Determinants of House Prices: A Quantile Regression Approach

**Joachim Zietz · Emily Norman Zietz ·
G. Stacy Sirmans**

Abstract OLS regression has typically been used in housing research to determine the relationship of a particular housing characteristic with selling price. Results differ across studies, not only in terms of size of OLS coefficients and statistical significance, but sometimes in direction of effect. This study suggests that some of the observed variation in the estimated prices of housing characteristics may reflect the fact that characteristics are not priced the same across a given distribution of house prices. To examine this issue, this study uses quantile regression, with and without accounting for spatial autocorrecation, to identify the coefficients of a large set of diverse variables across different quantiles. The results show that purchasers of higher-priced homes value certain housing characteristics such as square footage and the number of bathrooms differently from buyers of lower-priced homes. Other variables such as age are also shown to vary across the distribution of house prices.

J. Zietz (✉)
Department of Economics and Finance, Middle Tennessee State University,
Murfreesboro, TN 37132, USA
e-mail: jzietz@mtsu.edu

E. N. Zietz
Department of Economics and Finance, Middle Tennessee State University,
P.O. Box 27, Murfreesboro, TN 37132, USA
e-mail: ezietz@mtsu.edu

G. S. Sirmans
Department of Insurance, Real Estate and Business Law, The Florida State University,
Tallahassee, FL, USA
e-mail: gsirmans@cob.fsu.edu

Introduction

The published real estate literature has put forth a number of housing characteristics to explain house prices. Hedonic regression analysis is typically used to identify the marginal effect on house price of each of these housing characteristics. Sirmans et al. (2005) examine hedonic pricing models for more than 125 empirical studies and find that studies often disagree on both the magnitude and direction of the effect of certain characteristics. For example, their analysis shows that, of 40 empirical studies examining the number of bedrooms, 21 studies find that bedrooms have a positive impact on house price, nine studies identify a negative relationship, and ten studies report no significant relationship between house price and the number of bedrooms.

Different estimation results for a given variable, in particular disagreement on the direction of the effect, can be confusing to market participants. In addition, there may be reason to believe that housing characteristics are not valued the same across a given distribution of house prices. Malpezzi et al. (1980) acknowledge the problem in valuing individual house features and note that the impact on price of individual features cannot be easily quantified. Malpezzi (2003) also notes that different consumers may value housing characteristics differently. To alleviate some of the confusion, this study examines the extent to which conflicting results may be attributed to differences in the effect of housing characteristics across the distribution of house prices. For example, if a particular housing characteristic is priced differently for houses in the upper-price range as compared to houses in the lower-price range, the typical OLS regression may not provide useful information for either price range since it is based on the mean of the entire price distribution.

As an alternative to OLS regression, this study uses quantile regression to identify the implicit prices of housing characteristics for different points in the distribution of house prices. This explicitly allows higher-priced houses to have a different implicit price for a housing characteristic than lower-priced houses. Since quantile regression uses the entire sample, the problem of truncation is avoided (Heckman 1979). This will eliminate the problem of biased estimates that is created when OLS is applied to house price sub-samples (e.g., Newsome and Zietz 1992).

The Implicit Pricing of Housing Characteristics

Sirmans et al. (2005) review the hedonic pricing models of 125 empirical studies. Some of their results are summarized in Tables 1 and 2. As shown, there is some parameter uncertainty for even key housing characteristics. This parameter uncertainty manifests itself in signs that are opposite to expectations or estimates that are statistically insignificant. For example, age is the variable most often included in hedonic pricing models. Although age has a negative sign in most studies, it is positive in some. In contrast, the general expectation is that the number of bedrooms would have a positive effect on house price. Of 40 studies examining this variable, almost half (19 studies) show a negative or not-significant result.

A key question is the cause of this parameter uncertainty. Based on the findings of Sirmans et al. (2005), it seems unlikely that parameter variation for housing characteristics can be fully explained by regional differences, different specifica-

Table 1 Variables with predominantly consistent results across studies

Variable	Appearances	No. times positive	No. times negative	No. times non-significant
Lot size	52	45	0	7
Square feet	69	62	4	3
Brick	13	9	0	4
No. bathrooms	40	34	1	5
No. rooms	14	10	1	3
Full baths	37	31	1	5
Fireplace	57	43	3	11
Air-conditioning	37	34	1	2
Basement	21	15	1	5
Garage spaces	61	48	0	13
Pool	31	27	0	4

The results are from Sirmans et al. (2005)

tions, or alternative data sets. In addition, as suggested by Newsome and Zietz (1992), housing characteristics may not be valued the same across a given distribution of housing prices. Specifically, the marginal value, percentage contribution, or elasticity value of a certain housing characteristic may be different across the range of house prices. In fact, would one expect to find that owners of high-end houses and low-end houses attach the same value to every housing characteristic? This would require that the preference structure of all homeowners be identical and that the owners of low-end and high-end homes differ only in the income constraint they face.

As discussed by Rosen (1974), Epple (1987), and Bartik (1987), the demand and supply functions that underlie hedonic price equations can be very difficult to identify empirically. The general acceptance of hedonic pricing models in real estate application rests on the assumption that the underlying supply function of housing characteristics is vertical in price/quantity space. The supply of housing characteristics is fixed at any given point in time and is independent of the implicit price of a characteristic. The intersection of the downward sloping demand curve for a housing characteristic with the given vertical supply curve of that characteristic identifies the implicit price of the housing characteristic. This implicit price is identical to the one generated by the hedonic pricing model. Assuming that all consumers are equal, then the implicit price of a characteristic is the implied valuation of that characteristic by the representative consumer. OLS estimation fits nicely into this representative agent framework since it identifies those implicit prices that optimally predict the mean house price for a given sample.

A problem arises when the relevance of the representative agent paradigm is questioned.[1] For the sake of argument, assume that there are two consumers: a "poor" one who is income and credit constrained and a "rich" one who is not. The poor consumer is not in the market for an expensive house because no bank will underwrite the needed loan and the rich household would not think of buying a poor man's house because it does not provide the desired amenities and may negatively affect his/her desire for social status. Thus, in essence, there are two segmented

---

[1] See Kirman (1992) for a scathing critique of the representative agent paradigm.

