2008 V36 4: pp. 717–751
REAL ESTATE
ECONOMICS

Assessing the Market Value of Real Estate Property with a Geographically Weighted Stochastic Frontier Model

Stephen A. Samaha* and Wagner A. Kamakura**

In this study we consider the problem of sellers, buyers and real estate appraisers in determining the price for a house, taking into account the characteristics of the house and its location as well as the goals of these three different parties. The appraiser's job is to determine the fair market value of the house, while the buyer and seller want to find, respectively, the lowest and highest feasible price for it. We combine recent developments in geography and econometrics to develop an approach that determines local estimates of property values from the perspectives of the buyer, seller and appraiser, taking into account the characteristics of the house as well as its location. We illustrate our approach analyzing closing prices in one residential real estate market.

In this study we propose to take the perspectives of the seller and the buyer in uncovering the lowest price that the seller should accept or the highest price the buyer should pay for the real estate property, while simultaneously factoring in the critical aspects of market imperfections and location. Because real estate transaction prices are often settled in a less-than-perfect negotiation process, and, consequently, the estimates of property values based on past prices should leave room for negotiation, we develop a geographically weighted stochastic frontier model that takes the seller's perspective of determining the potential buyer's reservation value (*i.e.*, the highest possible price she should expect for the house) as well as the buyer's perspective of determining the seller's reservation value (*i.e.*, the lowest feasible price to bid on a house). We then apply and illustrate our model using data from one real estate market.

It is difficult to overstate the central role that the hedonic price regression framework has had in valuing residential and commercial real estate properties. This framework (Griliches 1961, Rosen 1974, Epple 1987) assumes that the prices sold for each property represent the market clearing prices, and therefore the results of the regression should provide an unbiased estimate for the fair market value of each house. However, there is an increasing body

*University of Washington, Seattle, WA 98195 or ssamaha@u.washington.edu.
**Duke University, Durham, NC 27708 or kamakura@duke.edu.

of research that questions whether these core assumptions of perfect market equilibrium truly hold in an empirical setting. It is now known that real estate properties can suffer from numerous market imperfections. These imperfections or inefficiencies can result from many sources, including transaction noise resulting from the imperfect matching of buyers and sellers, idiosyncratic bargaining outcomes and unique physical asset characteristics that imply an incomplete market in comparables (Childs, Ott and Riddiough 2002). In addition, imperfections can also occur because of liquidity, private demands of investors and bilateral bargaining outcomes (Quan and Quigley 1991). Harding, Knight and Sirmans (2003) also find that household wealth, gender and other demographic traits influence bargaining power and therefore negotiated prices.

Research has also shown that sellers are heterogeneous in their motivations to sell, and these motivations can also affect sales prices. Glower, Haurin and Hendershott (1998) find that home sellers who are motivated to sell more quickly will set a lower list price, have a lower reservation price and accept earlier, lower offers. Quan and Quigley (1991) present a model of heterogeneous buyers and sellers and conclude that the greater a seller's urgency is to sell, the lower the seller's reservation price. Miceli (1989) also notes that sellers may not want to simply maximize sale price but may also consider "minimizing the amount of time it takes to obtain a pre-agreed price." Geltner, Kluger and Miller (1991) discuss how variations in seller holding costs also affect the seller's reservation price and find that the greater the time cost of holding a property, the lower the seller's reservation price.

Because of these inefficiencies, the analyst has highly imperfect information about the seller's and buyer's reservation values; the analyst only knows that the observed transaction price is equal to or greater than the seller's reservation value and equal to or lower than the buyer's reservation value. Because the factors affecting the final transaction price are rarely observed, it is imperative to not only model them as errors in the hedonic setting, but also to account for these unobservable imperfections in estimating the buyer's and seller's reservation values. Sales efficiencies could be improved if buyers and sellers had not only a best guess of a property's true value, but also a sense for how idiosyncrasies in the real estate transaction process can affect the negotiation process and subsequent price range that the home could conceivably be sold for. This range can be modeled as stochastic due to the stochastic nature of these errors, and, indeed, Glower *et al.* (1998) hypothesize that variations in list prices of similar properties are attributed to random seller errors and explicitly introduce stochastic errors into list price as the sum of the optimal price and the seller's error in setting the list price.

However, it is not sufficient to simply incorporate these phenomena into the hedonic framework without also factoring in the critical aspects of location and the role that geographic location plays in valuing properties. It is well known that there are many unobservable characteristics that can affect market value: proximity to amenities (waterfront, mountain view) and disamenities (railroad, highways, power lines, telecom towers, traffic conditions and adjacent negative land uses) are but a few examples. Some of these unobservable factors are directly related to geographic location, while others, such as curb appeal, add errors to the model and also uncertainty in the valuation of the property by buyers and sellers. It is critical to be able to explicitly model these unobservable factors as well to allow for model realism.

The Efficiency of Real Estate Markets

Because hedonic price regressions assume that transaction prices for each property represent the market clearing prices, the results of the regression should provide an unbiased estimate for the fair market value of each house, under the perfect-market assumption. However, as stated earlier, there are certain inefficiencies present. For example, the set of houses available for purchase at any given time is limited, and it does not cover the entire characteristics space, giving each house some monopoly over those buyers whose equilibria lie in its vicinity in the characteristics space (Rosen 1974). Second, despite technological advances in the multiple listing service, search costs are still high for buyers, forcing them to make trade-offs between their search costs and the marginal benefit of efficient buying decisions (Ratchford 1980). Third, participants in real estate markets often have incomplete information regarding the home's attributes, and decisions to buy and sell must often be made on partial knowledge. Trades are also decentralized and market prices are the outcomes of pair-wise negotiations (Quan and Quigley 1991). These inefficiencies can consequently lead buyers and sellers to respectively pay more for and accept less than their respective optimal prices. Fourth, market inefficiencies may also result because of potential conflicts of interests between brokers and sellers. Although sellers hope that properties will sell at the highest possible prices, brokers have an incentive to sell the house quickly. In general, the lower the asking price, the lower the effort required on the part of the broker to sell the property and the easier it is for the property to sell (Geltner *et al.* 1991). Brokers are also heterogeneous with respect to their ability levels, which may enter into the determination of both selling price and time on the market (Miceli 1989, Miller 1978). Last, bargaining power can also affect the outcomes of real estate transactions. Previous research has found systematic evidence in the housing markets that weak buyers pay higher prices and weak sellers receive lower prices for their homes (Harding *et al.* 2003, Harding, Rosenthal and Sirmans 2003, Colwell and Munneke 2006).

Given these uncertainties and inherent market inefficiencies, the observed transaction price does not necessarily represent the equilibrium between the seller's and buyer's valuations. In an imperfect market such as real estate, the transaction price is only known to be higher (lower) than or equal to the seller's (buyer's) reservation value. Therefore, individual buyers and sellers would be better informed with valuations at the Pareto frontiers rather than at the least-squares line as defined by traditional hedonic price regression models. In other words, in order to obtain a better sense of potential buyers' valuations, the seller wants to estimate the buyer's hedonic price at the frontier, taking into account that the buyer's reservation value is at least as high as or higher than the observed transaction price. Similarly, when deciding on an offer of whether or not to buy a house, the buyer wants to estimate the seller's hedonic price, taking into account that sellers' reservation values in the past were at least as low as or lower than the observed transaction prices.

Stochastic frontier estimation originated under the contexts of production economics with the work done by Meeusen and van den Broeck (1977), Aigner, Lovell and Schmidt (1977) and Battese and Corra (1977). Conventional econometric production models treated producers as successful optimizers and utilized methods such as those from Cobb and Douglas (1928) to estimate production, cost and profit functions. Under this classic framework, deviations from the maximum possible output or profit or minimum cost, given a set of inputs, were attributed exclusively to random statistical noise. Stochastic frontier estimation was developed to incorporate a theory of producer behavior that explicitly incorporated the possibility of suboptimal performance in addition to random statistical noise. The efficient frontier under this context (from the standpoint of a producer, for example) would represent the maximum output that is possible for a given set of inputs and technology. Producers who are operating on this frontier are efficient, while those operating beneath this frontier are technically inefficient. In our application to hedonic pricing, the stochastic frontier represents the maximum price that the home could be sold for (*i.e.*, the buyer's reservation value) given its measurable characteristics. Similarly, the inverse stochastic frontier represents the minimum price that the home could conceivably be sold for (the seller's reservation value). For more details regarding stochastic frontier estimation, see Kumbhakar and Lovell (2000).

Accounting for Geographic Heterogeneity in Hedonic Price Analysis

Because properties near one another tend to be influenced by similar factors (size, age, architectural type and design, proximity to schools and city centers, *etc*.), the selling prices of these houses may be correlated over geography. There have been numerous attempts to incorporate spatial characteristics and correlation into hedonic models. Can (1992) and Can and Megbolugbe (1997) have allowed the marginal prices of property attributes to vary according to neighborhood characteristics. Similarly, Day (2003), Day, Bateman and Lake (2003) and Goodman and Thibodeau (2003) sought to separate hedonic price functions into sets of socioeconomically homogeneous neighborhoods. The idea of using submarkets (areas in which the prices or the quantities of attributes differ from those in different areas) to address spatial autocorrelation in the data is not new. Wilhelmsson (2004) used cluster analysis to define functional submarkets as dummy variables to address the problem of spatial autocorrelation, while Dale-Johnson (1982) used factor analysis to define submarkets for hedonic pricing. Goodman and Thibodeau (1998) used hierarchical methods to define submarkets, while Bourassa *et al.* (1999) used a combination of principal components with cluster analysis.

Some researchers have also incorporated polynomial expressions of latitude and longitude coordinates into their models (Dubin 1992, Pace and Gilley 1997, Pavlov 2000, Clapp 2001) or the interactions of the latitude and longitude coordinates (Fik, Ling and Mulligan 2003) to model spatial error dependence. Still other approaches have included location value signature models (Fik *et al.* 2003), local polynomial regressions (LPR) (Host 1999), LPR with Bayesian smoothing (Clapp, Kim and Gelfand 2002) and smooth-spatial effects estimators (Gibbons and Machin 2001, 2003). Others have approached the problem by implanting simultaneous autoregression and simultaneous temporal autoregression (Pace 1997, Pace and Barry 1997, Pace *et al.* 1998), as well as conditional autoregression, Kriging (Dubin 1992, Basu and Thibodeau 1998, Dubin, Pace and Thibodeau 1999) and spatial error dependence models (Kelejian and Prucha 1999, Bell and Bockstael 2000). Last, geographically weighted regression (GWR) (Fotheringham, Brunsdon and Charlton 2002) and space-varying regression coefficients (Pavlov 2000) have been used to analyze the spatially varying relationships of housing characteristics and prices across geography.

