ON THE CHOICE OF FUNCTIONAL FORM FOR HEDONIC PRICE FUNCTIONS

Maureen L. Cropper, Leland B. Deck and Kenneth E. McConnell*

Abstract—This study examines how errors in measuring marginal attribute prices vary with the form of the hedonic price function. In simulations, consumers with known utility functions bid for houses with given attributes. Various forms of the hedonic function are estimated using equilibrium housing prices. Errors in estimating marginal attribute prices are calculated by comparing each consumer's equilibrium marginal bid vector with the gradient of the hedonic function. When all attributes are observed, linear and quadratic Box-Cox forms produce lowest mean percentage errors; however, when some attributes are unobserved or are replaced by proxies, linear and linear Box-Cox functions perform best.

I. Introduction

THE fact that economic theory places few restrictions on the form of the hedonic price function has led most researchers to use a goodness-of-fit criterion in choosing an appropriate form for the hedonic function. If, however, one's goal is to value product attributes, the form of the hedonic price function that should be used is the one that most accurately estimates marginal attribute prices. The latter measure consumers' marginal willingness to pay for attributes and thus may be used directly to value small changes in attribute levels. Marginal prices also constitute the dependent variables in the estimation of marginal bid functions; hence errors in their measurement may bias the valuation of non-marginal attribute changes as well. This paper examines how errors in measuring marginal prices vary with the form of the hedonic price function.

Received for publication September 28, 1987. Revision accepted for publication March 14, 1988.

* University of Maryland and Resources for the Future; U.S. Environmental Protection Agency; and University of Maryland, respectively.

This research was supported by the University of Maryland Computer Science Center, which provided computer time for the project, and by the U.S. Environmental Protection Agency, through a cooperative agreement with the University of Maryland. We thank Larry Bodin for his comments on the assignment problem, Allen Goodman for providing us with data and Tim Phipps for his collaboration at earlier stages of the research. The comments of two anonymous referees are also appreciated.

Since computation of such errors requires that true marginal prices be known, errors in measuring marginal prices must be computed in a simulation context. Our results are based on simulations of housing market equilibria in which consumers bid for a fixed housing stock. Equilibrium housing prices, together with housing attributes, provide the data used to estimate hedonic price functions. Since each consumer's equilibrium marginal bid for each attribute is known, the true marginal price paid for each attribute is also known and can be compared with the gradient of the hedonic price function.

Errors in estimating marginal prices are first examined assuming that the researcher observes all product attributes without error, and then assuming that some attributes are unobserved or are measured by proxies. Whether or not all attributes are observed by the researcher significantly affects the performance of various forms of the hedonic price function. When all attributes are observed, linear and quadratic functions of Box-Cox transformed variables provide the most accurate estimates of marginal attribute prices: the goodness-of-fit criterion suggested by Rosen (1974), Goodman (1978) and Halvorsen and Pollakowski (1981) coincides with accurate measurement of marginal prices.

When certain variables are not observed, or when a variable is replaced by a proxy, a simple linear hedonic price function consistently outperforms the quadratic Box-Cox function, which provides badly biased estimates of "hard to measure" attributes. The misgivings of Cassel and Mendelsohn (1985) regarding the ability of the quadratic Box-Cox function to measure marginal attribute prices thus seem to be justified, at least when the hedonic price function is misspecified. A linear Box-Cox function, however, performs well in the presence of specification error. Since it also provides accurate marginal price estimates under perfect information, our simulations suggest that it is the functional form of choice.

II. Simulation of Housing Market Equilibria

Our housing market consists of N houses, each described by an attribute vector $\mathbf{Z}$, and N households, each endowed with a utility function, U_h, income, y_h, and a vector of personal characteristics, $\mathbf{C}_h$. Households bid against each other for the housing stock, with houses sold to the highest bidder (Wheaton, 1974).

