JOURNAL OF HOUSING RESEARCH VOLUME **15** ISSUE **2**

J
H
R

An Alternative Spatial Hedonic Estimation Approach

Clifford A. Lipscomb

Abstract

Housing studies typically use parcel level distance variables or some variant of the spatial weights matrix approach to incorporate spatial effects into hedonic regression models. In this paper, using very detailed data on household attitudes and parcel attributes, hedonic regression residuals are used in a structural equations framework to check for additional spatial effects in the hedonic coefficients beyond those captured in the hedonic regression itself. In this way, a ''nearest neighbors'' approach utilizing parcel level distance variables is compared directly to OLS estimation using spatial variables, showing the relative efficiency of the estimates in the former approach.

This paper received the award for the best paper on Real Estate Market Analysis (sponsored by Torto Wheaton Research) presented at the ARES 2006 Annual Meeting.

Hedonic price analysts frequently assume that households within a particular ''community'' (however defined) are homogeneous enough to be aggregated into a single group. Thereafter, the demand for changes in community amenities can be analyzed using a single-line regression model such as ordinary least squares (OLS) or generalized least squares (GLS). For hedonic price models designed to analyze the value of a single real estate amenity, aggregation is critical in order to translate the marginal implicit prices of neighborhood attributes into valid economic ''welfare'' information. If aggregation assumptions are not met, the analyst might be able to report expected changes in real estate values with some confidence, but cannot convey with theoretical confidence true welfare change estimates. Fortunately, demand aggregation does not require *perfect* homogeneity; under some conditions analysts can manage heterogeneity in hedonic goods through assumptions about the error structure of the hedonic demand estimate.[1]

Those who adopt aggregability under a single hedonic equation will be referred to as aggregation ''optimists.'' Based on their articles, Ekeland, Heckman, and Nesheim (2002), Sieg, Smith, Banzhaf, and Walsh (2002), and those who adopt the ''one neighborhood, one submarket'' housing market equilibrium, may be considered to be aggregation ''optimists'' for the purposes of this research. At the other end of the spectrum are the aggregation ''pessimists,'' who view aggregation more as a ''one household, one submarket'' equilibrium. Pessimists argue that household diversity or heterogeneity largely invalidates welfare analyses from hedonic price studies. Strong heterogeneity, they argue, in the packaging of goods translates into a market where each developed parcel

represents a unique supply decision, so unique from all others that each household demand corresponds to its own unique bundling technology, which obviates valid aggregation. Unfortunately, each perspective frequently relies on the same data. So, what kind of hedonic model might appease the extremists in both camps? One answer is ''a hedonic model that structurally accounts for households' different implicit marginal prices for an array of goods and amenities.''

Then, once these two camps can agree on the kind of hedonic model to estimate, the researcher must focus on the incorporation of spatial effects into the model. Several recent articles, including Anselin (2002) and Case, Clapp, Dubin, and Rodriguez (2004), highlight the use of new spatial econometric approaches to estimate housing hedonic models. Typically, these approaches are used when data do not contain information on the particular spatial features that are considered important to local residents (i.e., when spatial contiguity is used to account for spatial variation in prices, etc.). The common theme in these articles is that a spatial weights matrix (SWM) approach is used to correct the hedonic estimates for the impact of spatial phenomena. However, the SWM approach has its advantages and disadvantages, which have been well documented in the spatial literature (see, for example, Anselin 2002). When a researcher knows nothing about what spatial variables may impact a particular dwelling's sales price, the researcher may use a SWM approach to incorporate the assumed impacts of neighbors' characteristics on the dwelling of interest. This is particularly useful when the researcher does not have complete data for each parcel in a given area, suggesting that a nearest neighbor's approach is possible when missing parcel information prevents the use of a traditional spatial lag or spatial error model. The choice of SWM, then, is not an exact science, although the literature provides some guidance on the appropriate scenarios to use such approaches as a spatial lag or spatial error model (*Ibid.*).

In this paper, an alternative spatial estimation strategy is proposed that incorporates the residuals from a hedonic regression (with spatial variables) as explanatory variables in a structural equations model. Then, information from the structural model is used to weight the original hedonic regression in a White (1980) style heteroscedasticity correction, which improves the relative efficiency of the coefficient estimates.

Data from an 820-dwelling neighborhood in Atlanta, Georgia, well situated away from other residential areas by a major interstate highway on the east, a new 137-acre mixed use development on the north, a commercial corridor to the west, and Georgia Institute of Technology on the south, is used to estimate the models. The data include survey questions administered to Home Park (near the Georgia Tech campus) neighborhood residents, publicly available Multiple Listing Service (MLS) data, and geographical information system (GIS) data.

The rest of the paper follows. After a brief review of the hedonic pricing method, a theoretical model of aggregation is proposed that shows how spatial phenomena impact the development of ''submarkets'' of households. Then, the alternative estimation strategy is discusses. Next, the results of the ''nearest neighbors'' model are reported, alongside the results from other models in the hedonics literature. The paper closes with concluding remarks.

The Hedonic Pricing Method

In the early 1900s, research that considered the sales price of a good as a function of its characteristics was novel. This technique, the hedonic pricing method, treats the price of a good as the dependent variable and the characteristics of that good as the independent variables. Using regression equations, slope coefficients (which are interpreted as unobserved, implicit prices) on each independent variable are estimated. This method is commonly used to model the demand for goods (in a second-stage analysis) that do not have traditional economic markets (e.g., air quality or water clarity) using the prices of goods determined in other markets (e.g., the price of housing). Automobiles, water quality, trees, and houses are only a few of the goods that can be analyzed using this method of valuation.

Debates in the literature on the first use of this technique are common. Colwell and Dilmore (1999) suggest that an overlooked monograph by G. C. Haas of the Division of Agricultural Economics at the University of Minnesota Agricultural Experiment Station in 1922 was the first hedonic study. This conclusion is different from that of Goodman (1998) who argues that Andrew T. Court conducted the first hedonic study on automobiles in 1939. Regardless of its exact origins, Lancaster (1966) and Rosen (1974), who developed early theories of consumer behavior, popularized the hedonic method as a useful valuation exercise.

The usefulness of the hedonic pricing method is to estimate the demand for a good that does not have a traditional economic market. In the case of housing, which is a heterogeneous bundled good, researchers can use the hedonic pricing method to determine the demand for housing characteristics, such as square footage of living space, acreage, proximity to green space, etc. Houses are *listed* at a price determined by the owner and based on the sellers' characteristics, the dwelling's characteristics, the dwelling's amenities, the actual selling price of similar houses in the same neighborhood, the types of land-use designations surrounding the particular dwelling for sale, among other factors. But, the *actual* sales price of the dwelling is determined by a negotiation process that ultimately equates the purchaser's willingness to pay for the characteristics and amenities with the seller's willingness to accept a certain level of compensation. This produces a buying/selling decision in the housing *market*. For a dwelling, one does not directly observe a separate price for a fireplace, for a bay window, or for a bedroom on the main floor of the dwelling. Regression analysis, when conducted properly, can reveal a robust estimate of the marginal implicit price for a dwelling's structure characteristics and certain neighborhood features that can be used by urban planners, policy researchers, and economists. However, one must make sure that the hedonic theory has a very strong empirical translation.