Table 2 Variables with predominantly inconsistent results across studies

Variable	Appearances	No. times positive	No. times negative	No. times not significant
Age	78	7	63	8
Bedrooms	40	21	9	10
Distance	15	5	5	5
Time on market	18	1	8	9

The results are from Sirmans et al. (2005)

markets. Segmentation may not only imply that the rich and the poor occupy houses of different values but they may also develop group-specific likes and dislikes of certain housing characteristics.[2] Builders, aware of this situation, would build houses to fit the perceived needs of the groups. What results is not one set of supply curves of housing characteristics but two, one for the "rich" household and one for the "poor" household. Similarly, there are two sets of demand curves for each housing characteristic resulting in two sets of implicit prices for housing characteristics.

The above argument suggests that there may be marked differences in the elasticity of house price with respect to housing characteristics across the distribution of housing prices. A seemingly logical approach would be to tie the different segments to the house price. A high house price rations "poor" households out of the market intended for "rich" households and a low housing price is a sufficient deterrent for entry by a "rich" household. The major task is to identify the different market segments and their implicit prices. In this regard, the usefulness of OLS regression may be questioned and a more appropriate approach may be quantile regression.

Quantile Regression Methodology

Quantile regression is based on the minimization of weighted absolute deviations (also known as L_1 method) to estimate conditional quantile (percentile) functions (Koenker and Bassett 1978; Koenker and Hallock 2001). For the median (quantile= 0.5), symmetric weights are used, and for all other quantiles (e.g., 0.1, 0.2 ....., 0.9) asymmetric weights are employed. In contrast, classical OLS regression (also known as L_2 method) estimates conditional mean functions. Unlike OLS, quantile regression is not limited to explaining the mean of the dependent variable. It can be employed to explain the determinants of the dependent variable at any point of the distribution of the dependent variable. For hedonic price functions, quantile regression makes it possible to statistically examine the extent to which housing characteristics are valued differently across the distribution of housing prices.

One may argue that the same goal may be accomplished by segmenting the dependent variable, such as house price, into subsets according to its unconditional distribution and then applying OLS on the subsets, as done, for example, in

---

[2] The articles in Durlauf and Young (2001) provide a good idea of the social dynamics that may evolve and why they may evolve.

Newsome and Zietz (1992). However, as clearly argued by Heckman (1979), this "truncation of the dependent variable" may create biased parameter estimates and should be avoided. Since quantile regression employs the full data set, a sample selection problem does not arise.

Quantile regression generalizes the concept of an unconditional quantile to a quantile that is conditioned on one or more covariates. Least squares minimizes the sum of the squared residuals,

$$\min_{\{b_j\}_{j=0}^{k}} \sum_i \left(y_i - \sum_{j=0}^{k} b_j x_{j,i} \right)^2,$$

where y_i is the dependent variable at observation i, $x_{j,i}$ the jth regressor variable at observation i, and b_j an estimate of the model's jth regression coefficient. By contrast, quantile regression minimizes a weighted sum of the absolute deviations,

$$\min_{\{b_j\}_{j=0}^{k}} \sum_i \left| y_i - \sum_{j=0}^{k} b_j x_{j,i} \right| h_i,$$

where the weight h_i is defined as

$$h_i = 2q$$

if the residual for the ith observation is strictly positive or as

$$h_i = 2 - 2q$$

if the residual for the ith observation is negative or zero. The variable q ($0 < q < 1$) is the quantile to be estimated or predicted.

The standard errors of the coefficient estimates are estimated using bootstrapping as suggested by Gould (1992, 1997). They are significantly less sensitive to heteroskedasticity than the standard error estimates based on the method suggested by Rogers (1993).[3]

Quantile regression analyzes the similarity or dissimilarity of regression coefficients at different points of the distribution of the dependent variable, which is sales price in our case. It does not consider spatial autocorrelation that may be present in the data. Because similarly priced houses are unlikely to be all clustered geographically, one cannot expect that quantile regression will remove the need to account for spatial autocorrelation.

In this paper, spatial autocorrelation is incorporated into the quantile regression framework through the addition of a spatial lag variable. The spatial lag variable is defined as $\mathbf{W}y$, where $\mathbf{W}$ is a spatial weight matrix of size $T \times T$, where T is the number of observations, and where y is the dependent variable vector, which is of size $T \times 1$. Any spatial weight matrix can be employed, for example, one based on the ith nearest neighbor method, contiguity, or some other scheme. In the present application, a contiguity matrix is used.[4]

---

[3] The quantile regressions employ the "sqreg" command in Stata for seed 1001.

[4] The Matlab program xy2cont.m of J.LeSage's Econometrics Toolbox is employed, which is an adaptation of the Matlab program fdelw2.m of Kelley Pace's Spatial Statistics Toolbox 2.0.

Adding a spatial lag to an OLS regression is well known to cause inference problems owing to the endogeneity of the spatial lag (Anselin 2001). This is not any different for quantile regression than for OLS. We follow the approach suggested by Kim and Muller (2004) to deal with this endogeneity problem in quantile regression. As instruments we employ the regressors and their spatial lags.[5] However, instead of using a density function estimator for the derivation of the standard errors, we follow the well established route of bootstrapping the standard errors (Greene 2000, pp. 400–401).[6]

Data and Estimation Results

This study uses multiple listing service (MLS) data from the Orem/Provo, Utah area.[7] The data consist of 1,366 home sales from mid-1999 to mid-2000. Table 3 provides a description of the variables. Most are standard housing characteristics while some are specific to the region. The data also include a number of geographic and neighborhood variables, which are derived by geo-coding all observations. An objective is to measure the effect of quantile regression on a large number of diverse variables. Table 4 gives summary statistics for the explanatory variables and the dependent variable, sale price. The quantile values reported in Table 4 for the independent variables are averages of the values that are associated with the sale prices found in a 5% confidence interval around a given quantile point of the dependent variable (sp). For example, the sale price associated with quantile point 0.2 is \$123,000. A 5% confidence interval of this quantile point covers the price range from \$121,902 to \$124,526 and the houses with sale price in this range have on average square footage of 1,760.6.

The hedonic pricing model takes the form

$$ln(sp) = \alpha + \sum_i \beta_i X_i + \varepsilon,$$

where selling price (sp) is expressed in logged form, α is a constant term, β_i is the regression coefficient for the ith housing characteristic, X_i, and ε is the residual error term.