In order to account for the inefficiencies in real estate markets and the fact that real estate values are location dependent, we next propose a new model that combines the concept of geographically weighted estimation with stochastic frontier estimation. This new model allows us to assess the potential market value of homes at the Pareto frontiers, both from the perspective of the seller trying to find the highest (buyer reservation) value for the home and for the buyer seeking to find the lowest possible (seller reservation) value to pay. The differences between the unobservable buyer and seller reservation values define the range or latitude for negotiation between the two parties.

Assessing Property Values at the Pareto Frontier

As mentioned earlier, our goal is to extend the traditional hedonic price regression framework for real estate assessment in two ways. First, we acknowledge that real estate markets are imperfect and that the observed transaction prices are higher (lower) than the seller's (buyer's) unobserved reservation values. Therefore, we attempt to obtain better estimates of the seller's and buyer's reservation values by fitting the hedonic price regression at the Pareto frontiers, both from the perspective of the buyer attempting to find the lowest possible price to pay or the seller seeking the highest price that the market can bear. Second, we take into account that the physical features of a house may be valued differently depending on its location; buyers looking for an urban residence may attach different value for the features of a house than those interested in suburban or rural living.

A Pareto price frontier (Kumbhakar and Lovell 2000) can be written as

$$y_i = f(x_i; \beta) \cdot TE_i, \tag{1}$$

where y_i represents the transaction price for home i ($i = 1 \ldots n$), x_i is the vector of physical characteristics for that home, β is the vector of parameters to be estimated, $f(x_i; \beta)$ is the Pareto (or maximum price) frontier and TE_i defines the technical efficiency (the ratio of the observed price that the home sold for to the maximum feasible price). Thus, the observed selling price of a home, y_i, reaches its maximum feasible value of $f(x_i; \beta)$ only when $TE_i = 1$. Otherwise, per Kumbhakar and Lovell (2000), $TE_i < 1$ provides a measure of the shortfall of observed price from the maximum feasible price.

In order to augment the above deterministic frontier to incorporate house-specific random shocks, the above specification can be rewritten as

$$y_i = f(x_i; \beta) \cdot \exp\{v_i\} \cdot TE_i, \tag{2}$$

where $f(x_i; \beta) \cdot \exp\{v_i\}$ is the stochastic frontier in which $f(x_i; \beta)$ represents the deterministic portion common to all homes being sold and $\exp\{v_i\}$ represents the random shocks on each particular home. Assuming that $f(x_i; \beta)$ takes the log-linear Cobb-Douglas form, and letting $TE_i = \mathrm{Exp}\{-u_i\}$, the stochastic frontier can be rewritten (Aigner *et al.* 1977) as

$$\ln y_i = \beta_1 + \sum_k \beta_k \ln x_{ik} - |u_i| + v_i, \tag{3}$$

where $u \sim N(0, \sigma_u^2)$, that is, distributed as nonnegative half normal and $v \sim N(0, \sigma_v^2)$, where u_i represents the inefficiency term and v_i represents the idiosyncratic effects specific to each house. Then, letting $\varepsilon_i = v_i - |u_i|$, $\delta = \sigma_u/\sigma_v$, $\sigma = (\sigma_u^2 + \sigma_v^2)^{1/2}$, and assuming the half-normal model, the log-density (Aigner *et al.* 1977) can be written as

$$\ln h_i = -\ln \sigma - \frac{1}{2}\ln\frac{2}{\pi} - \frac{1}{2}\left(\frac{y_i - x_i'\beta}{\sigma}\right)^2 + \ln \Phi\left(\frac{-\delta(y_i - x_i'\beta)}{\sigma}\right), \tag{4}$$

where $\Phi(z)$ is the cumulative distribution function of the standard normal distribution. Then, a local maximum likelihood function (Fan, Farmen and Gijbels 1998) for a focal location j can be expressed as

$$\ln L_j = \sum_{i \neq j} W_{ij} \cdot \ln h_i, \tag{5}$$

and the local estimator is

$$\hat{\theta}_j = \arg\max_{\theta} \sum_{i \neq j} W_{ij} \cdot \ln h_i, \tag{6}$$

where

- $\hat{\theta}_j$ represents the vector of local parameters to be estimated by local maximum likelihood; that is, $\theta = \{\beta_1 \ldots \beta_k, \delta, \sigma\}$,
- $W_{ij} = \mathrm{Exp}[-\lambda_1^2(z_{j1} - z_{i1})^2 - \lambda_2^2(z_{j2} - z_{i2})^2]$ are the geographic weights,
- z_{i1} = longitude coordinate for location i,
- z_{i2} = latitude coordinate for location i and
- λ_1, λ_2 are the longitudinal and latitudinal weighting parameters to be estimated.

Equation (5) shows the log-likelihood function of a half-normal stochastic frontier regression (Kumbhakar and Lovell 2000), except that it is estimated at a given location j. The geographic weights W_{ij} define how relevant the data from a neighbor i is, depending on its distance from the focal location j along the East–West (longitude) and North–South (latitude) directions. We allow the geographic weights to decay differently in the two directions, leading to an anisotropic geographically weighted model, which depends on the parameters λ_1, λ_2 that are to be estimated.

Estimation of the two geographic decay coefficients λ_1 and λ_2 can be done by cross-validation (see Fotheringham *et al.* 2002 for a discussion of calibrating spatial weighting functions with cross-validation). This calibration of the decay coefficients can be attained by solving the following nonlinear optimization problem:

$$\hat{\theta} = \arg\max_{\theta} \sum_{j=1}^{N} \sum_{i \neq j} e^{-\lambda_1^2 (z_{j1} - z_{i1})^2 - \lambda_2^2 (z_{j2} - z_{i2})^2}$$
$$\times \left[-\ln \sigma_j - \frac{1}{2} \ln \frac{2}{\pi} - \frac{1}{2} \left(\frac{y_i - x_i' \beta_j}{\sigma_j} \right)^2 + \ln \Phi \left(\frac{-\delta_j (y_i - x_i' \beta_j)}{\sigma_j} \right) \right], \quad (7)$$

where $\theta = \{\lambda_1, \lambda_2, (\beta_{j1} \ldots \beta_{jk}, \delta_j, \sigma_j); j = 1, \ldots, N\}$.

Once the geographic decay coefficients λ_1 and λ_2 are determined for the calibration sample of sold houses, the coefficients for the local stochastic frontier for the buyer's reservation value at any new location A can be obtained by maximizing the log-likelihood below, which depends only on the data from calibration sample

$$\hat{\theta}_A^{Buyer\_RV} = \arg\max_{\theta} \sum_{i \neq A} e^{-\lambda_1^2 (z_{A1} - z_{i1})^2 - \lambda_2^2 (z_{A2} - z_{i2})^2}$$
$$\times \left[-\ln \sigma_A - \frac{1}{2} \ln \frac{2}{\pi} - \frac{1}{2} \left(\frac{y_i - x_i' \beta_A}{\sigma_A} \right)^2 \right.$$
$$\left. + \ln \Phi \left(\frac{-\delta_A (y_i - x_i' \beta_A)}{\sigma_A} \right) \right], \quad (8)$$

where $\theta = \{\beta_{A1} \ldots \beta_{Ak}, \delta_A, \sigma_A\}$.

Note that the global anisotropic parameters, $\{\lambda_1, \lambda_2\}$, which are estimated via cross-validation across the calibration sample of sold houses, are treated as constants in this local maximum likelihood estimation.

Hence, the local frontier for the buyer's reservation value is given by the maximization of the local log-likelihood function above, whereas the local frontier for the seller's reservation value is the negative half of the half-normal distribution, which requires a substitution from the above term $\Phi(z)$ to $[1 - \Phi(z)]$. Thus,

$$\hat{\theta}_A^{Seller\_RV} = \arg\max_{\theta} \sum_{i \neq A} e^{-\lambda_1^2 (z_{A1} - z_{i1})^2 - \lambda_2^2 (z_{A2} - z_{i2})^2}$$
$$\times \left[-\ln \sigma_A - \frac{1}{2} \ln \frac{2}{\pi} - \frac{1}{2} \left(\frac{y_i - x_i' \beta_A}{\sigma_A} \right)^2 \right.$$
$$\left. + \ln \left(1 - \Phi \left(\frac{-\delta_A (y_i - x_i' \beta_A)}{\sigma_A} \right) \right) \right]. \quad (9)$$

Table 1 ■ Data summary statistics.

Variable	Minimum	Mean	Median	Maximum	Std. Dev.
Total heated square feet	746	2,209.24	2,085	7,231	855.27
Total unheated square feet	0	187.39	0	3,360	397.65
Number of bedrooms	1	3.48	3	6	0.80
Total acres	0	1.00	0.46	33.5	1.93
Age of property (years)	0	17.58	13	92	16.58
Number of bathrooms	1	2.60	2.5	7	0.80
Number of garages	0	1.12	1	3	0.99
Number of fireplaces	0	0.96	1	4	0.53
Sold in spring	0	0.34	0	1	0.47
Sold in summer	0	0.31	0	1	0.46
Sold in fall	0	0.19	0	1	0.39
Number of days on the market	0	70.52	41	780	87.21
House is vacant	0	0.14	0	1	0.35
Seller adjusted list price	0	0.28	0	1	0.45
Seller granted concessions	0	0.14	0	1	0.34
Seller requires appointment	0	0.66	1	1	0.47

Assessing the Value of Residential Real Estate with Hedonic Price Analysis

In addition to estimating the geographically weighted stochastic frontiers, we also estimate a geographically weighted hedonic model to be used as the benchmark for fair market value. This will be useful for comparative purposes and interpretation later. Data were collected from a multiple listing service (MLS) database for a medium-sized county in the continental United States. The data collected included all detached single-family homes sold in that county for a particular year, their corresponding sales prices and predictor variables believed to influence the prices of the homes, as well as their geographic coordinates. A sample size of 1,035 homes was used in the study. We list below the variables used in our hedonic price analysis. These variables include both the characteristics of the home as well as factors believed to reflect whether the buyer was in a strong bargaining position at the time of sale. Summary statistics for them are reported in Table 1.

- *Log selling price*: The dependent variable, which is the natural logarithm of the closing price in dollars for the home.
- *Log heated square feet*: The natural logarithm of total heated square footage of the home.
- *Log unheated square feet*: The natural logarithm of total unheated square footage of the home.