A Theoretical Model of Aggregation

It is easy to see that a household's choice of residence simultaneously determines the distances that it must travel in order to access all goods, including private consumption (eggs, milk, apples, etc.) and non-private consumption (green space, national parks, etc.).

Then, depending on the mode of transportation chosen by each household to access these goods, the costs of traveling to the sites where these goods can be purchased (or consumed) might be different for each household. That the household location decision simultaneously determines the distances and costs of accessing private and non-private goods suggests that different households in different geographical areas pay different prices for the same types of goods. But, in the single neighborhood modeled here, it is reasonable that (on a more macro scale) each household approximately travels the same distances to access groceries, entertainment, work, etc. However, the distances each household must travel to certain *within* neighborhood features (like the Home Park green space or to the Georgia Tech campus) vary greatly, particularly when one considers the various modes of transportation available to local residents (students tend to walk to school, whereas full-time employees tend to drive to work) and the time required to travel those distances. This suggests that households will react differently (i.e., have different marginal price reactions) to structural changes in the local land use (altering the land use designations of particular parcels), changes to the transportation infrastructure, and others. But, if these marginal price reactions among households are similar enough in many dimensions, then it is possible that each household can be aggregated into groups with other households that have similar marginal price reactions. In fact, this kind of household aggregation based on marginal price reactions by Lipscomb and Farmer (2005) is used as a springboard in this work to propose an alternative hedonic estimation technique.

A household utility maximization model (with a Cobb-Douglas utility specification) similar to that described in Farmer (2003) in which the arguments of the utility function are assumed to be weakly separable (the marginal rate of substitution between two arguments is independent of the quantities of all other goods) is used to describe household aggregation.[2]

$$MaxU_i\,(\mathbf{x}_c, \mathbf{A}_j, s_H, L) = \alpha_i \ln \mathbf{x}_c + \beta \ln \mathbf{A}_j + \gamma_i \ln s_H + \delta_i \ln L$$

$$st$$

$$\mathbf{w}_{sl}\mathbf{I}_{sl} + \mathbf{w}_{usl}\mathbf{I}_{usl} + S_i = \mathbf{p}_c\,\mathbf{x}_{ci} + R_H\,[s_{Hi}, \mathbf{c}(\mathbf{x}_{ci}, \mathbf{A}_{ji}, I_J), \mathbf{d}(\mathbf{x}_{ci}, \mathbf{A}_{ji}, I_J)]$$

$$T = L_i + \sum_{J=(sl,usl)} I_J + \sum_{J=(sl,usl)} t_J d_J + \sum_{x=1}^{n} t_{\mathbf{x}} d_{\mathbf{x}i} * \mathbf{x}_{ci} + \sum_{A=1}^{m} t_{\mathbf{A}} d_{\mathbf{A}i} * \mathbf{A}_{ji}$$

$$+ \sum_{M=1}^{q} t_{Mi} d_{Mi} * f(\mathbf{x}_{ci}, \mathbf{A}_{ji}, I_J). \tag{1}$$

This theoretical model says that households maximize utility (satisfaction and happiness) through the consumption of four items: (1) consumption other than housing space and neighborhood amenities [$\mathbf{x}_c$]; (2) cultural and recreational amenities located outside the neighborhood that are not capitalized into the local housing market [$\mathbf{A}_j$]; (3) housing space [s_H] from which households receive services (including the monetary impacts of local amenities, such as sidewalk quality and proximity to open space); and (4) leisure time [L].[3]

The two main constraints in the model are income (endogenously determined in the model even though skill level is exogenously determined) and time (exogenous to the model). In the income constraint, assume that skilled-worker households (denoted by the ''sl'' subscript) and unskilled-worker households (denoted by the ''usl'' subscript) work different amounts of time [$\mathbf{I} = (I_{sl}, I_{usl})$] for different wage rates [$\mathbf{w} = (w_{sl}, w_{usl})$]. This total household income [$\mathbf{wI}$] is used to purchase a certain amount of housing space, amenities, consumables, and travel-related goods (car, gasoline, bus tickets, etc.). In the case of retired persons and students, who may not earn income in the same way as other households, S_i denotes saved monies (savings from work done in previous periods, college loans, Social Security, other endowments, etc.) used to purchase consumable goods, purchase housing space, travel to various neighborhood amenities, shopping opportunities, etc., and purchase neighborhood amenities. On the right-hand side of the income constraint, households pay a price $\mathbf{p_c}$ for goods consumed $\mathbf{x_{ci}}$, which include non-local, out-of-neighborhood amenities for which households pay user or entrance fees. They also pay a rental price R_H for a dwelling which is a function of square feet of housing space s_H, a vector of costs $\mathbf{c}$ associated with traveling distances $\mathbf{d}$ to consume goods, to consume within-neighborhood amenities $\mathbf{A_{ji}}$, and to work; R_H varies depending on the intensity with which each $\mathbf{A_{ji}}$ is consumed or ''accessed.''

In the time constraint, each household has only 24 hours a day in which to consume leisure L_i; to work $\sum_{J=(sl,usl)} I_J$; to commute between work and home $\sum_{J=(sl,usl)} t_J d_J$; to purchase goods $\sum_{x=1}^{n} t_{\mathbf{x}} d_{\mathbf{x}i} * \mathbf{x}_{ci}$; to consume amenities $\sum_{A=1}^{m} t_{\mathbf{A}} d_{\mathbf{A}i} * \mathbf{A}_{ji}$; and to make multiple purpose trips $\sum_{M=1}^{q} t_{Mi} d_{Mi} * f(\mathbf{x}_{ci}, \mathbf{A}_{ji}, I_J)$. $\mathbf{t}$ contains the times required to travel distances $\mathbf{d}$ at costs of $\mathbf{c}$ per unit of distance.[4]

Next, the Lagrangian formulation of the utility maximization model is differentiated with respect to each argument in the utility function, as well as the shadow prices to yield the first-order conditions. Then, a rearrangement of the first-order conditions yields the Marshallian demands, which are used to derive the general indirect utility function (IUF).[5]

$$V\,(\mathbf{p}_c, R_H\,[c_J, c_{\mathbf{x}}, c_{\mathbf{A}}, c_{Mi}, t_J, t_{\mathbf{x}}, t_{\mathbf{A}}, t_{Mi}, d_J, d_{\mathbf{x}}, d_{\mathbf{A}}, d_{Mi}], \mathbf{wI}; T, S). \tag{2}$$

Equation 2 suggests that 1) utility is not aspatial and that 2) the purchase or rental of a dwelling reflects the simultaneous determination of a household's driving/walking times to work, to shopping, to recreation, to entertainment, to healthcare services, and to others; noise from the nearest major road; and others. In other words, dwellings sales or rental prices reflect the many distances and therefore times and costs associated with traveling those distances that are ''chosen'' by the household location decision.[6] So, when the researcher attempts to model dwelling sales price as a function of a series of characteristics, it is important to account for determinants that vary spatially, as well.