The estimation results for the quantile regressions that do not account for spatial autocorrelation are presented in Tables 5 and 6. Table 5 gives the coefficient estimates and Table 6 provides the associated probability values (p values). P values of less than 0.05 indicate statistical significance of a coefficient estimate at the 5% level or better.[8] Both Tables 5 and 6 present the results of the standard OLS regression in the leftmost column and the estimates of the quantile regressions in the

---

[5]If **X** identifies the data matrix, then the spatial lags of the regressors are computed as **WX**, where **W** is the spatial weight matrix used for the construction of the spatial lag of the dependent variable.

[6]The bootstrap is based on 500 replications.

[7]The data used are similar to the data used in Zietz and Newsome (2002).

[8]Variance inflation factors (VIF) are calculated for all variables. The maximum VIF is 2.51, the mean VIF is 1.54. This does not suggest that the regressions suffer from multicollinearity.

Table 3 Variable definitions

Variable	Definition
sp	Sale price in 1,000 dollars; ln(sp)=dependent variable
lagh	Spatial lag variable, based on normalized contiguity weight matrix
sqft	Size of house in square feet, divided by 1,000
acres	Lot size in acres
year	Year in which the property was built
bedr	Number of bedrooms
bathf	Number of full bathrooms
batht	Number of 3/4 bathrooms (shower, no tub)
bathh	Number of half baths
deck	Number of decks
patio	Number of patios
garage	Number of garage places
basmt	Percentage of house covered by finished basement
pool	1 if pool is present, 0 otherwise
airevr	1 if air conditioning is evaporator, roof type, 0 otherwise
airevw	1 if air conditioning is evaporator, window type, 0 otherwise
airel	1 if air conditioning is electric, 0 otherwise
airgas	1 if air conditioning is gas, 0 otherwise
flhar	1 if hardwood flooring is present in house, 0 otherwise
fltil	1 if tile flooring is present in house, 0 otherwise
extu	1 if exterior is made of stucco
exbri	1 if exterior is made of brick
exalu	1 if exterior is made of aluminum
exfra	1 if exterior is of type frame
laful	1 if full landscaping
lapar	1 if partial landscaping
lotspr	1 if lot contains a sprinkler system
lotmtn	1 if lot has mountain view
di15	Distance to interstate Highway 15, in miles (US Topographical map)
dorem	Distance to city center of Orem, in miles (US Topographical map)
earthqk	Magnitude of largest earthquake, on Richter Scale (EPA data)
nwrate	Percentage of population classified as non-white, by census tract
forrent	Percentage of all vacant housing units for rent, by census tract

remainder of the tables.[9] The points on which the quantile regressions are centered are provided in the first row of Table 4. Tables 7 and 8 present the quantile regression results when spatial autocorrelation is taken into account.

Table 9 contains all variables for which marginal effects can be calculated. The marginal effects are the product of the coefficient reported in Table 7 and the relevant housing price from Table 4, multiplied by 1,000. The relevant price is the mean price for 2SLS and the associated quantile point for the quantile regressions. The marginal effects given in Table 9 reflect the prices of 1999/2000. Table 10 converts all percentage change effects reported in Table 7 into dollar values by multiplying the coefficients of Table 7 by the respective sale prices of Table 4, multiplied by 1,000. Table 11 reports price elasticities for square footage and acreage. The elasticities are derived as the product of the estimated coefficients of Table 7 and the associated mean or quantile values of variables sqft and acres from Table 4.

---

[9]The p values of the OLS estimates are based on an estimate of the variance–covariance matrix that is robust to heteroskedasticity.

Table 4 Basic statistics and quantiles of individual variables, 1,366 observations

	Mean	Min.	Max.	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
sp	146.649	90.000	247.000	115.000	123.000	129.900	135.980	141.350	148.000	158.990	169.900	188.590
5% conf.				113.574	121.902	128.049	134.000	139.900	146.500	156.000	166.805	185.000
Interval				116.961	124.526	131.500	137.675	142.500	150.000	160.000	173.000	191.000
sqft	2.2020	0.792	4.800	1.4709	1.7606	1.9829	2.0385	2.1117	2.1824	2.3931	2.5893	3.0604
acres	0.2478	0.010	2.100	0.2598	0.2228	0.2392	0.2193	0.2669	0.2388	0.2664	0.3096	0.2794
year	1975	1877	2000	1957	1960	1973	1976	1984	1986	1984	1985	1985
bedr	3.76	1	7	3.0455	3.5614	3.5077	3.8714	3.7385	3.7500	3.7705	3.9455	4.4000
bathf	1.63	0	5	1.1364	1.4561	1.4615	1.5000	1.6154	1.7143	1.8689	1.9818	2.1714
batht	0.37	0	3	0.2500	0.2281	0.3385	0.5000	0.3692	0.3393	0.4426	0.4364	0.4286
bathh	0.21	0	3	0.2500	0.1754	0.2615	0.0714	0.2462	0.1429	0.1639	0.2182	0.4857
deck	0.27	0	3	0.1818	0.1404	0.2154	0.2571	0.3385	0.2500	0.4098	0.2545	0.4000
patio	0.47	0	2	0.3636	0.4737	0.5077	0.4429	0.4615	0.5179	0.5902	0.4727	0.6000
garage	1.39	0	5	0.7500	0.7719	1.0615	1.3143	1.6923	1.7679	1.9016	1.7818	1.8286
basmt	0.44	0	1	0.2045	0.3660	0.3708	0.5067	0.4886	0.4609	0.4161	0.4695	0.4957
pool	0.01	0	1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0364	0.0000
airevr	0.39	0	1	0.4091	0.4035	0.3846	0.5714	0.3846	0.4107	0.3770	0.4000	0.2286
airevw	0.11	0	1	0.1136	0.2105	0.1538	0.1000	0.1077	0.0179	0.0492	0.0182	0.0286
airel	0.26	0	1	0.1591	0.1404	0.2000	0.1429	0.2615	0.3571	0.2787	0.2727	0.4571

