Available online at www.sciencedirect.com

SCIENCE DIRECT

Forest Policy and Economics 7 (2005) 732–744

Forest Policy and Economics

www.elsevier.com/locate/forpol

Spatial analysis of rural land development

Seong-Hoon Cho*, David H. Newman

University of Tennessee, Department of Agricultural Economics, 2621 Morgan Hall, Knoxville, TN 37996-4518, United States

Abstract

This article examines patterns of rural land development and density using spatial econometric models with the application of Geographical Information System (GIS). The cluster patterns of both development and high-density development indicate that the spatially continuous expansions of development and high-density development exist in relatively remote rural areas. The results also revealed that a closer distance to roads, a closer distance to cities, greater access to streams and rivers, higher elevations, and greater proportions of flat area are valued highly in rural land development.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Spatial; Cluster; GIS; Development; Density; Rural

The population of non-metropolitan counties grew by 5.3 million, or 10.3% in the 1990s, compared with an increase of just 1.3 million, or 2.7% in the 1980s. Net migration also shifted from an average annual outmovement of 269,000 in the 1980s to an average inmovement of 348,000 in the 1990s (Economic Research Service, 2004). The non-metropolitan population growth has slowed down recently but there are still rapidly growing counties with amenities that attract retired people. The Blue Ridge Mountains area is among the fastest growing rural areas in the country and Macon County, North Carolina, situated at the southern end of the Blue Ridge Mountains, is an area specifically experiencing this rapid development.

Macon County is classified as rural by the Census Bureau and "non-metro" by White House's Office of Management and Budget (OMB).[1] The county grew from 20,178 people to 29,811 in the 1990s, an increase of nearly 48%. At the same time, the number of housing units increased from 13,358 to 20,746, a gain of 55%. The higher increase of housing units relative to population growth reflects the impact of recreational second home developments in the mountains. For instance, in 2002, 45% of all new residences

* Corresponding author. Tel.: +1 865 974 7411; fax: +1 865 974 9492.

E-mail address: scho@mail.ag.utk.edu (S.-H. Cho).

1389-9341/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.forpol.2005.03.008

[1] The Census Bureau classifies urban area as a central city and the surrounding densely settled territory that together have a population of 50,000 or more and a population density generally exceeding 1,000 people per square mile. All others are considered rural. OMB classifies a metro area as one city with 50,000 or more inhabitants or an urbanized area (defined by the Census Bureau) with at least 50,000 inhabitants and a total metropolitan statistical area (MSA) population of at least 100,000 (75,000 in New England). Any area not included in an MSA is considered "non-metro".

built in the county were second homes. An increasing number of rural homeowners, interfacing with the unprecedented growth of the metropolitan Atlanta area's northern suburbs (e.g., the population of Cherokee County, Georgia ,grew 174% between 1980 and 2000), has expanded second home communities of the county at a rapid pace.

The rapid growth in Macon County has given rise to concerns over declining environmental quality. Scientific monitoring revealed that the water quality in some streams has declined significantly during the past two decades (N.C. Division of Water Quality, 2002). This rapid growth puts pressure on such public services as sewage treatment and overall water quality. Despite the common recognition of the consequences of the county's rapid growth, it has had difficulties adopting a land use plan. The county needs a systematic study to help decision makers propose land use development patterns that make the most efficient and feasible use of infrastructure and public services. Because development is tied to economic incentives, locational externalities, and geological features, spatial econometric models are needed to design development and conservation strategies that address specific environmental consequences. Macon County provides an excellent study site for testing our methodology because institutional factors such as land use regulations have only a minor influence on the area's development because the region contains no land use zoning or regulations.

While the process of urban growth and development has long been a focus of study, there has been increasing interest in non-metro and fringe area development (e.g., Irwin et al., 2003; Miller, 2003; Libby and Sharp, 2003; Irwin and Bockstael, 2002). The development of tests for spatial autocorrelation or dependence in linear regression models as well as the development of efficient and consistent estimators for these types of models have been an important part of the spatial econometric literature over the last few decades (e.g., McMillen, 2003; Tse, 2002; Leung et al., 2000; LeSage, 1997; McMillen, 1992; Anselin, 1988; Cliff and Ord, 1973). While land development models that account for spatial relationships have begun to emerge, such models have focused on development probability or stochastic processes (Dubin, 1988, 1992; Can, 1990, 1992; McMillen, 1992, 1995; Bockstael, 1996). Details of spatial pattern such as density or intensity have not been accommodated. To understand spatial processes and patterns, we must take both types into account (Cheng and Masser, 2003).

In this article, we examine the spatial patterns of land development and the density of land development of a rural county experiencing rapid change. It focuses on an empirical analysis that is useful in understanding rural growth in a spatial context. We also account for spatial dependence by using an integrated approach that combines Geographical Information System (GIS) and spatial econometric models. The spatial dependence with unknown disturbance error is diagnosed by creating spatial lagged variables that capture unobserved characters in regression models (Cliff and Ord, 1973). The GIS and spatial statistics allow for spatially explicit analysis by providing flexibility in specifying models and measuring variables (e.g., Ding, 2001; Lake et al., 2000; Geoghegan et al., 1997).

1. Empirical model

Land development decisions by a landowner at the parcel level have been modeled using discrete choice models. These models estimate the probability of land development as a function of parcel-level attributes (e.g., Bockstael and Bell, 1998; Bockstael, 1996). Because a priori returns from parcel development are unknown with certainty, Bockstael (1996) developed a hedonic model of land values to estimate predicted land values, which were then used as a proxy for the expected returns of development. Then, in a second stage, they modeled land development using a discrete choice model incorporating these predicted land values.