Equation 2, looked at another way, also suggests that submarkets are important. If a general indirect utility function can be partitioned, then the empirical translation of that partition idea might suggest that multiple housing submarkets can exist in a single

neighborhood; and that the analysis of this single neighborhood using single-line OLS regression may not be the exact empirical translation of the theory suggested. If utility partitions reflect differences in households' demographic and attitudinal characteristics, as well as travel times, distances, and their associated costs, then where a household lives sets in motion all sorts of constraints on retail purchases, on recreation, and on available groceries. Even income is a constraint that tends to tip the scale in favor of Tiebout-like (1956) similarity where residents, choosing from an array of similar overlapping constraints, tend to cluster vis-à-vis more extreme Samuelson-like (1954) markets where residents have very little in common beyond a few local amenities. *This is the link between the utility discussion and the aggregation optimist/pessimist discussion*. But these households, while they may have relatively similar incomes, cannot be distinguished on a single dimension; this is the mistake made and realized by Rosen (1974). Therefore, to model a neighborhood that questions the common ''one neighborhood, one submarket'' assumption, households must be distinguished into submarkets to segment the housing market. This paper takes the housing submarkets determined by Lipscomb and Farmer (2005) as given in order to test for any *additional* spatial dependencies that may not be captured with spatial independent variables in a regression framework.

Alternative Hedonic Regression Models

Baseline Hedonic Estimates (SUR and OLS)

The variables available for this particular neighborhood can be found in Exhibit 1. Using the household submarkets identified by Lipscomb and Farmer (2005), each submarket is treated as a separate line in seemingly unrelated regression (SUR) estimation that follows Zellner (1962). *Notice that the way in which submarkets are identified in this research allows the possibility for households of a certain submarket to be located in the parcel adjacent to a household of a different submarket*.[7] Exhibit 2 shows the spatial distribution of the three household submarkets.

As described in Lipscomb and Farmer (2005), the two steps in the iterative process are the identification of ''groups'' through principal components analysis (PCA) and ''submarkets'' through an iterative hedonic regression model. First, 24 candidate variables (demographic and attitudinal variables in Exhibit 1 that account for taste differences between households and allow the threshold that defines one potential group from another to lie anywhere along a single surface in R_+^{24} space) are used in a PCA, producing 24 factors, eight of which have eigenvalues greater than one. Then, households are matched to the factors that best describe them based on the minimization of $G_i = |(\ln \hat{f}_{in} - \ln \overline{f}_i)/\hat{\sigma}_i|$, where $\hat{f}_{in}$ (i = 1 to 8; n = 1 to 400) is the factor score for each observation, $\overline{f}_i$ is the mean of each factor i, and $\hat{\sigma}_i$ is the standard deviation. Now that observations have been aggregated into eight ''groups,'' each group is treated as a separate line in an eight equation Zellner-like (1962) SUR model that regresses the natural log of sales price against the structural characteristics of the dwellings (**S**), distances to various amenities such as open spaces (**D**), and adjacency variables **A**. The source data for this iterative PCA and hedonic regression model (see Exhibit 1) include publicly available data (dwelling structure characteristics), a housing survey administered by the

Exhibit 1. Variable List

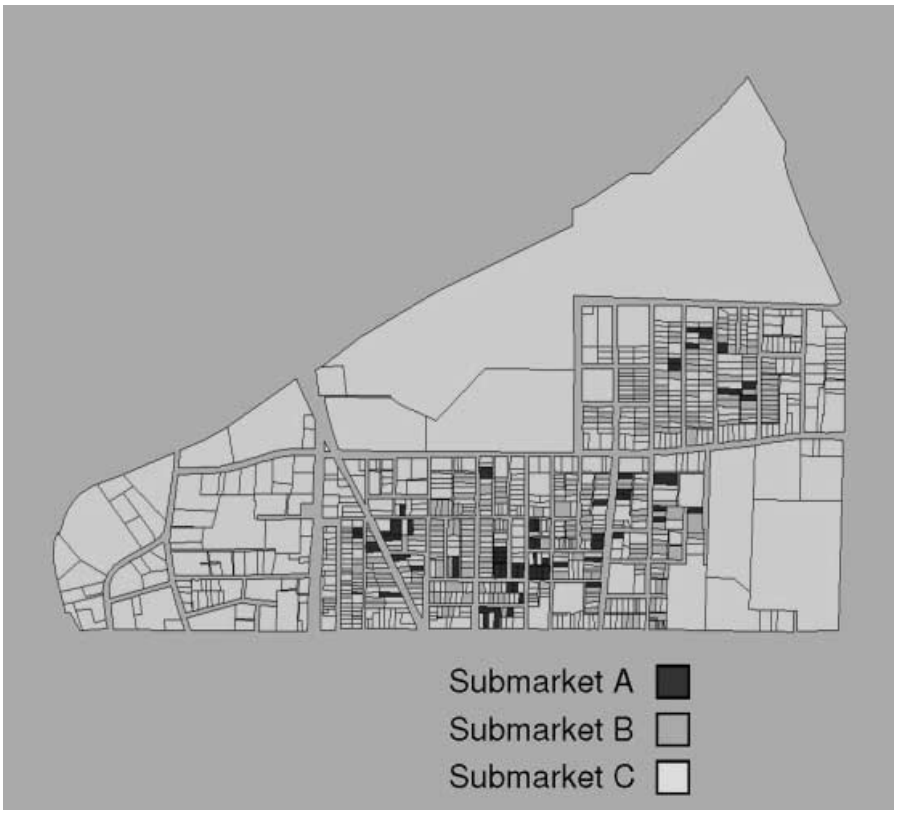
Demographic and Attitudinal Variables (from Housing Survey)
Are you a renter or owner?; dichotomous (0 = renter, 1 = owner)
Years lived at current residence; interval
Total household income (IV); multinomial
Race of the respondent (IV); multinomial
Number of adults in dwelling (IV); interval
Number of adults that work at least part-time; interval
Number of children in the household (IV); interval
Education level of the respondent (IV); multinomial
Age of the survey respondent (IV); continuous
Work status of survey respondent [retired (0,1), Student (0,1), Full-time employed (0,1)]
Sex of the survey respondent (IV); dichotomous (0 = female, 1 = male)
Is the house owner-occupied? (IV); dichotomous (0 = no, 1 = yes)
Is some part of the house rented? (IV); dichotomous (0 = no, 1 = yes)
What particular features attracted you to Home Park?; Close to Georgia Tech, Investment Potential, Tree Cover
What particular features attracted you to your dwelling?; Near Home Park green space, Near public transportation, Near work
Political tendency (conservative, moderate, liberal); dichotomous
Most important issue besides national security (Crime, Environment, Education, Social Security)

