JOURNAL OF URBAN ECONOMICS 8, 47–68 (1980)

Some Empirical Results on the Nature of the Hedonic Price Function for the Urban Housing Market[1]

PETER LINNEMAN

Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Received September 6, 1978; revised December 11, 1978

An attempt is made to develop a systematic statistical methodology for the analysis of the urban housing market. The standard estimation procedures used for fitting hedonic price functions for the urban housing market are reviewed, and several potentially serious sources of bias are noted. An alternative estimator which capitalizes property values into flows and also searches for the appropriate functional form which avoids these biases is developed. The capitalization rate for owner-occupied housing in 1973 is found to be about 0.03. The magnitudes and seriousness of several of the estimation biases are examined within the context of inappropriate policy decisions which can result from the use of the standard estimators. The importance of neighborhood site characteristics in the determination of local site valuations is also examined and it is found that they explain between 15 and 50% of the standardized variation in site valuations. Further, it is found that these traits are capable of inducing valuation differentials as large as 100% between structurally identical sites.

I. INTRODUCTION

Urban analysts have long recognized the value of hedonic price functions in understanding both locational and public policy issues as they provide consistent estimates of the marginal values of location-specific traits. This allows one to evaluate empirically the relevance of standard first-order conditions for suppliers (i.e., equating the marginal cost and marginal return of the jth trait) and demanders (i.e., equating the ratios of marginal utilities to marginal trait valuations) as well as to attach dollar valuations to various public policy actions.

Through the early history of applied hedonic analyses of urban housing markets the primary estimation problems have been (i) the interpretation of the estimated coefficients in terms of demand and supply functions, and (ii) coefficient bias introduced by omitted variables. Following Rosen

[1]The author is indebted to James J. Heckman, George S. Tolley, Dennis Carlton, Philip E. Graves, members of the Urban Economics Workshop at The University of Chicago, and three anonymous referees. Neil Lenhoff and Anthony Abowd provided useful computational assistance. This research was completed while the author was at the Center for the Study of the Economy and the State, University of Chicago.

0094-1190/80/040047-22$02.00/0

(1974), the partial derivatives of the hedonic function are interpreted here as the market clearing marginal trait prices.

Perhaps because of the influence of urban analysts on survey design, potential problems of missing variable bias have been reduced as more detailed information has become available on both the structural and neighborhood characteristics of a residence. This paper reports results obtained using a new source of urban housing market data, the 1973 Annual Housing Survey (AHS), conducted annually by the Census Bureau in conjunction with the Department of Housing and Urban Development. This data set, described in the next section, is rich in information about the structural and neighborhood characteristics of a residence, providing the necessary data to examine, in Section III, the relative importance of neighborhood characteristics in determining the value of a residence.

This paper also addresses two largely ignored but important sources of estimation bias in hedonic price studies of the urban housing market: (i) the choice of the functional specification of the hedonic price function, and (ii) the appropriate sample to use when estimating a hedonic function. In Section II, a statistical methodology for determining the best functional form for the urban housing market is suggested. In Sections III and IV evidence on the best form obtained from a restricted application of this technique to the 1973 AHS is presented. In Section V the potential for policy errors induced by functional form specification bias is considered for the cases of a private developer and a public clearance project of abandoned buildings.

The question of appropriate sample choice for estimating a hedonic price function has received little attention despite the potential for introducing sample selection bias into the estimated hedonic prices. There are two relevant dimensions of sample selection in the case of the urban housing market: (i) sample choice with respect to ownership status, and (ii) the geographic extent of the market. Housing market hedonic functions have been estimated using either exclusively owner-occupied samples (or their synthetic aggregate counterpart) or exclusively renter-occupied samples. Since owner-occupied housing tends to be relatively high in quality it is plausible that these samples yield biased estimates of the hedonic prices (see Heckman, 1979). Further, use of these subsamples makes it difficult to analyze issues such as condominium conversion and the relative attractiveness of owner versus rental dwellings. In Section IV, it is argued that since owning is a substitute for renting, the correct sample to use is all residences, where owner-occupied property values are converted to annual rental equivalents by an appropriate capitalization rate. That section presents results using such a sample and reports the average capitalization rate implied by the data. In Section V these estimates are applied to the condominium conversion problem.

In Section III, the appropriate geographical extent of the housing market is examined briefly. It is hypothesized that, contrary to common assumptions in the literature, there is a single national housing market, since both development capital and housing consumers are nationally mobile. This line of argument seems particularly appropriate to the hedonic framework in which housing is viewed as a composite of basic locational traits. This hypothesis is tested by comparing the estimates of the hedonic prices for Chicago and Los Angeles (a "traditional" and a "new" city, respectively) to those yielded by the sample of the largest 34 metropolitan areas.

II. MODEL SPECIFICATION AND DATA DESCRIPTION

The hedonic price model as developed by Griliches (1961, 1971) and Rosen (1974) when applied to the urban housing market (see, for example, Dewees, 1976; King, 1976; King and Mieszkowski, 1973; Straszheim, 1973, 1974; Kain and Quigley, 1975) argues that housing payments for a residence site are a function of the basic site-specific traits possessed by the site,

$$V_i = V(N_i, S_i) \qquad \text{for owners,} \tag{1}$$

$$R_i = R(N_i, S_i) \qquad \text{for renters,} \tag{1'}$$

where V_i is property value of the ith owner-occupied residence, R_i is the annual rental payment for the ith renter-occupied residence, N_i is the vector of neighborhood (or nonstructural) traits associated with site i, and S_i is the vector of structural traits associated with the ith site. The partial derivative of these hedonic functions with respect to any trait, say, S_j, describes the marginal change in the total site valuation associated with a change in that trait when all other trait levels are held constant. These partial derivatives reveal the same marginal valuation information as do prices in standard market analyses. For this reason the partial derivatives are often referred to as the shadow prices of the underlying locational traits.

Following Rosen (1974), we know that marginal trait prices are the result of the equilibration of demand and supply for traits. In particular the hedonic price functions (1) and (1′) represent the locus of supply-equal-demand intersections for various levels of each trait. If a trait is produced with constant returns to scale then the market clearing price in a competitive housing market is supply determined. Further, if each trait were produced by an independent production process we could a priori specify (1) and (1′) as linear equations. However, since many of the relevant housing traits are not producible (at least in a reasonable time horizon) and because the production of other traits tends to be a joint production process, it is difficult a priori to give a precise functional form to hedonic price functions. Instead we will examine simple functions and leave explicit modeling of trait interaction effects for future analysis.