We extend Bockstael and Bell's two-stage model into a three-stage model to accommodate the density of development. We estimate a hedonic model of land value in the first stage, a development model in the second stage, and a density of development model in the third stage. The first-stage hedonic model utilizes attributes of land values. The predicted land value is estimated from the hedonic model and used as a proxy for the expected return of development in the second and third stage estimations. The second stage estimations of the development model identify character-

istics that determine land development. The third stage estimations of the density of development model identify characteristics that determine the density of land development. We also attempt to identify how the patterning of landscape and spatial dependency affect land value, development, and the density of development.

1.1. First stage: the hedonic model

A hedonic model can integrate anything that affects land values. We include obvious variables from the literature, such as distance to city centers, major roads, proximity to desirable natural amenities (e.g., lakes, streams, and rivers), geological such as jurisdictional, development, and neighboring development indicators, and neighborhood features such as relative abundance of roads, streams, and rivers. Under the assumption that the hedonic model closely approximates what developers are likely to view as the market condition, the predicted land values from the model are used as proxies for expectations of returns from development in the second and third stages.

Economic theory generally places few restrictions on the functional form of the hedonic price form. The form is selected empirically according to the performance of the several functional forms, typically ad hoc, and goodness of fit criteria. Since the main focus of our analysis is on spatial econometric aspects, a detailed analysis of the role of different functional forms is not provided. Flexible functional forms, such as the Box–Cox transformation, have become increasingly prevalent for estimation of hedonic models but such specifications are not readily implemented in the presence of spatial dependence. Therefore, we considered some common functional forms (linear, semi-log, log–linear, and inverse-semi log) for our analysis. An analysis of the residuals from these functional forms revealed that linear and inverse-semi log provided a poor fit while the semi-log and log–linear fit the models relatively well. We judged that the best of the four alternatives was the log–linear form.

A number of studies using parcel data have found that the most common cause of heteroscedasticity in hedonic models is parcel size. This is largely due to the fact that fragmentation of land use and growth increase development density, which leads to increasing land values. We tested for heteroscedasticity using the Glejser test, which regresses acres of parcel against the absolute value of the residuals from the hedonic model and used the Hansen–White correction to estimate standard errors. Initially, the standard hedonic model is estimated in semi-log form. The initial ordinary least squares (OLS) model is then re-estimated with weighted least squares using the reciprocals of the normalized predicted values from the absolute residual model based on Goodman and Thibodeau (1995, 1997). This iterative procedure is then repeated until the largest change in any of the parameters is less than 0.0001. This model is therefore based explicitly on the assumption that heteroscedasticity is due to the size of the parcel.

To evaluate the prediction accuracy of the hedonic model, we estimated a hedonic regression and then use the estimated coefficients for out-of-sample prediction. For each out-of-sample prediction, we computed the absolute difference between the estimated value of the land and the actual assessed land value. We then calculated the percentage of differences within 10% and 20% of the actual assessed land value. This was done for both initial OLS model and weighted least squares model.

1.2. Second stage: binary model for land development decision

The simplest characterization of the development decision for a parcel of land is that the landowner of parcel j, which is currently in state a is converted to state i at time t if

$$R_{jit|a} \geq R_{jmt|a} \tag{1}$$

for all land uses $m=1,\ldots, M$ (including a). $Rjit|a$ represents the present value of the infinite stream of net returns to parcel j, in state i at time t, given that the parcel was in state a in time $t-1$. It is a function of observable variables and a random portion, η. The probability that land parcel j that is in land use a at time $t-1$ will be found in land use i at time t is given by

$$Pr\left(R_{jit|a} + \eta_{jit|a} > R_{jmt|a} + \eta_{jmt|a}\right) \tag{2}$$

for all $m=1\ldots, M$ (including a).

This may be an oversimplication because the infinite stream of net returns is not known with

certainty and most conversions are irreversible. Arrow and Fisher (1974) showed that over-development would likely occur when irreversibility is ignored. Titman (1985) examined the relationship between uncertainty and land values and showed that uncertainty about future land values decreases building activities in the current period. The conversion decision could be made temporarily dynamic by recognizing that the probability of conversion depends on the initial state and possibly the cumulative history of the parcel such as accumulation or depreciation of natural, human, and structural capital. The conversion decision could be made spatially dynamic as well by recognizing that the value of a parcel land in different uses may be a function of changes in the neighboring land uses, in the pattern of those land uses, and the existence and spatial pattern of infrastructure. However, in the general case with dynamic states, the discrete choice model of this kind would be difficult to estimate reliably and would preclude any spatial treatment of error structure. In addition, over 90% of the human induced-land use conversions that took place are from agriculture or forest to some density of residential development. The problem can be reduced to a dichotomous choice problem conditioned on the time period t, by estimating the probability that land will or will not be developed (Bockstael, 1996).

The probability that parcel j (which is currently in an undeveloped state), will be developed in time t is given by

$$\begin{aligned}\Pr(\text{develop}) &= \Pr\left(R_{jdt} + \eta_{jdt} > R_{jut} + \eta_{jut}\right) \\ &= \Pr\left(R_{jdt} - R_{jut} > \eta_{jut} - \eta_{jdt}\right),\end{aligned} \tag{3}$$

where d denotes a developed state and u denotes an undeveloped state. The certain returns from development of a parcel are not known when the developer makes her decision because the decision is made based on the expected value of net return. We assume that the landowner formulates a hedonic on land characteristics from her observations of market transaction following Bockstael (1996). If the hedonic model estimates what the developer is likely to view as the market condition at time t, then we can use the predictions from the hedonic model as proxies for the expected value of net return from conversion.