airgas	0.07	0	1	0.1136	0.0351	0.0769	0.0571	0.0615	0.0536	0.0984	0.1636	0.1143
flhar	0.28	0	1	0.2500	0.3158	0.2154	0.2429	0.1231	0.2321	0.1967	0.2364	0.5714
fltil	0.25	0	1	0.1364	0.2982	0.2154	0.2429	0.1385	0.2679	0.2787	0.2545	0.4571
extu	0.15	0	1	0.0000	0.0702	0.1538	0.0429	0.0462	0.1429	0.2295	0.2909	0.3429
exbri	0.69	0	1	0.4773	0.5965	0.5692	0.7286	0.8462	0.8214	0.7377	0.7455	0.7714
exalu	0.54	0	1	0.5227	0.4386	0.4000	0.5429	0.5692	0.7143	0.6557	0.5273	0.4857
exfra	0.06	0	1	0.1364	0.1053	0.1077	0.1143	0.0615	0.0179	0.0328	0.0727	0.0571
laful	0.71	0	1	0.7045	0.6842	0.6615	0.7571	0.7231	0.6607	0.6721	0.6000	0.7143
lapar	0.10	0	1	0.1591	0.1053	0.1077	0.1429	0.0769	0.1250	0.0656	0.1091	0.1143
lotspr	0.54	0	1	0.2955	0.3860	0.4000	0.5143	0.5231	0.6964	0.6393	0.5273	0.7429
lotmtn	0.64	0	1	0.4318	0.5614	0.6000	0.7143	0.7385	0.6964	0.7705	0.6727	0.6286
di15	1.56	0.01	11.13	1.1520	1.4553	1.8746	1.2814	1.4517	1.8157	1.6734	1.9460	1.7200
dorem	6.53	0.26	22.30	7.0700	6.5818	7.7612	4.6543	5.5723	6.5064	6.7964	6.5222	6.4486
earthqk	1.55	0.12	4.08	1.8268	1.5537	1.5674	1.6931	1.6769	1.4861	1.4915	1.3985	1.5229
nwrate	0.07	0.02	0.23	0.0966	0.0792	0.0721	0.0812	0.0806	0.0638	0.0633	0.0646	0.0646
forrent	0.23	0.00	0.76	0.3238	0.2768	0.2105	0.2420	0.2509	0.1995	0.1984	0.1953	0.1939

The quantile values of all variables other than sp are means of the variable values that are associated with those sp values that fall within a 5% confidence interval around any given quantile point of sp (as noted in the body of the table). In other words, the values of the explanatory variables are approximately tied to the values of the dependent variable for each quantile point; although not point for point to avoid unrepresentative values of the explanatory variables being associated with a particular quantile point of the dependent variable

Table 5 Coefficient estimates, OLS and by quantile

	OLS	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
constant	1.9733	0.5323	1.1311	1.3725	1.6984	1.8964	1.9134	2.2572	2.3973	2.7520
sqft	0.1179	0.0896	0.1025	0.1073	0.1083	0.1228	0.1254	0.1324	0.1346	0.1376
acres	0.1549	0.1362	0.1432	0.1400	0.1408	0.1858	0.1824	0.1917	0.2292	0.3223
year	0.0013	0.0020	0.0017	0.0015	0.0014	0.0013	0.0013	0.0011	0.0011	0.0009
bedr	0.0052	0.0090	0.0071	0.0137	0.0115	0.0088	0.0068	0.0043	0.0063	0.0020
bathf	0.0510	0.0372	0.0473	0.0457	0.0511	0.0479	0.0462	0.0468	0.0506	0.0694
batht	0.0221	0.0161	0.0169	0.0145	0.0210	0.0166	0.0202	0.0203	0.0261	0.0477
bathh	0.0268	−0.0041	0.0220	0.0207	0.0192	0.0215	0.0300	0.0390	0.0469	0.0441
deck	0.0051	0.0036	0.0076	0.0060	0.0045	0.0099	0.0111	0.0031	−0.0011	0.0031
patio	0.0050	0.0073	0.0041	0.0062	0.0042	0.0029	0.0035	0.0077	0.0015	0.0039
garage	0.0268	0.0284	0.0263	0.0268	0.0272	0.0265	0.0275	0.0254	0.0235	0.0246
basmt	0.0002	0.0015	0.0020	−0.0054	−0.0062	−0.0067	0.0023	0.0015	−0.0037	−0.0081
pool	0.0106	0.0542	0.0560	0.0428	0.0451	0.0090	0.0068	−0.0036	−0.0035	0.0086
airevr	−0.0045	0.0016	−0.0065	−0.0079	−0.0092	−0.0103	−0.0066	−0.0081	−0.0149	0.0006
airevw	−0.0060	0.0191	0.0073	0.0085	−0.0006	−0.0139	−0.0125	−0.0199	−0.0192	−0.0030
airel	0.0283	0.0400	0.0205	0.0175	0.0199	0.0189	0.0237	0.0247	0.0200	0.0359
airgas	−0.0023	−0.0247	−0.0117	−0.0013	−0.0025	−0.0046	0.0041	0.0103	0.0073	0.0015
flhar	0.0290	0.0261	0.0249	0.0255	0.0292	0.0296	0.0338	0.0351	0.0341	0.0319
fltil	0.0174	0.0062	0.0111	0.0158	0.0197	0.0191	0.0217	0.0258	0.0200	0.0069
extu	0.0724	0.0691	0.0640	0.0736	0.0744	0.0745	0.0722	0.0711	0.0652	0.0546
exbri	0.0119	0.0209	0.0150	0.0172	0.0157	0.0123	0.0112	0.0127	0.0081	−0.0128
exalu	0.0207	0.0303	0.0295	0.0254	0.0273	0.0264	0.0227	0.0227	0.0156	0.0092
exfra	0.0158	0.0304	0.0144	0.0097	0.0060	0.0048	0.0082	0.0178	0.0025	0.0154
laful	0.0023	0.0286	0.0152	0.0116	0.0108	−0.0016	−0.0071	−0.0064	−0.0090	−0.0250
lapar	−0.0114	0.0066	0.0096	−0.0064	−0.0148	−0.0398	−0.0422	−0.0236	−0.0142	−0.0109
lotspr	0.0238	0.0335	0.0236	0.0230	0.0231	0.0203	0.0175	0.0228	0.0237	0.0243
lotmtn	0.0136	0.0266	0.0214	0.0191	0.0135	0.0123	0.0090	0.0030	−0.0023	−0.0109
di15	0.0046	0.0062	0.0020	0.0030	0.0019	0.0033	0.0074	0.0100	0.0102	0.0086
dorem	−0.0018	−0.0027	−0.0028	−0.0021	−0.0018	−0.0025	−0.0019	−0.0016	−0.0018	−0.0017
earthqk	0.0025	−0.0011	−0.0017	−0.0030	−0.0041	−0.0041	−0.0001	0.0043	0.0072	0.0182
nwrate	−0.2315	−0.2552	−0.1894	−0.1462	−0.1695	−0.1832	−0.2398	−0.2114	−0.2206	−0.2992
forrent	−0.0199	−0.0189	−0.0179	−0.0237	−0.0135	−0.0125	−0.0131	−0.0084	−0.0091	−0.0316
R^2	0.7648	0.5225	0.5307	0.5428	0.5476	0.5539	0.5648	0.5684	0.5688	0.5485