Since the data used in housing market hedonic price studies are not designed for the purpose, a prime problem in estimating (1) and (1′) has been omitted variable bias. This bias has been particularly bothersome with respect to the N vector as standard data sets, such as the Census of Housing, collect little data on neighborhood characteristics. In 1973, the Census Bureau began conducting the AHS to provide a nationwide data base for use in analysis of topics such as the composition of the housing stock and changes in the neighborhood and structural qualities of the stock. In addition to data on a large set of structural housing characteristics the survey also contains data on a large set of neighborhood characteristics and socioeconomic information about the occupants.

Unfortunately, the survey provides relatively poor information about a much emphasized urban trait—accessibility. To the extent that market clearing prices of omitted measures of access are correlated with those of included traits the results reported here are biased. However, since the primary purpose of this paper is to develop a general statistical methodology for evaluating the hedonic price function of housing, this bias is ignored except as noted.

The public use computer tape file of the AHS contains approximately 52,000 observations on occupied housing units located in counties and independent cities constituting the 461 sampling areas used in current national surveys by the Census Bureau. The sampling rate is approximately 1 in 1350 housing units for dwelling units enumerated in 1970 and roughly twice that rate for units constructed since 1970.[2] In this paper we use a national subsample of 5694 observations on occupied housing units on less than 10 acres of land and located in the 34 largest metropolitan areas.[3] We also employ a subsample of 479 observations for the Chicago metropolitan area as well as a 541 observation subsample for the Los Angeles metropolitan area.

Table 1 reports summary statistics for variables utilized in this study. Of the 30 location-specific traits included, 12 refer to neighborhood characteristics such as street conditions and local traffic. As we shall see in the next section the results of our study indicate some missing variable biases in spite of the extensive set of location-specific traits available in this data set. The most notable missing variables are measures of accessibility to various work and recreational sites, room sizes, and lot sizes.

[2]For further detail regarding the AHS sampling structure the reader is advised to see "Annual Housing Survey: 1973 Technical Documentation," revised August 1976, Data User Services Division, Bureau of the Census.

[3]Unfortunately their location within the metropolitan areas is not available on the public use file.

TABLE 1
1973 AHS Summary Statistics

Variable name	Definition	Sample: National mean	Chicago mean	Los Angeles mean
BAG	Building age	28.71	31.26	25.696
BSIZ	Number of units in building	12.04	13.72	7.517
BFL	Number of floors in building	3.06	3.59	2.120
ELV	1 if building has an elevator	0.114	0.136	0.022
PRIV	1 if master bedroom is not relatively traffic free	0.094	0.046	0.063
KIT	1 if complete kitchen facilities	0.985	0.994	0.982
WABK	1 if a water breakdown recently	0.018	0.017	0.011
BATH	Number of bathrooms	1.201	1.118	1.319
PUSE	1 if connected to public sewer	0.976	1.000	0.980
CENT	1 if unit is centrally heated	0.377	0.290	0.275
AIRC	1 if central air conditioned	0.115	0.094	0.089
GCOL	1 if building provides garbage collection	0.929	0.956	0.869
BASE	1 if there is a basement	0.639	0.919	0.168
HOLEF	1 if there are holes in the floors	0.032	0.010	0.015
EXT	1 if building provides exterminator service	0.040	0.044	0.018
LIHL	1 if hall lighting in building is adequate	0.336	0.549	0.244
NOR	Number of nonbathrooms	3.545	3.575	3.263
GAR	1 if a garage facility available	0.593	0.468	0.824
ODOR	1 if neighborhood odor levels considered bad	0.104	0.113	0.790
CIAC	1 if neighborhood commercial activity considered bad	0.050	0.048	0.074
STRE	1 if street conditions in the neighborhood are bad	0.090	0.113	0.041
SLIT	1 if neighborhood street lighting is bad	0.098	0.079	0.142
CRIM	1 if neighborhood crime rates are considered bad	0.230	0.242	0.242
SCHQ	1 if neighborhood school quality is considered bad	0.071	0.075	0.072
ASHP	1 if neighborhood shopping facilities are considered bad	0.132	0.165	0.118
NOIS	1 if neighborhood traffic is considered bad	0.242	0.278	0.250
AIRN	1 if airplane noise in the neighborhood is considered bad	0.112	0.109	0.150
TRAF	1 if neighborhood traffic is considered bad	0.206	0.200	0.207
ABAN	1 if the neighborhood has abandoned buildings	0.058	0.071	0.037
TRAS	1 if neighborhood truck, junk, and litter are considered bad	0.154	0.125	0.133
V	Value of home (for owners)	26,832.	28,396.	36,942.
R	Annual rent (for renters)	1,710.	1,758.	1,759.
Percentage owners		0.609	0.320	0.400
Number of observations		5,694	479	541

A potentially serious source of bias in hedonic price studies is associated with functional form misspecification. Theory provides no strong restrictions on the functional forms of the hedonic functions (1) and (1′). Existing empirical work has specified these forms as either linear or log-linear and has not examined the statistical appropriateness of these forms or the bias they may induce in the marginal trait price estimates. Box and Cox (1964) develop a statistical model which determines the functional specification providing the best fit in terms of log likelihood. Their model specifies that some transformation of the dependent variable be normally distributed and linearly related to some set of transformations of the independent variables. Restriction to the family of power transformations implies that for the ith site,

$$\frac{V_i^{\lambda} - 1}{\lambda} = b_0 + \sum_{j=1}^{k} b_j n_i^j + \sum_{j=k+1}^{m} b_j s_i^j + e_i, \qquad \text{for owners,} \tag{2}$$

$$\frac{R_i^{\lambda} - 1}{\lambda} = a_0 + \sum_{j=1}^{k} a_j n_i^j + \sum_{j=k+1}^{m} a_j s_i^j + u_i, \qquad \text{for renters,} \tag{2'}$$

where $n_i^j \equiv ((N_i^j)^{v_j} - 1)v_j^{-1}$, $s_i^j \equiv ((S_i^j)^{v_j} - 1)v_j^{-1}$, e_i and u_i are normally distributed with means zero and variances J_V^2 and J_R^2, λ is the power transform factor on the dependent variable while v_j is the power transformation for the jth locational trait, and there are k neighborhood and $(m - k)$ structural site characteristics. Using the log likelihood function associated with this specification one can solve for the maximum likelihood estimates of the respective sets of coefficients and power transformation factors for both the owners' and renters' hedonic functions.[4] If a transformation factor equals 1 the associated functional form of the variable is linear. Similarly, using L'Hôpital's rule, it can be shown that as the power factor approaches zero the functional form approaches natural logarithmic form.

