In: Global Change and Forestry
Editors: J. Gan, S. Grado and I. A. Munn
ISBN: 978-1-60876-262-0

Chapter 3

LAND USE DYNAMICS ALONG URBAN-RURAL GRADIENT: A COMPARISON OF MODELING APPROACHES[1]

Maksym Polyakov[1] and Daowei Zhang[2*]
[1] School of Agricultural and Resource Economics, University of Western Australia, Australia
[2] School of Forestry and Wildlife Sciences, Auburn University, AL, USA

ABSTRACT

Modeling land use dynamics is an important component of the landscape level analysis of socio-ecological drivers at the urban-rural interface. We constructed a dataset consisting of 5313 sample points characterizing land cover in 1992 and 2001 across three contiguous counties in West Georgia. We examined two specifications of conditional logit model based the land rent theory: a model of land use allocation, which is similar to models of land use share that utilize aggregated data, and a model of land use change. The land use change model showed better goodness of fit as it controls for cross-sectional effects and allows parameters of explanatory variables to capture temporal effects. The results have implications on land use modeling at watershed or sub-watershed levels.

Keywords: Land use change; Urban-rural interface; Logit model; Georgia.

1. INTRODUCTION

Land use changes, while driven by landowners seeking maximization of economic benefits, sometimes produce negative externalities such as air and water pollution, loss of

---

[1] This study is supported by the National Research Initiative of the Cooperative State Research, Education and Extension Service, USDA, Grant #USDA-2005-3540015262 and the Auburn University Center for Forest Sustainability.

* Email: zhangd1@auburn.edu, phone: (334)844-1067, School of Forestry and Wildlife Sciences, Auburn University, Auburn, AL 36830.

biodiversity, wildlife habitat fragmentation, and increased flooding (Stavins & Jaffe, 1990). In places where the majority of the land base is privately owned, like the U.S. South, it is important to understand how socioeconomic and natural factors affect private landowners' decisions concerning land use.

Most existing studies of land use in the U.S. are based on the classic land use theory developed by David Ricardo and Johann von Thünen (Ricardo, 1817; von Thünen, 1826). This theory explains land use patterns in terms of relative rent to alternative land uses, which depends on land quality, location, and socio-economic factors. Due to data limitations, most econometric studies of land use utilized data describing areas or proportions of certain land use categories within a well defined geographic area, such as a county, as a function of socioeconomic variables and land characteristics aggregated at the level of geographic unit of observation (Alig & Healy, 1987; Plantinga, Buongiorno, & Alig, 1990; Stavins & Jaffe, 1990). Some studies modeled shares of an exhaustive set of land uses within a specified land base using a binomial or multinomial logit model, which allows restricting shares to unity (Parks & Murray, 1994; Hardie & Parks 1997; Ahn, Plantinga, & Alig, 2000; Nagubadi & Zhang, 2005; Zhang & Nagubadi, 2005).

While being widely used, the model of aggregated land use shares can have some limitations related to its ability to reveal the temporal effect of economic variables on land use and, thus, its usability for the inferences about land use dynamics. Comparing pooled, fixed effects, and random effects specifications of the cross-sectional time series land use share model at county level, Ahn, Plantinga, & Alig (2000) concluded that the pooled specification does not adequately control for cross-sectional variations in dependent variables. As a result, the models' parameters measure a combination of spatial and temporal effects and cannot be used for inferences regarding land use change or land use change predictions. They suggested that a specification with cross-sectional fixed effects provides a better measure of temporal relationship. However, the use of cross-sectional fixed effects requires relatively long time series, limits number of cross-sectional elements, and prevents the use of explanatory variables such as land quality and slope that do not have temporal variation.

A number of recent studies utilized disaggregate approach to modeling land use by using parcel level (Irwin & Bokstael, 2002; Carrión–Flores & Irwin, 2006), sample plots (Kline, Moses & Alig, 2001; Lubowski, Plantinga & Stavins, 2006), or remotely sensed (Chomitz & Gray, 1996; Turner, Wear and Flamm, 1996) data. This approach allows better using of physical attributes affecting land use decision by connecting them to specific units of observation. Furthermore, this approach allows modeling land use on a small scale, such as within one or several counties (Carrión–Flores & Irwin, 2006), which was not possible with aggregate data models.

