z_i = b_i P^{(1-\lambda)} Z_i^{(\gamma_i - 1)}$. For our constrained case where $\gamma_i$ equals $\lambda$, we have: $\partial P/\partial Z_i = b_i(P/Z_i)^{(1-\lambda)}$. Averages of the marginal trait prices are given in Table 4.3 for $\lambda$ equals 0.1. All of the traits, which are defined to be "goods," have positive trait prices except for NPK. The unexpected sign on NPK can be explained by measurement error. Air distance to the nearest park was used—a measure that is probably a highly inaccurate measure of the effective distance to a park. In addition, no weighting was assigned to parks with differing facilities. (The prices for DPP, NFF, NPK, and WHOS are based on hedonic coefficients that are not significant at any of the usual levels.)

Using the implicit prices calculated from the hedonic housing equation, demands are estimated for the neighborhood amenities and the most important structural characteristic—ROOM.

## III. DEMANDS FOR HOUSING CHARACTERISTICS

In keeping with our concern for proper specification, preliminary demand equations for housing characteristics were estimated, first assuming that the implicit prices of characteristics are taken as given by consumers and then, as an alternative, assuming that quantities of characteristics are fixed. The best functional form for the demand equation is determined by using a constrained Box–Cox procedure similar to that used for the hedonic equation. Table 4.4 shows the demand equations under both assumptions concerning the relative speed of adjustment of the supply of characteristics.

For each of the five demands for characteristics that is estimated, chi-square tests indicate that the "best" (λ) functional form is significantly different from the linear at the 1% level. Comparing the "best" functional form to the log–log form, we find that log–log is the "best" for NLAK and DSS (as dependent variables). Chi-square tests indicate that the log–log form could not be rejected for DPP (as a dependent variable) and NLAK (as an independent variable) but could be rejected for NFF (as a dependent variable) at the 10% level and ROOM (as a dependent variable) at the 5% level. All others could be rejected at the 1% level.

In the demand equations we observe several patterns. First, the coefficients for own-price are negative and significant at the 1% level for 4 of the 5 housing characteristics. For the fifth, ROOM, the coefficient is negative in the standard specification and positive in the inverse demand. (The results of the Hausman tests reported later suggest an explanation for the contradictory coefficients for ROOM.) Second, the structural characteristic ROOM appears to be a substitute for each of the locational amenities. The coefficient for PROOM is positive and significant at the 5% level for NLAK, DSS, NFF, and DPP in both the standard and inverse demand equations. Third, the income effect appears to be positive in that all 10 coefficients are positive. Only 4 of the income coefficients are significant at the 10% level, but this is not surprising with the number of other demand factors included in the demand equation. Age of family head appears to increase the demand for housing in that 4 of the coefficients are positive and significant at the 10% level and 4 of the remaining 6 are positive. There is a weak indication that family size increases the demand for housing in that 9 of the 10 coefficients are positive. There is some evidence that professional status decreases demand. Given these mixed but reasonable results for the demand equations, we turn our attention to determining the appropriate demand specification: standard, inverse, or one that accounts for simultaneity with supply.