A major difficulty with applying this Box–Cox functional form search is that 25 of our 30 independent variables are dichotomous (i.e., one if yes; zero otherwise). This is problematic in that the power transformation for these variables, which include all of our neighborhood traits, must necessarily be linear. Thus the Box–Cox functional form search was conducted only with respect to the five continuous independent variables and the dependent variables. Preliminary investigation indicated that the value of

[4]See Heckman and Polachek (1974) for the derivation of the likelihood function. Kendall and Stuart (1961) discuss the test properties of the likelihood function and show that for large samples twice the difference in the log likelihood between a null and an alternative hypothesis is distributed χ^2 with the degrees of freedom depending on the number of parameters specified in the null hypothesis.

the likelihood function was substantially more sensitive to changes in the specification of the dependent variable than to changes in the specification of independent variables. This (and limited computer funds) suggests that the parameter search process should focus mainly on the specification of the dependent variables. The search over the form for the five relevant independent variables was restricted to the linear and natural logarithmic forms, where it was further assumed that the same power transformation was appropriate for all five independent variables. Therefore, the results reported in this paper refer to local and not necessarily global maximums.

III. ESTIMATES OF THE PROPERTY VALUE AND RENTAL HEDONIC PRICE FUNCTIONS

Table 2 reports the maximum likelihood estimates of equations (2) and (2′) for the Chicago, Los Angeles, and full urban samples.[5] The explanatory power of the location-specific characteristics is quite good, particularly in view of the micro nature of the observations.

The limited functional form search reveals that in 1973 the maximum likelihood specification of the property value function is the same in both Los Angeles and Chicago. Specifically, a natural logarithmic transformation of property value and a linear transformation of the independent variables provide the best description of the data; however, one cannot reject (at the 95% level) that the true λ for the property value equation is between 0.2 to -0.2 when $v_j = 1$ and 0.3 to -0.3 when $v_j = 0$. In both samples the statistically preferred property value equation is significantly better than the linear specification ($\lambda = 1$, $v_j = 1$) at the 99% level.

For the annual rental equations the preferred functional form differs across our subsamples. For Chicago the maximum likelihood value of λ is 0.4 and $v_j = 0$. This estimate of λ is quite precise as we are able to reject all values of λ greater than 0.5 or less than 0.3 at the 95% level. The maximum likelihood λ value for Los Angeles is 0.2 with $v_j = 1$. We reject (at the 95% level) values of λ less than zero or greater than 0.3.

In general, the signs of the hedonic price function coefficients are as expected. Because of omitted accessibility measures some of our results reflect the net impact of the measured trait and the omitted accessibility variable. For example, high levels of airplane noise are found to significantly increase Chicago property values. Since noise levels will be highest near the airport (and its expressway linkages) the variable AIRN reflects the net impact of the desirable trait of airport accessibility and the negative trait of airplane noise. Similar arguments explain the findings that bad street conditions raise site values in Chicago and undesirable levels of local

[5]Variables for which no coefficients are reported are omitted from the regressions because they exhibited no sample variance.

TABLE 2
Estimated City-Specific Hedonic Function for Property Value and Annual Rent

	Chicago		Los Angeles		Full Urban Sample	
Variable	Property value	Annual rent[a]	Property value	Annual rent	Property value	Annual rent
BAG	−0.0670	−2.3197	−0.0038	−0.0365	−0.1472	−0.0534
	(0.0024)	(0.7451)	(0.0023)	(0.0074)	(0.0175)	(0.0050)
BSIZ		.0048		0.3883		0.0034
		(0.0406)		(0.0075)		(0.0048)
BFL		−4.0435		0.5234		−0.0215
		(2.1024)		(0.1247)		(0.0280)
ELV		8.3591		0.9428		1.6480
		(3.5788)		(0.8072)		(0.3007)
PRIV	0.0692	0.5128	−0.0841	−0.4603	−3.9746	−0.7951
	(0.1675)	(2.0118)	(0.1134)	(0.3041)	(0.7608)	(0.1888)
KIT		7.6196		−0.6667	−3.3101	0.2747
		(4.4796)		(0.4943)	(2.9634)	(0.4295)
WABK		−0.4412	0.0651	0.4484	1.0715	−0.5610
		(2.8070)	(0.1635)	(0.8307)	(2.2386)	(0.3774)
BATH	0.1262	17.9363	0.3178	0.8807	7.5110	3.8979
	(0.0513)	(3.0963)	(0.0409)	(0.2727)	(0.3255)	(0.2151)
PUSE			−0.1561	0.1402	0.1409	0.7726
			(0.1197)	(0.6911)	(0.8366)	(0.6420)
CENT	−0.0638	0.3470	0.1913	0.7407	−0.7030	−0.0460
	(0.0582)	(1.2873)	(0.0565)	(0.2445)	(0.4103)	(0.1605)
AIRC	0.2167	8.7467	−0.0399	1.3825	1.4859	1.8599
	(0.0685)	(2.7594)	(0.0633)	(0.5861)	(0.5325)	(0.2690)
GCOL		1.3449	0.1081	−0.4153	−2.0075	0.8835
		(1.8562)	(0.0547)	(0.3206)	(1.0198)	(0.2234)
BASE	0.0938	1.8052	0.0240	0.1023	1.3975	0.8424
	(0.1317)	(1.4419)	(0.0640)	(0.2289)	(0.4201)	(0.1562)
HOLEF		0.8370	−0.5224	0.4489	−4.9496	−0.0098
		(3.6701)	(0.2784)	(0.5718)	(2.1754)	(0.2868)
EXT	0.3234	−1.0742	−0.1070	−1.1114	−1.0733	−1.1241
	(0.2791)	(1.8312)	(0.1541)	(0.6090)	(1.1340)	(0.2760)
LIHL		2.2741		−0.2623		0.8638
		(1.0871)		(0.1918)		(0.1401)
NOR	0.0166	2.5155	0.0459	0.6374	1.4962	0.6354
	(0.0251)	(1.0372)	(0.0198)	(0.0749)	(0.1545)	(0.0531)
CAR	0.1145	0.0692	0.0308	0.6154	6.5396	0.1079
	(0.0583)	(1.0895)	(0.0876)	(0.2013)	(0.4363)	(0.1391)
ODOR	−0.0546	−4.6458	0.0424	0.2457	0.6602	−0.4961
	(0.0770)	(1.5849)	(0.0759)	(0.3264)	(0.5958)	(0.2107)
CIAC	−0.0232	0.9516	0.3338	−0.0844	2.0472	0.2187
	(0.1048)	(2.3194)	(0.0940)	(0.3278)	(0.8406)	(0.2806)
STRE	0.1625	2.5757	0.0124	0.1013	−0.5802	0.2381
	(0.0718)	(1.5298)	(0.0997)	(0.4180)	(0.6203)	(0.2223)
SLIT	−0.1221	0.5435	0.0087	−0.0103	0.8908	0.4040
	(0.0957)	(1.7188)	(0.0529)	(0.2600)	(0.5879)	(0.2101)
CRIM	0.0835	0.8512	−0.0185	−0.2085	0.7662	−0.0234
	(0.0710)	(1.0378)	(0.0484)	(0.2030)	(0.4762)	(0.1487)
SCHQ	−0.1399	2.4170	−0.0912	0.1447	0.7257	0.6128
	(0.0788)	(1.9279)	(0.0791)	(0.3165)	(0.6552)	(0.2458)
ASHP	−0.0382	2.1541	−0.0390	−0.2318	−0.3914	−0.5146
	(0.0665)	(1.2521)	(0.0656)	(0.2495)	(0.5251)	(0.1835)

[a]All five continuous independent variables are in natural logarithmic form.