The distribution for $\eta_{jut} - \eta_{jdt}$ is typically assumed to follow a logistic or normal distribution. If a logistic distribution is assumed, the probability of land development can be derived as a logit model. If a normal distribution is assumed, the probability of land development can be derived as a probit model. The logit model is similar to the probit model, and attempts practically the same mission. In this study, we assume the probability of land development follows a probit model:

$$\Pr(\text{develop}) = \Phi(\beta' X) \tag{4}$$

This land development model is estimated using parcel-level data. Given that we only know whether a particular parcel of land was developed or not at this stage, the probit model must be estimated using maximum likelihood methods. The likelihood function is

$$L = \prod_{i=I} F(\beta' X_i)^{y_i} [1 - F(\beta' X_i)]^{(1-y_i)}, \tag{5}$$

where $F(\cdot)$ is the standard normal cumulative distributive function, and $y_i = 1$ if parcel i is developed. The model can be used to evaluate the effect of alternative variables on land development. For example, the marginal effect of predicted land value on the probability of land development equals

$$\frac{\partial \text{Prob}[y = 1]}{\partial v} = \varphi(\beta' X)\beta_v \tag{6}$$

for the probit model where β_v is the coefficient on the predicted land value measure v.

1.3. Third stage: a binary model of high and low density development

A binary system of high and low density development is estimated in the third stage. The density reflects the strength of development in any developed parcel. High density development defines a parcel of land with more than one structure per acre while low density development defines a parcel of land with one structure or fewer per acre. The probability of high-density development is derived as a probit model:

$$\Pr(i|y) = \Phi(\beta' X) \tag{7}$$

Let $i|y$ represents the event of high-density development i when development, $y = 1$ occurs (conditional

probability); theoretically Pr $ob[iy]$=Pr $ob[i|y]$ (iy means the event of high-density development). Based on the two formulae (Eqs. (4) and (7)), we can calculate the probability value of high-density development (that is Pr $ob[iy]$).

1.4. Spatial autocorrelation

In order to detect spatial autocorrelation of land value for the hedonic model, we use Moran's index. The index can be applied to zones and points with continuous variables associated with them. The formula to calculate Moran's Index (I) for the case of the hedonic model is following:

$$I = \frac{N \sum_i \sum_j W_{ij}(X_i - \overline{X})(X_j - \overline{X})}{\left(\sum_i \sum_j W_{ij}\right) \sum_i (X_i - \overline{X})^2} \tag{8}$$

where N is the number of cases, X_i is the land value for parcel i, $\overline{X}$ is the mean of land value, and W_{ij} is the distance based-weight which is the inverse distance between parcel i and j. The value is similar to a correlation coefficient, varying between -1 and 1. When autocorrelation is high, the coefficient is correspondingly high so that a high I-value indicates positive autocorrelation.

The joint–count spatial statistics technique is used to diagnose the spatial dependence of development and density of development of binary variables. The join patterns of the parcels can be identified as rook's and queen's cases. The rook's case identifies the circumstance in which two polygons are adjacent to each other so that they share a common boundary. The queen's case identifies the circumstance in which two polygons share either a common boundary or a common vertex.

The null hypothesis of spatial dependence of development states that neighboring parcels are more likely to be of the same category, developed or undeveloped. The observed joint–count statistics n_{dd} and n_{uu} count the number of joint encounters in adjacent parcels having the same category; the corresponding n_{du} statistic counts the number of adjacent parcels not having the same category. Thus the n_{dd} and n_{uu} statistics assess the presence of positive spatial autocorrelation, while n_{du} assesses the presence of negative spatial autocorrelation. The count of developed–developed (1–1) joins can be calculated as

$$n_{dd} = \sum \sum \psi_{ij} Y_{ij}, \qquad \text{where } Y_{ij} = x_i x_j \tag{9}$$

with the observation x_i is 1 and x_j is 1, and the spatial weight matrix, ψ_{ij} is calculated as

$$\psi_{ij} = 1 \quad \text{if } d_{ij} = b\sqrt{\frac{\text{area}_i}{\pi}} \quad (b \text{ is a parameter}),$$

$$\text{and} \quad \psi_{ij} = 0 \quad \text{if } d_{ij} = b\sqrt{\frac{\text{area}_i}{\pi}} \tag{10}$$

where ψ_{ij} is the element of the spatial weight matrix; d_{ij} represents the distance between parcels i and j (air distance between centroids); b is a parameter. If adjacent parcels have the same (or similar) size and shape as the central one, b=2 represents the rook's case while b=3 represents the queen's case (Cliff and Ord, 1973). The count of developed–undeveloped (1–0) joins is calculated as

$$n_{du} = \sum \sum \Psi_{ij} Y_{ij}, \qquad \text{where } Y_{ij} = x_i x_j \tag{11}$$

with the observation x_i=1 and x_j=0 or vice versa. The spatial weight matrix is included to account for parcel size and the distance between parcels. The significance of the joint–count statistic is achieved by computing a standard normal deviate using a two-tailed test to detect positive or negative spatial autocorrelation. The same joint–count statistics have been applied to density of development.

As the spatial autocorrelations are detected in continuous or binary variables, spatial lagged variables are created and used to correct the problem. The variables are added as explanatory variables in the equations to capture the spatial pattern. Mean land value of the parcels adjacent to its own parcel is used for the variables capturing the spatial patterns of land value. A dummy variable indicating whether an adjacent parcel is developed or undeveloped is used for the variable capturing the spatial pattern of development. A dummy variable indicating whether an adjacent parcel has high density or low density development is used for the variable capturing the spatial pattern of the density of development.

2. Data

The land records division of the Macon county tax administration department provided us with the January 2003 tax assessment records (updated every 4 years) in a shape file that included information about land parcel size, number and type of structures in a parcel, and the assessed land and structure values. The assessed land value is used for the proxy for land price. In 1967, there were 40,075 vacant polygons in the records. The records of 40,039 polygons were used for our study after eliminating missing observations. Structures existing prior to 1967 were not considered in our model due to unavailability of built years in the records. Each polygon represents one land parcel. Out of the 40,039 parcels, 15,725 parcels are predicted to be developed by the probit model for the land development decision. These 15,725 parcels are then used to estimate the probit model for high-density development because the probability of high-density development is conditional on parcel development.