Dwelling Characteristics (from Housing Survey and Multiple Listing Service) used in iterative SUR process
Rental/Selling price of house (DV); continuous in dollars; rental multiplier of 120 used to convert rental prices to sales prices
Square footage of the house (IV); continuous in square feet
Number of bedrooms (IV); discrete
Number of baths (IV); discrete
Year of last sale (IV); discrete
Age of the dwelling (IV); continuous in years
Is the dwelling at or above street level? (IV); dichotomous

Spatial Features (from GIS) used in iterative SUR process
Road network distance to nearest brown industry (IV); continuous in miles
Road network distance to child care facility (IV); continuous in miles
Road network distance to local churches (IV); continuous in miles
Road network distance to most distant parcel from Home Park (IV); continuous in miles
Road network distance to Piedmont Park (IV); continuous in miles
Road network distance to local elementary school (IV); continuous in miles
Road network distance to local convenience and ethnic grocery stores (IV); continuous in miles
Road network distance to 14th Street commercial/retail opportunities (IV); continuous in miles
Road network distance to nearest public transportation bus stop (IV); continuous in miles
Road network distance to Georgia Tech (IV); continuous in miles
Road network distance to Muslim school (IV); continuous in miles
Distance interaction terms ($D<200$ = ln distance to Home Park green space if dwelling is within 200 meters of the park, 0 otherwise; $D>200$ = ln distance to Home Park if dwelling is more than 200 meters from the park and within South Home Park [south of 14th Street], 0 otherwise; DNorth = ln distance to Home Park if dwelling is located in North Home Park [north of 14th Street], 0 otherwise)

Adjacency Variables (Yes or No dummy variables created from GIS) used in iterative SUR process
Do you live adjacent to a renter?
Do you live adjacent to a homeowner?
Do you live adjacent to a student?
Do you live adjacent to a college graduate?
Do you live adjacent to a household that wishes (to move/not to move) in the next two years?
Do you live adjacent to a household of a different race than your own?
Do you live adjacent to a household that has made home improvements in the last two years?
Do you live adjacent to a household that chose Home Park for its proximity to Georgia Tech?
Do you live adjacent to a household that is of the same/different submarket than you?

Exhibit 2. Spatial Distribution of Household Submarkets



authors (demographic and attitudinal variables, adjacency variables), and geographic information system analysis (distance variables).

This hedonic estimation strategy essentially models the competition between households for space in this market. The SUR estimator is $\hat{\beta}^*(\hat{\Sigma}^*) = (\mathbf{X}^*\hat{\Sigma}^{*-1}\mathbf{X}^*)^{-1}\mathbf{X}^*\hat{\Sigma}^{*-1}\mathbf{y}^*$, where $\hat{\beta}^*$ is the estimated coefficient vector, $\hat{\Sigma}^*$ is the estimated error variance-covariance matrix, $\mathbf{X}^*$ is the matrix of independent variables, and $\mathbf{y}^*$ is the vector of observations on the dependent variable. Baseline SUR estimates for each submarket can be found in Exhibit 3.

Generalized least squares (GLS) estimation proceeds as the error variance-covariance matrix Σ is estimated using OLS and then used as a weight to obtain the coefficient vector $\boldsymbol{\beta}^*$. Then, to determine whether the SUR estimator realizes more efficiency gains than equation-by-equation OLS, the off-diagonal elements of the error variance-covariance matrix are tested for their statistical difference from zero with a Breusch-Pagan (1980) Lagrangian multiplier test which, if significant, indicates that the off-diagonal elements are non-zero and that the SUR estimator is more efficient than OLS. Without a lot of information about the interrelationships between submarkets, SUR has the advantage of correcting for an unmodeled relationship between equation estimates, which makes it the more efficient estimator of the two. In fact, as the correlations between the

Exhibit 3. Iterative SUR Estimates

	Submarket A	Submarket B	Submarket C
Constant	4.78	5.50	2.19
	(9.90)	(11.04)	(9.60)
Structural Variables			
Natural log of Dwelling Square Feet	−1.28**	0.56	0.92
	(0.65)	(0.77)	(0.72)
Natural log of Number of Beds	3.60***	−4.93***	1.40*
	(0.68)	(0.81)	(0.76)
Natural log of Number of Baths	−1.18*	−1.17	2.45***
	(0.69)	(0.81)	(0.77)
Natural log of Number of Acres	−0.19	−0.64	0.72*
	(0.34)	(0.40)	(0.38)
Condominium (yes = 1)			−0.17*
			(0.09)
Spatial Variables			
Ln Dist away from 14th St. Commercial Site	−0.40	1.60	−1.78
	(1.48)	(1.63)	(1.44)
Ln Dist away from Piedmont Park	0.44	−1.85***	1.37***
	(0.37)	(0.43)	(0.41)
Ln Dist away from Meineke Service Station	1.85**	−3.97***	2.17***
	(0.74)	(0.86)	(0.79)
Ln Dist away from Georgia Tech	−1.76***	4.00***	−2.52***
	(0.68)	(0.78)	(0.73)
D<200	0.50***		
	(0.10)		
D>200m	0.11	0.40***	−0.02
	(0.10)	(0.11)	(0.09)
DNorth		0.50***	
		(0.10)	
Is Dwelling Above Street Level?	−1.14**	−0.57	1.84***
	(0.45)	(0.53)	(0.50)
Located on State Street?	1.22	−6.23***	5.00***
	(0.96)	(1.13)	(1.07)
Above Street Level * State Street	−0.40	5.42***	−5.27***
	(1.40)	(1.66)	(1.55)
Adjacency Variables			
Lives Adjacent to a Renter	1.74***	−2.29***	0.24
	(0.67)	(0.79)	(0.74)
Lives Adjacent to Undergrad Student	1.25**	1.39**	−2.65***
	(0.56)	(0.67)	(0.63)
Lives Adjacent to a College Educated Person	−0.19	−0.82	1.17*
	(0.63)	(0.75)	(0.71)
Lives Adjacent to a Person of a Different Race Than Them	−1.92***	−2.69***	4.58***
	(0.47)	(0.55)	(0.52)
Lives Adjacent to Someone that Made Home Improvements	0.69	−3.98***	3.61***
	(0.48)	(0.57)	(0.53)
Lives Adjacent to Someone who Lives in Home Park Because it is Near Georgia Tech	1.47***	2.61***	−4.13***
	(0.57)	(0.67)	(0.63)
Root MSE	3.88	4.59	4.34
R^2	.23	.38	.43

Exhibit 3. Iterative SUR Estimates (continued)

Notes: The dependent variable is the Natural Log of Sales Price. Standard errors are in parentheses beneath the coefficients. The Breusch-Pagan test of independence of all submarkets is χ^2 (3) = 305.558, prob. = .0000. The χ^2 for Submarket A is 112.87; for Submarket B is 245.56; and for Submarket C is 312.12; p = .0000 for these submarket values.
* Significant at the .10 level.
** Significant at the .05 level.
*** Significant at the .01 level.

independent variables get stronger, the SUR estimator provides even more efficient estimates than equation-by-equation OLS. To see the differences, Exhibit 4 shows the less-efficient OLS estimates.