To define a housing market equilibrium let $B_{hj}(u_h)$ denote household h's bid for house j when its utility level is u_h. $B_{hj}(u_h)$ is defined implicitly by

$$u_h = U_h\left(y_h - B_{hj}, \mathbf{Z}_j; \mathbf{C}_h\right). \tag{1}$$

Let $X_{hj} = 1$ if household h occupies house j and $X_{hj} = 0$ otherwise. An equilibrium in the housing market is a set of utilities $u^* = (u_1^*, u_2^*, \ldots, u_N^*)$, prices $P^* = (P_1^*, P_2^*, \ldots, P_N^*)$, and an allocation matrix $[X_{hj}]$ such that (2)–(4) hold,

$$\begin{aligned} B_{hj}(u_h^*) &= P_j^* \quad \text{if} \quad X_{hj} = 1 \\ B_{hj}(u_h^*) &\le P_j^* \quad \text{if} \quad X_{hj} = 0 \end{aligned} \tag{2}$$

$$\sum_{h=1}^{N} X_{hj} = 1, \qquad j = 1, \ldots, N \tag{3}$$

$$\sum_{j=1}^{N} X_{hj} = 1, \qquad h = 1, \ldots, N. \tag{4}$$

Equation (2) states that the equilibrium rent on house j equals the maximum willingness to pay (at utility u_h^*) of the household occupying j. It implies, furthermore, that no household is willing to pay more for house j than the household buying the house (houses are sold to the highest bidder); moreover, the household could not receive higher utility by purchasing any other house. Conditions (3) and (4) specify that each house must be occupied and that each household must buy a house.[1] Equilibrium prices may be computed by iteratively solving an assignment problem (Koopmans and Beckmann, 1957; Wheaton, 1974) until the shadow prices attached to buyers (the side-payments necessary to maintain current utility levels) are zero. The housing shadow prices, P_j, then constitute equilibrium rents.

[1] This definition of equilibrium corresponds to a "closed city" in which the number of buyers and sellers is fixed, and utility levels adjust to equate the supply and demand for houses.

For our results to be robust, alternative housing market equilibria must be computed. The true hedonic price function, i.e., the set of pairs $\{P_j, \mathbf{Z}_j\}$, can be altered by varying either (i) the form of U_h; (ii) the distribution of parameters of U_h; (iii) the attributes included in $\mathbf{Z}$; (iv) the distribution of those attributes; (v) the distribution of buyer characteristics. A set of assumptions about (i)–(v) is called a scenario. Our results are based on six scenarios, summarized in the chart, and described below.[2]

HOUSING MARKET SCENARIOS

Form of Utility Function	Housing Stock: Baltimore City	Housing Stock: Baltimore County
Translog	Attribute List #1	Attribute List #1
Diewert	Attribute List #1 Attribute List #2	Attribute List #1 Attribute List #2

A. *The Housing Stock*

To make our simulations realistic, houses are drawn from homes sold in Baltimore City or Baltimore County in 1977–78. Baltimore City and County are treated as distinct housing markets, the former representing an older, urban, heterogeneous housing stock and the latter a more homogeneous, suburban area. The attributes of houses come from Multiple Listing data, and the attributes of neighborhoods from the 1980 Census of Housing and Population.

The attributes selected are those that commonly appear in empirical studies of housing demand (see table 1).[3] In selecting neighborhood attributes we have purposely chosen two that are highly correlated, *PERCENT PROFESSIONAL* and *PERCENT HIGH SCHOOL*, to see how various forms of the hedonic price function handle collinearity.[4] We have also chosen housing attributes that are discrete (*NO. ROOMS, BATHROOMS, DETACHED, AIR-CONDITIONED, FIREPLACE, GARAGE*) to compare the accu-

[2] A more complete description of our simulations is contained in an appendix, available from the authors upon request. The appendix includes detailed descriptions of the housing stock, of housing buyers and of the procedures used to select utility function parameters.

[3] The variables used in our simulations, with the exception of *PERCENT OF HOUSEHOLDS WITH CHILDREN*, are a subset of the variables used by Palmquist (1984).