Some of the studies utilizing disaggregate approach modeled the probability of a parcel or sample plot to be allocated to a particular land use as a function of physical and economic variables (Chomitz & Gray, 1996; Nelson et al., 2005). This model is similar to pooled models of aggregate land use shares. Another group of studies used plot- or parcel-based observation of land characteristics over at least two points in time in order to directly measure land use transitions, or, in other words, probability of a parcel to be allocated to a particular land use was modeled as a function of physical and economic variables, as well as previous land use (Bockstael, 1996; Lubowski, Plantinga & Stavins, 2006; Polyakov & Zhang, 2008). This allowed isolating temporal effects of the factors driving land use dynamics, similar to the fixed effect model of aggregate land use shares.

In this chapter, we apply a discrete choice model to disaggregate data and model land uses at a sub-watershed level in a three-county region. We compare performance of two approaches used to model land use. Our results show that a model that takes into account previous land use of sample plots (land use change model) performs better than a land use allocation model. In the next section we lay out a discrete choice model of land use allocation and a discrete choice model of land use change as well as their corresponding econometric models. The third section describes the study area and data. The fourth section provides estimation results of conditional logit models of land use allocation and land use change. The final section concludes.

2. The Models

We assume that spatial patterns of land use and their changes are results of decisions of the owners of individual land parcels or cells in the landscape. A landowner chooses to allocate a parcel of land of uniform quality to one of several possible alternative uses. We further assume that the landowner's decision is based on the maximization of net present value of future returns generated by the land. The landowner's expectation concerning future returns generated by different land uses are drawn from the characteristics of the parcel and historical returns.

Let W_{nj} be the net present value of parcel n in use j, which is not directly observable for individual parcels. However, there are observable attributes of a plot, $\mathbf{x}_n$, that are related to, and depend on, characteristics of a parcel such as land quality and location, as well as economic conditions. The landowner's utility of keeping parcel n in use j is $U_{nj} = W_{nj}(\mathbf{x}_n) + \varepsilon_{nj}$, where ε_{nj} captures factors that affect utility and assumed to be random.

The parcel will be allocated to land use with greatest utility to the landowner. Depending on assumptions about the density distribution of random components of utility, several different discrete choice models could be derived from this specification (Train 2003). Assuming random components are independent and identically distributed (IID) with a type I extreme value distribution, we obtain a conditional logit model (McFadden 1973):

$$P_{njt} = \frac{\exp(W_{nj})}{\sum_{k=1}^{J} \exp(W_{nk})} = \frac{\exp(\alpha_j + \boldsymbol{\beta}_j{}'\mathbf{x}_{n,t-1})}{\sum_{k=1}^{J} \exp(\alpha_k + \boldsymbol{\beta}_k{}'\mathbf{x}_{n,t-1})} \tag{1}$$

where P_{njt} is the probability of allocating parcel n to land use j at time t, $\boldsymbol{\beta}_j$ is a vector of land use specific parameters, and α_j are land use specific constants ($\boldsymbol{\beta}_{Jn} = 0$, $\alpha_{Jn} = 0$ to remove an indeterminacy in the model). This is a model of land use allocation similar to one used by Nelson et al. (2005).

However, the situation when the owner allocates to use j parcel of land, which is not already allocated to another productive use, is very uncommon. More frequent is the situation when the land is changed from one productive use to another, for example, converting forestland to development or agricultural land to forestry use. Let us consider the utility of converting parcel n from land use i to land use j. This involves return to old and new land uses W_{ni} and W_{nj}, as well as one time conversion cost C_{nij}, all of which depend on the particular land uses that the parcel is being converted from and to, as well as the characteristics of the parcel:

$$U_{nj|i} = W_j(\mathbf{x}_n) + \varsigma_{nj} - W_i(\mathbf{x}_n) - \upsilon_{ni} - C_{ij}(\mathbf{x}_n) - \sigma_{nij}$$