The coefficient of determination (R^2) for the quantile regressions are pseudo R^2, calculated as 1 minus (sum of deviations about the estimated quantile/sum of deviations about the raw quantile)

There is very little difference in the results of Tables 5 and 7, although the spatial lag variable of Table 7 is statistically significant for most but not for all quantiles. In comparing the *p* values of Tables 6 and 8, it appears that those of Table 8 are on average slightly lower, especially for some variables, such as *airel*, *exbri*, *laful*, or *dorem*. The similarity in results between Tables 5 and 7 for the regression coefficients and between Tables 6 and 8 for the *p* values of the coefficient estimates suggests that the quantile effects dominate the spatial autocorrelation effects. Put differently, for the given model and data set, it is more important for the results to account for quantile effects than for spatial autocorrelation effects. Whether this result holds in general awaits further research on other models and data sets..

Tables 5 and 7 both show that the coefficients of a number of variables vary considerably across quantiles. For example, there is more than a 50% difference between the square footage coefficient for the 0.1 quantile and the 0.9 quantile. This is economically significant. The dollar price effects reported for variable *sqft* in

Table 6 *P* values of coefficient estimates, OLS and by quantile

	OLS	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
constant	0.000	0.352	0.009	0.001	0.000	0.000	0.000	0.000	0.000	0.000
sqft	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
acres	0.000	0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
year	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.007
bedr	0.171	0.040	0.080	0.001	0.013	0.043	0.020	0.312	0.193	0.812
bathf	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
batht	0.000	0.051	0.021	0.040	0.000	0.028	0.022	0.047	0.008	0.000
bathh	0.000	0.717	0.041	0.003	0.006	0.006	0.002	0.001	0.000	0.000
deck	0.362	0.669	0.242	0.257	0.547	0.150	0.035	0.592	0.868	0.704
patio	0.310	0.270	0.536	0.371	0.533	0.633	0.486	0.160	0.823	0.560
garage	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
basmt	0.982	0.817	0.750	0.447	0.441	0.289	0.765	0.876	0.740	0.493
pool	0.581	0.068	0.112	0.187	0.088	0.742	0.815	0.860	0.906	0.812
airevr	0.561	0.909	0.555	0.418	0.244	0.227	0.538	0.484	0.138	0.957
airevw	0.555	0.261	0.590	0.412	0.948	0.076	0.274	0.138	0.110	0.817
airel	0.000	0.003	0.082	0.109	0.026	0.004	0.016	0.028	0.059	0.004
airgas	0.841	0.324	0.412	0.916	0.835	0.731	0.830	0.617	0.618	0.911
flhar	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
fltil	0.002	0.406	0.143	0.035	0.001	0.007	0.006	0.002	0.054	0.547
extu	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
exbri	0.052	0.043	0.016	0.001	0.000	0.002	0.139	0.175	0.416	0.326
exalu	0.001	0.001	0.001	0.000	0.000	0.000	0.001	0.010	0.070	0.466
exfra	0.105	0.016	0.333	0.494	0.506	0.569	0.487	0.188	0.854	0.486
laful	0.768	0.061	0.139	0.143	0.303	0.864	0.521	0.470	0.415	0.107
lapar	0.323	0.771	0.509	0.472	0.197	0.001	0.002	0.206	0.459	0.589
lotspr	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.002	0.003	0.035
lotmtn	0.013	0.009	0.002	0.000	0.066	0.047	0.103	0.666	0.800	0.253
di15	0.075	0.215	0.540	0.473	0.662	0.411	0.042	0.001	0.001	0.028
dorem	0.021	0.018	0.002	0.003	0.036	0.002	0.030	0.059	0.099	0.294
earthqk	0.535	0.860	0.690	0.510	0.397	0.411	0.986	0.326	0.267	0.017
nwrate	0.010	0.007	0.037	0.022	0.013	0.004	0.048	0.150	0.138	0.091
forrent	0.177	0.253	0.302	0.181	0.491	0.545	0.313	0.596	0.575	0.078

Probability values are presented for the hypothesis that the estimated coefficient is equal to zero. A *p* value of 0.05 or less means that it is highly unlikely (a 5% chance or less) that the estimated parameter is statistically insignificant

Table 9 attest to that: the marginal price of a square foot for quantile point 0.9 is close to 150% above that of quantile point 0.1; yet, the sale price for quantile point 0.9 (Table 4) is only 64% above that of quantile point 0.1. Table 11 shows a similar effect for the price elasticity of square footage: the price elasticity for the 0.9 quantile of housing prices is three times as high as that for the 0.1 quantile. The 2SLS estimate of variable *sqft* clearly overstates the contribution of a square foot to the sale price of lower-price houses but understates the contribution for higher-priced houses. The results are very similar, although more dramatic, for the variable *acres*.

The variable *year* is a proxy for age.[10] A 1-year increase reduces the age of the house by 1 year. The positive sign reported in Tables 5 and 7 suggests that newer

---

[10]The variable *year* can be converted to measure the age of a house by simply subtracting the value of *year* from 2000 for a given observation. This linear transformation does not affect the coefficients of any variable other than *year* or *age* and the constant.