Spatial Dependence Checks and an Alternative Spatial Estimator

At this point, most researchers might cease the search for any additional spatial dependency remaining in the model. To search for additional spatial dependence, error terms from the SUR model are incorporated into a structural model (simultaneous equations model) to determine the appropriate weights that can be used in a White-style (1980) heteroscedasticity correction of the SUR model estimated previously.[8] This is accomplished by following Guilkey and Schmidt (1973), who devise an efficient estimator for SUR models with vector autoregressive errors that use a transformation matrix **R** to weight both sides of the regression equation. While the original article was intended for time-series data, their method is translatable to the cross-section data context where one wishes to model the simultaneous dependencies of error terms in a type of *spatial* conditional heteroscedasticity error approach.[9] The goal of using this method is to correct for heteroscedasticity in the error terms, not household heterogeneity, *per se*.

In this structural model, the hypothesis is that at least one neighbor's residuals (residuals of the nearest Submarket A household, residuals of the nearest Submarket B household, or residual of the nearest Submarket C household) will be a significant predictor of the residuals of a dwelling of interest, no matter in which submarket it has been classified. Coefficient significance will indicate that the residuals of ''neighbors,'' regardless of submarket, are indeed correlated, suggesting that additional spatial dependence can be captured using the structural model. Then, the incorporation of this additional information into the transformation matrix **R** will result in more efficient estimates of the marginal prices for each household submarket.[10]

To test the null hypothesis that no additional spatial dependence is present in the residuals from the basic SUR model, we model three simultaneous equations in the style of Hendry (1971) and Guilkey and Schmidt (1973):

$$\varepsilon_{1i}^2 = \alpha_0 + \alpha_1{}^*\varepsilon_{1i,ADJ}^2 + \alpha_2{}^*\varepsilon_{2i,ADJ}^2 + \alpha_3{}^*\varepsilon_{3i,ADJ}^2 + \boldsymbol{\alpha_4}{}^*\mathbf{Z} + \mu_{1i}. \tag{3a}$$

$$\varepsilon_{2i}^2 = \alpha_5 + \alpha_6{}^*\varepsilon_{1i,ADJ}^2 + \alpha_7{}^*\varepsilon_{2i,ADJ}^2 + \alpha_8{}^*\varepsilon_{3i,ADJ}^2 + \alpha_9{}^*\mathbf{Z} + \mu_{2i}. \tag{3b}$$

$$\varepsilon_{3i}^2 = \alpha_{10} + \alpha_{11}{}^*\varepsilon_{1i,ADJ}^2 + \alpha_{12}{}^*\varepsilon_{2i,ADJ}^2 + \alpha_{13}{}^*\varepsilon_{3i,ADJ}^2 + \boldsymbol{\alpha_{14}}{}^*\mathbf{Z} + \mu_{3i}. \tag{3c}$$

Exhibit 4. Pooled and Submarket-Specific OLS Models

	Pooled OLS	Submarket A	Submarket B	Submarket C
Constant	8.35***	5.30	8.65***	12.33***
	(1.46)	(5.97)	(1.46)	(1.53)
Structural Variables				
Natural log of Dwelling Square Feet	0.24***	0.21	0.39***	−0.07
	(0.06)	(0.18)	(0.08)	(0.11)
Natural log of Number of Beds	0.09	0.17	0.11	−0.03
	(0.06)	(0.19)	(0.09)	(0.13)
Natural log of Number of Baths	0.08	−0.17	−0.05	0.13
	(0.06)	(0.18)	(0.10)	(0.11)
Natural log of Number of Acres	−0.12***	0.28*	−0.23***	−0.08
	(0.03)	(0.16)	(0.04)	(0.05)
Condominium	−0.16*			−0.37***
	(0.09)			(0.13)
Spatial Variables				
Ln Dist to 14th St. Commercial Site	0.33	1.42	0.03	0.32*
	(0.27)	(0.97)	(0.23)	(0.18)
Ln Dist to Piedmont Park	−0.02	−0.44	−0.01	−0.29
	(0.03)	(0.28)	(0.03)	(0.18)
Ln Dist to Meineke Service Station	−0.14	−0.79	−0.05	−0.48**
	(0.09)	(0.50)	(0.10)	(0.19)
Ln Dist to Georgia Tech	0.17	0.49	0.03	0.49**
	(0.13)	(0.60)	(0.08)	(0.20)
*Ln Dist to Home Park Green Space*HP <200m*	−0.13	0.01		
	(0.16)	(0.10)		
*Ln Dist to Home Park Green Space*HP >200m*South Home Park*	−0.14	−0.02	−0.02	−0.02*
	(0.16)	(0.09)	(0.03)	(0.01)
*Ln Dist to Home Park Green Space*North Home Park*	−0.11		−0.01	
	(0.15)		(0.03)	
Is Dwelling Above Street Level?	0.11**	0.13	0.07	0.04
	(0.04)	(0.14)	(0.05)	(0.07)
Located on State Street?	0.04	0.25	−0.01	−0.43**
	(0.09)	(0.24)	(0.15)	(0.18)
*Above Street Level*State Street*	−0.22	−0.69**	−0.28	0.30
	(0.13)	(0.32)	(0.23)	(0.24)
Adjacency Variables				
Are you Adjacent to a Renter?	−0.31***	0.19	−0.37***	−0.26**
	(0.06)	(0.34)	(0.07)	(0.12)
Are you Adjacent to Undergrad Student?	−0.02	0.12	0.004	−0.14
	(0.05)	(0.13)	(0.07)	(0.15)
Are you Adjacent to a College Educated person?	0.14**	0.01	0.13*	−0.15
	(0.06)	(0.16)	(0.07)	(0.17)
Are you Adjacent to a Person of a Different Race Than You?	−0.04	−0.02	−0.12*	−0.15*
	(0.04)	(0.16)	(0.06)	(0.17)
Are you Adjacent to a Dwelling that Made Home Improvements?	0.31***	0.38**	0.28***	0.36***
	(0.04)	(0.16)	(0.07)	(0.07)
Are you Adjacent to Someone who Lives in Home Park Because it is Near Georgia Tech?	−0.05	−0.33	0.06	−0.000
	(0.05)	(0.24)	(0.08)	(0.08)

Exhibit 4. Pooled and Submarket-Specific OLS Models (continued)

	Pooled OLS	Submarket A	Submarket B	Submarket C
F	9.32[a]	1.21[b]	7.12[a]	5.33[a]
Root MSE	.37	.42	.33	.34
Adjusted R^2	.30	.05	.38	.36

Notes: The dependent variable is the Natural Log of Sales Price. Standard errors are in parentheses beneath the coefficients. Pooled OLS = 400 observations; Submarket A = 70 observations; Submarket B = 184 observations; Submarket C = 146 observations.
[a] $p = .0000$
[b] $p = .2900$
* Significant at the .10 level.
** Significant at the .05 level.
*** Significant at the .01 level.