where ς_{nj}, υ_{ni}, and σ_{nij} capture factors that affect utility, and $\varepsilon_{nj|i} = \varsigma_{nj} + \upsilon_{ni} - \sigma_{nij}$ is assumed to be random. The parcel could be converted to land use j if $U_{nj|i}$ is positive. Furthermore, a parcel will be converted to the land use for which utility of conversion is greater. The parcel will remain in its current land use ($C_{nii} = 0$; $U_{ni|i} = 0$) if $U_{nj|i} < 0 \; \forall \; j \neq i$. The conditional logit model of the probability of converting parcel n from land use i to land use j is (see Polyakov & Zhang (2008) for details):

$$P_{nj,t|i,t-1} = \frac{\exp\left(\alpha_{ij} + \boldsymbol{\beta}_j{}'\mathbf{x}_{n,t-1} - \boldsymbol{\beta}_i{}'\mathbf{x}_{n,t-1}\right)}{\sum_{k=1}^{J}\exp\left(\alpha_{ik} + \boldsymbol{\beta}_k{}'\mathbf{x}_{n,t-1} - \boldsymbol{\beta}_i{}'\mathbf{x}_{n,t-1}\right)} = \frac{\exp\left(\alpha_{ij} + \boldsymbol{\beta}_j{}'\mathbf{x}_{n,t-1}\right)}{\sum_{k=1}^{J}\exp\left(\alpha_{ik} + \boldsymbol{\beta}_k{}'\mathbf{x}_{n,t-1}\right)} \quad (2)$$

where α_{ij} is a transition specific parameter.

Both models (1) and (2) predict probability of parcel n being allocated to land use j in period t. However, in model (1) this probability is conditional on the characteristics of parcel n in period t–1, while in model (2) it is conditional on the characteristics of parcel n in period t–1 and the particular land use of parcel n in period t–1. Formally, the difference between these models is the categorical variable indicating previous land use. This leads to two conceptual differences in these models. First, model (2) takes into account transition costs. As we mentioned, all the changes (or non-changes) of land use are influenced by the conversion costs, which are ignored by the model (1). Second, in both models, there are unobserved characteristics of plots, which influence returns to alternative land uses, and assumed to be random. However, in model (2), we compare differences of returns to two alternative land uses, and effects of unobserved characteristics of plots on two land uses are partially cancelled. This is similar to the fixed effect model. Thus, model (2) has a potential to perform better than model (1) in predicting future land use.

Although both these models predict probability of future land use, because model (2) takes into account previous land use, in this chapter it will be referred to as the "model of land use change", while model (1) will be referred to as the "model of land use allocation".

Due to the size of the study area, the prices and costs do not vary and therefore do not affect relative rents and land use choice behavior (Turner, Wear, & Flamm, 1996; Bockstael, 1996). The factors that are variable within the study area and influence relative rents are location of sample point relative to employment, market and population centers, transportation networks, and physical site characteristics. We hypothesize that by affecting relative rents to alternative land uses, location (accessibility to jobs, markets, and population) influences changes both between rural and developed land uses, and between agricultural and forestry uses.

Model evaluation of discrete choice models is challenging because estimated models yield probabilities of land uses, while the actual outcome is represented by a set of categorical values (Bockstael, 1996). To account for these issues, we evaluated the forecasting performance of the conditional logit model using information indices and statistics developed for evaluating performance of discrete choice models by Hauser (1978). The information index, $I(\mathbf{A};\mathbf{X})$, quantifies the additional information provided by the explanatory variables through the estimated model in comparison to a null model:

$$I(\mathbf{A};\mathbf{X})=\frac{1}{N}\sum_{n=1}^{N}\sum_{j=1}^{J}\delta_{nh}\ln\left(\frac{p(a_j \mid \mathbf{x}_n)}{p(a_j)}\right),$$

where $p(a_j)$ is the prior likelihood of the land use j (based on the null model), $p(a_j \mid \mathbf{x}_j)$ is the land use h predicted by the model, and δ_{nj} is the binary variable indicating land use j observed at sample plot n. The information index can be compared to the expected information provided by the model:

$$EI(\mathbf{A};\mathbf{X})=\frac{1}{N}\sum_{n=1}^{N}\sum_{j=1}^{J}p(a_j \mid \mathbf{x}_n)\ln\left(\frac{p(a_j \mid \mathbf{x}_n)}{p(a_j)}\right).$$