[4] The simple correlation coefficient between these variables is 0.76 in the City and 0.87 in the County.

racy of their marginal prices with those of continuous attributes.

B. *Consumer Preferences*

In all simulations utility is a function of housing attributes and all other goods, x, and has the general form

$$U_h = g(x) + \sum_i a_{ih}(\mathbf{C}_h) g(z_i) + 0.5 \sum_i \sum_j b_{ij} g(z_i) g(z_j). \quad (5)$$

Two specific forms are used: the translog, in which $g(x) = \ln(x)$, and the Diewert, in which $g(x) = \sqrt{x}$. In both cases parameters a_{ih} depend on a vector of measured buyer characteristics, $\mathbf{C}_h$, which includes race, family size, whether the household has children, and the education and occupation of the household head. Preferences also reflect unmeasured, individual-specific taste factors, $\boldsymbol{\alpha}_h = (\alpha_{1h}, \ldots, \alpha_{nh})$, which are assumed to be identically normally distributed for all buyers, independently of income and $\mathbf{C}_h$, with mean vector $\bar{\alpha}$ and diagonal covariance matrix Σ. Formally,

$$a_{ih} = \alpha_{ih} + \boldsymbol{\delta}_i' \mathbf{C}_h. \quad (6)$$

Randomness in preferences captures the notion that observationally equivalent persons may have different tastes. The diagonal covariance matrix implies that persons with a strong preference for interior space need not have a strong preference for outdoor space. Whereas $\boldsymbol{\alpha}_h$ varies among households, the $\{b_{ij}\}$, which allow for complementarity between attributes, are the same for all buyers.[5] The choice of utility function parameters is described in an appendix available from the authors. Table 1 indicates which buyer characteristics and which housing attributes enter the marginal bid function for each attribute.

One implication of randomness in $\{a_{ih}\}$ is that simulation results hinge on the realization of $\{\alpha_{ih}\}$. For each scenario, 20 Monte Carlo simulations were run, each corresponding to a different draw from the distribution of $\boldsymbol{\alpha}_h$. In all runs the joint distribution of y_h and $\mathbf{C}_h$ comes from the

[5] If only $\{\alpha_{ih}\}$ vary across buyers, the marginal bid function of a person with a higher value of α_{ih} must lie above the marginal bid function of a person with a lower value of α_{ih}. We believe this captures the notion of a person having strong preferences for an attribute. If $\{b_{ij}\}$ varied across persons independently of $\{\alpha_{ih}\}$ marginal bid functions would cross and this property would be destroyed.

TABLE 1.—VARIABLES ENTERING MARGINAL ATTRIBUTE BID FUNCTIONS

Bid Function for	Other Attributes Entering Function	Buyer Characteristics Entering Function
NO. BATHROOMS[a,b]	Interior Space	Number in household
INTERIOR SPACE	No. Bathrooms	Number in household
SQ. FT.[a]	Lot Size	
NO. ROOMS[b]		
LOT SIZE[a,b]	Interior Space Median Income	Whether children
YEAR BUILT[a,b]	None	None
SQ. FT. PER ROOM[b]	None	None
AIR-CONDITIONING[a,c]	None	None
DETACHED[b]	None	None
FIREPLACE[a]	None	None
GARAGE[a,b]	None	None
% CHILDREN[a]	None	Whether children
% HIGH SCHOOL[a]	Median income	Whether High School Degree Whether Some College
% PROFESSIONAL[a,b]	None	Whether Technical/Sales Whether Manager/Professional
MEDIAN INCOME[a,b]	% High School Lot Size, % White	None
OWNER-OCCUPIED[a,b,d]	None	None
% WHITE[a,b,e]	Median Income	Race
MEDIAN AGE[b]	None	None

[a]Attribute List #1.
[b]Attribute List #2.
[c]County only.
[d]City only, attribute list #1.
[e]City only.