- *Number of bedrooms*: The total number of bedrooms in the home.
- *Total acres*: The total size of the lot in acres that the home sits on.
- *Age of property*: The age of the property in years.
- *Number of bathrooms*: The effective number of full bathrooms in the house. Note that half-bathrooms count as 0.5 full bathrooms. Hence, a home with two full bathrooms and one half-bathroom has effectively 2.5 full bathrooms.
- *Number of garages*: The total number of garage spaces.
- *Number of fireplaces*: The total number of fireplaces.
- *Days on the market*: The number of days the home was listed before being sold.
- *Closed in spring*: A seasonal indicator variable that designates whether the home was sold in the spring quarter (April, May or June).
- *Closed in summer*: A seasonal indicator variable that designates whether the home was sold in the summer quarter (July, August or September).
- *Closed in fall*: A seasonal indicator variable that designates whether the home was sold in the fall quarter (October, November or December).
- *House is vacant*: An indicator variable that designates whether the home was vacant at the time of sale.
- *Seller adjusted list price*: An indicator variable that designates whether the list price at the time of sale was different from the original listing price.
- *Seller granted concessions*: An indicator variable that designates whether the seller at the time of sale granted any buyer concessions not directly related to the sales price of the home. Examples include the seller paying for all of the closing costs, agreements to make repairs or coverage for any inspections and warranties.
- *Seller requires appointment*: An indicator variable that designates whether the seller required a prior appointment for the prospective buyer to view the property.

Accounting for the Endogeneity Between Days on the Market and Selling Price

The relationship between days on the market and selling price has been studied quite extensively in the real estate literature (Miller 1978, Taylor 1999, Knight

2002, Anglin, Rutherford and Springer 2003). Under standard search theory, we should expect a positive relationship between the number of days on the market and selling price, because, by waiting, a seller increases the probability of encountering a buyer with a high reservation price. On the other hand, homes that remain unsold may become stigmatized in the eyes of potential buyers who regard a long market duration as evidence of some defect. Taylor (1999) explains this phenomenon as "negative herding" and suggests that a longer time on market may result in a lower sales price. In addition, sellers with high search costs should be more impatient for a sale and thus more willing to revise a price downward if an early sale is not achieved (Knight 2002). Under either theory, however, the days on the market and selling price are likely to be jointly determined; consequently, we use an instrumental variables approach via a two-stage least squares regression to correct for simultaneity between the selling price and days on the market, following an approach similar to that taken by Harding *et al*. (2003). Specifically, we include four indicator variables measuring atypicality: whether the list price is on the top and bottom deciles of list prices in this market and whether the home is very new (less than 3 years old) or very old (more than 35 years old). We also decided to omit the two other potential indicators of atypicality used by Harding *et al.* (2003) (whether the house has unusually high numbers of bedrooms and bathrooms), as they were determined to be weak instruments for our particular sample.

In addition to atypicality, we also use five other indicators believed to influence selling time on the market. These variables indicate whether the property being sold was (1) in an urban or rural area, (2) located in one town deemed to be quite popular in the area, (3) had a dining room, (4) had a family room and (5) had vinyl siding. The results from the first stage ordinary least squares (OLS) regression are presented in Table 2.

As can be seen from the results, homes located in urban areas or in the popular town tend to sell faster. This may be reflective of the density level of potential buyers for properties in urban locations: properties situated in denser markets have a faster exposure rate and shorter expected time to sale (Geltner *et al*. 1991). The measures for atypicality also indicate that homes with very low or high prices or homes that are either very new or very old also take longer to sell. Hence, this is consistent with the findings of Haurin (1988) in that the marketing times of atypical houses should be relatively longer than those of standard houses. As expected, homes with dining rooms, family rooms and those sided with vinyl also take longer to sell.

In the second stage, we substitute these predicted values as the explanatory variable for days on the market for our analysis below and proceeded with estimation.

Table 2 ■ First-stage regression results: Days on market.

Variable	Parameter Estimate	t Value	Pr > |t|
Intercept	56.11	6.2	<.0001
House is very new	42.57	5.1	<.0001
House is very old	40.93	5.4	<.0001
Listing price is very high	33.89	3.8	0.0002
Listing price is very low	22.15	2.3	0.0231
House is located in urban area	−20.70	−2.9	0.0038
House is located in desirable city	−23.84	−3.0	0.003
House has dining room	23.18	3.0	0.0028
House has family room	13.32	2.4	0.0172
House is sided with vinyl	12.22	1.4	0.1582
R^2		0.113	
Adjusted R^2		0.105	

Testing for Functional Form Misspecifications and Structural Breaks

Before continuing, we also check our model for evidence of omitted variables, functional form misspecifications and structural breaks. As a first step, indicator variables were created for the number of bedrooms and bathrooms and run in place of the actual number of bedrooms and bathrooms. After examining the results, the adjusted R^2 remained unchanged (0.883) between the two specifications (dummy coding *vs.* using actual numbers of bedrooms and bathrooms). In addition, many of the indicator variables in the dummy coding specification were found to not be statistically significant, and, at the same time, this specification required an additional 14 parameters to be estimated. Because the original specification is more parsimonious and yields the same adjusted R^2, we retain the original specification of using the actual numbers of bedrooms and bathrooms rather than using indicator variables.

In addition, we also tested for functional form misspecifications by employing an auxiliary regression to check for additional possible nonlinearities in the covariates. Specifically, we use the Regression Equation Specification Error Test based on Ramsey (1969) to check for possible higher order terms in the model (quadratic and cubic) for the given covariates. The results of the F test are nonsignificant: the p values are 0.728 for a quadratic functional form and 0.386 for a cubic form. Hence, we do not reject the null hypothesis and retain our original functional form assuming linear predictors for subsequent analysis.

Last, because a nontrivial number of homes in our sample (15.5%) sold above their respective list prices, we conducted a Chow test for structural change

(Chow 1960) to determine whether the regression coefficients differed across the homes that sold above their respective list prices and those that did not. We failed to find a significant difference at the 5% significance level (p value = 0.085). Hence, we failed to determine that a substantial difference exists. We also regressed the sales price against the usual housing characteristics but also included a dummy variable (1 = sold above list price, 0 = otherwise) which was interacted with our original covariates. None of the interactions were significant at the $\alpha = 0.05$ level.

The Appraiser's Perspective: The Fair Market Value

In order to arrive at the fair market value of a house at any given location, we applied Fotheringham *et al*'s. (2002) GWR to our data, resulting in localized estimates of the hedonic weights for various characteristics of the house. GWR is the equivalent of our geographically weighted stochastic frontier model in (8) with normal, rather than half-normal errors (*i.e.*, setting $\delta = 0$), although estimation is much simpler as it can be done through (geographically) weighted least squares estimation. See Fotheringham *et al*. (2002) for details about the GWR model and its estimation. The first step in this process is to estimate the optimal anisotropic spatial weighting function, which determines the relevance of each observed data point for the hedonic price regression, based on how close each observation is to the focal location.

The anisotropic spatial weighting function was calibrated using geographically weighted cross-validation (Cleveland 1979) leading to estimates of $\hat{\lambda}_1 = 12.43$ and $\hat{\lambda}_2 = 24.99$, thereby indicating that the influence of neighboring locations on a particular home's value decays more rapidly in the North–South than in the East–West direction. In other words, neighboring homes that are offset longitudinally from a particular home have a correspondingly greater influence in determining the value of that home than those located the same distance away in the North–South direction. This might be because the region we analyzed is more elongated in the north–south direction, as shown in the map. Because the geographic distribution of homes varies for each county and our data are based on a cross-section of homes sold in a particular county at a particular time, our results do not necessarily generalize to other counties or to other time periods, which suggest the critical roles that distance, direction and time can play in local markets for defining and creating a list of market comparables to base an appraisal on.

A natural question to ask is whether the additional computational complexity required for GWR models actually results in significantly better predictive performance and fit over a standard OLS model. Our results show a higher cross-validated R^2 value for the GWR model (0.908) than for the simple OLS model

(0.881). While at first glance the differences may not appear to be substantial, one must also take into account that they reflect predictive performance, rather than just mere fit.

Examination of the Geographically Weighted Parameter Estimates

We present a summary of the GWR parameter estimates along with those from a standard OLS model in Table 3. Because the GWR model produces one set of parameter estimates for each location, Table 3 reports the means and standard deviations for these estimates across all locations. The coefficients for *Log Heated* and *Log Unheated Square Feet* show the respective estimated percentage increases in price for a given percentage increase in heated and unheated square footage, respectively, for the home. The fact that these estimates are smaller than one show that the cost per square foot decreases with house size. The intercept for the GWR model accounts for unobservable geographic heterogeneity in house prices not captured by the geographic changes in valuation for the physical characteristics of the houses. When looking at the OLS model, two noteworthy results are that age of property has a positive impact on price and that, as the number of days on the market increases, selling price decreases. As we shall see shortly, when the coefficients are allowed to vary geographically, the parameter estimate for number of days on the market varies from positive to negative, depending on location. The positive impact of age on prices may be due to other unobservable factors (such as quality of construction or historical value) associated with age.

Examination of the Bargaining Position Variables

As mentioned earlier, we have also included a series of indicator variables to indicate the degree to which a buyer may have been in a strong bargaining position at the time of sale.[1] As illustrated in Table 3, homes that are vacant at the time of sale can be expected to sell for less than homes that are occupied. This result seems to be consistent with previous research in that sellers of vacant homes are in a weaker bargaining position than sellers whose homes are occupied (Harding *et al*. 2003). As noted by Harding *et al.* (2003), fully furnished homes are often more appealing to buyers than vacant homes, and the sellers of vacant homes bear the full cost of carrying the home with no offsetting benefits from occupancy or rental income. The fact that we allow our estimates to vary as a function of geographic space implies that the impact

[1] We originally also interacted the indicators for bargaining position with the housing characteristics but found an overwhelming majority of these interactions to be statistically insignificant. Consequently, we proceeded with using only the main effects for these indicator variables.

Table 3 ■ Parameter estimates from geographically weighted (GWR) and ordinary least squares (OLS) regressions.

