Spatial Relationships in Rural Land Markets with Emphasis on a Flexible Weights Matrix

Patricia Soto, Lonnie Vandeveer, and Steve Henning

Department of Agricultural Economics and Agribusiness
Louisiana Agricultural Experiment Station
Louisiana State University Agricultural Center, Baton Rouge

Paper prepare in conjunction with
Paper session of the American Agricultural Economics Association
Long Beach, Florida, July 28-31, 2002

Spatial Relationships in Rural Land Markets with Emphasis on a Flexible Weights Matrix

The value of rural real estate is determined by a number of factors including its inherent productive capacity, location, accessibility, and alternative uses. Continued economic and population growth increases the need for land, which puts upward pressure on rural land market values. With more and more rural land acres being converted at the urban fringe, buyers, sellers, planners, appraisers, tax assessors, and others are expected to have an increasing need for information related to the effect of location and economic development on rural land values. Important questions relate to the magnitude of these influences and to the spatial extent of these influences in rural land markets. Generally, research aimed at identifying the effects of location and economic development on rural land market values is expected to provide improved information for both private and public decisions.

This study differs from other rural land market studies in that Geographical Information Systems (GIS) and spatial econometric procedures are used to model a rural land submarket in Louisiana. GIS procedures are necessary for determining the spatial component in the data and spatial econometric procedures are necessary for diagnostic tests for spatial autocorrelation and for estimating spatial econometric models. These procedures are important because modeling the real estate market in the presence of spatial autocorrelation using traditional OLS procedures may result in models with less than desirable statistical characteristics. Similarly, Pace et al. indicate that real estate and spatial statistics complement each other, and employing spatial estimators provide benefits over ignoring dependencies in the data. The benefits include improved prediction, better statistical inference through unbiased standard errors, and better estimates because of the way that location is handled within the modeling procedure.

Additionally, hedonic models are estimated using maximum likelihood procedures that are adjusted for autocorrelation using two types of weight matrices. The Delaunay, a rigid form of a weight matrix, is used to estimate the hedonic model. An alternative spatial matrix is a flexible form of the nearest neighbor asymmetric method as specified by Pace and Barry. To determine the nature of the spatial structure, this study includes a decay parameter *r* that lies between 0.4 and 1, along with different number of neighbors *m* ranging from 6 to 30 as suggested by Pace et al. Likelihood ratio tests are used to test for statistical fit between spatial and OLS models.

Model and Data

An empirical procedure that has been used to analyze rural land markets includes the hedonic pricing model. Rosen defined hedonic prices as implicit prices of attributes and notes that they are revealed to economic agents from observed prices of differentiated products and the specific amounts of characteristics associated with them. Prices of these characteristics are implicit because there is no direct market for them. Palmquist provided a discussion of the theoretical basis for using hedonic analysis in rural land value studies and Danielson, in the same year, used the procedure to empirically analyze the rural land market in North Carolina.

This study follows the approach used by Danielson. Value in a rural land submarket is specified by the following transcendental function:

$$\text{Price} = ß_0 Z_1^{ß_1} \exp\left[\sum_{i=1}^{m} a_i X_i + \sum_{j=1}^{n} ?_j D_j + e \right], \qquad (1)$$

where Price is the per acre price of land, Z_1 is the size of tract in acres, m is the number of additional continuous variables (X_i), n is the number of discrete (dummy) variables (D_j), and e is a random disturbance term. Taking the natural logarithm of both sides of equation (1) gives:

$$\ln \text{Price} = \ln \beta_0 + \beta_1 \ln Z_1 + \sum_{i=1}^{m} a_i X_i + \sum_{j=1}^{n} ?_j D_j + e. \tag{2}$$

Because the price of land is hypothesized to decline as the size of tract (Z_1) increases, but at a decreasing rate, nonlinearities were incorporated for Z_1. Therefore, β_1 is hypothesized to be negative. To conduct the hedonic model, equation (2) was fit with simultaneous autoregression (SAR) using Pace and Barry's spatial statistics toolbox 1.1 and MATLAB computer software. Likelihood ratio tests were conducted to select the model that better explains variation in prices.

Data for this study are based on rural land market sales for the southeast area of Louisiana that were collected using mail survey techniques. These data represent a subset of a larger data base collected for the state for the period January 1993 through June 1998. The rural land market survey was mailed to state certified appraisers, officers in commercial banks, Farmers Service Agency personnel, Federal Land Bank personnel, Production Credit personnel, members of the Louisiana Chapter of the American Society of Farm Managers and Rural Appraisers, and members of the Louisiana Realtors Land Institute.

The study area consists of eight parishes in southeast Louisiana. The primary general soils of the area are Southern Mississippi Valley Silty Uplands, Coastal Plain, and Gulf Coast Flatwoods. The area is characterized by rolling hills with pine tree, nursery crop, dairy farm, and other animal production activities. A total of 257 rural land sales were included in this study. Each rural land sale is ten acres or more in size, includes attachments to the surface such as buildings and other improvements, and lies outside major metropolitan areas in Louisiana.