**TABLE 4.4**
Demands and Inverse Demands for Housing Characteristics Using the Best Functional Form[a]

| | Dependent variable, best λ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Regressors[b] | (1) NLAK 0.0 | (2) PNLAK[c] −.2 | (3) DSS 0.0 | (4) PDSS[c] −.1 | (5) NFF .1 | (6) PNNF[c] .2 | (7) DPP .1 | (8) PDPP[c] −.4 | (9) ROOM −.3 | (10) PROOM[c] −.4 |
| PNLAK | *−.8585[d] | *−1.8326 | −.0438 | .0634 | .0078 | .0068 | .0209 | 1.2896 | .0043 | .0076 |
| | (18.44) | (16.25) | (1.11) | (.76) | (.16) | (.44) | (.63) | (1.13) | (1.18) | (.73) |
| PDSS | −.0286 | −.0947 | *−.9176 | *−1.6390 | *−.1515 | .0298 | −.0687 | −1.4081 | — | — |
| | (.41) | (.52) | (15.43) | (12.55) | (2.01) | (1.28) | (1.34) | (.68) | | |
| PNFF | .0614 | −.3160 | *−.1943 | .1601 | *−1.7707 | *−.2877 | −.1223 | −.2758 | — | — |
| | (.63) | (1.04) | (2.36) | (.75) | (14.99) | (8.22) | (1.51) | (.08) | | |
| PDPP | .1559 | *1.8302 | *.2431 | .0553 | *.6281 | −.0411 | *−2.2623 | *−40.025 | *.0045 | *.0896 |
| | (.98) | (2.64) | (1.80) | (.13) | (2.19) | (.85) | (11.54) | (4.10) | (2.66) | (1.67) |
| PROOM | *.8279 | *9.1286 | *1.3314 | *5.0675 | *.8700 | *1.3955 | *1.5393 | *84.541 | *−.4554 | *.5250 |
| | (2.23) | (5.78) | (4.23) | (5.04) | (3.73) | (4.83) | (9.64) | (7.65) | (2.09) | (7.30) |
| PNPK | .0275 | −1.3845 | .0925 | .6028 | .0850 | −.0002 | .0422 | *63.579 | — | — |
| | (.42) | (.75) | (1.64) | (.69) | (1.09) | (.00) | (.79) | (2.78) | | |
| PWHOS | .0053 | .8209 | −.0171 | −.6024 | .0485 | .0862 | .0232 | 11.256 | −.0027 | .3285 |
| | (.13) | (.43) | (.50) | (.64) | (1.04) | (.79) | (.73) | (.48) | (1.60) | (1.58) |
| PHBAS | — | — | — | — | — | — | — | — | *.0041 | .0108 |
| | | | | | | | | | (1.78) | (.26) |
| PHPLB | — | — | — | — | — | — | — | — | *.1740 | *.7521 |
| | | | | | | | | | (6.30) | (2.91) |
| INC | .0771 | *1.3205 | .0929 | .3000 | .1358 | *.0659 | *.1042 | 10.061 | .0106 | *.1198 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| FAMSZ | .0091 | 1.4067 | .1532 | .5903 | .2135 | .1446 | .2639 | *8.8788 | −.0677 | .0412 |
| | (.02) | (1.61) | (.39) | (.91) | (.51) | (.79) | (.92) | (1.82) | (1.48) | (1.00) |
| AGE | .1726 | *2.0444 | −.0734 | .7436 | .2461 | *.2024 | .0994 | *24.189 | −.0907 | *.2290 |
| | (.42) | (1.72) | (.21) | (1.09) | (.80) | (1.90) | (.47) | (2.19) | (1.12) | (2.52) |
| GSPER | .0039 | *.1303 | .0053 | .0009 | −.0080 | −.0188 | −.0011 | *1.1622 | −.0085 | *.0143 |
| | (.26) | (1.88) | (.42) | (.04) | (.38) | (.98) | (.08) | (1.77) | (1.42) | (2.53) |
| HSPER | .0060 | *−.0941 | .0075 | −.0215 | .0200 | .0127 | *.0087 | *−.5911 | *.0040 | *−.0046 |
| | (.68) | (2.95) | (1.00) | (1.57) | (1.58) | (.98) | (2.70) | (2.56) | (1.91) | (2.37) |
| PROF | .1051 | *−.3933 | −.0316 | −.1370 | −.0428 | −.0299 | −.0633 | *−2.3103 | *.0218 | −.0130 |
| | (1.54) | (2.22) | (.55) | (1.28) | (.77) | (1.28) | (1.66) | (1.75) | (1.99) | (1.11) |
| SMHOS | −.0060 | .1506 | −.0015 | .0122 | .0044 | *.0144 | −.0056 | 2.6566 | .0006 | .0069 |
| | (.53) | (1.18) | (.16) | (.34) | (.34) | (2.65) | (.64) | (1.26) | (.05) | (.39) |
| Constant | −.9465 | −13.513 | −1.8560 | −8.3632 | −2.2226 | −4.7946 | −13.592 | −280.47 | 2.9204 | −1.7626 |
| $R^2$ | .9756 | .9612 | .9792 | .9550 | .9718 | .9742 | .9672 | .9452 | .9756 | .9214 |
| F | 65.61 | 40.73 | 77.16 | 34.90 | 56.63 | 62.07 | 48.48 | 28.33 | 41.73 | 21.65 |
| SEE | .2052 | .3200 | .1744 | .2425 | .2131 | .1072 | .1456 | 1.6138 | .0173 | .0141 |
| ln L | −132.68 | 56.09 | −120.14 | 45.21 | −127.79 | 105.84 | −136.11 | 243.96 | 25.77 | −123.41 |

*Note:* Absolute $t$ values are shown below each coefficient.

[a] The sample consists of 38 census block groups.