Fit Statistics	OLS Model	GWR Model
Cross validated R^2	0.881	0.908
Full data R^2	0.885	0.928
Full data adjusted R^2	0.883	0.927

Variable Name	OLS Parameter Est.	OLS t Value	GWR Mean Parameter Est.	GWR Parameter Std. Dev.	GWR Parameter Coeff. Variation
Intercept	**5.40**	**28.7**	5.95	0.458	0.077
Log heated square feet	**0.88**	**30.3**	0.80	0.081	0.102
Log unheated square feet	**0.01**	**6.9**	0.01	0.002	0.240
Number of bedrooms	0.001	0.1	0.005	0.013	2.610
Total acres	**0.01**	**4.8**	0.03	0.012	0.444
Age of property (years)	**0.003**	**8.1**	0.001	0.001	1.067
Number of bathrooms	**0.12**	**11.1**	0.10	0.032	0.334
Number of garages	**0.04**	**5.0**	0.04	0.013	0.334
Number of fireplaces	**0.10**	**8.3**	0.08	0.018	0.234
Sold in spring	0.02	1.4	0.03	0.022	0.734
Sold in summer	**0.04**	**2.7**	0.04	0.014	0.375
Sold in fall	0.01	0.3	0.01	0.026	2.789
Number of days on the market	**−0.0017**	**−8.6**	0.0005	0.001	1.180
House is vacant	**−0.06**	**−3.5**	−0.06	0.021	−0.347
Seller adjusted list price	**−0.03**	**−2.3**	−0.03	0.024	−0.849
Seller granted concessions	**−0.05**	**−3.5**	−0.04	0.022	−0.490
Seller requires appointment	−0.02	−1.4	−0.01	0.023	−3.129

Note: Bolded OLS estimates and t values are statistically significant at the 0.01 level.

of vacancy on selling price changes, depending on location. For example, if a seller's relative weakness is due to a high carrying cost but the demand for homes in that location is high, then this higher demand could offset some of the weakness in bargaining position due to the home being vacant.

We also observe that sellers who adjust their list prices tend to sell their homes for less than sellers who maintain the same list price throughout. One explanation may be that sellers with higher search costs should be more impatient for a sale and therefore more willing to revise a list price if an early sale is not achieved. Indeed, it has previously been established that such impatience results in lower bargaining power (Binmore 1992). In addition, seller motivations are likely to increase over time, as for example when a planned moving date approaches or when an unmatched seller finds a desirable new home (Glower *et al*. 1998). Knight (2002) looks at the impact of changes to list price and shows that homes whose list prices were revised took longer to sell and sold for less than homes that were initially priced correctly. Last, homes with list price changes may have been mispriced initially and consequently may have been on the market a relatively long time. In addition to greater seller impatience, a stigma effect as explained by Taylor (1999) could cause the selling price to be adjusted downward.

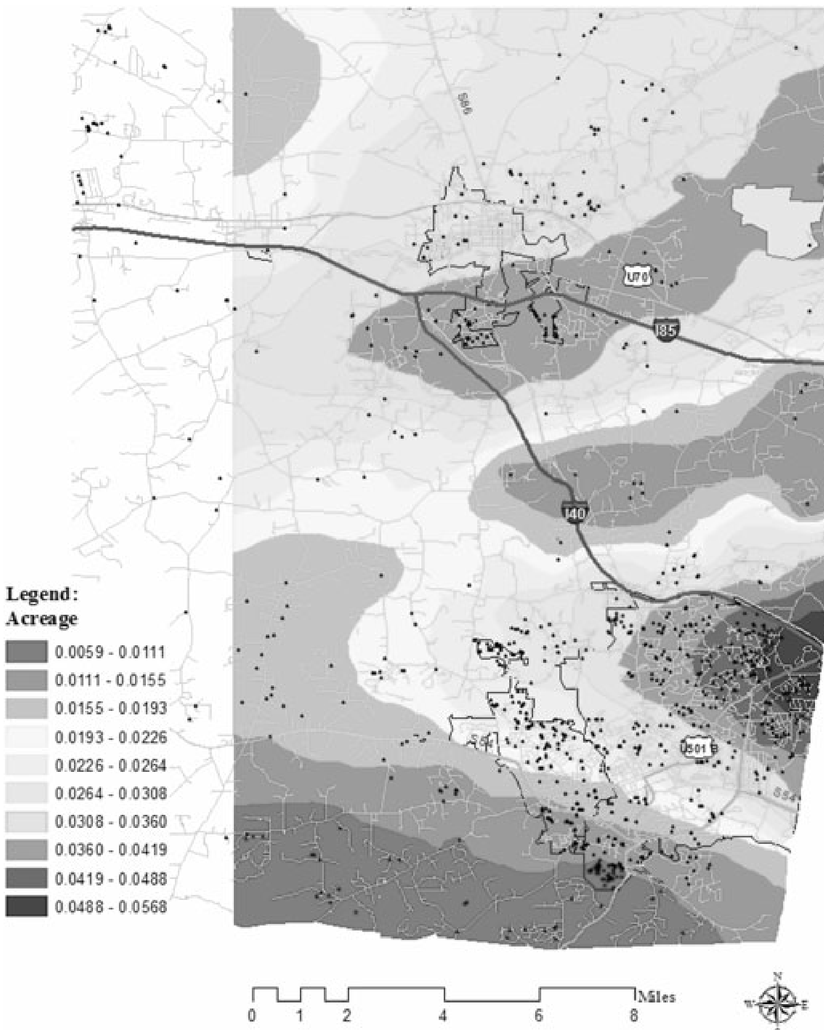
In addition, we also note that sellers who grant concessions not directly related to decreasing the sales price of the home also tend to sell their home for less than sellers who did not grant concessions. These concessions varied from closing cost dispensations to coverage of home inspections and warranties, agreements to pay for certain repairs and more. All other things being equal, sellers who agreed to concessions may have had a greater motivation to sell and therefore may have been more impatient for a sale, which resulted in a lower selling price for reasons mentioned earlier.

Finally, we also note that sellers who require that their property be shown with a prior appointment only also tended to sell their homes for less than other sellers. Although the reason for this is unclear, one possible explanation may be that sellers who require prior appointments may reduce the rate at which their property is inspected by prospective buyers. Although the parameter estimate is marginally significant in the OLS regression, 9% of the local GWR parameter estimates for prior appointment are statistically significant at the 5% significance level.

As mentioned earlier, one of the more valuable features of geographically weighted models is that they can produce local estimates of the parameters for any location, even when data are not available for that location, as it uses only the geographic coordinates for the focal location, as shown in Equation 8. A direct consequence of this feature is the ability to produce continuous surface plots for a particular area to understand how attribute values change as a function of geography. Figures 1 and 2 illustrate the GWR surface plots for the coefficients of two such attributes: *total acres* and *log heated square feet*. Each dot in these maps represents the location of one of the houses in our sample.

From the maps shown in Figures 1 and 2, one can see that the highest values for *total acres* appear in the southeast and, to a lesser extent, in the north to northeastern portions of the county. *Log heated square feet* illustrates that the southwest portions of the county generally have the highest elasticities. Thus,

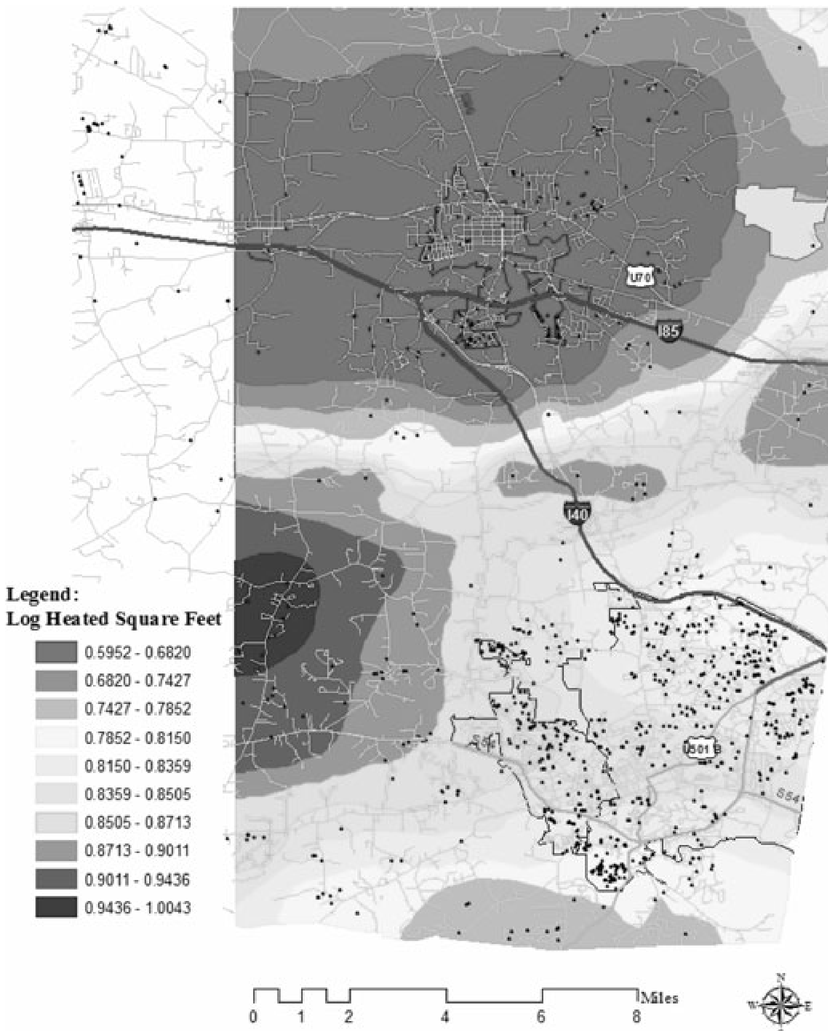
Figure 1 ■ Geographic distribution of the coefficients for acreage.



in these areas, absolute additions in acres and relative additions in heated square feet are likely to bring about the greatest relative increases in home prices.

Another important point of interest relates to number of days on the market. Because homes do not sell in a perfectly competitive environment, time is required to match sellers of this heterogeneous product with appropriate buyers (Knight 2002). Search theory is most frequently used to explain this matching

Figure 2 ■ Geographic distribution of the elasticities for heated square footage.



process as well as the presumed trade-off relationship between the seller's choice of price and time on the market. However, whether a longer search duration actually translates into a higher ultimate selling price is a subject of controversy. As mentioned earlier, sellers who are willing to wait longer increase the probability of encountering a buyer with a high reservation price. At the same time, however, we have also seen evidence that waiting too long could signal that there is some possible defect or problem with the home, which could actually lower the sales price. A chief advantage of a GWR model is that we can examine the changes of the local parameter estimates for time on the market to obtain a greater understanding of the different local effects. In our data set, approximately 88% of the homes had positive parameter estimates for days on the market, indicating that, for a majority of homes being sold, waiting longer generally translates into a higher selling price. However, in 12% of the homes being sold, we found negative parameter estimates, thereby lending at least some evidence that, in certain areas, the potential for negative herding exists and hence it is important for both buyers and sellers to be aware of this possibility. Moreover, we find that 49% of these local parameter estimates for days on the market are statistically significant at the 5% significance level.