The information index $I(\mathbf{A};\mathbf{X})$ is normally distributed with a mean of $EI(\mathbf{A};\mathbf{X})$ and a variance of $V(\mathbf{A};\mathbf{X})$:

$$V(\mathbf{A};\mathbf{X})=\frac{1}{N}\sum_{n=1}^{N}\left\{\sum_{j=1}^{J}p(a_j \mid \mathbf{x}_n)\left[\ln\left(\frac{p(a_j \mid \mathbf{x}_n)}{p(a_j)}\right)\right]^2 - \left[\sum_{j=1}^{J}p(a_j \mid \mathbf{x}_n)\ln\left(\frac{p(a_j \mid \mathbf{x}_n)}{p(a_j)}\right)\right]^2\right\},$$

which allows testing the accuracy of the model.

The index of the prior entropy,

$$H(\mathbf{A};\mathbf{X}) = -\sum_{j=1}^{J} p(a_j)\ln(p(a_j))$$

defines the uncertainty inherent in the null model and allows measuring proportion of uncertainty explained by the estimated model:

$$U^2 = \frac{I(\mathbf{A};\mathbf{X})}{H(\mathbf{A};\mathbf{X})}.$$

Furthermore, the log–likelihood ratio $LLR = 2n \times I(\mathbf{A};\mathbf{X})$ is χ^2 distributed with degrees of freedom equal to the number of coefficients in the model and allows testing the significance of the empirical model.

3. Study Area and Data

Our study area is in the Georgia Piedmont, a region that is developing rapidly and ranks among the highest regions in terms of percentage increase in developed land area during the 1990s. Within this region we study land use change in three contiguous counties: Muscogee, Harris, and Meriwether. Despite being contiguous, these counties exhibit broad ranges of population pressure and patterns of land uses and land use change from urban and moderately developing (Muscogee county) to suburban and fast developing (Harris county), and to rural (Meriwether county). Columbus, located in Muscogee County, is the third largest city in Georgia. Muscogee County accounts for 80% of the population of the three county region. However, it only had a moderate population growth during the 1990s. The population of Harris County, which is located north of Muscogee County and is becoming its bedroom community, increased by one-third during the same period, while the population of Meriwether County increased only slightly (Table 1).

Table 1. Population and land use statistics for Harris, Meriwether and Muscogee counties 1990-2000 (U.S. Census Bureau, 2002; Natural Resources Conservation Service, 2000)

Characteristics		County			Total
		Harris	Meriwether	Muscogee	
Population:	Person, 2000	23,695	22,534	186,291	232,520
	Person/km2, 2000	19	17	325	75
	Annual % change, 1990-2000	3.3%	0.1%	0.4%	0.6%
Agricultural lands:	% of land base, 1997	6.3%	10.2%	5.5%	7.8%
	Annual % change, 1992-1997	-0.3%	-3.1%	-4.7%	-2.5%
Forest lands:	% of land base, 1997	78.3%	80.5%	24.8%	69.3%
	Annual % change, 1992-1997	-0.4%	0.8%	-2.1%	0.0%
Developed lands:	% of land base, 1997	6.9%	5.9%	29.8%	10.7%
	Annual % change, 1992-1997	4.6%	4.1%	3.8%	4.1%

To develop a model of land use transitions, we need information about land uses for a set of sample points for at least two points in time. We used USGS National Land Cover Datasets (NLCD) 1992, generated from leaf-on Landsat TM data centered on a nominal collection date 1992, and NLCD 2001, generated from multi-season Landsat 5 and Landsat 7 imagery centered on a nominal collection year of 2001 (USGS, 1999, 2003; also see Homer et al. (2004) for details). However, these two datasets cannot be used directly to model land use transition on a point (pixel) basis for several reasons.[2] First, these datasets use slightly different classification schemes; many land cover types of 1992 cannot be matched with land cover types of 2001. Second, the accuracy is not good enough to model land use/cover transition on a pixel basis. Finally, NLCD land cover classification does not identify clearcuts and young plantations among other (non-forest) barren/grasses/shrubs land cover types. Clearcuts and young plantations have different land cover change patterns than non-forestry barren land, grasses, or shrubs.