GIS is utilized to generate the spatial variables. The variables of sum of road per acre and sum of stream and river length per acre are calculated by using an ArcView script.[2] The variable of sum of road per acre is created to measure the effects of both relative abundance and existence of roads in a parcel which cannot be measured by the variable of distance to roads. Similarly, the variable of sum of stream and river length per acre is created to measure the effects of both relative abundance and existence of streams and rivers in the neighborhoods.

The variables of distance to the nearest city center (the cities of Franklin or Highlands), distance to the closest stream and river, and distance to the closest road are calculated using the ArcView script, "Nearest Features, with Distances and Bearings".[3] Although travel distance is a more accurate measure, air-distance can be used as a proxy for accessibility[4] (McMillen, 1989; Hushak, 1975). Thus, the air distances are used for the proxies of all the distance variables in our study.

The variables of flat ratio and median elevation are used for the geological characteristics. Because many second-home owners seek locations at higher elevations with better views in the mountainous region of the county, the ratio of flat area to total area is suspected to be an influential factor for both land value and development decisions. The flat ratio and median elevation are measured using the data set from digital elevation models. The data set consists of a raster grid of regularly spaced elevation values that have been primarily derived from the topographic map of the U.S. Geological survey series. The flat ratio and median elevation for polygons of the area are calculated using an ArcView script, "Surface Tools for Points, Lines, and Polygons".[5]

3. Estimation and results

Definitions and descriptive statistics of the variables used in the estimation of the models are shown in Tables 1 and 2, respectively. The results in Table 3 show that statistical tests reject the null hypothesis of no spatial dependence among developments of adjacent parcels in both rook's and queen's cases, supporting the notion that spatial dependence exists in the land development. This implies that a parcel is more likely to be developed if the parcel is closer to a developed parcel. The negative value of the standard normal deviation of *du* (a case of a developed lot adjacent to an undeveloped lot) suggests that developed and undeveloped lots are less likely to be neighbors. The results in Table 4 also show that statistical tests reject the null hypothesis of no spatial dependence among density of developments in both rook's and queen's cases, supporting the notion that spatial dependence exists among density of develop-

[2] The script was developed with help of Tripp Lowe, Information Analyst at the School of Forest Resources, University of Georgia based on Census 2000 TIGER/Line shape files that are downloaded from Environmental System Research Institute, Redlands, California.

[3] The script was developed by Jeff Jenness, GIS Analyst at US Forest Service, Rocky Mountain Research Station.

[4] Travel distance can be estimated using accessibility analysis which integrates road category as a main classifier along with traffic flow data, physical barrier information, and transportation network data.

[5] The script was developed by Jeff Jenness, GIS Analyst at US Forest Service, Rocky Mountain Research Station.

Table 1
Descriptions of the variables

Variables	Descriptions
Land value	Natural log of land value per acre
Adjacent land value	Mean land value of adjacent parcels
Land development	Dummy variable, 1 if a parcel is developed; 0 otherwise
Land development density	Dummy variable, 1 if a parcel is developed in high density; 0 otherwise
Adjacent development	Dummy variable, 1 if an adjacent parcel is developed; 0 otherwise
Adjacent density	Dummy variable, 1 if an adjacent parcel is developed in high density; 0 otherwise
Area	Area of a parcel in acre
Sum of road[a] per acre	Sum of road length in mile per acre
Sum of stream and river[b] length per acre	Sum of stream and river length in mile per acre
Distance to city center	Distance from a center of a parcel to the nearest city, Franklin or Highlands in mile
Franklin dummy	Dummy variable, 1 if a parcel is within a jurisdiction of city of Franklin; 0 otherwise
Highlands dummy	Dummy variable, 1 if a parcel is within a jurisdiction of Highlands; 0 otherwise
Distance to stream and river	Distance from a center of a parcel to the nearest stream or river in mile
Flat ratio	Ratio of flat area to total area
Median elevation	Median elevation in mile
Distance to road	Distance from a center of a parcel to the nearest road

[a] Road includes primary road, secondary road, and local, neighborhood, and rural road defined by U.S. Census Bureau.
[b] Stream and river include basic hydrograph and natural flowing water defined by U.S. Census Bureau.

ments as well. A positive and strong Moran's index of 0.4035 suggests existence of spatial dependence in the value of land. These results regarding spatial dependence support the notion of rural cluster development and a spatial pattern of housing value.

The initial estimates of the hedonic model, reported in Table 5, include *t*-statistics from both the original specification and re-run using the White correction. The results from the Glejser test, reported in Table 6, have acres as the independent variable and the residuals from the aggregate hedonic model including the square root of acres and one of acres terms as the dependent variable. The test shows that heteroscedasticity is present in the hedonic model. Both the initial and adjusted *t*-statistics with respect to acres are highly significant. The estimates with weighted least squares using the reciprocals of the normalized predicted values from the absolute residual model are reported in Table 7. The accuracies of the predicted values in Tables 5 and 7 show that the model corrected for the heteroscedasticity of parcel size improves accuracy at the 20% level by about 8%.