Table 7 Coefficient estimates of spatial lag model, 2SLS and by quantile

	2SLS	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
constant	1.9014	0.4768	1.1333	1.3286	1.6242	1.8207	1.8231	2.2428	2.4257	2.7359
lagh	0.0137	0.0077	0.0158	0.0159	0.0127	0.0107	0.0096	0.0062	0.0097	0.0148
sqft	0.1173	0.0905	0.1025	0.1052	0.1078	0.1217	0.1256	0.1318	0.1328	0.1370
acres	0.1533	0.1363	0.1420	0.1340	0.1420	0.1854	0.1830	0.1968	0.2301	0.3190
year	0.0013	0.0020	0.0016	0.0015	0.0014	0.0013	0.0013	0.0011	0.0010	0.0009
bedr	0.0051	0.0087	0.0075	0.0132	0.0109	0.0086	0.0065	0.0046	0.0058	0.0017
bathf	0.0513	0.0382	0.0464	0.0486	0.0512	0.0488	0.0471	0.0476	0.0534	0.0696
batht	0.0230	0.0177	0.0181	0.0171	0.0221	0.0177	0.0203	0.0215	0.0281	0.0482
bathh	0.0275	−0.0031	0.0215	0.0232	0.0198	0.0204	0.0304	0.0401	0.0488	0.0432
deck	0.0053	0.0049	0.0073	0.0063	0.0042	0.0107	0.0109	0.0031	−0.0015	0.0048
patio	0.0051	0.0047	0.0047	0.0061	0.0050	0.0037	0.0028	0.0069	−0.0008	0.0041
garage	0.0265	0.0279	0.0261	0.0263	0.0264	0.0261	0.0266	0.0259	0.0248	0.0253
basmt	−0.0001	0.0007	0.0044	−0.0057	−0.0056	−0.0059	0.0011	−0.0001	−0.0017	−0.0082
pool	0.0109	0.0567	0.0583	0.0391	0.0482	0.0115	0.0065	−0.0037	−0.0041	0.0040
airevr	−0.0047	−0.0004	−0.0088	−0.0107	−0.0097	−0.0090	−0.0062	−0.0100	−0.0157	0.0032
airevw	−0.0065	0.0185	0.0041	0.0043	−0.0015	−0.0142	−0.0125	−0.0221	−0.0204	−0.0032
airel	0.0274	0.0369	0.0185	0.0141	0.0182	0.0192	0.0235	0.0229	0.0183	0.0359
airgas	−0.0034	−0.0274	−0.0130	−0.0058	−0.0033	−0.0032	0.0020	0.0081	0.0072	0.0004
flhar	0.0287	0.0284	0.0241	0.0282	0.0287	0.0293	0.0347	0.0342	0.0338	0.0300
fltil	0.0174	0.0049	0.0112	0.0145	0.0195	0.0174	0.0205	0.0263	0.0205	0.0085
extu	0.0722	0.0667	0.0707	0.0766	0.0744	0.0717	0.0725	0.0704	0.0634	0.0534
exbri	0.0121	0.0195	0.0152	0.0159	0.0159	0.0139	0.0132	0.0121	0.0084	−0.0110
exalu	0.0206	0.0268	0.0310	0.0268	0.0272	0.0259	0.0234	0.0233	0.0156	0.0085
exfra	0.0152	0.0279	0.0145	0.0108	0.0042	0.0060	0.0091	0.0184	0.0050	0.0132
laful	0.0027	0.0262	0.0160	0.0162	0.0086	−0.0011	−0.0054	−0.0051	−0.0109	−0.0260
lapar	−0.0110	0.0046	0.0036	−0.0037	−0.0156	−0.0395	−0.0382	−0.0195	−0.0150	−0.0082
lotspr	0.0235	0.0344	0.0253	0.0229	0.0234	0.0186	0.0174	0.0227	0.0249	0.0245
lotmtn	0.0140	0.0267	0.0219	0.0191	0.0145	0.0120	0.0103	0.0042	−0.0035	−0.0127
di15	0.0043	0.0057	0.0016	0.0026	0.0008	0.0039	0.0061	0.0102	0.0100	0.0071
dorem	−0.0017	−0.0026	−0.0023	−0.0018	−0.0019	−0.0023	−0.0020	−0.0016	−0.0020	−0.0017
earthqk	0.0027	−0.0022	−0.0002	−0.0030	−0.0038	−0.0038	−0.0018	0.0047	0.0073	0.0191
nwrate	−0.2198	−0.2269	−0.1529	−0.1385	−0.1637	−0.1627	−0.2331	−0.1800	−0.2180	−0.3411
forrent	−0.0188	−0.0195	−0.0202	−0.0245	−0.0100	−0.0131	−0.0118	−0.0090	−0.0085	−0.0294

2SLS stands for Two-Stage Least Squares. The quantile estimates are based on two-stage quantile regressions as discussed by Kim and Muller (2004). The variable *lagh* identifies the spatial lag

houses sell for relatively more. This is a standard result. However, the coefficients of Table 7 and the corresponding marginal effects of Table 9 reveal that there is a lower premium for newness for higher-priced homes. Lower-priced homes have the highest premium for newness (or discount for age).

The 2SLS coefficient for the number of bedrooms, *bedr*, is not significant in Table 7, which is not surprising given what is reported in Table 2. However, the quantile regressions provide a somewhat different picture. The regression coefficients for *bedr* are statistically significant primarily in the lower and middle price ranges and are not significant in the upper price range. The underlying economic reason for this result may be tied to the fact that lower- and medium-priced houses tend to have fewer bedrooms than expensive houses, yet will often contain as many or more occupants. As a result, an additional bedroom will have a higher marginal value in the lower-priced ranges.