In order to overcome these problems, we used subset of the data, which had been visually validated using high resolution aerial photographs. The aerial photographs allow much better precision of identification of land cover types than Landsat images. We performed a systematic sampling of the original data by placing a 750–meters rectangular grid over the study area, yielding 5313 30×30 meter sample points. These sample points were assigned land cover values from the NLCD 1992 and NLCD 2001 datasets. A GIS layer with 30×30 meter polygons representing sample points was overlaid with black and white aerial orthophotos with 1 m resolution dated 1992, and with color aerial orthophotos with 0.8 m resolution dated 2003. The land cover values of sample points were then visually validated, if necessary corrected or reclassified according to the NLCD 2001 classification scheme of 21 types. For this study, we aggregated these 21 land use/cover types to Developed, Forest, Agriculture, Wetlands, Water Bodies, and Others.

There are several ways to quantify the effect of location when multiple populated places of different sizes influence each parcel of land simultaneously. Regional scientists traditionally evaluate and compare the influence of multiple populated places using gravity potential, which is proportional to the population of the populated place and inversely proportional to the squared distance between the populated place and the parcel of interest. Because the influences of multiple populated places on a given parcel are additive, Hoover (1971) suggests aggregating gravity potentials into a single index. This approach has been used by a number of land use change studies (Kline et al., 2001, 2003; Polyakov & Zhang, 2008). There are various specifications of gravity index (Song, 1996). We calculated Population Gravity Index (PGI) using the traditional specification suggested by Hoover (1971):

$$G_i = \sum_k \frac{P_k}{D_{ki}^2} \quad \forall\, k : D_{ki} \le 50,$$

where G_i is the population gravity index for sample point i, P_k is the population of census block k, and D_{ki} is the distance between sample point i and census block k in miles. The

[2] Direct comparison of NLCD 1992 and NLCD 2001 is not recommended by USGS, see

1990 and 2000 Censuses of Population census block data were taken from ESRI Data and Maps (ESRI, 1999, 2005).

Other explanatory variables used in our models are per capita income, slope and proximity to transportation network. Per capita income on the census tract level was obtained from US Census 2000 (U.S. Census Bureau, 2002). The slope affects suitability of land for development and agriculture. The value for the slope attribute is derived from the Digital Elevation Model (DEM) obtained from the Georgia Spatial Data Clearinghouse (GSDI, 2006). The factor that affects accessibility of a parcel is its proximity to a transportation network. We hypothesized that proximity to roads and highways might have different effects on relative rents to different land uses. In particular, proximity to highway may be irrelevant for rural land uses, but for the developed (residential) use could have both positive effect (need for access to transportation) and negative externality effect (noise nuisance and bad air quality). Distances from each sample plot to the nearest road and to the nearest highway are calculated using TIGER/Line spatial data from the US Census Bureau (2006).

4. Results

We modeled land use in these three counties in 2001 using the land use allocation and land use change models. Because there is virtually no transition to and from such land uses as wetlands, and water bodies, we excluded them from the consideration in both models. Transition to developed land is practically irreversible; therefore developed lands were excluded from the list of initial land use types in the second model. As a result, in our land use allocation model we consider four (*j*) land use types (developed, forest, agricultural, other), while in our land use change model we consider three initial (*i*) and four final (*j*) land use types.

The models were estimated using SAS 9.1 (SAS Institute, Inc., 2004). The results are presented in Table 2. "Other" land use type is a reference type ($\boldsymbol{\beta}_{Jn} = 0$ to remove indeterminacy in the model). The regression coefficients for the rest of the land use types indicate effects of particular variables on allocation/conversion to respective land use types relative to the reference type ("Other lands"). As expected, the land use change model has a substantially better fit than the land use allocation model, as indicated by McFadden pseudo-R^2.

To evaluate forecasting performance of our model, we conducted within-sample predictions of land uses for 2001 using both land use allocation and land use change models. We utilized two *a priori* null models of the land uses: $p_0(a_j)$ is equal probability for all land uses, and $p_1(a_j)$ is a probability proportional to the occurrence of land uses in the sample, which is equivalent to the probability predicted by the choice model using only an intercept. Information indices and statistics are presented in Table 3. The U^2 values suggest that the proportion of uncertainty explained by the empirical model is relatively high and in tests of full models versus equal probability null these indices are comparable with McFadden pseudo-R^2 (Table 2). Proportion of uncertainty explained by the land use change model vs.

http://www.mrlc.gov/multizone.php.

intercept-only null is much higher than proportions of uncertainty explained by the land use allocation model vs. intercept-only null. The analysis of information index, expected information index and its variance suggest that the empirical models are accurate, while the log-likelihood ratios (LLR) indicate that models are statistically significant.