Table 2
Statistics of the variables

	Mean	Maximum	Minimum	Standard deviation
Land value	6.30	10.13	0	8.64
Adjacent land value	7.31	10.07	0	9.27
Land development	0.49	1	0	0.39
Land development density	0.27	1	0	0.45
Adjacent development	0.77	1	0	0.42
Adjacent density	0.23	1	0	0.41
Area	5.98	7071	0.01	79.91
Sum of road per acre	0.08	1.86	0	0.77
Sum of stream and river length per acre	0.10	3.96	0	0.30
Distance to city center	22.20	91.22	0	19.35
Franklin dummy	0.06	1	0	0.24
Highlands dummy	0.07	1	0	0.26
Distance to stream and river	1.22	7.98	0	0.95
Flat ratio	0.98	1	0.81	0.01
Median elevation	0.82	1.63	0.57	0.21
Distance to road	0.31	5.39	0	0.46

Table 3
Spatial dependencies of land development

	$b=2$			$b=3$		
	dd[a]	*du*[b]	*uu*[c]	*dd*	*du*	*uu*
Observations	16,215	9412	34,935	38,367	31,238	75,621
Expectations	8813	2768	23,182	23,213	73,659	51,326
Variance	400	307	724	1161	744	1598
t-value	16.33	−47.26	19.25	14.54	−49.78	17.88

[a] A case of a developed parcel is adjacent to a developed parcel.
[b] A case of a developed parcel is adjacent to an undeveloped parcel.
[c] A case of an undeveloped parcel is adjacent to an undeveloped parcel.

Table 4
Spatial dependencies of density of land development

	$b=2$			$b=3$		
	hh[a]	hl[b]	ll[c]	hh	hl	ll
Observations	2574	1430	5946	5945	4271	12871
Expectations	1430	409	3964	3716	1630	7987
Variance	65	52	162	199	112	325
t-value	22.18	−35.71	12.51	19.75	−39.71	13.19

[a] A case of a high-density developed parcel is adjacent to a high-density developed parcel.
[b] A case of a high-density developed parcel is adjacent to a low-density developed parcel.
[c] A case of a low-density developed parcel is adjacent to a low-density developed parcel.

The 63% goodness of fit and 81% of accuracy with 20% deviation (Table 7) verify that the predictability of the hedonic model with weighted least squares is reasonably good and can be used as a proxy for the expected returns of development for the development models in the second and third stages.

Table 7 shows that all the variables are significant at the 5% level. The natural log of the value per acre increases as a parcel of land is developed, as an adjacent parcel's land value increases, and as parcel size decreases. While a decrease in distance from a land parcel to the closest road increases land value, an increase in the sum of road per acre decreases land value. This implies that people value the convenience of being closer to roads but they do not like to be crowded by roads. Both a decrease in distance from a land parcel to the closest stream and river and an increase in sum of stream and river length per acre increase land value. This result reflects that streams and rivers are valued positively on both their own land as well as on neighboring land. Finally, land parcels with a greater proportion of flat area are valued higher and land parcels at higher elevations also have greater

Table 5
Hedonic model with ordinary least squares

Variable	Parameter estimate	t-statistics	White's t-statistics
Intercept	−22.5897	−44.64	−30.56
Land development	−0.0059	−0.50	−0.66
Adjacent land value	0.1233	45.95	52.12
Area	−0.0177	−37.81	−69.27
Sum of road per acre	−0.5988	−12.95	−8.64
Sum of stream and river length per acre	0.0509	1.95	1.85
Distance to city center	−0.0086	−25.22	−29.04
Franklin dummy	0.9455	37.22	39.22
Highlands dummy	1.1016	40.91	36.45
Distance to stream and river	−0.0386	−6.06	−6.96
Flat ratio	20.9053	41.46	46.41
Median elevation	1.5427	46.80	48.30
Distance to road	−0.1669	−12.97	−16.14

Dependent variable=natural log of land price per acre.
Observation=40,039.
F-value=3328.36.
Prob>F=0.0001.
R-square=0.5338.
Adjusted R-square=0.5336.
Percentage of prediction within 10% of land value=43.8.
Percentage of prediction within 20% of land value=73.7.

Table 6
Glejser heteroscedasticity tests on acres

	α	β
Glejser test on original specification	0.6928	0.0061
t-statistic	159.28	22.33
White t-statistic	169.63	27.4
Glejser test on square root acres	0.7207	−0.0013
t-statistic	118.95	18.33
White t-statistic	121.43	21.10
Glejser test on one over acres	0.0041	0.0005
t-statistic	152.12	76.78
White t-statistic	158.64	81.81

Table 7
Hedonic model with weighted least squares

Variable	Parameter estimate	t-statistic
Intercept	3.0417	144.85
Land development	0.1765	13.08
Adjacent land value	0.0904	53.83
Area	−0.0929	−46.47
Sum of road per acre	−0.8890	−16.90
Sum of stream and river length per acre	0.1746	6.11
Distance to city center	−0.0039	−10.43
Franklin dummy	0.3291	11.37
Highlands dummy	1.0648	35.14
Distance to stream and river	−0.1113	−15.71
Flat ratio	4.8424	120.22
Median elevation	1.1331	31.69
Distance to road	−0.0793	−5.53

Observation=40,039.
F-value=3554.24.
R-square=0.6269.
Adjusted R-square=0.6258.
Percentage of prediction within 10% of land value=49.5.
Percentage of prediction within 20% of land value=81.4.

Table 8
Probit model results of land development

Variable	Parameter estimate	Standard error	Chi-square	Prob>Chi
Intercept	−8.8723	0.7423	106.23	0.0001
Predicted land value	0.0596	0.0124	20.31	0.0001
Adjacent development	0.6601	0.0248	680.21	0.0001
Area	−0.0154	0.0072	5.03	0.0361
Sum of road per acre	0.0006	0.0001	199.28	0.0001
Sum of stream and river length per acre	−0.0001	0.0003	0.45	0.6013
Distance to city center	−0.0024	0.0009	22.58	0.0001
Franklin dummy	0.0527	0.0549	0.58	0.4374
Highlands dummy	0.1925	0.0428	11.09	0.0010
Distance to stream and river	−0.0193	0.0861	4.25	0.0561
Flat ratio	9.6351	0.8627	127.46	0.0001
Median elevation	−0.3294	0.0528	33.44	0.0001
Distance to road	−0.4121	0.0224	371.31	0.0001

Dependent variable=1 if land parcel is developed; =0 if not.
Log likelihood for normal=−12,901.23.
Number of observation=40,039.
Percent correct predictions=65.23.

value confirming that second-homeowners value sites at higher elevations with better views.