In this model, which maintains the submarkets described earlier (e.g., the residual equation for Segment A is 3a, etc.), the independent variable matrix **Z** includes spatial and attitudinal variables that are hypothesized to explain the squared residuals from the basic SUR model described in Exhibit 3. These six variables are the natural log of distance to the Muslim school (MUSLIM); the natural log of the distance to Home Park (HP); yes/no—Are you politically conservative?; yes/no—Did you choose to live in Home Park because it is near public transportation?; yes/no—Did you purchase/rent your particular dwelling because of its windows?; and yes/no—Did you purchase/rent your particular dwelling because of its proximity to entertainment opportunities? Of the possible explanatory variables listed in Exhibit 1, this subset was determined through a series of F tests where each spatial and attitudinal variable in the dataset was tested for joint significance across the three equations at the .10 level. Also, $\varepsilon^2_{1i}. \; . \; .\varepsilon^2_{3i}$ are the squared residuals from the basic SUR model, corresponding to Submarkets A, B, and C, respectively; $\varepsilon^2_{1i,ADJ}. \; . \; .\varepsilon_{3i,ADJ}$ are the squared residuals of the closest neighboring households of Submarkets A, B, and C, respectively, from the basic SUR model;[11] and $\mu_{1i}. \; . \; .\mu_{3i}$ are independently and identically distributed error terms with no assumed correlation between them. The idea in Equations 3a–3c is that the squared residuals (dependent variables) can be explained by the squared residuals of the neighboring households, as well as certain spatial variables.[12] As an example, suppose that a Submarket A household lives in an area surrounded by households classified as other submarkets. The first line in Equation 3 says that the unexplained sales price variation of this Submarket A household is dependent on the unexplained sales price variation of the adjacent households of all submarkets, including other Submarket A households, and on other variables (**Z**) that are hypothesized to influence the residuals.

The main concern with the structural model is equation identification, to make sure that order and rank conditions are satisfied. The literature on the identification problem concludes that the order condition is the first concern econometrically. Harvey (1990; p. 328) says that ''Fortunately, the order condition is usually sufficient to ensure identifiability, and although it is important to be aware of the rank condition, a failure to verify it will rarely result in disaster.'' So, the fact that the model depicted in Equation 3 is exactly identified in Equations 3a–c satisfies the order condition.

Since 3SLS and SUR provide identical estimates in this case, Equation 3 is treated as seemingly unrelated. The Breusch-Pagan test statistic of 304 (probability = .0000) tells us that the error term variance-covariance matrix is not diagonal (i.e., there are non-zero elements in the off-diagonal elements), which indicates significant dependence among the error terms in this regression; indeed the ''residuals of the residuals'' are dependent upon each other.

In Exhibit 5, the independent variables explain 13%, 3%, and 6% of the variation in the squared residuals for Submarket A, B, and C households, respectively. The basis for choosing this functional form was to increase the explanatory power of the model for Submarket A households, as it had the weakest explanatory power of the three submarkets in Exhibits 3 and 4. It can be seen that the independent variables significantly explain variation in the residuals of all three submarkets even though the focus was on

Exhibit 5. Structural Model Residual Analysis

	Dependent Variables		
	Submarket A Residuals^2	**Submarket B Residuals^2**	**Submarket C Residuals^2**
Constant	46.21***	10.12	31.24***
	(11.61)	(12.91)	(10.91)
Independent Variables			
Residuals^2 of Nearest Submarket A	0.35***	0.02	0.06
	(0.07)	(0.11)	(0.09)
Ln Dist to Home Park	−5.62***	2.15	−1.72
	(1.86)	(2.07)	(1.75)
Conservative Politically?	5.56**	2.54	−2.30
	(2.77)	(3.08)	(2.60)
Near Public Transportation?	8.94**	−4.11	−1.74
	(3.65)	(4.06)	(3.43)
Windows?	−5.53*	−8.96**	−5.17*
	(3.16)	(3.51)	(2.96)
Near Entertainment Opportunities?	1.94	−4.82	−3.26
	(2.78)	(3.10)	(2.61)
χ^2	42.06	11.59	8.21
	(p = .00)	(p = .07)	(p = .22)
Root MSE	22.43	24.95	21.08
R^2	.09	.02	.02
Durbin-Wu-Hausman tests of endogeneity[a]	F (3,390) = 3.23	F (3,390) = 0.97	F (3,390) = 3.55
	(p = .02)	(p = .40)	(p = .01)

Notes: Standard errors are in parentheses beneath the coefficients.
[a]The squared residuals of the nearest Submarket B and C households were also included in preliminary runs of the structural model. Also, all three predicted residual variables jointly are significantly different from zero, meaning that an OLS regression of the SUR residuals on the adjacent residuals, etc., gives inconsistent estimates, meaning that either maximum likelihood or generalized moments estimation is required.
* Significant at the .10 level.
** Significant at the .05 level.
*** Significant at the .01 level.

Exhibit 6. Weighted SUR Estimates (double log specification)