Geographic Distribution of the Fair Market Value of a Typical Home

For a better sense of the geographical distribution of property values across our sample, we used our local regression parameter estimates to compute the expected value of a typical house on a grid of locations covering the county under study. We took the median attributes of the houses in our sample as the attributes for the typical house:

- 2,085 heated square feet ($\ln(2{,}085) = 7.6425$)
- 0 unheated square feet
- Three bedrooms
- 0.46 acres
- 13 years old in age
- 2.5 full bathrooms
- One garage
- One fireplace
- Home was on the market for 41 days
- Home was sold in the spring
- Home was occupied
- List price was not revised (original list price was used throughout sales process)
- Seller did not grant any concessions
- Seller did not require appointments prior to showing of home

Figure 3 ■ Geographic distribution of the fair market price estimate for a standard house.

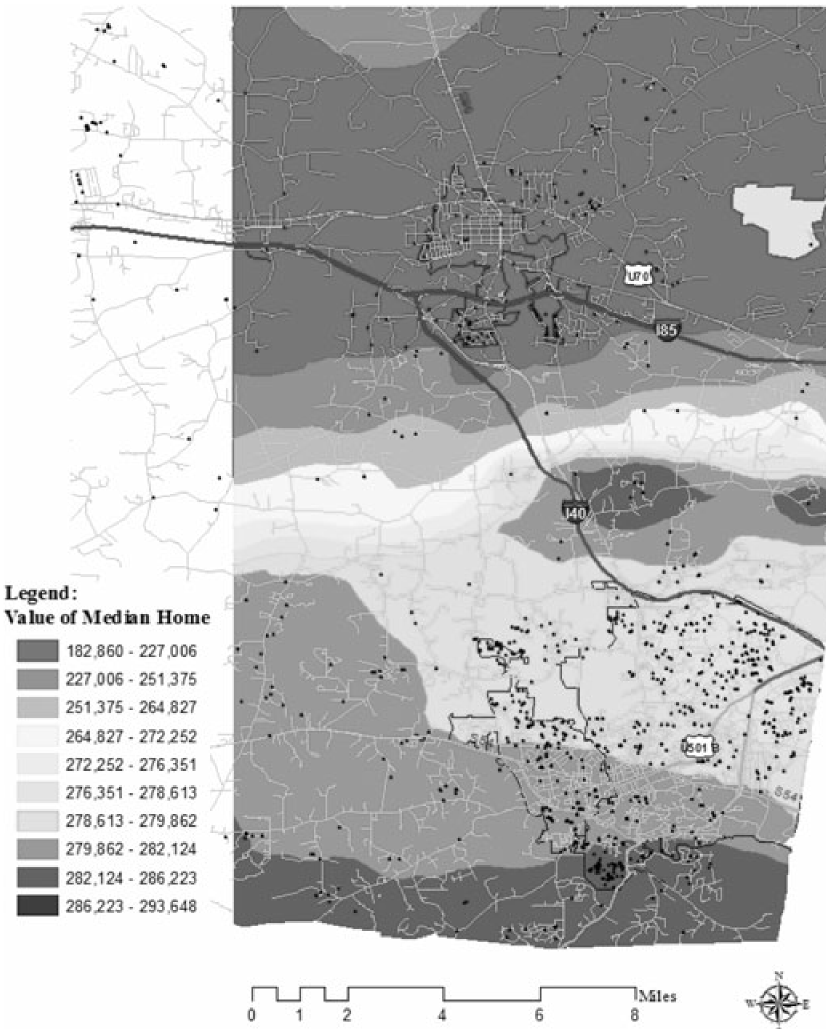


Figure 3 displays the contour map of the fair market value of this typical house in the sampled region, confirming that there is substantial geographic variation in property values and that the same house would be worth more if located in the southern region of the county.

While it is well known that correlations in errors can suggest omitted variables that have not been exploited in a model (Chatterjee, Hadi and Price 2000),

Table 4 ■ Average value of the standard median home by designated middle school.

Middle School	Mean	N	Std. Dev.
Stanford	$213,204	180	$18,003
Stanback	$248,196	136	$27,820
McDougle	$279,770	170	$607
Smith	$279,397	126	$1,263
Phillips	$278,174	199	$542
Culbreth	$281,778	224	$1,499
Total	$264,127	1,035	$28,581

geographically weighted models can account for these errors by allowing them to be modeled through the local parameter estimates. If this is true, then the geographic differences in property values shown in Figure 3 should be able to better reflect those unobservable characteristics that affect market value. In order to test whether this is the case, we consider an important factor known to influence residential property values: access to good-quality education (Jud and Watts 1981, Hayes and Taylor 1996, Bogart and Cromwell 1997, Crone 1998, Black 1999). To test this hypothesis, we compare the average prices of the same standard median home shown in Figure 3 by their designated middle school. Results from this comparison, shown in Table 4, strongly support the hypothesis; there is a clear and statistically significant difference in average prices for the same standard median home across the designated schools, and middle school alone explains 81% of the variance in standard prices across all 1,035 locations in our sample. Similar results were obtained using the designated elementary schools and high schools.

The benefits of allowing for geographic variation in hedonic price analysis are numerous. As illustrated in Figure 3, it is possible to take the same home with identical attributes and understand how the expected price changes solely as a function of location. New housing contractors could use this method to determine the best tracts of land to build on to maximize expected profit. It is also possible to determine which specific housing attributes add the greatest profits by location, and contractors could even tailor construction in different areas to emphasize those features that are most highly valued at a given location. Following is an illustrative example.

The county used in this study had a majority of home sales in eight different ZIP codes: 27517, 27514, 27516, 27243, 27705, 27302, 27510 and 27278. Eight specific home locations (one in each of the above ZIP codes) were randomly selected and the value for the same typical house described earlier was computed for each of these eight locations. The “Appraiser’s Price” column of Table 7 reflects the fair market price for the same standard (median) home at each of these specific locations. One can see that location alone can account for substantial variations in this fair market price for the same standard home, ranging from $189,902 (in ZIP code 27705) to $279,831 (in ZIP code 27510).

The Buyer and Seller Perspectives: Geographically Weighted Stochastic Frontier

In order to sell homes more effectively, however, more information is needed than simply the appraiser's estimate of home value. For example, in one area, perhaps the estimated differential between the buyer and seller reservation values is minimal. This might suggest that brokers have less work to do in selling properties in these areas, compared to other areas where the differentials are large. A property with a tighter expected distribution range between buyer and seller reservation values could potentially decrease conflicts of interest between the seller and broker. These properties likely require less negotiations and efforts on the parts of the broker, who can allow the properties to speak for themselves.

The geographically weighted stochastic and inverse stochastic frontiers for the same eight median locations are presented below. Table 5 contains the local stochastic frontiers (buyer's reservation value) and Table 6 contains the inverse local stochastic frontier estimates (seller's reservation value) for these same eight locations. Using these estimates, it is possible to calculate predicted values for the maximum and minimum expected selling prices, respectively, for the identical home located in different areas; recall that, in the previous example, the expected selling price (from the appraiser's standpoint) of a hypothetical home located in ZIP code 27510 was $279,831. Then, continuing with this example and using Tables 5 and 6, we see that the estimated buyer's reservation value for this same hypothetical home in ZIP code 27510 is $308,313, while seller's reservation value of the same home in 27510 is $279,753.

Table 7 presents a summary of the three price estimates by location. It is interesting to see how the expected selling prices and ranges of price negotiations for the same house vary as a function of location. By looking at the location in ZIP code 27243, for example, one can see that the expected appraisal estimate ($244,700) is much closer to the maximum frontier of the property ($244,727) than to the minimum frontier ($212,708). This might indicate that this location (all other things being equal) is in more of a local buyer's market, in that the estimated fair market value is close to the maximum expected price that the buyer could expect to pay for the home. In other words, using the fair market value as a benchmark, the buyer will likely not pay much more than the

Table 5 ■ Best feasible price from the seller's perspective.

Variable Name	Median Values Hypoth. Home	ZIP: 27517 Local Betas	ZIP: 27514 Local Betas	ZIP: 27516 Local Betas	ZIP: 27243 Local Betas	ZIP: 27705 Local Betas	ZIP: 27302 Local Betas	ZIP: 27510 Local Betas	ZIP: 27278 Local Betas
Intercept	1	5.83	5.64	5.34	6.14	6.11	6.25	5.75	7.18
Log heated square feet	7.64	0.83	0.84	0.89	0.74	0.74	0.71	0.85	0.59
Log unheated square feet	0	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01
Number of bedrooms	3	0.009	−0.032	−0.015	0.053	0.017	0.000	0.001	0.019
Total acres	0.46	0.04	0.02	0.01	0.03	0.04	0.04	0.02	0.03
Age of property (years)	13	0.001	0.004	0.004	0.000	−0.001	0.001	0.002	0.002
Number of bathrooms	2.5	0.09	0.17	0.13	0.11	0.06	0.17	0.08	0.15
Number of garages	1	0.03	0.04	0.04	0.03	0.06	0.08	0.03	0.07
Number of fireplaces	1	0.06	0.10	0.11	0.08	0.04	0.16	0.08	0.09
Sold in spring	1	0.02	0.01	0.02	0.14	0.01	−0.03	0.03	0.02
Sold in summer	0	0.04	0.03	0.03	0.09	0.05	−0.04	0.04	0.03
Sold in fall	0	−0.03	0.02	0.02	0.11	0.05	0.00	0.00	0.04
Number of days on the market	41	0.001	−0.002	−0.002	−0.001	0.002	0.000	0.000	0.000
House is vacant	0	−0.05	−0.08	−0.06	−0.07	−0.08	−0.06	−0.08	−0.09
Seller adjusted list price	0	−0.02	−0.05	−0.04	0.05	−0.04	−0.13	−0.03	−0.04
Seller granted concessions	0	−0.05	−0.08	−0.05	−0.11	0.02	0.06	−0.05	−0.04
Seller requires appointment	0	0.01	−0.07	−0.05	−0.05	0.00	0.06	0.01	−0.07
δ		0.76	0.00	0.00	0.00	0.00	0.00	0.96	0.86
σ		0.19	0.14	0.15	0.11	0.11	0.12	0.18	0.14
Predicted selling price for each location:		$305,004	$267,832	$279,796	$244,727	$189,926	$226,925	$308,313	$236,783

Table 6 ■ Best feasible price from the buyer's perspective.