[b] In none of the demand equations does the best $\lambda$ equal one, the transformation that is perhaps the most easily interpreted. To compare the coefficients reported in this table to those in the linear ($\lambda = 1.0$) equations, the values of the retransformed variables can be plotted using the Box–Cox coefficients, and the slope calculated at the mean. For example, for NFF where $\lambda = .1$, the own-price coefficient is approximately equal to a linear coefficient of −416.78, which can be compared to the estimated linear coefficient of −223.1.

[c] In each regression where the price is the dependent variable, all regressors are quantities. For example, PNLAK is regressed on NLAK, DSS, NFF, DPP, ROOM, NPK, WHOS, INC, FAMSZ, AGE, GSPER, HSPER, PROF, and SMHOS.

[d] The asterisk indicates significance for $\alpha = .10$ for a two-tail test except for own-price and INC where a one-tail test is performed.

## IV. THE RELATIVE EXOGENEITY OF PRICES AND QUANTITIES

The question of the appropriate specification can be posed in terms of exogenity. If supply adjusts slowly, then the quantities of housing characteristics are, in a sense, exogenous to consumers and the implicit prices are endogenous. If supply adjusts rapidly, then the implicit prices are, in a sense, exogenous to the consumer and the quantities are endogenous. If supply adjusts at a moderate pace, then both quantities and prices are endogenous.

Although theoretical arguments may be sufficiently convincing for some housing characteristics, it is not always clear which specification is appropriate. As Freeman noted, the question of whether quantities and–or prices are endogenous is an empirical one, with the answer depending upon the relative speed of supply-side adjustment.

An implication of the endogeneity of a regressor is that the regressor is correlated with the error term. This correlation results in OLS yielding biased estimates of the model's parameters and will exist even for large samples. Hence OLS estimates are inconsistent in this case. Hausman (1978) derived an asymptotic test of the null hypothesis that the regressor is uncorrelated with the error term; the test may be used in this context. It tests whether or not the regressor of interest—say, price of the housing characteristic—is correlated with the error term in the demand equation. If the null hypothesis is not rejected, it indicates that the regressor is exogenous and that OLS can yield consistent estimates of demand. If the null hypothesis is rejected, then endogeneity of a regressor is one possible explanation.[4]

Since it is contradictory to maintain that we have a demand system in equilibrium and at the same time accept $H_0$: $E(Z_i\epsilon) = 0$ and $H_0$: $E(P_{Z_i}\epsilon) = 0$, a relative test of significance is employed. If exogeneity of both quantity and price can be accepted at the standard significance levels, the decision is made to reject the null hypothesis for the variable with the higher significance. If the null hypothesis is rejected at the standard significance levels for quantity and price, it indicates simultaneity.

A strong case can be made for using significance levels that are lower than standard levels. Given that 2SLS and IV estimation procedures are consistent in the presence or absence of an endogenous explanatory variable and that OLS is consistent only in the absence of endogenous explanatory variables, one would prefer to minimize the probability of Type II

---

[4] Endogeneity of a regressor is not the only circumstance that will result in rejection of the null hypothesis. Errors in variables and the omission of a relevant regressor are other possible causes of correlation between the regressors and the error term.

error, even if this means allowing more Type I errors. This is the case here. If we reject $H_0$ when it is actually true (Type I error), all we will have lost is some efficiency of estimation. However, if we do not reject $H_0$ when it is actually false (Type II error), our estimation will no longer be consistent.

A Hausman test is performed on the best functional form of the demand equation under the two alternative assumptions concerning exogeneity. Let $X_j$ be the variable (either $P_{Z_j}$ or $Z_j$) to be tested for endogeneity and **X** be the vector of either price or quantity variables other than $P_{Z_j}$ or $Z_j$. Let $Y_j$ be the demand variable opposite to $X_j$, that is, $Z_j$ if $X_j = P_{Z_j}$. Let **W** be the vector of additional variables used to estimate the instrumental variable for the Hausman test. Following Hausman, the alternative estimator should be consistent under both the null and alternative hypotheses but inefficient under the null hypothesis. **W** will include the average price of housing $\bar{\mathrm{P}}$ in the block group and the population of the block group.[5] We regress:

$$X_j = c + b\mathbf{X} + a\mathbf{W} \tag{4.2}$$

to obtain $\hat{X}_j$. Then to test $H_0$: $E(X_j\epsilon) = 0$ we regress:

$$Y_j = c + B(\mathbf{X}X_j) + \alpha\hat{X}_j. \tag{4.3}$$

The significance of $\hat{\alpha}$ is then tested, with significance being taken as reason to reject the $H_0$: $E(X_j\epsilon) = 0$.