	Submarket A	Submarket B	Submarket C
Constant	−0.01	−0.01	−0.06
	(0.02)	(0.02)	(0.04)
Structural Variables			
Natural log of Dwelling Square Feet	−0.12	1.49***	0.30
	(0.39)	(0.57)	(0.59)
Natural log of Number of Beds	5.60***	−6.19***	1.01
	(0.50)	(0.68)	(0.70)
Natural log of Number of Baths	−1.25**	−1.43**	2.27***
	(0.49)	(0.63)	(0.69)
Natural log of Number of Acres	−0.05	−0.87**	0.79**
	(0.23)	(0.35)	(0.33)
Condominium (yes = 1)			0.25
			(0.44)
Spatial Variables			
Ln Dist to 14th St. Commercial Site	−1.13**	2.46***	−1.01
	(0.50)	(0.89)	(0.84)
Ln Dist to Piedmont Park	0.50**	−2.60***	1.94***
	(0.23)	(0.47)	(0.41)
Ln Dist to Meineke Service Station	3.30***	−5.34***	1.95***
	(0.49)	(0.78)	(0.71)
Ln Dist to Georgia Tech	−3.20***	4.78***	−2.32***
	(0.47)	(0.71)	(0.68)
*Ln Dist to Home Park Green Space*HP <200m*	0.81***		
	(0.16)		
*Ln Dist to Home Park Green Space*HP >200m*South Home Park*	0.15**	0.74***	−0.09
	(0.06)	(0.09)	(0.07)
*Ln Dist to Home Park Green Space*North Home Park*		0.86***	
		(0.09)	
Is Dwelling Above Street Level?	−0.44	−0.74*	1.50***
	(0.28)	(0.42)	(0.45)
Located on State Street?	−0.42	−4.55***	4.57***
	(0.69)	(0.97)	(0.98)
*Above Street Level*State Street*	−0.37	5.69***	−5.22***
	(1.43)	(1.33)	(1.48)
Adjacency Variables			
Lives Adjacent to a Renter	1.12**	−2.04***	0.34
	(0.45)	(0.65)	(0.66)
Lives Adjacent to Undergrad Student	2.66***	0.01	−2.26***
	(0.38)	(0.58)	(0.56)
Lives Adjacent to a College Educated person	−0.69	−1.00*	1.53**
	(0.45)	(0.58)	(0.61)
Lives Adjacent to a Person of a Different Race Than You	−1.71***	−3.38***	5.16***
	(0.31)	(0.46)	(0.48)
Lives Adjacent to a Dwelling that Made Home Improvements	0.72**	−4.05***	3.63***
	(0.29)	(0.45)	(0.47)
Lives Adjacent to Someone who Lives in Home Park Because it is Near Georgia Tech	1.19***	3.62***	−4.90***
	(0.33)	(0.52)	(0.54)

Exhibit 6. Weighted SUR Estimates (double log specification) (continued)

	Submarket A	Submarket B	Submarket C
χ^2	446.11 (p = .00)	753.04 (p = .00)	534.81 (p = .00)
Root MSE	.30	.22	.23
Adjusted R^2	.32	.51	.50

Notes: Standard errors are in parentheses beneath the coefficients. Submarket A = 70 observations; Submarket B = 184 observations; Submarket C = 146 observations.
*Significant at the .10 level.
**Significant at the .05 level.
***Significant at the .01 level.

Submarket A residuals. So, in the spirit of Guilkey and Schmidt (1973), the square root of the predicted values of the dependent variables will be used as weights in the original SUR model as a White-style heteroscedasticity correction. So, the transformation matrix **R** takes the form $(\sqrt{\hat{\varepsilon}_{ji}^2})^{-1}$, where j = 1 to 3 and i = 1 to 400, and will be used to weight both sides of the SUR model as a GLS correction for heteroscedasticity. Essentially, this means that each dependent and independent variable in Exhibit 3 is divided by $\sqrt{\hat{\varepsilon}_{1i}^2}$ for the Submarket A regression line, by $\sqrt{\hat{\varepsilon}_{2i}^2}$ for the Submarket B regression line, and by $\sqrt{\hat{\varepsilon}_{3i}^2}$ for the Submarket C regression line.

The weighted SUR estimates are reported in Exhibit 6. Compared to Exhibit 3, the standard error estimates have decreased throughout the model, exactly as theory predicts. Also, no significant coefficient in Exhibit 3 maintains significance *and* changes signs in Exhibit 6, which shows the ability of the residual structural model to correct the standard errors (a second GLS correction that adds efficiency without affecting the consistency of estimates) but not change the directions of influence of the independent variables on sales prices. Additional evidence, such as the generally higher χ^2 statistics and lower Root Mean Square Errors, indicates the strength of the model and relative efficiency of the estimates. Particularly, the Root MSE statistics are lowest for the model represented in Exhibit 6, which adds weight to the argument for this kind of exhaustive search for additional spatial dependencies beyond those captured in traditional neighborhood-level hedonic models in the literature. From these results, one can see that non-spherical disturbances associated with spatial autocorrelation do not appear. And, only a modest form of heteroscedasticity appears, which suggests that the model summarized in Exhibit 6 is an efficient representation of the characteristics that describe sales prices in this neighborhood.

Conclusion

What does this alternative estimation strategy mean for urban economists, policy analysts, and regional scientists? First, it is data intensive, which suggests that a lot more survey work on households' attitudes be incorporated into housing hedonic studies. Second, it provides a way to handle non-spherical disturbances in a straightforward manner, which

is consistent with econometric theory. While the application of Guilkey and Schmidt (1973) to cross-sectional data may be somewhat provocative, it is a novel approach that provides an alternative solution to the old problem of how to model spatial heteroscedasticity in an effective way. In the end, this technique may not be necessary for all hedonic property value models. But, in this paper it is instructive as it shows relative efficiency gains vis-à-vis the traditional first-stage hedonic models.

Endnotes

[1] For clarity, the ''aggregation problem'' in this paper is the assumption that unique consumers (households in this case) with unique preferences can be grouped together for analysis purposes. This is different from the ''identification problem that dominates discussion in the applied [hedonics] literature'' described by Ekeland, Heckman, and Nesheim (2002, p. 304); the former arises because households' unique preferences for unique bundles of goods cannot be measured ''uniquely;'' the latter concerns the identification of the second-stage hedonic equation, which is not estimated estimate here.

[2] In this model, the implicit assumption is made that $\alpha_i = \beta_i = \gamma_i = \delta_i = 1$, which assumes increasing returns to scale.

[3] Anas, Arnott, and Small (1998; p. 1436), in their synthesis of the urban economics literature, argue that more realistic models of urban areas require the derivation of the shadow value of time endogenously by ''adding leisure and a time budget to the model.'' This is consistent with the model presented here.

[4] For simplicity, this paper assumes that trips to consume goods other than housing and/or amenities *on the way to or on the way from work* are negligible and will not be included in either constraint in the econometric model. Also, $\mathbf{c_i}$ vary because it is assumed that the monetary cost per unit of distance for short trips is greater than the cost per unit of distance for longer distance trips. An example is the cost per unit of distance for a single-purpose trip to the local supermarket during rush-hour traffic versus the cost for a multiple-purpose trip to the hair salon, grocery store, and the bank.

[5] The specification of the IUF is consistent with the specifications of other IUFs that include the prices of housing services and the prices of composite goods (Rapaport, 1997), income (Ibid., Haab and Hicks 1997; and Sieg, et al., 2000), household and community characteristics (Rapaport, 1997; and Chattopadhyay, 2000), and public and private goods (Rapaport, 1997; and Sieg, et al., 2000).

[6] Individuals who maximize utility will rearrange their purchases of a particular commodity or urban amenity until the marginal rate of substitution (or the slope of an indifference curve) between a composite commodity and each desired characteristic of a dwelling is equal to the implicit price of that characteristic (Hanley and Spash, 1993).

[7] The statistical algorithm in Lipscomb and Farmer (2005) does not force parcels in the same submarket to be contiguous.