Table 8 *P* Values of coefficients of spatial lag model, 2SLS and by quantile

	2SLS	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
constant	0.000	0.177	0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000
lagh	0.010	0.008	0.003	0.005	0.064	0.021	0.037	0.306	0.006	0.000
sqft	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
acres	0.000	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
year	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
bedr	0.142	0.018	0.051	0.000	0.000	0.003	0.016	0.197	0.070	0.669
bathf	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
batht	0.000	0.001	0.004	0.000	0.000	0.001	0.000	0.001	0.000	0.000
bathh	0.000	0.671	0.002	0.000	0.000	0.004	0.000	0.000	0.000	0.000
deck	0.334	0.395	0.142	0.134	0.391	0.035	0.021	0.534	0.753	0.440
patio	0.301	0.354	0.286	0.125	0.255	0.349	0.431	0.158	0.884	0.481
garage	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
basmt	0.983	0.922	0.520	0.376	0.299	0.203	0.825	0.993	0.816	0.330
pool	0.735	0.007	0.017	0.073	0.083	0.574	0.695	0.831	0.863	0.892
airevr	0.512	0.954	0.139	0.044	0.110	0.178	0.351	0.178	0.027	0.796
airevw	0.490	0.145	0.665	0.555	0.814	0.070	0.111	0.034	0.026	0.813
airel	0.000	0.000	0.002	0.020	0.008	0.005	0.000	0.001	0.008	0.002
airgas	0.756	0.221	0.332	0.555	0.722	0.741	0.834	0.467	0.427	0.973
flhar	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
fltil	0.002	0.423	0.038	0.001	0.000	0.000	0.000	0.000	0.000	0.313
extu	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
exbri	0.035	0.010	0.001	0.000	0.000	0.001	0.005	0.043	0.087	0.140
exalu	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.019	0.235
exfra	0.137	0.004	0.087	0.116	0.604	0.439	0.201	0.092	0.564	0.467
laful	0.723	0.014	0.034	0.022	0.227	0.890	0.432	0.476	0.148	0.011
lapar	0.266	0.740	0.667	0.608	0.065	0.000	0.000	0.106	0.162	0.585
lotspr	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
lotmtn	0.008	0.000	0.000	0.000	0.003	0.007	0.012	0.396	0.518	0.050
di15	0.032	0.020	0.333	0.185	0.759	0.127	0.027	0.002	0.000	0.074
dorem	0.013	0.000	0.001	0.002	0.001	0.001	0.002	0.032	0.003	0.072
earthqk	0.506	0.628	0.948	0.361	0.329	0.244	0.620	0.182	0.065	0.000
nwrate	0.013	0.036	0.080	0.027	0.024	0.011	0.002	0.088	0.005	0.011
forrent	0.242	0.112	0.149	0.043	0.507	0.371	0.362	0.521	0.526	0.025

The *p* values of the quantile regressions are bootstrapped from the two-stage quantile estimator of Kim and Muller (2004). Five hundred replications are employed. The variable *lagh* identifies the spatial lag

The bathroom variables show a similar result: additional bathrooms have a much higher value-added impact in higher-priced homes than in lower-priced ones.

Estimating quantile regressions as shown in Table 7 gives an opportunity to measure the relationship of selling price to a large number of variables. As shown above, defining the relationship between the typical hedonic pricing variables (square footage, lot size, age, bedrooms, bathrooms) and selling price is improved by using quantile regression. The Table 7 results show this is true for a number of variables in the model, i.e., that the relationship changes over different price ranges. For some variables the quantile regression results confirm that their relationship with selling price remains relatively stable across different price ranges. For other variables that are not statistically significant in the 2SLS estimation, the quantile regression results confirm them to be not significant over different price ranges. Table 12 provides a summary of the relationships between the explanatory variables and selling price as defined by the quantile regressions.

Table 9 Price effect of unit increase in characteristic, 2SLS and by quantile

	2SLS	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
sqft	17,206	10,410	12,608	13,672	14,661	17,205	18,594	20,953	22,558	25,836
acres	22,478	15,672	17,471	17,407	19,306	26,208	27,081	31,290	39,090	60,163
year	185	227	200	199	190	185	195	177	174	162
bedr	747	996	917	1,715	1,483	1,220	967	724	985	322
bathf	7,526	4,396	5,703	6,311	6,966	6,892	6,967	7,568	9,071	13,133
batht	3,370	2,030	2,226	2,218	3,012	2,508	3,010	3,411	4,781	9,093
bathh	4,030	−355	2,650	3,010	2,694	2,880	4,500	6,368	8,283	8,147
deck	775	558	901	819	569	1,508	1,615	498	−247	904
patio	743	546	583	790	682	522	416	1,097	−143	775
garage	3,883	3,209	3,211	3,420	3,589	3,692	3,936	4,122	4,207	4,771
basmt	−22	79	546	−737	−759	−840	158	−9	−294	−1,541
di15	630	657	202	342	110	550	908	1,626	1,703	1,344
dorem	−243	−300	−279	−230	−255	−332	−295	−262	−334	−327
earthqk	390	−247	−28	−389	−520	−544	−261	740	1,247	3,603
nwrate	−32,232	−26,089	−18,809	−17,993	−22,254	−23,004	−34,494	−28,615	−37,040	−64,320
forrent	−2,756	−2,247	−2,490	−3,184	−1,359	−1,845	−1,747	−1,429	−1,440	−5,549

The marginal effects are expressed in dollar values by multiplying the estimated coefficients of Table 7 by 1,000 times the corresponding value of variable *sp*, as reported in Table 4 for the mean and the quantiles, respectively

Conclusions

One of the most popular areas of research in real estate economics and finance has been the pricing of residential real estate. Empirical research has primarily focused on identifying house characteristics that most influence selling price. The results from this body of literature have often been in conflict regarding the impact of a variable on selling price. This study seeks to clarify some of the confusion by using

Table 10 Price effect of characteristic, 2SLS and by quantile

	2SLS	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
pool	1,599	6,520	7,172	5,079	6,550	1,628	968	−592	−695	761
airevr	−688	−50	−1,080	−1,391	−1,317	−1,269	−919	−1,594	−2,667	597
airevw	−957	2,125	506	555	−197	−2,009	−1,846	−3,512	−3,458	−609
airel	4,015	4,239	2,276	1,830	2,474	2,712	3,477	3,633	3,110	6,779
airgas	−496	−3,155	−1,595	−757	−444	−454	298	1,291	1,219	71
flhar	4,208	3,265	2,962	3,668	3,899	4,144	5,129	5,436	5,738	5,654
fltil	2,556	566	1,372	1,889	2,653	2,458	3,029	4,187	3,482	1,596
extu	10,584	7,665	8,692	9,944	10,122	10,137	10,736	11,189	10,769	10,064
exbri	1,780	2,247	1,875	2,064	2,166	1,960	1,951	1,920	1,421	−2,077
exalu	3,025	3,086	3,811	3,484	3,705	3,665	3,469	3,711	2,655	1,610
exfra	2,230	3,203	1,784	1,405	565	854	1,342	2,920	854	2,487
laful	397	3,013	1,964	2,100	1,173	−151	−792	−805	−1,853	−4,897
lapar	−1,607	527	439	−486	−2,116	−5,589	−5,655	−3,099	−2,541	−1,547
lotspr	3,451	3,954	3,115	2,974	3,182	2,632	2,579	3,609	4,228	4,615
lotmtn	2,047	3,067	2,688	2,481	1,970	1,691	1,519	662	−592	−2,397

The percentage change effects are expressed in dollar values by multiplying the estimated coefficients of Table 7 by 1,000 times the corresponding value of variable *sp*, as reported in Table 4 for the mean and the quantiles, respectively

Table 11 Price elasticities of square footage and acres, 2SLS and by quantile

	2SLS	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
sqft	0.258	0.133	0.180	0.209	0.220	0.257	0.274	0.315	0.344	0.419
acres	0.038	0.035	0.032	0.032	0.031	0.049	0.044	0.052	0.071	0.089

The elasticities are calculated as the product of the coefficient estimates of Table 7 and the associated values of variables *sqft* and *acres* from Table 4

quantile regression to measure the effect of various housing characteristics on house prices.