Variable Name	Median Values Hypoth. Home	ZIP: 27517 Local Betas	ZIP: 27514 Local Betas	ZIP: 27516 Local Betas	ZIP: 27243 Local Betas	ZIP: 27705 Local Betas	ZIP: 27302 Local Betas	ZIP: 27510 Local Betas	ZIP: 27278 Local Betas
Intercept	1	5.73	5.88	5.21	5.58	6.11	6.29	5.60	7.11
Log heated square feet	7.64	0.83	0.78	0.89	0.80	0.73	0.70	0.86	0.60
Log unheated square feet	0	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01
Number of bedrooms	3	0.010	−0.023	−0.014	0.045	0.015	−0.004	0.003	0.017
Total acres	0.46	0.04	0.03	0.01	0.02	0.04	0.04	0.02	0.03
Age of property (years)	13	0.001	0.002	0.004	0.001	−0.002	0.000	0.002	0.001
Number of bathrooms	2.5	0.09	0.16	0.12	0.13	0.07	0.15	0.08	0.15
Number of garages	1	0.03	0.03	0.03	0.03	0.06	0.09	0.03	0.07
Number of fireplaces	1	0.07	0.12	0.11	0.05	0.05	0.19	0.08	0.09
Sold in spring	1	0.02	0.02	0.02	0.09	0.03	−0.02	0.03	0.02
Sold in summer	0	0.04	0.03	0.03	0.07	0.06	−0.04	0.04	0.03
Sold in fall	0	−0.03	0.03	0.02	0.07	0.06	0.02	0.00	0.04
Number of days on the market	41	0.001	−0.002	−0.002	−0.001	0.002	0.000	0.000	0.000
House is vacant	0	−0.05	−0.05	−0.06	−0.02	−0.07	−0.07	−0.08	−0.08
Seller adjusted list price	0	−0.02	−0.07	−0.04	0.04	−0.04	−0.14	−0.03	−0.04
Seller granted concessions	0	−0.05	−0.07	−0.05	−0.14	0.00	0.08	−0.05	−0.04
Seller requires appointment	0	0.01	−0.06	−0.05	−0.05	0.00	0.05	0.00	−0.06
δ		0.00	3.62	1.35	3.13	1.79	1.39	0.00	0.00
σ		0.16	0.23	0.20	0.17	0.16	0.16	0.15	0.12
Predicted selling price for each location:		$279,179	$223,453	$247,176	$212,708	$174,285	$205,985	$279,753	$221,307

Table 7 ■ Price comparisons from the appraiser's, seller's and buyer's perspectives.

ZIP Code	Buyer's Best Price	Appraiser's Price	Seller's Best Price	Negotiation Range	Neg. Range % Appr. Price
27517	279,179	279,257	305,004	25,825	9.2%
27514	223,453	267,803	267,832	44,379	16.6%
27516	247,176	279,742	279,796	32,620	11.7%
27243	212,708	244,700	244,727	32,020	13.1%
27705	174,285	189,902	189,926	15,640	8.2%
27302	205,985	226,898	226,925	20,941	9.2%
27510	279,753	279,831	308,313	28,560	10.2%
27278	221,307	221,356	236,783	15,476	7.0%

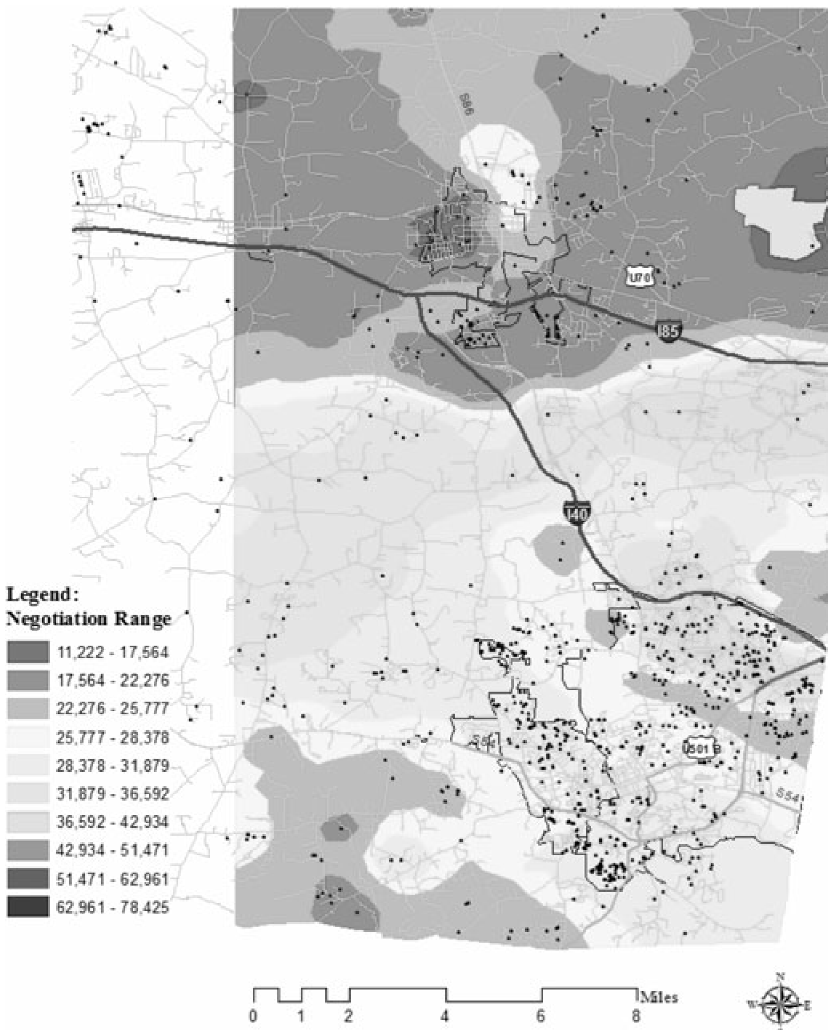
appraisal price, but stands to pay substantially less, particularly if the buyer is able to bargain effectively with the seller. In Figure 4 we present the range of negotiations (difference between the buyer and seller reservation values) across the whole county, which suggests that the range of price negotiations between buyers and sellers is largest in the eastern parts of the county, where property values are generally in the relatively high-priced ranges. However, this is not always the case. It can be seen from comparing Figure 3 to Figure 4 that homes with higher fair market valuations do not always correspond to having largest negotiation ranges. Hence, the estimated fair market value of the home is not the only driver of how large the potential negotiation range may be.

Application in a Bargaining Context

As mentioned earlier, bargaining power can materially affect the outcomes of real estate transactions. In a hedonic framework, bargaining can be thought of as a multifaceted activity, in which bargaining over the price of the home partially consists of bargaining over the implicit shadow prices of each characteristic, each of which contributes to the overall bargaining outcome (Song 1995).

Our model presents a straightforward interpretation of this perspective. Each variable in our model has three different estimated shadow prices; two of the shadow prices correspond to the respective buyer and seller reservation values and one shadow price corresponds to the fair market value. Note that if all the shadow prices from the buyer and seller perspectives converge to the shadow prices for the fair market value, then the estimated reservation values would also converge to the fair market value and bargaining would dissipate as homes would be expected to sell for their appraisal prices. Hence, the overall negotiation range between the buyer and seller (which defines the latitude under which bargaining takes place) is a direct consequence of how

Figure 4 ■ Geographic distribution of the estimated negotiation ranges for a standard house.



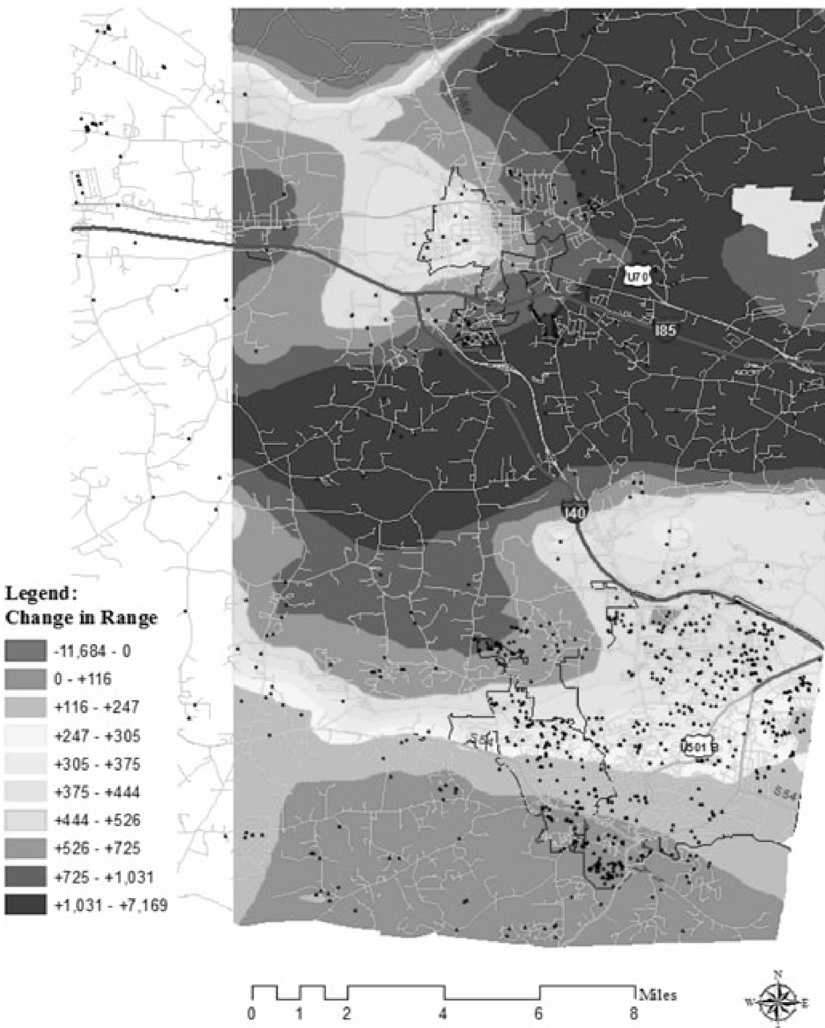
large the differences are between the buyer's and seller's valuations for each of the shadow prices. If the buyer and seller shadow prices are close to each other for a certain housing component, then this component is likely to have little effect on the negotiation range, and therefore bargaining is not likely to take place over that attribute. However, this relationship is not necessarily constant over geographic space. As an example, buyers and sellers may have different shadow prices for the age of a home in an old neighborhood than in an area that is predominantly new construction. As a result, our model also allows for the shadow prices to vary as a function of location. Hence, in some locations, certain housing attributes may be relatively more important in driving the negotiation process than in other areas. This can provide the analyst with a greater understanding of what characteristics are relatively more important in the bargaining process as a function of location.