[8] Generally, autocorrelation is a problem in time-series data while heteroscedasticity is a problem in cross-sectional data. For autocorrelation and heteroscedasticity corrections, Kmenta (1986; p. 296) suggests that the residuals and residuals squared, respectively, be used in the calculation of the weights to be employed in a weighted least squares approach similar to White (1980). However, a White test (as described in Gujarati) suggests that no pure heteroscedasticity exists in the model presented thus far. So, different forms of the residuals are fit before settling on the residuals squared as the appropriate dependent variables in the structural model.

[9] Other spatial lag variables were tested and rejected as insignificant predictors of the squared residual system. These variables included the nearest neighbor's living space, nearest

neighbor's distances to various neighborhood attributes, and others. Including these characteristics of nearest neighbors seems like a spatial weights matrix approach. The difference is that these variables are included as explanatory variables directly, whereas the use of this information in a spatial weights matrix would also include the multiplication of the matrix **W** times the error term in the spatial multiplier. The implications of this approach versus the spatial weights matrix approach will be the subject of future research.

[10] In matrix notation, the SUR (sandwich) estimator allowing for spatial corrections is $\hat{\beta} = [\mathbf{X}'\mathbf{R}'(\hat{\Sigma}^{-1} \otimes \mathbf{I})\mathbf{R}\mathbf{X}]^{-1}\mathbf{X}'\mathbf{R}'(\hat{\Sigma}^{-1} \otimes \mathbf{I})\mathbf{R}\mathbf{Y}$, where **R** is the transformation matrix and $(\hat{\Sigma}^{-1} \otimes \mathbf{I})$ is the SUR error variance covariance matrix.

[11] Most spatial econometrics models define contiguous parcels based on vertex adjacency, side adjacency, or both (called rook, bishop, and queen adjacency, respectively). Here, households assumed to be on either side (left or right) or across the street (straight or diagonally) are considered ''neighbors'' and that backyard neighbors have no influence on the households on other streets. Also note that some households' adjacent neighbors did not respond to the survey, meaning that the closest respondents classified into a particular submarket may be several dwellings away. However, only 9% of the observations in the data have nearest respondents classified into a particular submarket that are more than five dwellings away. While this might seem problematic in practice, these kinds of problems occur with survey non-responses.

[12] In preliminary runs of this model, an exactly identified six-equation model was used to account for the potentially endogenous land uses in Home Park: the Muslim day school, the 14th Street commercial strip the (14th Street), and the Home Park green space. These three variables were treated endogenously in the residual system (Equation 3) because they are not predetermined factors in the neighborhood. The Muslim day school was constructed in reaction to the high concentration of Muslim residents in the neighborhood, which then caused another influx of Muslim residents; the commercial strip was built as residents moved into the neighborhood and required nearby restaurants and shopping opportunities; and the Home Park green space was built as a place for local residents to walk their pets and enjoy recreational opportunities without having to travel several miles to other green spaces like Piedmont Park. A joint F test of the last three equations in this model showed that these land uses were not endogenous to the system and were omitted from the final model presented as Equation 3.

References

Anas, A., R.J. Arnott, and K.A. Small. Urban Spatial Structure. *Journal of Economic Literature*, 1998, 36, 1426–64.

Anselin, L. Under the Hood Issues in the Specification and Interpretation o Spatial Regression Models. *Agricultural Economics*, 2002, 27, 247–67.

Breusch, T.S. and A.R. Pagan. The Lagrangian Multiplier Test and its Applications to Model Specification in Econometrics. *Review of Economic Studies*, 1980, 47, 239–53.

Case, B., J. Clapp, R. Dubin, and M. Rodriguez. Modeling Spatial and Temporal House Price Patterns: A Comparison of Four Models. *Journal of Real Estate Finance and Economics*, 2004, 29, 167–91.

Chattopadhyay, S. The Effectiveness of McFadden's Nested Logit Model in Valuing Amenity Improvement. *Regional Science and Urban Economics*, 2000, 30, 23–43.

Colwell, P.F. and G. Dilmore. Who was First? An Examination of an Early Hedonic Study. *Land Economics*, 1999, 75, 620–26.

Ekeland, I., J.J. Heckman, and L. Nesheim. Identifying Hedonic Models. *American Economic Review*, 2002, 92, 304–09.

Farmer, M.C. Re-Investing in the Built Environment: A Spatial Network Approach. *European Planning Studies*, 2003, 11, 57–76.

Goodman, A.C., Andrew Court and the Invention of Hedonic Price Analysis. *Journal of Urban Economics*, 1998, 44, 291–98.

Guilkey, D.K. and P. Schmidt. Estimation of Seemingly Unrelated Regressions with Vector Autoregressive Errors. *Journal of the American Statistical Association*, 1993, 68, 642–47.

Haab, T.C. and R.L. Hicks. Accounting for Choice Set Endogeneity in Random Utility Models of Recreation Demand. *Journal of Environmental Economics and Management*, 1997, 34, 127–47.

Hanley, N. and C.L. Spash. *Cost-Benefit Analysis and the Environment*. Cheltenham: Edward Elgar, 1993.

Harvey, A. *The Econometric Analysis of Time Series*. Second Edition. Cambridge, MA: The MIT Press, 1990.

Hausman, J.A. Specification Tests in Econometrics. *Econometrica*, 1978, 46, 1251–71.

Hendry, D.F. Maximum Likelihood Estimation of Systems of Regression Equations with Errors Generated by a Vector Autoregressive Process. *International Economic Review*, 1971, 12, 257–72.

Kmenta, J. *Elements of Econometrics*. Second Edition. New York, NY: Macmillan Publishing Co., 1986.

Lancaster, K. A New Approach to Consumer Theory. *Journal of Political Economy*, 1966, 74, 132–57.

Lipscomb, C.A. and M.C. Farmer. Household Diversity and Market Segmentation Within a Single Neighborhood. *The Annals of Regional Science*, 2005, 39, 791–810.

Rapaport, C. Housing Demand and Community Choice: An Empirical Analysis. *Journal of Urban Economics*, 1997, 42, 243–60.

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

Samuelson, P.A. The Pure Theory of Public Expenditures. *Review of Economics and Statistics*, 1954, 36, 387–89.

Sieg, H., V.K. Smith, H.S. Banzhaf, and R. Walsh. Interjurisdictional Housing Prices in Locational Equilibrium. *Journal of Urban Economics*, 2002, 52, 131–53.

Tiebout, C. A Pure Theory of Local Expenditures. *Journal of Political Economy*, 1956, 64, 416–24.

White, H. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, 1980, 48, 817–38.

Zellner, A. An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association*, 1962, 57, 348–68.

This paper has benefited from a Faculty Research Development grant from Valdosta State University. The author also wishes to thank George Dell, Graeme Newell, and other participants at the 2006 American Real Estate Society meetings in Key West, FL, for their comments and advice. The author also thanks Elaine Worzala and the ARES manuscript awards committee.

Clifford A. Lipscomb, Valdosta State University, Valdosta, GA 31698-0075
email: calipscomb@valdosta.edu