It is difficult to interpret the coefficients in a conditional logit model because the effect of the variable on a particular transition probability is jointly determined by all the coefficients for this variable. In Table 4 we presented marginal effects of the explanatory variables on the transition probabilities.[3]

Table 2. Estimation Results of Conditional Logit Models of Land Use in West Georgia

Variables	Model of land use change				Model of land use allocation			
	Final land use (*j*):				Final land use (*j*):			
	Developed	Forestry	Agricult	Other	Developed	Forestry	Agricult	Other
Intercept					5.4216	11.0925†	25.2582‡	
Initial Forestry	-24.0534‡		-1.6864	-20.9026†				
Initial Agriculture	-27.6295‡	-6.0344		-23.1974†				
Initial Other		22.9644‡		9.9049				
Log (PGI)	0.4561	-0.5868†	-0.1611		0.4867‡	-0.6855‡	-0.7654‡	
Log (PGI rate of change)	4.0989†	-1.1492	-1.0554		0.0826	-0.8115	-0.6939	
Log Per Capita Income	0.7651	-0.9774	-1.8630*		-0.1089	-0.4960	-2.0320‡	
Log Distance to Road	-0.6473‡	-0.2786	0.0461		-0.4410‡	0.0346	0.1883	
Log Distance to Highway	-0.1613	-0.0039	-0.0380		-0.6501‡	0.3071‡	-0.0327	
Slope	-0.1780*	0.0281	-0.1922*		-0.0668	0.0905*	-0.1941‡	
McFadden Pseudo R^2	0.8706				0.5961			

* significant at 10%; † significant at 5%; ‡ significant at 1%.

Table 3. Information Indices and Statistics

Model	Test	$H(A;X)$	U^2	$I(A;X)$	$EI(A;X)$	$V(A;X)$	LLR
Land use allocation	Full model vs. equal prob.	1.39	0.60	0.82632	0.82632	0.00016	7761
	Full model vs. intercept only	0.73	0.23	0.16867	0.16864	0.00005	1584
Land use change	Full model vs. equal prob.	1.39	0.88	1.21975	1.22022	0.00010	11456
	Full model vs. intercept only	0.73	0.77	0.56210	0.56202	0.00006	5279

[3] The marginal effect of attribute m of a sample plot on the probability of transition to land use type j is $\partial P_j / \partial x_m = P_j \left(\beta_{jm} - \sum_{k=1}^{J} \beta_{km} P_k \right)$.

Table 4. Marginal Effects of Explanatory Variables in Conditional Logit Models of Land Use in West Georgia

Variables	Model of land use change				Model of land use allocation			
	Final land use:				Final land use:			
	Developed	Forestry	Agricult	Other	Developed	Forestry	Agricult	Other
Intercept					-0.2091†	-0.7723‡	1.2026‡	-0.2212†
Initial Forestry	-0.3287‡	0.4903‡	-0.0105	-0.1511†				
Initial Agriculture	-0.2962‡	0.3634‡	0.0571	-0.1243*				
Initial Other	-0.3118‡	0.6030‡	-0.1991‡	-0.0921				
Log (PGI)	0.0142‡	-0.0220‡	0.0036	0.0042*	0.0359‡	-0.0367‡	-0.0111	0.0119‡
Log (PGI rate of change)	0.0721‡	-0.0801‡	0.0001	0.0079	0.0267*	-0.0465	0.0058	0.0140
Log Per Capita Income	0.0240‡	-0.0230*	-0.0081	0.0071	0.0161	0.1012‡	-0.1286‡	0.0113
Log Distance to Road	-0.0051‡	0.0002	0.0029†	0.0021	-0.0151‡	0.0016	0.0141‡	-0.0006
Log Distance to Highway	-0.0022	0.0024	-0.0003	0.0000	-0.0283‡	0.0575‡	-0.0247‡	-0.0044†
Slope	-0.0028‡	0.0049‡	-0.0019‡	-0.0002	-0.0040‡	0.0279‡	-0.0229‡	-0.0011

* significant at 10%; † significant at 5%; ‡ significant at 1%.