The results of the Hausman tests are shown in Table 4.5, where we see evidence that the appropriate specification of demand varies from one housing characteristic to another. The most striking results are for ROOM, where the null hypotheses are rejected at the .01% level for each price and quantity. The supply of houses with more rooms appears to adjust at a rate similar to that of housing prices. At least these speeds are close enough that estimation of the demand for rooms must take account of the simultaneity with supply (e.g., 2SLS, IV, or estimating a simultaneous system).

For NFF, proximity to Interstate 55, there is a clear indication that quantity is more exogenous than price. The null hypothesis that price is exogenous can be rejected at the 11% level, whereas that for quantity can be rejected only at the extreme 98% level. This means that the supply of housing with quick access responds slowly relative to housing prices and that the inverse demand shown in Column 6 of Table 4.4 is the more appropriate demand equation.

---

[5] The results subsequently reported are robust with respect to using $\bar{\mathrm{P}}$ alone and $\bar{\mathrm{P}}$ and population together but may be sensitive to the choice of other alternative estimators or other forms of the Hausman test.

**TABLE 4.5**
**Relative Exogenity of Prices and Quantities of Housing Characteristics**

| | Assumption | | | | |
|---|---|---|---|---|---|
| Characteristics | Exogenous variable | Dependent variable | $t$ value on $\alpha$ | Level at which $H_0$ is rejected | Appropriate exogenous variable |
| Proximity to Lake Springfield | PNLAK | NLAK | −.21 | 84 | PNLAK |
| | NLAK | PNLAK | 1.04 | 31 | |
| Distance from Highway 66 | PDSS | DSS | −.31 | 76 | PDSS |
| | DSS | PDSS | −.59 | 44 | |
| Proximity to Interstate 55 | PNFF | NFF | 1.64 | 11 | NFF |
| | NFF | PNFF | .03 | 98 | |
| Distance from electric power plant | PDPP | DPP | 1.21 | 24 | PDPP |
| | DPP | PDPP | 3.39 | .3 | |
| Rooms per house | PROOM | ROOM | 11.18 | .01 | Simultaneous |
| | ROOM | PROOM | 32.45 | .01 | |

For NLAK, proximity to Lake Springfield, and DSS, distance from Highway 66, and DPP, distance from the power plant, it appears that prices are relatively exogenous, meaning that supplies of these locational characteristics respond quickly relative to housing prices. This might occur through new housing developments on open land in these areas. For NLAK the null hypothesis can be rejected at the 31% level, whereas that for the price, PNLAK, can be rejected only at the 84% level. The standard demand equation shown in Column 1 of Table 4.4 is the more appropriate of the two. For DSS neither null hypothesis can be rejected at anything close to a usual level, and the difference between the levels is smaller than that for NLAK. Nonetheless it appears that the standard demand shown in Column 3 is the more appropriate. For DPP, the difference in rejection levels for the null hypothesis is approximately the same as that for DSS. Again the evidence is that the price is the relatively more exogenous variable and that the standard demand equation shown in Column 7 is the more appropriate. The levels of significance hint that one might consider a simultaneous estimation procedure.

For each housing characteristic the results are indicative of relative exogeneity–exogeneity of implicit characteristic prices compared to the exogeneity of the quantities of the characteristics.[6] It appears that the results are not an indication of errors in variables or stochastic regressors.

---

[6] It should be recognized that the evidence on relative exogeneity is based on equations in which either all implicit prices of complements and substitutes or all quantities of housing characteristics are regressors. More extensive testing might entail consideration of specifications with various combinations of prices and quantities of characteristics.

The relative exogeneity results are not sensitive to the functional form of the demand equation in that the list of appropriate exogenous variables reported in Table 4.5 is the same for the best ($\lambda^*$), log($\lambda = 0$) and linear ($\lambda = 1$) functional forms.

The variables for the implicit prices are clearly stochastic since they are derived from the estimated hedonic price function, and it would seem that we have a built-in bias for rejecting $H_0$ when testing the exogeneity of price variables. However, in three cases we accepted $H_0$ for price variables, which gives some indication that this effect is minor.

Generally the results indicate there is little simultaneity in the implicit market for the four locational housing characteristics. However, they strongly indicate that one must carefully consider the appropriate specification of the demand function, that is, the standard or inverse demand. In particular, our results are inconsistent with specifying that all markets have either completely inelastic supply or perfectly elastic supply.