Baltimore Travel Demand Dataset (1980), a study of transportation mode choice conducted in Baltimore County and City during 1977. The 200 consumers in our Baltimore City (County) housing market are employed homeowners with incomes above \$10,000 (\$1976) interviewed in the survey.

III. The Effect of the Form of the Hedonic Price Function on Errors in Estimating Marginal Bids

Data from each of the 120 Monte Carlo runs described above are used to estimate six forms of the hedonic price function: linear, semi-log (ln P_j on $\mathbf{Z}_j$), double-log, quadratic, and linear and quadratic functions of Box-Cox transformed variables. For Box-Cox functions, independent variables are constrained to have the same transformation, which is allowed to differ from the transformation of the dependent variable. Dummy variables are not transformed.[6]

For each estimated hedonic price function we calculate the error in estimating household h's bid for attribute i on trial t, e_{ith}, as the difference between the derivative of the hedonic price function and the household's true marginal bid,

$$e_{ith} = \partial h_t/\partial z_i - \partial B_{ht}/\partial z_i, \qquad i = 1, \ldots, n. \tag{7}$$

To summarize the empirical frequency distribution of errors across buyers, we calculate for each trial, the mean, $\bar{e}_{it}$, and standard deviation, s_{it}, of errors for each attribute. β_{it} and S_{it} express the mean and standard deviation of errors as a fraction of the mean true bid for each attribute,

$$\begin{aligned} \beta_{it} &= \bar{e}_{it}\Big/\left[N^{-1}\sum_h \partial B_{ht}/\partial z_i\right], \\ S_{it} &= s_{it}\Big/\left[N^{-1}\sum_h \partial B_{ht}/\partial z_i\right]. \end{aligned} \tag{8}$$

β_{it} and S_{it} are referred to as the normalized mean and normalized standard deviation of errors in estimating marginal price.

A. *Errors in Measuring Marginal Prices When All Attributes Are Observed*

When all attributes are observed, the linear and quadratic Box-Cox functions perform best based on the normalized mean and standard deviation of error criteria. We substantiate this with detailed results for the Baltimore City, Diewert Utility, Attribute List #1 scenario (see table 2). The β_{it} and S_{it} have been averaged over 20 Monte Carlo runs to produce β_i and S_i.

No function produces the lowest $|\beta_i|$ for all attributes, although the quadratic Box-Cox function has the lowest normalized error for 6 out of 12 attributes. If $|\beta_i|$ is averaged across all attributes (see table 3) the linear and quadratic Box-Cox functions produce the lowest ratio of mean error to mean true bid, 0.1369 and 0.1289. The linear, semi-log and double-log functions produce ratios of mean error to mean true bid that are 4 times as large as those of the Box-Cox functions.

The linear Box-Cox function has the lowest error variance, producing an average S_i of 0.3290 (see table 3). The quadratic function produces the highest average S_i, 0.9149. The large spread of the errors exhibited by the quadratic function reflects the tendency of that function to miss badly in predicting marginal prices at extreme values of $\mathbf{Z}$.

Table 2 also shows how accurately the marginal prices of different attributes are measured. The marginal prices of "important" attributes, ones that account for a high percentage of total utility from housing, tend to be measured with greater accuracy than those of "unimportant" attributes.[7] *PERCENT CHILDREN* and *FIREPLACE* have purposely been made unimportant to the housing decision. Marginal bids for both variables are consistently estimated with errors in excess of 100% by the linear and semi-log hedonic price functions, and by the quadratic function. The two Box-Cox functions, however, consistently avoid large errors in estimating marginal bids for "minor" attributes.

[6] If λ, the parameter used to transform independent variables, is constrained to be nonzero, the two Box-Cox functions can be estimated using Kenneth White's SHAZAM program. The restriction that $\lambda \neq 0$ should not greatly affect results since the Box-Cox transformation is a continuous function of λ.