The estimated results of the probit models of land development and density of land development are reported in Tables 8 and 9, respectively. The percent correct prediction for the land development model and the density of land development model are 65.23 and 84.17, respectively. The two models yield similar outcomes in terms of significance of the variables, but the marginal effect is quite different. To examine the effects of the variables, we estimate the empirical relationship between the continuous variables and the percent of land development and density of land development. This is done by first substituting the mean values of all the other variables and then calculating the probabilities of land development and density of land development when the continuous variable varies.

The probabilities of development and high density development increase similarly with an increase of the

Table 9
Probit model results of density of land development

Variable	Parameter estimate	Standard error	Chi-square	Prob>Chi
Intercept	−1.7231	0.2612	17.91	0.0001
Predicted land value	0.0427	0.0185	4.53	0.0384
Adjacent development	0.0315	0.0226	3.15	0.0720
Area	−2.5833	0.2433	4.01	0.0356
Sum of road per acre	0.7158	0.0619	142.51	0.0001
Sum of stream and river length per acre	−0.0671	0.1264	0.42	0.4327
Distance to city center	−0.0062	0.0011	12.39	0.0001
Franklin dummy	0.1188	0.0541	3.14	0.0554
Highlands dummy	0.0959	0.0242	9.62	0.0015
Distance to stream and river	−0.0284	0.0166	3.17	0.0741
Flat ratio	1.8858	0.1764	72.33	0.0001
Median elevation	−0.4634	0.1829	8.97	0.0022
Distance to road	−0.1973	0.0641	10.25	0.0017

Dependent variable=1 if land parcel is developed in high-density; =0 if not.
Log likelihood for normal=−3731.25.
Number of observation=15,725.
Percent correct predictions=84.17.

predicted land value per acre. The probability of development increases with a decrease in parcel size because land conversion occurs after lots are subdivided for subdivisions. Likewise, the probability of high density development also increases with a decrease in parcel size reflecting the fact that land use fragmentation and growth increase development density. The result is also consistent with previous literature in which population density is shown to increase the fragmentation of development (Carrion and Irwin, 2004). We also find that parcel sizes greater than 10 acres have close to 0% probability of high density development. This low density development in larger parcels provides evidence of the fact that more than half of total new single-family home acreage in the period from 1994 to 1997 is associated with new homes built on lots of 10 acres and larger (Lang, 2000).

Both the probabilities of development and high density development increase with an increase in the sum of roads per acre but decrease with an increase in distance to roads. The marginal effect of the predicted sum of roads per acre on the probability of land development indicates that 0.01 miles or 52.8 ft of sum of road per acre increases the probability of development by 1.6%. The probabilities of both development and high-density development are close to zero in land parcels 10 miles or more from the closest road. This implies that road accessibility of a given land parcel is a necessary element for development.

Land parcels closer to streams and rivers are more likely to be developed, and they are more likely to be developed in high density. The probability of high-density development of the land parcels adjacent to a stream or river is around 20%. They are close to zero for land parcels 10 miles or more from any stream and river. This shows that both development and high-density development are highly correlated with adjacency to streams and rivers.

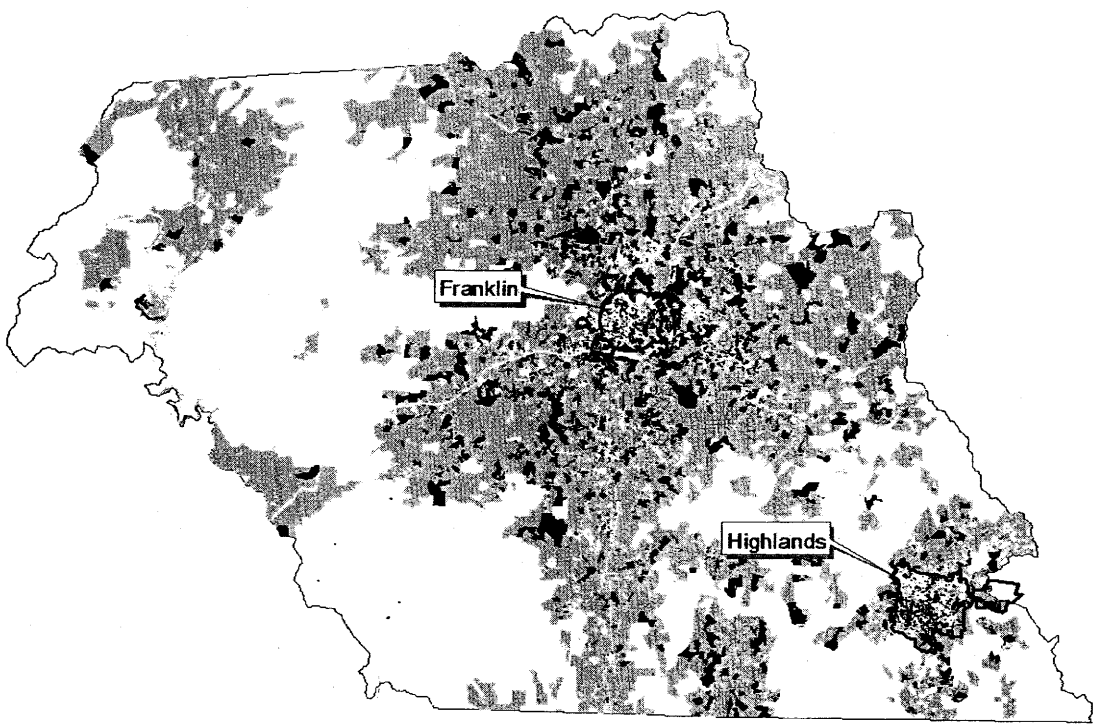

Fig. 1. A Map of the predicted probabilities of development.