Variables hypothesized to influence per acre rural land values are defined in Table 1. PRICE in Table 1 is the dependent variable used in the hedonic model and represents the per acre selling price for each tract of rural land and improvements. Continuous variables expected to have an inverse relationship with per acre selling price include size of tract (SIZE), actual distance to nearest town (ADNT), and travel time to nearest city (TTNC). There is generally a negative relationship between size of tract and per acre selling price because fewer buyers compete in markets for larger tracts; whereas, many buyers compete in markets for smaller tracts. For locational variables including travel time, location theory generally suggests an inverse relationship between distance to markets and per acre selling prices. TTNC and ADNT were computed using the Street Atlas USA computer software.

Table 1. Variables Used in Hedonic Model Estimation, Southeast Area, Louisiana Rural Land Market Survey, 1993- 1998.

Variable	Description	Expected Sign
Continuous Variables		
PRICE	Per acre price of land($)	(-)
SIZE	Size of the tract (acres)	(-)
VALUE	Value of improvements ($)	(+)
ROADFT	Road frontage	(+)
ADNT	Actual distance to nearest town (miles)	(-)
TTNC	Travel time to nearest city	(-)
TIME	Month of sale	(+)
Discrete Variables (1,0)		
RPREC	Reason for purchase: recreational	(-)
NORLMSA	New Orleans MSA	(+)

Continuous variables expected to positively influence rural land values include value of improvements (VALUE), road frontage (ROADFT), time of sale (TIME), and if the tract is located in the New Orleans metropolitan statistical area (NORLMSA). These variables represent positive attributes of rural land and hence are hypothesized to have a positive influence on per acre rural land values. The discrete recreational variable (RPREC) is hypothesized to have a negative relationship with per acre land values because much of the data in this analysis represent marginal marshland and upland well suited for hunting, trapping, and other outdoor uses.

OLS Estimation and Diagnostics

With spatial autocorrelation in the data, hedonic model estimation using OLS procedures could produce estimates that are not efficient. Inefficient estimates could produce misleading inferences from the model. Following Anselin, spatial autocorrelation is the situation where the dependent variable or error term at each location is correlated with observations for the dependent variable or error term at other locations. This means that for neighboring locations i and j:

$$E(y_i y_j) \ldots 0 \tag{3}$$

or

$$E(e_i e_j) \ldots 0 \tag{4}$$

where (3) is defined as a spatial lag situation (Anselin). When the dependent variable exhibits spatial autocorrelation, the simultaneous spatial autoregression estimator corrects the usual prediction off the dependent variable, $y = X\beta + e$, by a weighted average of the values on nearby observations, Dy. The spatial lag situation is specified by the following model:

$$y = \alpha Dy + X\beta + e \tag{5}$$

where:

y = vector dependent observations,
α = spatial autoregressive coefficient,
Dy = spatially lagged dependent variable,
X = matrix of explanatory variables,
β = vector of regression coefficients, and
e = vector of error terms.

D is an n by n weighting matrix with 0's on the diagonal and the rows of D sum to 1. In this spatial autoregressive model, if α is not equal to zero, then ordinary least square estimates will be biased and inefficient.

When spatial dependence occurs in the error, as defined in (6), a regression specification with a spatial autoregressive error term is used to develop model estimates. The spatial error model is:

$$y = X\beta + e \tag{6}$$

$$e = \alpha De + ? \tag{7}$$

where:

y = vector of dependent observations,
X = matrix of explanatory variables,
ß = vector of regression coefficients,
e = vector of error terms,
De= spatial lag for error terms,
α= autoregressive coefficient, and
? = error term iid.

Again, D is an n by n weighting matrix with 0's on the diagonal. In this spatial autoregressive model α is restricted to lie within the interval [0,1), and the errors ? are independently and normally distributed.

The weight matrix can be created based upon nearest neighbors and Delaunay triangle matrix. The Delaunay spatial weight matrix is a symmetric matrix that leads to a variance covariance matrix that depends upon only the autoregressive parameter α. On the other hand, the nearest neighbor variance covariance matrices depend upon three parameters α, the autoregressive parameter; m, the number of neighbors; and r, the rate weight decline with the order of the neighbors with the closest given by the highest weighting, the second closest given a lower weighting, and so forth. Using three parameters should make the nearest neighbor matrix more flexible for different applications.

Empirical results

Although space limitations do not allow presentation of the likelihood ratio (LR) tests performed to choose the decay parameter and the number of neighbors that better fit the ML spatial model, comparison of LR results indicate that more flexible form of the nearest neighbor weight matrix is obtained using a decay parameter of 0.5 along with 12 neighbors. These are the results presented in Table 2 for the ML spatial model using the nearest neighbor weight matrix.