[7] If the total utility from housing is approximated linearly about the mean $\mathbf{Z}$ vector, the shares of housing utility accounted for by the variables in table 2 are: *BATHS* 0.0958, *SQ FT* 0.1963, *LOT* 0.0911, *YR BLT* 0.1142, *% CHILD* 0.0138, *% HS GRAD* 0.0692, *% PROF* 0.0360, *MED INCOME* 0.1833, *% WHITE* 0.0988, *% OWNER* 0.0918, *FIREPLACE* 0.0200, *GARAGE* 0.0399.

TABLE 2.—ERRORS IN MEASURING MARGINAL PRICES, BALTIMORE CITY, DIEWERT UTILITY FUNCTION
Mean Error/Mean True Price
(Standard Deviation of Error/Mean True Price)

	All Attributes Observed						*SQ. FT.* Replaced by *NO. ROOMS*					
	Linear	Semi-Log	Log-Log	Box-Cox Linear	Quadratic	Box-Cox Quadratic	Linear	Semi-Log	Log-Log	Box-Cox Linear	Quadratic	Box-Cox Quadratic
BATHS	−0.1130 (0.3336)	−0.0150 (0.2567)	0.0817 (0.3266)	0.1109 (0.2164)	0.0167 (0.2879)	−0.0039 (0.2615)	0.3389 (0.3321)	0.3559 (0.2660)	0.7544 (0.5270)	0.6905 (0.3657)	0.5555 (0.8589)	0.5709 (0.8340)
SQ. FT.	−0.2410 (0.3238)	−0.3418 (0.2707)	0.1922 (0.3378)	0.0081 (0.2092)	0.0409 (0.1962)	0.0179 (0.1933)	This variable has been replaced by *NO. ROOMS*					
LOT	−0.2989 (0.3158)	−0.3699 (0.3499)	0.8162 (0.9876)	0.2822 (0.3768)	−0.0793 (0.2613)	−0.0045 (0.2554)	−0.1874 (0.3128)	−0.2855 (0.3567)	1.3381 (1.2946)	0.4687 (0.3801)	0.1028 (0.6836)	0.3043 (0.6975)
YEAR BUILT	−0.0098 (0.9137)	0.0711 (0.9081)	−0.2713 (0.3905)	0.0093 (0.5031)	−0.1362 (0.7129)	0.0269 (0.5490)	−0.4662 (0.9122)	−0.2925 (0.9035)	−0.6345 (0.5745)	−0.3187 (0.6467)	−0.5470 (1.5282)	−1.1186 (1.6190)
% CHILD	−1.3054 (0.9108)	−1.0573 (0.9160)	−1.6137 (1.1184)	0.0475 (0.8396)	1.4365 (4.1594)	0.5934 (2.2306)	−3.1192 (0.9065)	−2.4918 (1.0794)	−3.6538 (2.1917)	−1.6894 (1.1304)	−7.0016 (19.5682)	−5.8140 (22.3485)
% HIGH SCHOOL	1.0266 (0.3397)	1.5677 (0.5891)	−1.1890 (0.3714)	−0.0608 (0.2834)	0.0612 (0.8698)	−0.1977 (0.6654)	0.3612 (0.3429)	1.0155 (0.5014)	−2.3672 (0.6413)	−0.7064 (0.3115)	0.6720 (4.8613)	0.4739 (6.2633)
% PROFES-SIONAL	−0.3007 (0.3539)	−0.6097 (0.3628)	0.4466 (0.5082)	0.1502 (0.2476)	−0.0055 (0.6926)	0.1378 (0.6099)	−0.1129 (0.3464)	−0.4511 (0.3631)	0.9474 (0.7308)	0.3221 (0.2872)	−0.6599 (3.7797)	−0.2388 (5.0951)
MEDIAN INCOME	0.2648 (0.2784)	−0.1535 (0.2118)	0.1828 (0.2951)	−0.1402 (0.1914)	0.0121 (0.4132)	0.0678 (0.3681)	0.6172 (0.2754)	0.1178 (0.2213)	0.4202 (0.3258)	0.0185 (0.1930)	0.1113 (1.8312)	0.1252 (2.3421)
% WHITE	−0.3319 (1.3419)	−0.3501 (1.3771)	0.5855 (2.6831)	−0.0249 (0.2748)	−0.4772 (1.0360)	0.0096 (0.4991)	−0.3609 (1.3415)	−0.3727 (1.3748)	0.3840 (2.2032)	−0.1451 (0.4495)	−0.6942 (1.1184)	−0.6312 (0.9437)
% OWNER OCCUPIED	−0.1938 (0.6059)	0.3697 (0.6630)	−0.1167 (0.3251)	0.2288 (0.4018)	0.0953 (0.6017)	−0.0405 (0.5179)	−0.8265 (0.6050)	−0.1400 (0.6246)	−0.4957 (0.4056)	−0.0382 (0.4612)	0.1706 (3.2786)	0.0702 (3.4229)
FIREPLACE	1.1666 (0.3347)	1.1410 (0.2623)	−0.0102 (0.2044)	0.2657 (0.2079)	0.2656 (1.1259)	0.1840 (0.8446)	3.0095 (0.3354)	2.6969 (0.5432)	1.5112 (0.3213)	1.8824 (0.7501)	1.3171 (6.6329)	1.4029 (6.3900)
GARAGE	0.4671 (0.3317)	0.4615 (0.1973)	−0.3431 (0.2339)	−0.1129 (0.1966)	0.0419 (0.6200)	0.0008 (0.4939)	0.5906 (0.3285)	0.5667 (0.2035)	−0.1743 (0.2131)	0.0798 (0.2255)	0.2168 (1.3357)	−0.0013 (2.0304)