TABLE 2—*Continued*

Variable	Chicago: Property value	Chicago: Annual rent[a]	Los Angeles: Property value	Los Angeles: Annual rent	Full Urban Sample: Property value	Full Urban Sample: Annual rent
NOIS	−0.0697	0.4284	0.0147	−0.0145	−0.8285	−0.5841
	(0.0603)	(1.0542)	(0.0470)	(0.2224)	(0.4608)	(0.1591)
AIRN	0.3128	1.0582	−0.0623	0.1669	3.4856	0.1770
	(0.0632)	(2.0076)	(0.0532)	(0.2514)	(0.5397)	(0.2003)
TRAF	0.0385	−1.1188	−0.4369	−0.1361	−0.5113	0.3590
	(0.0705)	(1.1487)	(0.0553)	(0.2291)	(0.4924)	(0.1629)
ABAN	0.0398	2.1332	−0.03792	−0.5375	−2.9976	−0.2813
	(0.1681)	(1.6667)	(0.1498)	(0.4234)	(0.9138)	(0.2606)
TRAS	−0.0555	−2.0310	−0.0854	−0.3945	−2.7073	−0.4274
	(0.0942)	(1.3484)	(0.0643)	(0.2858)	(0.5703)	(0.1796)
Constant	9.8802	40.2141	9.7733	15.6699	51.3221	19.6839
λ	0	0.4	0	0.2	.3	.3
v_j	1	0	1	1	1	1
R^2	0.6626	0.5827	0.7843	0.7259	0.6759	0.5473
ln L	−1565	−2549	−2554	−2472	−23526	−27019

commercial activity significantly raise Los Angeles property values. The finding that inadequate local shopping facilities raise rents in Chicago is consistent with Linneman's (1977) finding that people prefer neighborhood homogeneity, which is absent in neighborhoods with major shopping facilities.

Table 3 displays the mean marginal trait valuations associated with the hedonic price functions reported in Table 2. For the dichotomous traits these trait prices reflect the mean marginal valuation of changing from zero to one while for the five continuous variables the marginal trait prices describe the change in total valuation associated with a small change in the trait. The trait prices shown in Table 3 are calculated as

$$P \equiv \partial R/\partial Z_j = b_j Z_j^{(v_j-1)} K^{(1-\lambda)}, \tag{3}$$

where K is equal to the mean rent for the renter-only sample and mean property value for the owner-only sample and Z_j is the jth location trait.

To address the question of the relative importance of structural and neighborhood traits in the determination of site prices we calculate the ratio of the sum of the absolute β coefficient values for the neighborhood traits to the same sum for all of the traits included in the regression for each sample.[6] For Chicago the absolute-value β-coefficient ratios with

[6] The β coefficient is a measure of the standardized impact of a variable and is defined as $b_j(\sigma_j/\sigma_y)$, where b_j is the estimated coefficient for the jth trait, σ_j is the standard deviation of trait j, and σ_y is standard deviation in the dependent variable.

TABLE 3

Marginal Trait Prices Evaluated (at Means) for City-Specific Property Value and Annual Rent Hedonic Functions

	Chicago		Los Angeles		Full Urban Sample	
Variable	Property value	Annual rent	Property value	Annual rent	Property value	Annual rent
BAG	−1902.53	−6.25	−140.38	−14.40	−185.34	−9.79
BSIZ		0.02		153.24		0.62
BFL		−82.73		206.56		−3.94
ELV		739.90		372.08		302.03
PRIV	1965.00	45.39	3106.82	−181.66	−5004.38	−145.72
KIT		674.44		−263.11	−4167.72	50.34
WABK		39.05	2404.92	176.96	1349.12	−102.82
BATH	3583.57	1565.69	11740.17	347.57	9457.03	714.38
PUSE			−5766.65	55.33	177.06	141.60
CENT	−1811.66	30.71	7067.00	292.32	−885.14	−8.43
AIRC	6153.41	774.20	−1473.98	545.60	1870.88	340.87
GCOL		119.04	3993.43	−163.90	−2527.62	161.92
BASE	2663.54	159.78	886.61	40.37	1759.58	154.39
HOLEF		74.09	−19298.50	177.16	−6232.00	−1.80
EXT	9183.27	−95.08	3952.79	−438.61	−1351.38	−206.02
LIHL		201.29		−103.52		158.31
NOR	471.37	73.07	1695.64	251.55	1883.85	125.61
CAR	3251.34	6.12	1137.81	242.87	8233.95	19.77
ODOR	−1550.42	−411.22	1566.34	96.96	831.25	−90.92
CIAC	−658.79	84.23	12331.24	−33.31	2577.61	40.08
STRE	4614.35	227.98	458.08	39.98	−730.52	43.64
SLIT	−3467.15	48.11	321.39	−4.06	1131.67	74.04
CRIM	2371.06	75.34	−683.43	−82.28	964.71	−4.29
SCHQ	−3972.60	213.94	−3369.11	57.10	913.72	112.31
ASHP	−1084.73	190.67	−1440.74	−91.48	−492.81	−94.31
NOIS	−1979.20	37.92	543.05	−5.72	−1043.16	−107.05
AIRN	8882.27	93.66	−2301.47	65.87	4388.69	32.44
TRAF	1093.25	−99.03	−16139.96	−53.71	−643.77	65.79
ABAN	1130.16	188.82	−1400.84	−212.12	−3774.25	−51.55
TRAS	−1575.98	−179.77	−3154.85	−155.69	−3408.74	−78.33

respect to the neighborhood traits are 0.48 for owners and 0.28 for renters while the ratios for Los Angeles are 0.28 and 0.17 for owners and renters, respectively.[7] These ratios indicate that neighborhood traits are nontrivial determinants of a location's valuation as they explain between 17 and 48% of the standardized variation of the site valuation. Care must be exercised, however, as the ratio is sensitive to the omission of independent variables from the regression. For example, if only neighborhood traits are included in the regression the value of this ratio will equal unity. For this reason we also use as an indicator the difference between the valuations of two hypothetical sites, which are identical with respect to structural traits, one

[7]If the independent variables had equal β coefficients (in absolute value) then these ratios would be 0.57, 0.41, 0.48, and 0.40, respectively.

site having the best possible and the other having the worst possible set of neighborhood traits.[8] The site in the hypothetical best neighborhood has values of one for all neighborhood variables which positively affect site values and zero values for all other neighborhood variables. Similarly the worst site has a value of unity for all neighborhood traits which reduce site values and zero for all other neighborhood traits. For the Chicago samples the differences between annual rent and property values of these two hypothetical sites are \$1964 and \$18,889, respectively. These represent 95 and 103% changes, respectively, relative to the average of the two valuations. For Los Angeles renters the difference between the annual rents of these two sites is \$1853, which is a 35% change relative to their average rent. Finally, the property values of these owner-occupied two sites in Los Angeles differed by \$28,373, or 106%.