Hedonic model coefficient estimates are presented in Table 2. Results indicate that hypothesized variables explain 54 percent of the variation in per acre rural land values. When using the OLS model procedures all the variables are statistically significant at the five percent level and all variables were estimated to have the correct expected sign. However, these results differ from those obtained using a ML spatial model. Specifically, the variable reason for purchase recreational (RPREC) was found to be statistically significant in the OLS model, while this was not the case for the ML model. Results suggest that using OLS model results could lead to incorrect conclusions regarding the effect of reason for purchase recreational on rural per acre land values.

Moreover, results also suggest that the SAR model based on the Delaunay matrix and nearest neighbor matrix performed better than the OLS model. Likelihood ratio tests indicate a statistically significant difference between the OLS model and the ML model estimated using both the Delaunay matrix and the nearest neighbor matrix. Significant difference between the OLS and the spatial ML models suggest the presence of spatial autocorrelation in the model.

Table 2. Estimated coefficients[a] for Hedonic OLS and ML models with Delaunay matrix and nearest neighbor matrix, southeast Louisiana rural land market area, 1993-1998.

Item	OLS model	Spatial weight matrix for ML model	
		Delaunay	Nearest Neighbor
Variable			
Ln SIZE	-0.2234	-0.2372	-0.2539
	(62.5121)***	(79.0148)***	(85.9905)***
TIME	0.0081	0.0093	0.0100
	(22.1438)***	(34.5698)***	(37.7033)***
NORLMSA	0.6563	0.7310	0.7255
	(53.1229)***	(24.0079)***	(26.5820)***
ROADFT	0.1776	0.1609	0.1458
	(10.7683)***	(10.8837)***	(9.1921)***
RPREC	-0.1687	-0.0923	-0.0959
	(3.5690)**	(1.3821)	(1.5296)
ADNT	-0.0100	-0.0145	-0.0130
	(10.9627)***	(11.1011)***	(8.8455)***
TTNC	-0.4336	-0.3631	-0.3602
	(15.9144)***	(5.1631)**	(4.7276)**
VALUE	0.1833	0.2128	0.2362
	(12.9911)***	(20.9486)***	(25.4714)***
INTERCEPT	8.6784	8.6913	8.7093
	(642.1644)***	(75.5003)***	(150.2954)***
LR test	-482.7292	-463.1123	-461.6046

[a] In the variables section, Likelihood Ratios are in parentheses, ***denotes statistical significance at the 0.01 level, ** denotes statistical significance at the 0.05 level, and * denotes statistical significance at the 0.10 level.

Summary and Conclusions

The general objective of this discussion was to demonstrate research procedures and modeling results when spatial autocorrelation is believed to exist within the data. Simultaneous autoregression (SAR) using Pace and Barry's spatial statistics toolbox 1.1 and MATLAB computer software were used to test for spatial autocorrelation within the rural land market in southeast Louisiana.

According to the results, hypothesized variables were used to explain over one-half of the variation in rural land values. For the analysis, using traditional OLS procedures would have led to the wrong conclusions regarding factors which influence rural land values. OLS results indicate that reason for purchase recreational (RPREC) was statistically significant in explaining per acre rural land values in the study area. However, when adjustments were made for spatial autocorrelation in the ML spatial models, RPREC was not significant. This suggests spatial autocorrelation in the data could have caused one to make erroneous conclusions concerning the effect of location on per acre land values in the study area.

Adjusting for spatial autocorrelation improves estimates. However, with this particular set of data, one cannot conclude that the flexible nearest neighbor matrix outperforms the use of a more rigid spatial weight matrix. Further research should continue to test for other forms of spatial weight matrices.

References

Anselin, Luc. “SpaceStat Version 1.80 User’s Guide.” Regional Research Institute, West Virginia State University, 1995. Danielson, Leon E. "Using Hedonic Pricing to Explain Farmland Prices." The Farm Real Estate Market, Proceedings of a Regional Workshop, Southern Natural Resource Economic Committee, p. 57-74, 1984.

Lesage, J. and R. K. Pace. "Using Matrix Exponentials to Explore Spatial Structure in Regression Relationships." Manuscript available at : http://www.spatial-statistics.com. 2002.

Pace, R. K., R. Barry, and C. F. Sirmans. “Spatial Statistics and Real Estate,” *Journal of Real Estate Finance and Economics,* 17(1998):5-13.

Pace, R. K. and R. Barry. Spatial Statistics Toolbox 1.1. Available: http://www.spatial-statistics.com. 2002.

Pace, R. K., R. Barry, O.W. Gilley, and C.F. Sirmans. "A Method for Spatial-Temporal Forecasting with an application to Real Estate Prices, *"International Journal of Forecasting,"* 16(2000):229-246.

Palmquist, R. B. "Heterogeneous Commodities, Hedonic Regressions, and the Demand for Characteristics." The Farm Real Estate Market. Proceedings of a Regional Workshop, Southern Natural Resource Economics Committee, 1984:45-56.

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