TABLE 3.—PERFORMANCE OF VARIOUS FORMS OF THE HEDONIC PRICE FUNCTION
BALTIMORE CITY HOUSING MARKET

	Form of the Hedonic Price Function					
Criterion	Linear	Semi-Log	Log-Log	Box-Cox Linear	Quadratic	Box-Cox Quadratic
	Diewert Utility, Attribute List #1					
Maximum $\|\beta_i\|$	1.3347	0.5677	1.6169	0.3266	1.4365	0.6783
Average $\|\beta_i\|$	0.4781	0.5432	0.4929	0.1369	0.2382	0.1289
Average S_i	0.5319	0.5304	0.6485	0.3290	0.9149	0.6241
	Translog Utility, Attribute List #1					
Maximum $\|\beta_i\|$	1.4866	1.9912	0.9948	0.3405	2.5435	0.8197
Average $\|\beta_i\|$	0.6940	0.6236	0.2552	0.1380	0.4394	0.1711
Average S_i	0.7147	0.6521	0.4516	0.3854	1.4029	0.7353
	Diewert Utility, Attribute List #2					
Maximum $\|\beta_i\|$	0.6839	3.3845	7.1235	0.3744	4.2906	1.5277
Average $\|\beta_i\|$	0.2830	0.5051	0.8781	0.1154	0.4546	0.1992
Average S_i	0.4356	0.4591	0.6317	0.2583	1.0534	0.6138

The other attribute whose marginal price is estimated with errors in excess of 100% is *PERCENT HIGH SCHOOL*, which is highly correlated with *PERCENT PROFESSIONAL* ($r = 0.76$). Collinearity presents problems for the linear, semi-log and double-log functions. The fact that marginal bids depend on several coefficients in the Box-Cox and quadratic cases may explain why these functions handle collinearity problems better: although individual coefficients may be unreliable due to collinearity, a linear combination of these coefficients need not be.

Since economists are often interested in estimating the marginal prices of attributes that may be measured with large error, we have ranked the six hedonic functions according to the maximum value of $|\beta_i|$ that each produces. By this criterion the linear Box-Cox function performs the best and the quadratic and double-log functions the worst.