Both measures of the importance of neighborhood traits indicate they can induce changes as large as the mean site valuation. Since most neighborhood traits are either directly controlled by public officials (e.g., local public goods such as street conditions and school quality) or indirectly influenced by their policy decisions (e.g., the passage and enforcement or ordinances with respect to noise pollution and abandoned structures), this is an important empirical finding as it establishes the fact that the impact of public policy actions may be significant. In Section V we examine the impact of an ordinance prohibiting dangerously abandoned buildings.

An issue which has received little empirical consideration is the geographic extent of a housing market. Generally, the boundaries of a housing market have been assumed to coincide with metropolitan area borders (see, for example, Kain and Quigley, 1975) or established arbitrarily in terms of local submarkets (see King, 1976, or Straszheim, 1975). Despite the presence of high geographic mobility of both development capital for housing and housing consumers the possibility of a national housing market has not been investigated. By national housing market we mean that suppliers and demanders flow across geographic locations so as to arbitrage all differences in the hedonic price function across location. In terms of Eqs. (2) and (2′), the national housing market hypothesis suggests that the relevant sample to employ in their estimation is a national sample of urban housing units. If this hypothesis is correct the use of local samples to estimate the hedonic price functions will induce sample selection bias to the extent that the local sample is not a random sample of the national sample.

[8]The structural traits for these calculations are: BAG = 10, NOR = 3, BSIZ = 1 for owners and 10 for renters, BFL = 1 for owners and 3 for renters, KIT = BATH = PUSE = GCOL = BASE = LIHL = GAR = 1, and all other structural traits equal zero.

The last two columns of Table 2 report the maximum likelihood estimates obtained by fitting Eqs. (2) and (2′) to a sample of observations from the largest 34 cities. For both the owner and renter samples the value of λ is 0.3 and v_j is unity. These estimates are extremely precise and are significantly better than all other combinations of λ and v_j.[9] In Table 3 we see the mean marginal trait prices for the full sample. To compare these prices across the sample specifications we calculate how many (city specific) standard errors the city-specific mean price for trait j deviates from the national sample point estimate.[10] One-third of the coefficients in the property value and 14% of those in the annual rent regressions for Chicago are more than 1.9 standard errors different from the full urban sample estimates. For Los Angeles these percentages are 24 and 27 for owners and renters, respectively. Only 4 (BATH, GCOL, NOR, and GAR) of the 30 independent variables have mean prices which are more than 1.9 standard errors in more than half of the regressions. Using the city-specific estimates of the standard errors of λ and v_j one cannot reject (at the 95% level) the functional forms of the hedonic regressions, which are the same for all three geographic samples. The impacts of neighborhood traits on site values are somewhat less in the full sample than in the local samples. The ratios of absolute β coefficients in the national sample are 0.25 for owners and 0.19 for renters.[11] The differential in annual rents between our two hypothetical sites is \$851 for owners (a 44% difference) and \$19,185 for owners (a difference of 80%). While not conclusive support of the national housing market hypothesis, this evidence does suggest its relevance. Further analysis of the hypothesis is warranted.[12]

IV. AN ANALYSIS OF ANNUALIZED HOUSING EXPENDITURES

In the last section we presented evidence of the hedonic price functions for owners and renters in 1973. The major problem associated with the use of these functions is that the property value equation gives us hedonic

[9]The search over λ is conducted in steps of 0.1.

[10]A Chow test was not performed as the generating equations are different between the samples.

[11]If the variables' β coefficient were equal these ratios would be 0.46 for owners and 0.40 for renters.

[12]One difficulty in determining whether there is a national housing market, i.e., whether a national sample should be used to estimate the hedonic functions, is that if the estimates are different for national and local samples the cause may be either that (i) the housing market is not national in scope and each locality has its own hedonic function and associated trait prices, or that (ii) there is a national housing market and estimates using local samples are biased as they reflect both the nonrandom selection rule and the true population coefficients. The identification of these two alternative interpretations lies beyond the scope of the current paper.

prices associated with the stock of housing while the annual rental hedonic prices refer to annual flow prices. The noncomparability of these prices is unfortunate as it prohibits one from directly addressing questions such as the price of conversion of an apartment building to a condominium. Of course, if one knew the appropriate capitalization rate to apply to the results such conversions would be possible, however, the results will be sensitive to the capitalization rate which is chosen. A related problem is that the estimated annualized hedonic prices obtained in this manner may differ from estimated annual rent prices, although theory suggests that suppliers and consumers will shift between these closely substitutable forms of obtaining housing services until the trait prices are equated across the owner and renter markets. This problem arises because the owner- and renter-only samples will be nonrandom samples (at least in terms of certain traits) of the entire housing services market. Therefore, parameter estimates obtained from these nonrandom samples will be biased estimates of the true trait prices in the housing market.[13]

To eliminate the owner–renter sample selection bias as well as the noncomparability problem one should pool the owner and renter samples and use annualized housing expenditures, H, as the dependent variable in a single hedonic price function for the urban housing market. Annualized housing expenditures are defined as annual rental payments plus utility payments for renters and property taxes plus the property value times the capitalization rate, r, plus utility payments for owners,

$$H \equiv \text{property value} \cdot r + \text{property taxes} + \text{annual rent} + \text{utilities}. \quad (4)$$

Utility payments are included in H as some rental units will include utilities in the contract rent; hence, for comparability, utilities must be added to those for which it is not included. Similarly, property taxes are included as annual rents implicitly include the renters' property tax payments so, for comparability, they too must be included for the remaining observations.

Applying the power transformation of the Box–Cox model to the hedonic price function for annualized housing expenditures, H, we obtain

$$\frac{H_i^{\lambda} - 1}{\lambda} = c_0 + \sum_{j=1}^{k} c_j n_i^j + \sum_{j=k+1}^{m} c_j s_i^j + w_i, \quad (5)$$

where the c_j's are the hedonic coefficients and w_i is a normally distributed error term with mean zero and variance J_H^2. Using the definition of H shown in (4) and its functional specification summarized in (5) one can

[13]See Heckman (1979) for a more complete discussion of the problem of sample selectivity bias.