The probabilities of both development and high density development are greater in land parcels closer to Franklin or Highlands. The land parcels within the jurisdictions of Franklin and Highlands are more likely to be developed and they are more likely to be developed in high density than the land parcels outside of the jurisdictions.

Land parcels need to have a flat ratio of at least 60% to have positive probability of development. The probability of development increases sharply as the flat ratio increases. The probabilities of both development and high density development increase with decreases in elevation. Despite the higher values of land parcels at higher elevation, the probabilities of development and high density development are still higher at lower elevations. This seemingly incongruous result may be explained by the scarcity of available land and low affordability of houses at higher elevations because of the additional building costs. Though many prefer to buy houses on top of mountains for better views, few can afford to buy those houses. Thus, the probability of development is higher at lower elevation despite the higher values of land parcels at higher elevation. The high density at lower elevations is expected since the smaller flat ratio at higher elevations does not allow high density development.

A map of the predicted probabilities of land development (Fig. 1) and a map of the predicted probabilities of density of development (Fig. 2) are

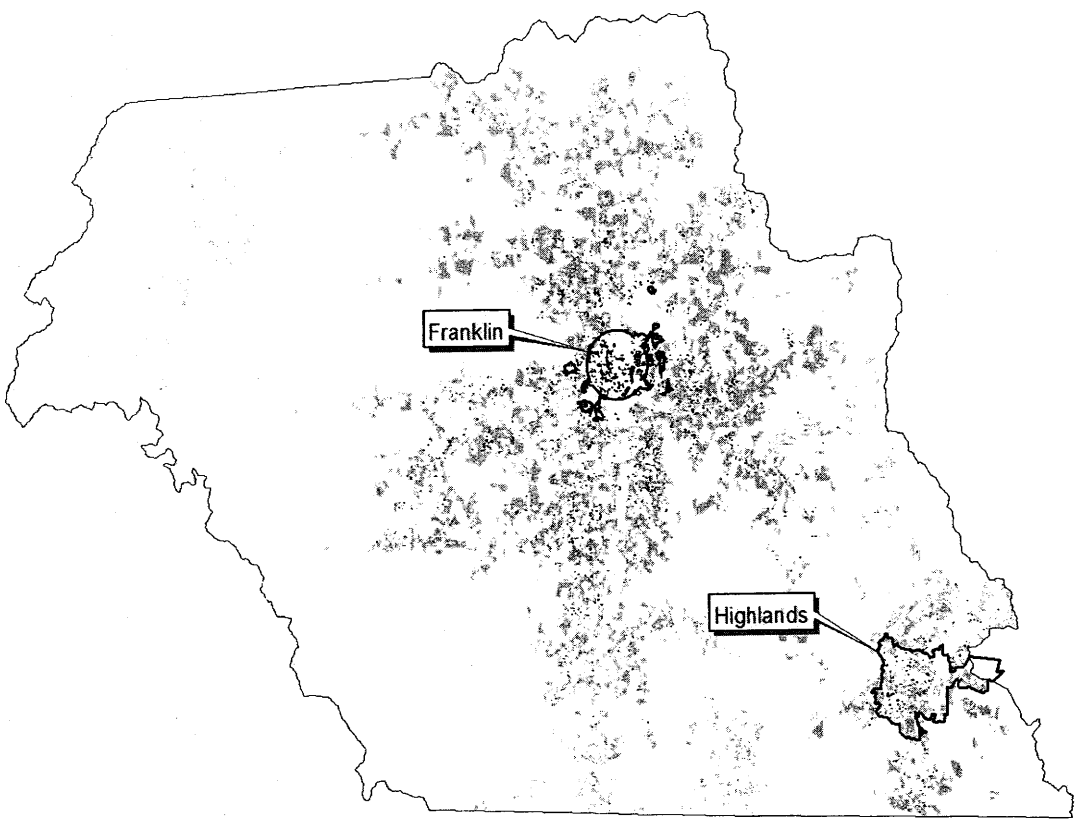

Fig. 2. A Map of the predicted probabilities of density development.

produced using the estimates from the models. These maps are not meant to predict which parcels will be developed, but rather highlights those areas which will most likely be developed. Fig. 1 shows that 15,725 or 39% of the land parcels have predicted probabilities of conversion equal to or greater than 50%. The large white area identifies missing observations and National Forest area. The predicted high density development in Fig. 2 shows that 3213 or 20% of the land parcels are likely to be high density conversion.

4. Conclusions

The spatial pattern of development at a multiscale perspective has important environmental consequences that range from change of water quality to biodiversity. In fact, land use change at one scale or another is perhaps the single greatest factor affecting ecological resources (Hunsaker and Levine, 1995). The pattern of land use is also tied to economic incentives, locational externalities, and geological features. These findings, along with our land-use projections, suggest how spatial models could be used to design development and conservation strategies that address specific environmental consequences in specific places.

For the evaluation of future environmental impacts or conservation strategies, predicting a specific land use in a particular place is as important as understanding where the relative probabilities and densities of change are. Land-use forecasting models in a spatial framework can be used to draw land-use maps which show the linkage between risk assessment and environmental impact models. In this kind of analysis, we can understand where human activities generate significant environmental consequences, focusing future research and planning by linking the spatial dynamics of human populations to potential environmental impacts that are the most critical in supporting environmental health.

The most obvious audience to benefit from this research will be decision-makers in Macon and the surrounding counties. The detailed projection of development and density pattern can be used to highlight the effects of local policy decisions. These include direct land use regulation, such as zoning, and more indirect land use policies such as the provision or expansion of public infrastructure or other public services. The projected density pattern provides background for the need for zoning the density of houses per acre in the various residential areas. The changes expected to be made in the various densities are intended to assist the general policy of expanding the residential population of the study area and facilitate clarity in development control. Decisions for changes can then be proposed to the current residential zones based on the projected density patterns.