Figure 5 illustrates the impact that the attribute age of home can have on the overall negotiation range. To construct Figure 5, we first set the buyer and seller local parameter estimates for age equal to their respective local fair market value estimates and then recalculate the negotiation range. Note that this new negotiation range corrects for the incremental impact that age has in determining how far apart the buyer and seller reservation values are because we have set the local parameter estimates for age in both cases equal to one another. We then subtract this newly calculated negotiation range from the original negotiation range to obtain an estimate for how the differences between the buyer's and seller's shadow prices for age impact the total negotiation process. Figure 5 plots this change.

As illustrated in Figure 5, we see that, for a large majority of the county, the differences between the buyer's and seller's local parameter estimates for age are positive and can increase the overall negotiation range up to approximately \$7,000. However, for a small number of cases (2.9%), the sellers' shadow prices actually exceed the buyers' shadow prices. It is interesting to note that, while the overall reservation value for the buyer must be at least as large as the overall reservation value for the seller (in order for a sale to occur), this does not necessarily constrain the individual shadow prices for each attribute to also be greater for the buyer than the seller. If the seller's shadow price for a particular attribute exceeds the buyer's shadow price for that same attribute, then this implies that no negotiations occur on that attribute, which can adversely affect the overall negotiation process if the disparity is large enough. The degree to which this adversely affects or decreases the overall negotiation range can be seen by the locations in Figure 5 that are shaded in dark gray which occur in a small area in the far north of the county.

Sellers and buyers, however, are likely most interested in obtaining the overall best price for a sale. To the extent that certain markets or pockets may favor one party over another would likely be of great interest to the sellers and buyers as well as the brokers who are interested in representing them. Using the fair market value as a benchmark, we compute a quantity that is designed to assess the degree to which the seller may have the upper hand in a potential sale,

Figure 5 ■ Geographic distribution of the change in negotiation range for a standard house when shadow prices for age are set to fair market value.



which is defined as follows:

$$\omega = \frac{(Buyer.RV - Fair.Market.Value)}{(Buyer.RV - Seller.RV)}. \tag{10}$$

The intuition behind the above quantity is as follows: If the buyer's reservation value is significantly above the fair market value, while at the same time the seller's reservation value is only slightly lower than the fair market value, then the home is likely to sell for a price that is above the fair market value, which is beneficial to the seller. On the other hand, if the buyer's reservation value lies close to the fair market value but the seller's reservation value is substantially below the fair market value, then the home is likely to sell for a price that is lower than the fair market value, which is advantageous for the buyer. The quantity ω above is constrained to lie between 0 and 1, and a value of 0.5 implies that the fair market value lies equidistant between the buyer and seller reservation values. Values closer to 1.0 favor the seller while values closer to 0.0 favor the buyer.

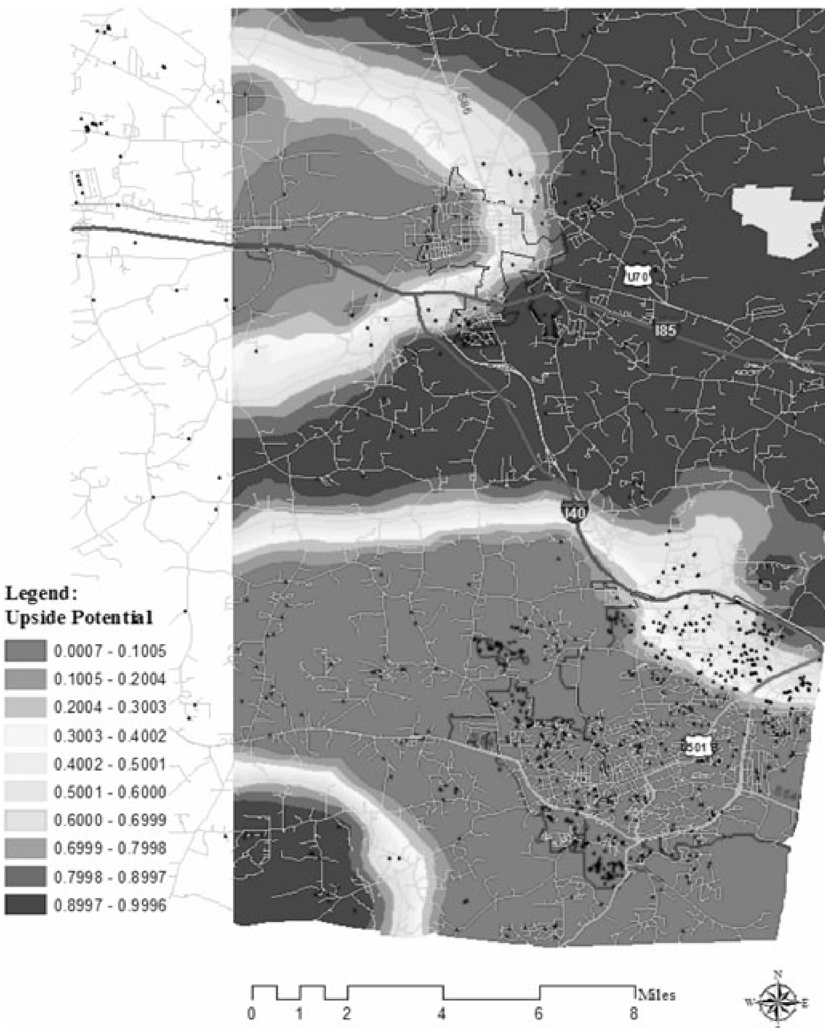
Figure 6 plots the quantity ω, which reflects the seller's upside potential of being able to sell the home for a price that is above the fair market value. As illustrated in Figure 6, much of the lower half of the county favors buyers while the upper half tends to favor sellers. There are also bands of areas (in particular in the center of the country near the interstate highway) where buyers and sellers appear to have roughly the same upside potential for negotiating a good price, relative to the appraisal price.

The mean fair market value for the sample of the typical home across our sample of locations was approximately \$264,127. Hence, the mean range of negotiations between the buyer and seller (\$27,790) divided by the expected fair market value sales price of \$264,127 implies a 10.5% average range of negotiation in our sample, suggesting that homes on average tend to sell for up to about 5.25% below or above their fair market values. This result seems to be consistent with other empirical findings regarding the effects that bargaining power has on the selling price. For example, using data from the American Housing Survey, Harding *et al.* (2003) find that a household's bargaining power can seasonally vary on the order from 5% to 7%.

Assessing the Reliability of the Frontier Estimates

In order to assess the reliability of the stochastic frontier estimates, the buyer's reservation value can be compared to the list price. Because the list price is often seen as a contractual upper bound on the selling price (Horowitz 1992, Yavas and Yang 1995, Knight 2002), we should expect to see a close correspondence between the buyers' estimated reservation values and the list prices for actual homes sold. Hence, we also compute the stochastic frontiers for our original data (not the typical median home) and compare the estimated buyers' reservation values to the actual list prices. Overall, we find that the buyers' estimated reservation values across our sample are approximately 6.4% higher than the list price. Hence, the empirical evidence supports the proposition that, while the list price is relatively close to the buyer's reservation value, the reservation values are slightly higher than the list prices, suggesting that reservation values

Figure 6 ■ Geographic distribution of the seller's upside bargaining potential for a standard house.



do in fact represent a theoretical ideal. It should also be noted that, in our sample, 15.5% of the homes were sold for amounts greater than the list price. Hence, in a scenario where more than a small proportion of homes are sold for amounts above their theoretical contractual upper bounds, it is not surprising to see frontiers that are slightly larger than what might be expected under more regular conditions.

As a further check on our estimated reservation values, we also regressed the list price against the housing characteristics and computed residuals. We then regressed these residuals against the estimated buyer reservation values computed earlier and found a strong, statistically significant, positive relationship between the two (p value < 0.0001). This seems to imply that, once we have controlled for housing characteristics, sellers with higher-than-predicted list prices are accompanied by higher buyer reservation values, while sellers with lower-than-predicted list prices are accompanied by lower buyer reservation values. Hence, we seem to have additional evidence that movements in the buyer's estimated reservation values mirror movements in the list price set by the seller.

Conclusions and Areas for Further Research

In this article, we provide a new model that can be used to obtain a greater understanding of buyer and seller reservation values as well as the fair market value. Our geographically weighted stochastic frontier model provides estimates of the fair market value and also the maximum and minimum expected price for a house, given its characteristics and location, based on available MLS data in the same real estate market. By doing so, this model provides a price range within which transaction prices can be negotiated, which can then be used to enhance decision-making processes along a multitude of dimensions. By estimating local frontiers using geographically weighted data, our model also accounts for the fact that different property features are valued differently, depending on location.

Our proposed model is useful in practical applications where there is large market data available but for which there is still imperfect information. For example, despite the fact that MLS data are now widely available, Yavas (1992) points out that real estate markets can best be described by imperfect information: the players do not know the locations of the reservation prices of their potential trading partners and that this lack of information compels each player to engage in search activity in order to find a trading partner. In our model, however, players can now have location-specific estimates for the expected reservation values for all parties involved. This could be used to reduce market uncertainties and improve negotiations. Sellers who are getting ready to list their homes for sale could, for example, compare the suggested list price that the brokers recommend to the expected buyer reservation value to get a feel for how reasonable the suggested broker list price is and also how closely aligned the broker's motivations are with the seller's motivations.

This model could also be used to match potential buyers and sellers more successfully. Astute buyers, for instance, could use this information to efficiently search for local submarkets or homes that offer lower-than-expected market prices or search for homes where their reservation values overlap with the expected seller reservation values.

We have also demonstrated how reservation values and fair market values can change across geography, which can have obvious consequences to urban planning, sales management and more.

As far as opportunities for future research, there are a number of areas where this model could be extended. One such example is to readjust the kernel to allow the anisotropic weightings to vary by location. There is no reason to believe that the decay factors are globally fixed throughout the study area. Using an anisotropic kernel that allows for the rates of decay to change as a function of location might incorporate additional realism and flexibility into the model. Another possibility lies with modeling changes in home prices temporally as well as spatially. Hence, it would be interesting to understand how the expected prices (and frontiers) of the median home evolve over time, in addition to location. For instance, in some areas, the reservation values may be expected to change much more quickly than in others, indicating a wider range of negotiations (and perhaps more market uncertainty) going forward. This might have important implications for long-term housing development projects.