TABLE 4
Estimated Hedonic For Annualized Housing Expenditures

Variable	Chicago	Los Angeles	Largest 34 cities
BAG	−0.9038	−0.2064	−0.0519
	(0.3183)	(0.0592)	(0.0053)
BSIZ	−0.0060	0.0089	0.0217
	(0.0206)	(0.0035)	(0.0053)
BFL	−0.7538	0.3934	0.0018
	(1.0579)	(0.5236)	(0.0320)
ELV	3.4670	−0.1105	0.4400
	(1.8306)	(0.6626)	(0.3423)
PRIV	0.5529	−0.2352	−0.4630
	(0.9122)	(0.1304)	(0.1858)
KIT	4.4501	−0.2606	1.2914
	(2.2978)	(0.2315)	(0.4476)
WABK	0.7427	0.0613	−0.2688
	(1.4222)	(0.2862)	(0.4014)
BATH	1.6080	0.8356	2.5179
	(0.8119)	(0.1239)	(0.1264)
PUSE		−0.1517	1.2992
		(0.2167)	(0.3495)
CENT	−1.1034	0.3769	−0.5288
	(0.4657)	(0.0915)	(0.1325)
AIRC	2.3094	0.0557	0.9288
	(0.6979)	(0.1290)	(0.1963)
GCOL	1.334	0.1167	0.3935
	(0.9484)	(0.0995)	(0.2345)
BASE	0.5564	0.0210	0.7857
	(0.6686)	(0.0895)	(0.1331)
HOLEF	0.8265	0.0050	0.8618
	(1.8533)	(0.2556)	(0.3116)
EXT	−0.5221	−0.2316	−0.7609
	(0.8940)	(0.2268)	(0.2735)
LIHL	2.6640	−0.0467	2.6083
	(0.4554)	(0.0920)	(0.1486)
NOR	0.2845	0.5161	0.0754
	(0.4880)	(0.0781)	(0.0439)
GAR	−0.3447	0.3216	0.6763
	(0.4290)	(0.0879)	(0.1245)
ODOR	−1.6638	0.1391	−0.1536
	(0.6028)	(0.1183)	(0.1844)
CIAC	0.0017	0.3424	0.2547
	(0.8749)	(0.1268)	(0.2513)
STRE	1.4682	−0.0528	−0.1485
	(0.5861)	(0.1543)	(0.1935)
SLIT	−0.1562	−0.0095	0.2707
	(0.1018)	(0.0886)	(0.1828)
CRIM	0.2765	−0.0185	0.1868
	(0.4560)	(0.0749)	(0.1366)

TABLE 4 —*Continued*

Variable	Chicago	Los Angeles	Largest 34 cities
SCHQ	0.7094	−0.0057	0.5875
	(0.7075)	(0.1186)	(0.2099)
ASHP	0.9641	−0.0499	−0.3505
	(0.4992)	(0.0966)	(0.1610)
NOIS	−0.2401	−0.0230	−0.3531
	(0.4403)	(0.0771)	(0.1407)
AIRN	1.6366	−0.0121	0.5724
	(0.6237)	(0.0878)	(0.1714)
TRAF	−0.2742	−0.0391	0.1565
	(0.4833)	(0.0839)	(0.1460)
ABAN	1.2902	−0.1570	−0.5007
	(0.7590)	(0.1766)	(0.2464)
TRAS	−1.0577	−0.2466	−0.6678
	(0.5901)	(0.1044)	(0.1645)
Constant	20.6510	10.6634	19.3328
λ	0.3	0.1	0.3
v_j	0	0	1.0
r	0.02	0.04	0.03
R^2	0.5481	0.7107	0.4927
ln L	−3727	−4180	−44382
Mean	1545.06	1937.13	1527.01

identify the maximum likelihood estimates of λ, v_j, and c_j, as well as the appropriate mean capitalization rate, r. Table 4 reports these maximum likelihood estimates for the two local samples as well as the national sample.

The estimates for all three samples are more significant and, as expected, their signs are more consistent with a priori expectations than is the case for the separate samples. The only notable unexpected signs are that central heat decreases annual housing expenditures significantly in the Chicago and full samples, a result which also held in half of the regressions shown in Table 2. The maximum likelihood estimates of λ, v_j, and r are: $\lambda = 0.3$, $v_j = 0$, and $r = 0.02$ for Chicago; $\lambda = 0.1$, $v_j = 0$, and $r = 0.04$ for Los Angeles; and $\lambda = 0.3$, $v_j = 1$, and $r = 0.03$ for the national sample.[14] Using standard likelihood ratio tests one cannot reject for Chicago (at the 95% level) λ in the range 0.2 to 0.4 when $v_j = 0$ and $r = 0.02$; λ in the range 0.2 to 0.5 when $v_j = 1$ and $r = 0.02$; λ in the range 0.2 to 0.5 when $r = 0.03$ and $v_j = 1$; and λ equal to 0.3 or 0.4 when $v_j = 0$ and $r = 0.03$. Similarly, for the Los Angeles sample the maximum likelihood estimates

[14] λ was once again varied in increments of 0.1 while r was changed by steps of 0.01.

are significantly superior to all specifications except: λ in the range 0.0 to 0.2 for $r = 0.04$ and $v_j = 0$; λ in the range -0.1 to 0.3 when $r = 0.04$ and $v_j = 1$; and λ equal to 0.2, 0.1, 0, or -0.1 when $r = 0.3$, and $v_j = 1$. The estimates for the full sample are significantly superior to all other specifications.

The maximum likelihood capitalization rates for these 1973 samples are 0.02 for Chicago, 0.04 for Los Angeles, and 0.03 for the national sample. These estimates are much lower than the 10% rate which is generally suggested. One possible explanation of these low estimates is in line with the Fama and Schwert (1977) finding that housing is the only effective portfolio hedge against inflation. This suggests that housing returns will command a negative risk premium in periods of inflationary expectations and hence their rate of return will be relatively low. Our estimates are consistent with their finding to the extent that 2 to 4% capitalization rates are below the average 1973 nominal rate of return and that 1973 was most likely a period of relatively high inflationary expectations. Neither local capitalization rate is significantly different from 3% for appropriate specifications of λ and v_j.[15]

The mean marginal trait prices obtained from the annualized housing expenditure hedonic price functions are displayed in Table 5. These prices are interpreted as the mean change in total annual housing expenditures when one of the continuous traits is varied slightly or when one of the dichotomous traits changes from zero to one when all other trait levels are held constant. Thus, for example, annual expenditures in Chicago fall by an average of \$4.94 for each additional year in a building's age. For renters this is directly interpreted as the reduction in their annual rental payments while for owners with a 20-year horizon, using the 2% capitalization rate implies a \$7.40 reduction per year in the property value of the mean owner's site.

Two questions are empirically interesting in terms of the hedonic price estimates contained in Table 5: (i) How important are neighborhood characteristics in the determination of the total annual housing expenditures which are necessary to obtain a site? (ii) Are the results of this hedonic specification consistent with the national housing market hypothesis?