## V. CONCLUSIONS

In this chapter we have addressed the issue of proper specification of the demand function for housing characteristics. After estimating the hedonic equation for housing prices using a limited Box–Cox search to determine the best functional form and calculating the implicit marginal prices for various housing characteristics, the demands for these characteristics were estimated under alternative assumptions about the supply conditions. First it was assumed that supply adjusted slowly relative to demand prices, and then it was assumed that supply adjusted relatively quickly. The purpose was to provide information on whether demands for characteristics should be estimated inversely, ordinarily, or simultaneously and also on the nature of the supply of each characteristic. The same constrained Box–Cox procedure was used to determine the best functional form of the demand equations.

A Hausman test was employed in its instrumental variable form, first to test the exogeneity of characteristic prices when demand is estimated with price as a regressor and then to test the exogeneity of quantity when the inverse demand is estimated. For the locational amenities—lakeside proximity (NLAK), distance from busy commercial area (DSS) and distance from an electric power plant (DPP)—there is evidence that demands can be estimated using OLS, assuming prices are exogenous. For access to a major interstate highway (NFF), demand can be estimated by OLS, assuming quantities are exogenous. For the structural characteristic, rooms per house (ROOMS), there is strong evidence of simultaneity.

Perhaps the most useful result is that proper specification of demands depends upon the particular housing characteristic and that alternatives to initial assumptions warrant at least some investigation. Caution should be exercised in applying these results to other housing markets since they may be quite sensitive to the instrument used in the exogeneity test and possibly relevant housing characteristics omitted (because of data limitations) from the hedonic and demand equations. Nonetheless, the results illustrate the advantages of using an empirical test along with theoretical arguments concerning the supply of housing characteristics in estimating the demand for housing amenities and the benefits of amenity provision.

## ACKNOWLEDGMENTS

We gratefully acknowledge the constructive comments of Douglas Diamond, Jr., Peter Linneman, Rati Ram, and David Spencer. The authors take sole responsibility for any errors.

## REFERENCES

Bender, B., Gronberg, T. J., and Hwang, H.-S. "Choice of Functional Form and the Demand for Air Quality." *Rev. Econ. Stat.* 62 (November 1980): 638–643.

Blomquist, G., and Worley, L. "Hedonic Prices, Demands for Urban Housing Amenities, and Benefit Estimates." *J. Urban Econ.* 9 (March 1981): 212–221.

Brown, G. M., Jr., and Pollakowski, H. O. "Economic Value of Shoreline." *Rev. Econ. Stat.* 59 (August 1977): 272–278.

Freeman, A. M., III, "Hedonic Prices, Property Values and Measuring Environmental Benefits: A Survey of the Issues." *Scandinavian J. Econ.* (1979): 154–173.

Goodman, A. C. "Hedonic Prices, Price Indices and Housing Markets." *J. Urban Econ.* 5 (October 1978): 471–484.

Harrison, D., Jr., and Rubinfeld, D. L. "Hedonic Housing Prices and the Demand for Clean Air." J. Environ. Econ. Manag. 5 (March 1978): 81–102.

Hausman, J. A. "Specification Tests in Econometrics." *Econometrica* 46 (November 1978): 1251–1270.

Linneman, P. "An Analysis of the Demand for Residence Site Characteristics." Ph.D. dissertation, University of Chicago, 1977.

Linneman, P. "Some Empirical Results on the Nature of the Hedonic Price Function for the Urban Housing Market." *J. Urban Econ.* 8 (July 1980): 47–68.

Mood, A. M., Graybill, F. C., and Boes, D. C. *Introduction to the Theory of Statistics*. 3 ed. New York: McGraw-Hill, 1974.

Nelson, J. "Residential Choice, Hedonic Prices and the Demand for Urban Air Quality." *J. Urban Econ.* 5 (October 1978): 357–369.

Ridker, R., and Henning, J. "The Determinants of Residential Property Values with Special Reference to Air Pollution." *Rev. Econ. Stat.* 49 (May 1967): 246–257.

Rosen, S. "Hedonic Prices and Implicit Markets." *J. Polit. Econ.* 82 (January–February 1974): 34–55.

Smith, B. "Demand for Housing: A Reexamination of Methodological Issues." Paper presented at the Conference on Urban Economics held June 24–25, 1980, at the University of Chicago, Chicago, Illinois.

Witte, A. D., Sumka, H. J., and Erekson, H. "An Estimate of a Structural Hedonic Price Model of the Housing Market: An Application of Rosen's Theory of Implicit Markets." *Econometrica* 47 (September 1979): 1151–1173.