Similarly, quantification of the effects of economic, locational, and geological features on residential development and density patterns should also help decision makers establish a land use development pattern that makes the most efficient and feasible use of existing infrastructure and public services. It also provides a guideline for new developments that maintain or enhance the quality of the study area. For example, policy makers could utilize the existing infrastructure and public services more efficiently by developing a program that encourages growth toward locations where development clusters exist and development is predicted.

The next logical step of our analysis is to improve the model specification by adopting unobserved site-specific characteristics. About 80% of the study area is forested land, so that forest site characteristics such as size, shape, and accessibility, and timber characteristics such as the number of trees per acre, species, and size are likely to be important to the land valuation decision. Inclusion of these characteristics would reduce biases caused by omitted variables.

Another extension of this research would be to combine the results with aggregate-type land analysis. Aggregate land analysis examines patterns of land use from a macro viewpoint. The analysis generally use counties, county groupings, census blocks, and census-block groups as units to highlight how socio-economic factors and physical landscape features influence land use allocations. This type of analysis can capture broader physical and social phenomena which landowner-specific analyses may miss. However, this type of analysis does not capture information in a spatially explicit framework. Spatial joining of the aggregate and landowner-specific datasets would allow us to bridge the analyses of two different scales.

References

Anselin, L., 1988. Spatial Econometrics: Methods and Model. Kluwer Academic, Dordrecht.

Arrow, K.J., Fisher, A.C., 1974. Environmental preservation, uncertainty, and irreversibility. Quarterly Journal of Economics 88, 312–319.

Bockstael, N., 1996. Modeling economics and ecology: the importance of a spatial perspective. American Journal of Agricultural Economics 78, 1168–1180.

Bockstael, N., Bell, K., 1998. Land use patterns and water quality: the effect of differential land management controls. In: Just, R., Netanyahu, S. (Eds.), International Water and Resource Economics Consortium, Conflict and Cooperation on Trans-Boundary Water Resources. Kluwater Academic Publishers, Dordrecht.

Can, A., 1990. The measurement of neighborhood dynamics in urban house prices. Economic Geography 66, 254–272.

Can, A., 1992. Specification and estimation of hedonic housing price models. Regional Science and Urban Economics 22, 453–474.

Carrion, C., Irwin, G., 2004. Determinants of residential land-use conversion and sprawl at the rural–urban fringe. American Journal of Agricultural Economics 86, 889–904.

Cheng, J., Masser, I., 2003. Modeling urban growth patterns: a multiscale perspective. Environmental & Planning A 35, 679–704.

Cliff, A., Ord, J., 1973. Spatial Autocorrelation. Pion Limited, London.

Ding, C., 2001. An empirical model of urban spatial development. Review of Urban and Regional Development Studies 13, 173–186.

Dubin, R., 1988. Estimation of regression coefficients in the presence of spatially autocorrelated error terms. Review of Economics and Statistics 70, 466–474.

Dubin, R., 1992. Spatial autocorrelation and neighborhood quality. Regional Science and Urban Economics 22, 433–452.

Economic Research Service, United States Department of Agriculture (USDA), 2004. Rural population and migration: rural population change and net migration.

Geoghegan, J., Weinger, L., Bockstael, N., 1997. Spatial landscape indices in a hedonic framework: an ecological economics analysis using GIS. Ecological Economics 23, 251–264.

Goodman, A., Thibodeau, T., 1995. Dwelling age heteroscedasticity in hedonic house price equations. Journal of Housing Research 6, 25–42.

Goodman, A., Thibodeau, T., 1997. Dwelling age heteroscec in hedonic house price equations: an extension. Jou Housing Research 8, 299–317.

Hunsaker, C.T., Levine, D.A., 1995. Hierarchical approache study of water in rivers. Bioscience 45, 193–203.

Hushak, L., 1975. The urban demand for urban–rural frin Land Economics 51, 112–123.

Irwin, E.G., Bockstael, N.E., 2002. Interacting agents, externalities and the evolution of residential land use p Journal of Economic Geography 2, 31–54.

Irwin, E.G., Bell, K.P., Geoghegan, J., 2003. Modeli managing urban growth at the rural–urban fringe: a parc model of residential land use change. Agricultural and R Economics Review 32, 83–102.

Lake, I., Bateman, I., Day, B., Lovett, A., 2000. Improvi compensation procedures via GIS and hedonic p Environment and Planning. C, Government and Pol 681–696.

Lang, R., 2000. Forest Fragmentation and Urban Sprawl lenges for Land Managers. The Ohio Hetuch.

LeSage, J.P., 1997. Regression analysis of spatial data. Jou Regional Analysis and Policy 27, 83–94.

Leung, Y., Mei, C., Zhang, W., 2000. Testing for spatial aut lation among the residuals of the geographically w regression. Environmental & Planning A 32, 871–890.

Libby, L.W., Sharp, J.S., 2003. Land-use compa change, and policy at the rural–urban fringe: i from social capital. American Journal of Agricultura nomics 85, 1194–1200.

McMillen, D., 1989. An empirical model of urban fringe la Land Economics 65, 138–145.

McMillen, D., 1992. Probit with spatial autocorrelation. Jou Regional Science 3, 335–348.

McMillen, D., 1995. Selection bias in spatial econometric r Journal of Regional Science 3, 417–436.

McMillen, D., 2003. Spatial autocorrelation or model misspe tion? International Regional Science Review 26, 208–2

Miller, A.P., 2003. Rural development considerations for management. Natural Resources Journal 43, 781–801.

N.C. Division of Water Quality, 2002. 2002 Little Tennesse Basinwide Water Quality Plan.

Titman, S., 1985. Urban land prices under uncertainty. An Economic Review 75, 505–514.

Tse, R.C., 2002. Estimating neighborhood effects in house towards a new hedonic model approach. Urban Stud 1165–1180.