The previously described ratios of the absolute β coefficients for these samples are 0.30, 0.18, and 0.18 for the Chicago, Los Angeles, and national samples, respectively. The differences in annualized housing expenditures are \$2000 (a 93% change), \$932 (a 51% change), and \$830 (a 43% change) for Chicago, Los Angeles, and the national samples. These impacts are

[15]An interesting extension of this approach would be to estimate r for many locations and then regress these estimates on a set of variables such as income and income variance in an attempt to ascertain the determinants of the mean capitalization rate.

TABLE 5
Marginal Trait Prices (at Means) for Annualized Housing Expenditures

	Chicago	Los Angeles	Largest 34 cities
BAG	−4.94	−7.30	−8.79
BSIZ	− .07	1.07	3.67
BFL	−35.84	168.63	0.30
ELV	591.85	−100.42	74.50
PRIV	94.39	−213.74	−78.39
LIT	759.68	−236.82	218.65
WABK	126.79	55.71	−45.51
BATH	245.53	575.70	426.31
PUSE		−137.86	219.97
CENT	−108.36	342.51	−89.53
AIRC	394.24	50.62	157.26
GCOL	227.80	106.05	66.62
BASE	94.98	19.08	133.03
HOLEF	141.09	4.54	145.91
EXT	−89.13	−210.47	−128.83
LIHL	625.48	−42.44	441.62
NOR	13.59	143.73	12.77
GAR	−58.85	292.25	114.51
ODOR	−284.03	126.41	−26.00
CIAC	.29	311.15	43.12
STRE	250.64	−47.98	−25.14
SLIT	−26.66	−8.63	45.83
CRIM	47.20	−16.81	31.62
SCHQ	121.10	−5.18	99.45
ASHP	164.58	−45.35	−59.34
NOIS	−40.99	−20.90	−59.77
AIRN	279.38	−11.00	96.91
TRAF	−46.81	−35.53	26.50
ABAN	220.25	−142.67	−84.77
TRAS	−180.56	−224.10	−113.07

once again quite large, with 18 to 30% of the standardized variation in H determined by neighborhood traits, and with these traits capable of inducing site payment differences of 43 to 93%.

With respect to the national housing market hypothesis 27% of the Chicago and 23% of the Los Angeles mean hedonic price estimates are more than 1.9 standard errors different from the full sample point estimates. Of these only 3% (those for BSIZ, LIHL, and GAR) were more than 1.9 standard errors different in both samples. Further, in the Chicago sample one cannot reject (at the 95% level) the hypothesis that the functional specification is the same as that for the national sample. This evidence is once again suggestive of the appropriateness of the national housing market hypothesis and indicates that a more thorough analysis of this hypothesis would be productive.

V. POTENTIAL POLICY ERRORS INDUCED BY SPECIFICATION BIAS

In the first four sections of this paper we have discussed and empirically examined four significant sources of estimation bias in hedonic price studies of the urban housing market. To eliminate these biases we used a sample which contains many relevant neighborhood and structural traits, conducted statistical tests to determine the appropriate functional form, constructed a measure of annual housing expenditures which is applicable to both owners and renters, and examined the geographic extent of an urban housing market. In this section we show how these corrections may lead to significantly different decisions for both public and private urban decision makers. Implicit throughout this discussion is the assumption that for the period of analysis we are discussing, the 1973 trait prices are the best estimators of prevailing trait prices.[16] We will ignore issues of forecast errors, returns to risk, and risk preference.

Suppose a consultant is hired by the owner of a Chicago apartment building, which has the structural characteristics described in footnote 8 and a zero amenity vector, to ascertain the maximum price he can expect from condominium developers.[17] If the consultant uses the linear specification estimates (say ours) for a sample of Chicago property owners, the estimated asking price for the building will be $283,127.[18]

This estimate is biased for three reasons: (i) some relevant trait prices were not estimable for the owner-only sample, because of inadequate sample variation; (ii) the owner-only sample estimates are not obtained from a random sample of the housing market; and (iii) the functional form of the hedonic function is incorrectly specified.[19] The maximum likelihood estimates of the hedonic price function for annualized housing expenditures in Chicago (column 1, Table 4) are capable of eliminating these sources of bias. Using these estimates, a 20-year horizon, and a 10% opportunity rate of return, the estimated maximum price for the same

[16] A more complete approach would be to replicate the analysis in this paper for each year's AHS and then make period t's expected trait price some polynomial function of past prices and a trend.

[17] The assumptions here are that (i) condominium developers are relatively more efficient than he is at converting buildings from apartments to condominiums; and (ii) for whatever reasons, he must sell his building in a very short period of time, that is, he does not have much time to search for the maximum developer bid.

[18] The predicted value per unit is $28,312.70 and there are 10 units in the building.

[19] For brevity, we will ignore the possibility of sample selection bias induced by the geographic definition of the market.

building is \$125, 352, or over 50% less than the other estimate.[20] If the consultant were to form his estimate using the Chicago sample maximum likelihood property value equation (Table 2, column 1) the estimated maximum asking price would be \$146,853. This suggests that 86% of the overestimate obtained using the linear hedonic property value function is due to functional form misspecification, with the remaining 14% due to sample selectivity bias associated with using an owner-only sample. If, as theory suggests, the maximum likelihood estimates using H as the dependent variables most accurately describe the world, then by using the linear form and an inappropriate dependent variable the consultant will grossly mislead his client into holding out for an unreasonably high price.[21]

A similar problem is faced by developers in deciding what type of residence site to supply. Consider the case of a Los Angeles developer who is building a tract of single-family residences in a neighborhood with a zero vector of neighborhood traits and who knows that the cost of building a unit with two bathrooms is \$2000 greater than the cost of building a single-bathroom residence.[22] If the developer's analysts use a linear specification of the Los Angeles property values (say ours) they will report that the second bathroom will raise values by \$13,493 per unit or an annual average rate of return of 29%. Using this estimate the developer will supply his units with two bathrooms only to find out that this was an incorrect policy, owing to the previously described estimation biases. The "true" estimated annualized expenditure hedonic function for Los Angeles (column 2, Table 4) implies a 2.1% annual average rate of return, which is clearly less than the alternative asset's 8%. Hence our best estimates of the state of the market suggest that the developer should supply one-bathroom residences. Using the maximum likelihood Los Angeles property value estimates we see that 29% of the overestimate associated with the linear estimate is due to functional form, with the remainder associated with sample selectivity bias induced by using an owner-only sample.

Our final example of incorrect policy conclusions which can result from inappropriate specification considers a scenario in which a city is considering a bill to have the city pursue an ongoing program of demolishing

[20]This is obtained by: \$1696.4 per unit per year annual expenditures on 10 units (\$16,964 for the building). At a 10% discount rate for a 20-year horizon this gives a current valuation of \$125,352. Of course, this estimate is sensitive to the discount rate and horizon which are chosen.