Table 3 summarizes mean $|\beta_i|$ and S_i and maximum $|\beta_i|$ for 60 Monte Carlo runs using the Baltimore City housing market. (The rankings of the various forms of the hedonic price function are similar in the county, although average errors are generally lower in that more homogeneous housing market.) The results are clear: the linear and quadratic Box-Cox functions consistently outperform all other functional forms. The simpler functions (linear, semi-log and double-log) generally do the worst, although, as noted above, the quadratic function often does poorly according to the maximum bias criterion.

B. *Bias in Measuring Marginal Price When Some Attributes Are Not Observed*

Results, however, change when attributes are omitted from the hedonic price function, or are replaced by proxies. Table 2 summarizes the results of 10 Monte Carlo runs of the Baltimore City, Diewert Utility, Attribute List #1 scenario, in which *SQ. FT.* has been replaced by *NUMBER OF ROOMS*. In contrast to the perfect information case it is now the quadratic and Box-Cox quadratic functions that produce the largest normalized bias, and this is sizeable: the quadratic functions underestimate the marginal bid for *PERCENT CHILDREN* by approximately 600%! The variance of the errors produced by the misspecified quadratic functions is also larger than in the perfect information case. The value of S_i, averaged over all attributes, is 4.1342 for the quadratic function and 4.7260 for the quadratic Box-Cox function, compared with 0.9149 and 0.6241 in the perfect information case.

The linear Box-Cox function avoids the extremely large mean errors that characterize the two quadratic functions, and also produces the smallest average value of S_i of all six functions, 0.4728. The linear and semi-log functions, too,

TABLE 4.—MARGINAL BID ESTIMATION WITH MISSPECIFICATION
AVERAGE PERCENT BIAS AND MAXIMUM PERCENT BIAS

	Form of the Hedonic Price Function					
Scenario[a]	Linear	Semi-Log	Log-Log	Box-Cox Linear	Quadratic	Box-Cox Quadratic
Baltimore City						
Lot Omitted	0.8439[b]	0.8705	0.5717	0.7061	1.0749	0.9639
List #1	2.8175	2.6612	1.7518	3.4694	6.8242	6.5446
Lot Omitted	0.7598	0.7484	0.4605	0.5299	1.1950	0.9547
List #1	1.9748	2.8725	1.1551	1.9879	7.5980	5.1141
Lot Omitted	0.4689	0.5104	0.8491	0.3262	1.8677	1.2971
List #2	1.4806	2.9129	5.5532	0.8818	14.4860	10.0485
Rooms for Sq. Ft.	0.9082	0.7988	1.1528	0.5807	1.0958	0.9815
List #1	3.1522	2.7232	3.6538	1.9411	7.0016	5.8140
Rooms Omitted	0.2915	0.5692	0.9018	0.2055	0.2418	0.2837
List #2	0.6198	4.2589	7.1543	1.0200	1.1780	2.1560
Baltimore County						
Lot Omitted	1.0144	1.1289	1.1217	1.1588	1.9195	2.9586
List #1	2.9486	4.4704	4.7085	5.0610	5.6506	10.6906
Lot Omitted	0.4021	0.4778	0.5452	0.6497	1.0426	0.8151
List #2	1.1045	1.2941	1.7217	2.3419	3.3259	2.6983
Rooms for Sq. Ft.	1.9157	1.3358	0.9971	0.9738	1.3483	1.3662
List #1	5.1226	3.5515	2.8767	2.6644	5.1555	2.5498
Detached for Lot	0.6374	0.6256	0.7659	0.8166	0.5845	2.5071
List #1	1.7626	2.2633	2.6191	2.8534	1.1628	16.7425
Rooms Omitted	0.3174	0.3531	0.2636	0.1726	0.5759	0.6411
List #2	0.7634	0.8135	0.5489	0.3019	2.4468	3.7201

[a] In the second scenario the utility function is translog; in all other cases it is Diewert.
[b] The first row contains $|\beta_i|$, averaged over all attributes i, the second row the maximum $|\beta_i|$ over all i.

produce errors with a narrow spread: the variance of the errors in the linear case is equal to the variance in the true marginal prices (bids); in the semi-log case the variance of e_{ith} is fairly insensitive to misspecification of the hedonic price function.