Results of this study show that the effect of housing characteristics on selling price can be better explained by estimating quantile regressions across price categories. For example, previous studies that have examined the effect of characteristics such as square footage or age on selling price have found mixed results in terms of both the level and the direction of change. This study shows that some of those differences may be explained by differences in house prices. In

Table 12 The relationship between explanatory variables and selling price as shown by the quantile regressions

Regression coefficient increases as selling price increases	Regression coefficient decreases as selling price increases	Regression coefficient remains relatively constant as selling price increases	Regression coefficient shows no definite pattern as selling price increases	Regression[a] coefficient is not significant as selling price increases
Square feet	Year built	Garage	Bedrooms	Deck
Acres	Mountain view lot	Electric AC	Percent of population nonwhite	Patio
Full baths		Stucco exterior		Basement
Three-quarter baths		Brick exterior		Pool
Half baths		Aluminum exterior		Evaporator AC
Hardwood floors		Sprinkler system		Window evaporator AC
Tile floors		Distance to interstate		Gas AC
		Distance to city center		Frame exterior
				Full landscaping
				Partial landscaping
				Earthquake magnitude
				Percent rental houses

[a] All variables in this column are also not significant in the 2SLS model. Frame exterior and earthquake magnitude are significant in one quantile regressions and partial landscaping is significant in two quantile regressions

particular, the regression coefficients of some variables behave differently across different house price levels, or quantiles. Buyers of higher-priced homes appear to price certain housing characteristics differently from buyers of lower-priced homes.

For the given data set, it is shown that the quantile effects dominate any effects on coefficient size and statistical significance that arise from spatial autocorrelation. In fact, taking explicit account of spatial autocorrelation in the quantile regressions, adds very little information. Whether this is a general result or particular to the data set that is being used in this study is an open question that awaits further research.

This study produces some interesting results. For example, square footage is often used to determine the appraised value of a home since it is expected to have a significant effect on the selling price. While previous studies bear this out, it is interesting to see how buyers in different price ranges value this variable. This is shown by the significant difference between the coefficients at the lowest and the highest quantiles where the additional price of a square foot for the highest priced homes is two and a half times the additional price per square foot for the lowest-priced homes. Clearly, traditional methodologies such as OLS or models that take into account spatial autocorrelation can overstate the value of a marginal square foot for lower-priced homes but understate the effect on higher-priced homes.

The quantile results provide some valuable insights to the different relationships that the explanatory variables have with selling price. For example, some variables such as square footage, lot size, bathrooms, and floor type have a greater impact as selling price increases. Other variables have a relatively constant effect on selling price across different price ranges. These include garage, exterior siding, sprinkler system, and distance to city center. Some other variables such as bedrooms and percentage of nonwhite population have a significant effect on selling price but there is no clear pattern of the effect across different price ranges. Lastly, the quantile regressions confirm that most variables showing no statistical significance under OLS or 2SLS remain not significant across the different price ranges.

These results add to the body of research explaining house prices. Even though variations in the value of housing characteristics across different price ranges may have been considered intuitive beforehand, quantile regression provides a way to confirm these expectations.

References

Anselin, L. (2001). Spatial econometrics. In B. H. Baltagi (Ed.), *A companion to theoretical econometrics* (pp. 310–330). Malden, MA and Oxford: Blackwell.

Bartik, T. J. (1987). The estimation of demand parameters in hedonic price models. *Journal of Political Economy, 95*, 81–88.

Durlauf, S. N., & Young, H. P. (Eds.) (2001). *Social Dynamics*. Cambridge, MA: MIT.

Epple, D. (1987). Hedonic prices and implicit markets: estimating demand and supply functions for differentiated products. *Journal of Political Economy, 95*, 59–80.

Gould, W. W. (1992). Quantile regression with bootstrapped standard errors. *Stata Technical Bulletin, 9*, 19–21.

Gould, W. W. (1997). Interquantile and simultaneous-quantile regression. *Stata Technical Bulletin, 38*, 14–22.

Greene, W. H. (2000). *Econometric Analysis*, 4th ed., Prentice Hall, Saddle River, N.J.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica, 47*, 153–161.

Kim, T.-H., & Muller, C. (2004). Two-stage quantile regression when the first stage is based on quantile regression. *Econometrics Journal, 7*, 218–231.

Kirman, A. P. (1992). Whom or what does the representative individual represent? *Journal of Economic Perspectives, 6*, 117–136.

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica, 46*, 33–50.

Koenker, R., & Hallock, K. F. (2001). Quantile regression. *Journal of Economic Perspectives, 15*, 143–156.

Malpezzi, S. (2003). Hedonic pricing models: A selective and applied review. In T. O. Sullivan & K. Gibbs (Eds.), *Housing economics and public policy: Essays in honor of Duncan Maclennan*. Oxford, UK: Blackwell.

Malpezzi, S., Ozanne, L., & Thibodeau, T. (1980). *Characteristic prices of housing in fifty-nine metropolitan areas, Research Report*. Washington, DC: The Urban Institute, December.

Newsome, B., & Zietz, J. (1992). Adjusting comparable sales using MRA—The need for segmentation. *Appraisal Journal, 60*, 129–135.

Rogers, W. H. (1993). Calculation of quantile regression standard errors. *Stata Technical Bulletin, 13*, 18–19.

Rosen, S. M. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy, 82*, 34–55.

Sirmans, G. S., Macpherson, D. A., & Zietz, E. N. (2005). The composition of hedonic pricing models. *Journal of Real Estate Literature, 13* (1),3–46.

Zietz, J., & Newsome, B. (2002). Agency representation and the sale price of houses. *Journal of Real Estate Research, 24*, 165–191.