[21]The building owner may quickly realize the inaccuracy of the consultant's linear property value estimate, because developers say it is an unreasonable price, because there is an absence of interest by condominium developers, or because he makes a crude comparison with the selling prices of similar buildings. In this case the consultant will also be made to pay for his error as the owner will fire him and pass the word that the consultant does low-quality work.

[22]The other optimal trait levels are assumed to be those described in footnote 8.

abandoned buildings in a neighborhood where abandoned buildings (and their associated externalities) are particularly prevalent. The neighborhood's occupied housing stock is predicted to be composed of 100 ten-unit rental structures and 2000 owner-occupied units whose respective trait vectors are described in footnote 8. The neighborhood has a zero neighborhood trait vector with the exception that ABAN = 1. City engineers have estimated that the average annual cost of the proposed ordinance, which changes ABAN from 1 to 0, is $100,000. The city hires a consultant to answer two questions: (i) Will the project be politically feasible in the sense that the monetary benefits exceed the costs of the program? (ii) Will property taxes have to be increased in order to pay for the program (they are currently at 4%)?

Since we are not told what city the project is proposed for we will treat our national sample estimates as the best estimators. If the consultant uses a linear form of annualized housing expenditures with a 10% capitalization rate he estimates that total neighborhood annual expenditures, his best estimate of the maximum amount the community is willing to pay annually to have the city eliminate the externality of abandoned buildings, will rise by $565,710. Further, using an 8% discount rate and a 20-year horizon this estimator implies a total rise in local tax revenues of $112,079. Therefore, this estimator implies not only that additional revenues will be sufficient to cover the costs of the program but also that a surplus of approximately $12,000 will be generated by the program.

These estimates are biased, due to both measurement error associated with an incorrect choice of the capitalization rate and functional form specification bias. Using the full-sample maximum likelihood estimator (column 3, Table 4) we see that the best statistical estimate of the maximum benefit of the program is $283,100, about half of the alternative estimate. However, both estimators indicate that the project is politically feasible as the maximum annual increase exceeds the annual cost estimate. The maximum likelihood estimator implies an additional $56,088 annually in property tax revenues at the current tax rate. Thus, while the biased estimator indicates an annual fiscal surplus of $12,000 for the project, the maximum likelihood estimator predicts a $44,000 annual fiscal deficit.

In this section we have presented three different, although somewhat artificial, cases in which urban policy makers may be seriously misled if their technical advisors use inappropriate estimators of the marginal trait valuations. In each case the appropriate maximum likelihood estimator is compared to a simple linear estimator and in each case the biased, linear estimator overpredicted the impact of the policy under consideration. Further, we saw that the costs of these incorrect policy decisions may be nontrivial.

VI. SUMMARY AND CONCLUSIONS

We have reviewed the standard estimation procedures for hedonic price functions of the urban housing market. Several sources of bias associated with these procedures were noted including those induced by missing variables, functional form misspecification, sample selection with respect to geographic extent of the market, sample selection bias with respect to tenure, and incorrect choice of the capitalization rate. It was noted that the AHS is quite extensive with respect to location-specific traits and these data were used together with the Box and Cox methodology to obtain the maximum likelihood estimates for the hedonic functions for annual rent, property value, and annualized housing expenditures. The latter is argued as the appropriate theoretical structure as it explicitly incorporates and identifies the effective capitalization rate. This rate was found to be quite low in our sample, which is consistent with the hypothesis that housing is an effective inflationary hedge. We also examined some crude evidence on the empirical relevance of the national housing market hypothesis and suggested that it merits future examination. We further found empirically that neighborhood-specific traits are important determinants of a site's valuation, explaining 15 to 50% of the standardized variation in valuations and inducing differential valuations as large as 100% between structurally identical sites. In the last section we presented evidence of the magnitude of policy errors which may be induced by the use of biased hedonic price function estimators and noted that the costs of these errors may be large.

In closing it should be stated that the purposes of this paper have been to report the best estimates of the state of the urban housing market in 1973 as well as to present a systematical methodology which may be applied to other data or urban housing markets. Since there is little reason to believe that the results found for 1973 will generalize to other years (i.e., relative prices vary over time) the main contribution of this paper is to help in establishing a consistent and systematic statistical procedure to be used by urban economists in the study of urban housing markets and in evaluating relevant public and private policy decisions.

REFERENCES

Box, G., and Cox, D. (1964), An analysis of transformations, *J. Roy. Statist. Soc., Ser. B*, **26**, No. 2, 211–252.

Dewees, D. N. (1976), The effect of subway improvement on residential property values in Toronto, *J. Urban Econ.*, **3**, No. 4, 357–369.

Fama, E., and Schwert, G. W. (1977), Asset returns and inflation, *J. Financ. Econ.*, **5**, 115–146.

Griliches, Z. (1961), Hedonic price indices for automobiles: An econometric analysis of quality change, *in* "Models of Income Determination, Studies in Income and Wealth," Vol. 28, Columbia Univ. Press for NBER, New York.

Griliches, Z., Ed. (1971), "Price Indexes and Quality Change," Harvard Univ. Press, Cambridge, Mass.

Heckman, J. J. (1979), Sample selection bias as a specification error, *Econometrica* (January), 153–162.

Heckman, J. J. and Polachek, S. (1974), Empirical evidence on the functional form of the earnings–schooling relationship, *J. Amer. Statist. Assoc.*, **69**, No. 346, 350–354.

Kain, J., and Quigley, J. M. (1975), "Housing Markets and Racial Discrimination: A Microeconomic Analysis," Columbia Univ. Press for NBER, New York.

Kendall, M. D., and Stuart, A. (1961), "The Advanced Theory of Statistics," Vol. 2, Hafner, New York.

King, A. T. (1976), The demand for housing: Integrating the roles of journey-to-work, neighborhood quality, and prices, *in* "Household Production and Consumption," (N. E. Terleckyj, Ed.) Columbia Univ. Press for NBER, New York.

King, A. T., and Mieszkowski, P. (1973), Racial discrimination, segregation and the price of housing, *J. Pol. Econ.* (May/June).

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

Rosen, S. (1974), Hedonic prices and implicit markets: Product differentiation in pure competition, *J. Pol. Econ.* (January/February), 34–55.

Straszheim, M. (1973), Estimation of the demand for urban housing services from household interview data, *Rev. Econ. Statist.* (February).

Straszheim, M. (1974), Hedonic estimation of housing market prices: A further comment, *Rev. Econ. Statist.* (August).

Straszheim, M. (1975), "An Econometric Analysis of the Urban Housing Market," Columbia Univ. Press for NBER, New York.