In general, when variables are omitted or replaced by proxies it is the simpler forms—the linear, semi-log, double-log—and the Box-Cox linear that do best. Table 4 contains summary bias measures (averaged across all trials) for each of 10 omitted variable scenarios. Regardless of the criterion used, the quadratic and Box-Cox quadratic functions perform the worst or second-worst of all forms of the hedonic function in the majority of the omission scenarios. These functions are especially likely to produce the highest maximum normalized bias, with predicted marginal prices sometimes off, on average, by an order of magnitude. The quadratic functions also produce the largest average bias in a majority of the cases examined. The quadratic forms may perform poorly when variables are omitted because each marginal price depends on more coefficients than in the linear cases. Omitting variables thus biases more coefficients, and there is no reason to expect these biases to cancel.

There is some analytical evidence to support this result. Assume that the true hedonic price function is quadratic and that attributes are orthogonal. It can be shown that, for certain ranges of attribute values, a linear hedonic price function with one variable omitted produces unbiased estimates of mean marginal bids. A quadratic function that omits a variable (including its square and cross products) does not yield unbiased estimates of mean marginal bids.

Of the six forms of the hedonic function considered, the linear and the Box-Cox linear perform the best in the presence of misspecification, with the Box-Cox linear function arguably the best of the six. Although it does not do quite as well as the linear function in producing the smallest maximum bias, it wins more often than the linear function according to the average bias criterion. Based on the results of our limited simulations, the linear Box-Cox function appears to be the functional form of choice when estimating hedonic price functions.

REFERENCES

Cassel, Eric, and Robert Mendelsohn, "The Choice of Functional Forms for Hedonic Price Equations: Comment," *Journal of Urban Economics* 18 (Sept. 1985), 135–142.

Goodman, Allen C., "Hedonic Prices, Price Indices and Housing Markets," *Journal of Urban Economics* 5 (Oct. 1978), 471–484.

Halvorsen, Robert, and Henry O. Pollakowski, "Choice of Functional Form for Hedonic Price Equations," *Journal of Urban Economics* 10 (July 1981), 37–47.

Koopmans, Tjalling C., and Martin Beckmann, "Assignment Problems and the Location of Economic Activities," *Econometrica* 25 (Jan. 1957), 53–76.

Palmquist, Raymond B., "Estimating the Demand for the Characteristics of Housing," this REVIEW 66 (Aug. 1984), 394–404.

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

U.S. Department of Transportation, *Baltimore Travel Demand Dataset User's Guide* (Washington, D.C., 1980).

Wheaton, William C., "Linear Programming and Locational Equilibrium: The Herbert-Stevens Model Revisited," *Journal of Urban Economics* 1 (July 1974), 278–287.

LINKED CITATIONS

- Page 1 of 1 -

You have printed the following article:

On the Choice of Funtional Form for Hedonic Price Functions
Maureen L. Cropper; Leland B. Deck; Kenenth E. McConnell
The Review of Economics and Statistics, Vol. 70, No. 4. (Nov., 1988), pp. 668-675.
Stable URL:
http://links.jstor.org/sici?sici=0034-6535%28198811%2970%3A4%3C668%3AOTCOFF%3E2.0.CO%3B2-8

References

Assignment Problems and the Location of Economic Activities
Tjalling C. Koopmans; Martin Beckmann
Econometrica, Vol. 25, No. 1. (Jan., 1957), pp. 53-76.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28195701%2925%3A1%3C53%3AAPATLO%3E2.0.CO%3B2-3

Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition
Sherwin Rosen
The Journal of Political Economy, Vol. 82, No. 1. (Jan. - Feb., 1974), pp. 34-55.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28197401%2F02%2982%3A1%3C34%3AHPAIMP%3E2.0.CO%3B2-U