Extending the model to account for temporal shifts in property values would also have the added benefit of increasing the sample of locations, leading to more accurate local measures of property value. With this increased accuracy, it might be possible to extend our analyses and look at the impact of other factors such as proximity to highways, lakes, power plants, business districts and so forth on property values, with obvious consequences to urban planning.

Finally, in the estimation of the buyer/seller valuations and fair market value, we used only the available market data on final transaction prices and property attributes, not considering the fact that some sellers may have decided not to take any of the offers they received and some buyers may have decided not to make a purchase. In other words, we only considered cases where the buyer's valuation was greater than the seller's and the final transaction price was somewhere in between; for this reason, our estimates are based on truncated data, because we do not observe the cases where a compromise was not reached.

References

Aigner, D.J., C.A. Knox Lovell and P. Schmidt. 1977. Formulation and Estimation of Stochastic Frontier Production Function Models. *Journal of Econometrics* 6: 21–37.

Anglin, P.M., R. Rutherford and T.M. Springer. 2003. The Trade-Off Between the Selling Price of Residential Properties and Time-on-the-Market: The Impact of Price Setting. *The Journal of Real Estate Finance and Economics* 26: 95–111.

Basu, S. and T.G. Thibodeau. 1998. Analysis of Spatial Autocorrelation in House Prices. *The Journal of Real Estate Finance and Economics* 17: 61–85.

Battese, G.E. and G.S. Corra. 1977. Estimation of a Production Frontier Model: With Application to the Pastoral Zone Off Eastern Australia. *Australian Journal of Agricultural Economics* 21: 169–179.

Bell, K.P. and N.E. Bockstael. 2000. Applying the Generalized Method of Moments Approach to Spatial Problems Involving Micro-Level Data. *Review of Economics and Statistics* 82: 72–82.

Binmore, K.G. 1992. *Fun and Games: A Text on Game Theory*. Lexington, MA: Houghton Mifflin.

Black, S.E. 1999. Do Better Schools Matter? Parental Valuation of Elementary Education. *Quarterly Journal of Economics* 2: 577–599.

Bogart, W.T. and B.A. Cromwell. 1997. How Much More Is a Good School District Worth? *National Tax Journal* 50: 215–232.

Bourassa, S.C., F. Hamelink, M. Hoesli and B.D. MacGregor. 1999. Defining Housing Submarkets. *Journal of Housing Economics* 8: 160–183.

Can, A. 1992. Specification and Estimation of Hedonic Housing Price Models. *Regional Science and Urban Economics* 22: 453–474.

Can, A. and I. Megbolugbe. 1997. Specification and Estimation of Hedonic Housing Price Models. *The Journal of Real Estate Finance and Economics* 14: 203–222.

Chatterjee, S., A.S. Hadi and B. Price. 2000. *Regression Analysis by Example*. New York: Wiley.

Childs, P.D., S.H. Ott and T.J. Riddiough. 2002. Optimal Valuation of Noisy Real Assets. *Real Estate Economics* 30: 385–414.

Chow, G. 1960. Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica* 28: 591–605.

Clapp, J.M. 2001. A Semi Parametric Method for Estimating Local House Price Indices. Working Paper. University of Connecticut: Storrs, CT.

Clapp, J.M., H.J. Kim and A. Gelfand. 2002. Predicting Spatial Patterns of House Prices Using LPR and Bayesian Smoothing. *Real Estate Economics* 30: 505–532.

Cleveland, W.S. 1979. Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association* 74: 829–836.

Cobb, C.W. and P.H. Douglas. 1928. A Theory of Production. *American Economic Review* 18: 139–165.

Colwell, P.F. and H.J. Munneke. 2006. Bargaining Strength and Property Class in Office Markets. *The Journal of Real Estate Finance and Economics* 33: 197–213.

Crone, T.M. 1998. House Prices and the Quality of Public Schools: What Are We Buying? *Business Review (Federal Reserve Bank of Philadelphia)* Sept/Oct: 3–14.

Dale-Johnson, D. 1982. An Alternative Approach to Housing Market Segmentation Using Hedonic Price Data. *Journal of Urban Economics* 11: 311–332.

Day, B.H. 2003. Submarket Identification in Property Markets: A Hedonic Housing Price Model for Glasgow. CSERGE Working Paper EDM 03-09. University of East Anglia: Norwich, UK.

Day, B.H., I.J. Bateman and I. Lake. 2003. What Price Peace? A Comprehensive Approach to the Specification and Estimation of Hedonic Housing Price Models. CSERGE Working Paper EDM 03-08. University of East Anglia: Norwich, U.K.

Dubin, R.A. 1992. Spatial Autocorrelation and Neighborhood Quality. *Regional Science and Urban Economics* 22: 433–452.

Dubin, R.A., K. Pace and T.G. Thibodeau. 1999. Spatial Autoregression Techniques for Real Estate Data. *Journal of Real Estate Literature* 7: 79–95.

Epple, D. 1987. Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products. *Journal of Political Economy* 95: 59–80.

Fan, J., M. Farmen and I. Gijbels. 1998. Local Maximum Likelihood Estimation and Inference. *Journal of the Royal Statistical Society B* 60: 591–608.

Fik, T.J., D.C. Ling and G.F. Mulligan. 2003. Modeling Spatial Variation in Housing Prices: A Variable Interactions Approach. *Real Estate Economics* 31: 623–646.

Fotheringham, A.S., C. Brunsdon and M. Charlton. 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. John Wiley & Sons: West Sussex, U.K.

Geltner, D., B. Kluger and N. Miller. 1991. Optimal Price and Selling Effort from the Perspectives of the Broker and Seller. *Journal of the American Real Estate and Urban Economics Association* 19: 1–24.

Gibbons, S. and S. Machin. 2001. Valuing Primary Schools. Working Paper. University College London and Centre for Economic Performance: London.

——. 2003. Valuing English Primary Schools. *Journal of Urban Economics* 53: 197–219.

Glower, M., D.R. Haurin and P.H. Hendershott. 1998. Selling Time and Selling Price: The Influence of Seller Motivation. *Real Estate Economics* 26: 719–740.

Goodman, A.C. and T.G. Thibodeau. 1998. Housing Market Segmentation. *Journal of Housing Economics* 7: 121–143.

——. 2003. Housing Market Segmentation and Hedonic Prediction Accuracy. *Journal of Housing Economics* 12: 181–201.

Griliches, Z. 1961. Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change. *Price Statistics of the Federal Government*, General Series 73: 137–196. New York: National Bureau of Economic Research.

Harding, J.P., J.R. Knight and C.F. Sirmans. 2003. Estimating Bargaining Effects in Hedonic Models: Evidence from the Housing Market. *Real Estate Economics* 31: 601–622.

Harding, J.P., S.R. Rosenthal and C.F. Sirmans. 2003. Estimating Bargaining Power in the Market for Existing Homes. *Review of Economics and Statistics* 85: 178–188.

Haurin, D. 1988. The Duration of Marketing Time of Residential Housing. *Journal of the American Real Estate and Urban Economics Association* 16: 396–410.

Hayes, K.J. and L.L. Taylor. 1996. Neighborhood School Characteristics: What Signals Quality to Homebuyers? *Economic Review (Federal Reserve Bank of Dallas*) Fourth Quarter: 2–9.

Horowitz, J.L. 1992. The Role of List Price in Housing Markets: Theory and an Econometric Model. *Journal of Applied Econometrics* 7: 115–129.

Host, G. 1999. Kriging by Local Polynomials. *Computational Statistics and Data Analysis* 29: 295–312.

Jud, G.D. and J.M. Watts. 1981. Schools and Housing Values. *Land Economics* 57: 459–470.

Knight, J.R. 2002. Listing Price, Time on the Market, and Ultimate Selling Price: Causes and Effects of Listing Price Changes. *Real Estate Economics* 30: 213–237.

Kelejian, H.H. and I.R. Prucha. 1999. A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model. *International Economic Review* 40: 509–533.

Kumbhakar, S.C. and C.A. Knox Lovell. 2000. *Stochastic Frontier Analysis*. New York: Cambridge University Press.

Meeusen, W. and J. Van Den Broeck. 1977. Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error. *International Economic Review* 18: 435–444.

Miceli, T. 1989. The Optimal Duration of Real Estate Listing Contracts. *Journal of the American Real Estate and Urban Economics Association* 17: 267–277.

Miller, N.G. 1978. Time on the Market and Selling Price. *Journal of the American Real Estate and Urban Economics Association* 6: 164–174.

Pace, R.K. 1997. Performing Large Spatial Regressions and Autoregressions. *Economics Letters* 54: 283–291.

Pace, R.K. and R. Barry. 1997. Quick Computation of Regressions with a Spatially Autoregressive Dependent Variable. *Geographical Analysis* 29: 232–247.

Pace, R.K., R. Barry, J. Clapp and M. Rodriguez. 1998. Spatiotemporal Autoregressive Models of Neighborhood Effects. *The Journal of Real Estate Finance and Economics* 17: 15–33.

Pace, R.K. and O.W. Gilley. 1997. Using the Spatial Configuration of the Data to Improve Estimation. *The Journal of Real Estate Finance and Economics* 14: 333–340.

Pavlov, A.D. 2000. Space-Varying Regression Coefficients: A Semi-Parametric Approach Applied to Real Estate Markets. *Real Estate Economics* 28: 249–283.

Quan, D.C. and J.M. Quigley. 1991. Price Formation and the Appraisal Function in Real Estate Markets. *The Journal of Real Estate Finance and Economics* 4: 127–146.

Ramsey, J.B. 1969. Tests for Specification Errors in Classical Linear Least Squares Regression Analysis. *Journal of the Royal Statistical Society B* 32: 350–371.

Ratchford, B.T. 1980. The Value of Information for Selected Appliances. *Journal of Marketing Research* 17: 14–25.

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

Song, S. 1995. Determinants of Bargaining Outcome in Single-Family Housing Transactions: An Empirical Examination. *Urban Studies* 32: 605–614.

Taylor, C. 1999. Time-on-the-Market as a Sign of Quality. *Review of Economic Studies* 66: 555–578.

Wilhelmsson, M. 2004. A Method to Derive Housing Sub-Markets and Reduce Spatial Dependency. *Property Management* 22: 276–288.

Yavas, A. 1992. A Simple Search and Bargaining Model of Real Estate Markets. *Journal of the American Real Estate and Urban Economics Association* 20: 533–548.

Yavas, A. and S. Yang. 1995. The Strategic Role of Listing Price in Marketing Real Estate: Theory and Evidence. *Real Estate Economics* 23: 347–368.