Population pressure (as reflected by the marginal effects of PGI) and population growth (as indicated by the marginal effects of PGI rate of change) increased the probability of allocation and conversion to developed use, and decreased the probability of allocation and conversion to forestry use. The "elasticity" of developed lands with respect to population change in these three counties is close to 7 (4.1%/0.6%; See table 1).

Steeper slope negatively affected both allocation and conversion to agricultural land use. It also negatively affected the probability of conversion to developed use. Slope positively affected the probability of conversion to forest. Development was found to be closer to highways and roads. Proximity to roads decreases the probability of forestry use and increased the probability of agricultural use, as expected. Finally, higher per capita income was related to higher probability of development and lower probability of agricultural use/conversion to agricultural use.

The directions of marginal effects of all explanatory variables were similar in both models. However, there was a substantial difference in their magnitudes. Consider variables that represent level and change of urbanization. The marginal effects for log of population gravity index have greater absolute values in the land use allocation model than in the land use change model. The coefficients for log of PGI rate of change are greater in the land use change model, and in the land use allocation model, only the coefficient for developed use is significant. This is likely due to the fact that coefficients (and marginal effects) in allocation model reflect both spatial and temporal influences of the factors driving land use change, while coefficients (and marginal effects) in the change model reflect only temporal effects. Similar phenomenon can be seen in Ahn et al. (2000), when in model with fixed effects the magnitude of variables driving land use change (rents to alternative land uses) had lower magnitudes than in pooled model, which does not control for cross-sectional effects. Thus, model of land use allocation could potentially yield biased results when used to predict future land use.

To highlight the differences in performance of analyzed models, we mapped the forecast errors at a sub-watershed level. For this, we aggregated our in-sample projections to 12–digits hydrologic units,[4] calculated predicted proportion of land uses (based on predicted probabilities of land uses for each sample plot within hydrologic unit), and calculated actual proportions of land uses for these units. Figure 1 presents deviation of proportions of developed land use predicted by each of the models from the actual proportions in 2000 for each hydrologic unit. The map of population density is shown as a reference. Comparison of two maps shows that our model of land use allocations under-predicts developed land use close to population centers (for example, Columbus/Bibb city and Warm Springs/Woodburry/Greenville). The land use allocation model and, to a lesser extent, the land use change model over-predict developed land use on the outskirts of population concentrations (east and southeast of Columbus). Overall, the land use change model shows much better results, which is consistent with the goodness of fit statistics of conditional logit models and our expectations.

5. CONCLUSION

The use of disaggregate data in land use modeling becomes more common with the availability of quality remote sensing data because it brings a possibility to incorporate spatially referenced socioeconomic and biophysical information and allows analysis on a small scale. Unfortunately, time series of disaggregate data are often unavailable, and researchers sometimes make inferences about allocation of land to alternative land uses in the future or effects of various policies on future land use based on models that utilize cross-sectional or pooled cross-sectional data. However, extreme caution should be exercised when interpreting results of the land use modeling where cross-sectional variation is not adequately controlled for. Due to a large number of cross-sectional elements, the use of fixed effect for the purpose of controlling for cross-sectional effects is not feasible; the alternative way is to incorporate information about previous land use.

In this study we compared application of conditional logit model of land use with and without using information about previous land use at each sample plot ("land use allocation" and "land use change" models). The results have shown that the model that incorporates previous land use thus controlling for cross-sectional variation, does a much better job predicting land use. The information theory statistics reveal that the land use allocation model explained about one quarter of uncertainty, while the model of land use change explained about of three-quarters of uncertainty in in-sample prediction of 2001 land use in the study region. Furthermore, the model of land allocation that does not control for cross-sectional effects could produce biased prediction of future land use because of changing compositions of land uses.

[4] A hydrologic unit is a topographically defined area of land, the boundaries of which are ridge tops. A 12-digit hydrologic unit is a level 6 sub-watershed. The sizes of the 12-digit hydrologic units within the study area range between 3 to 12 thousand hectares.

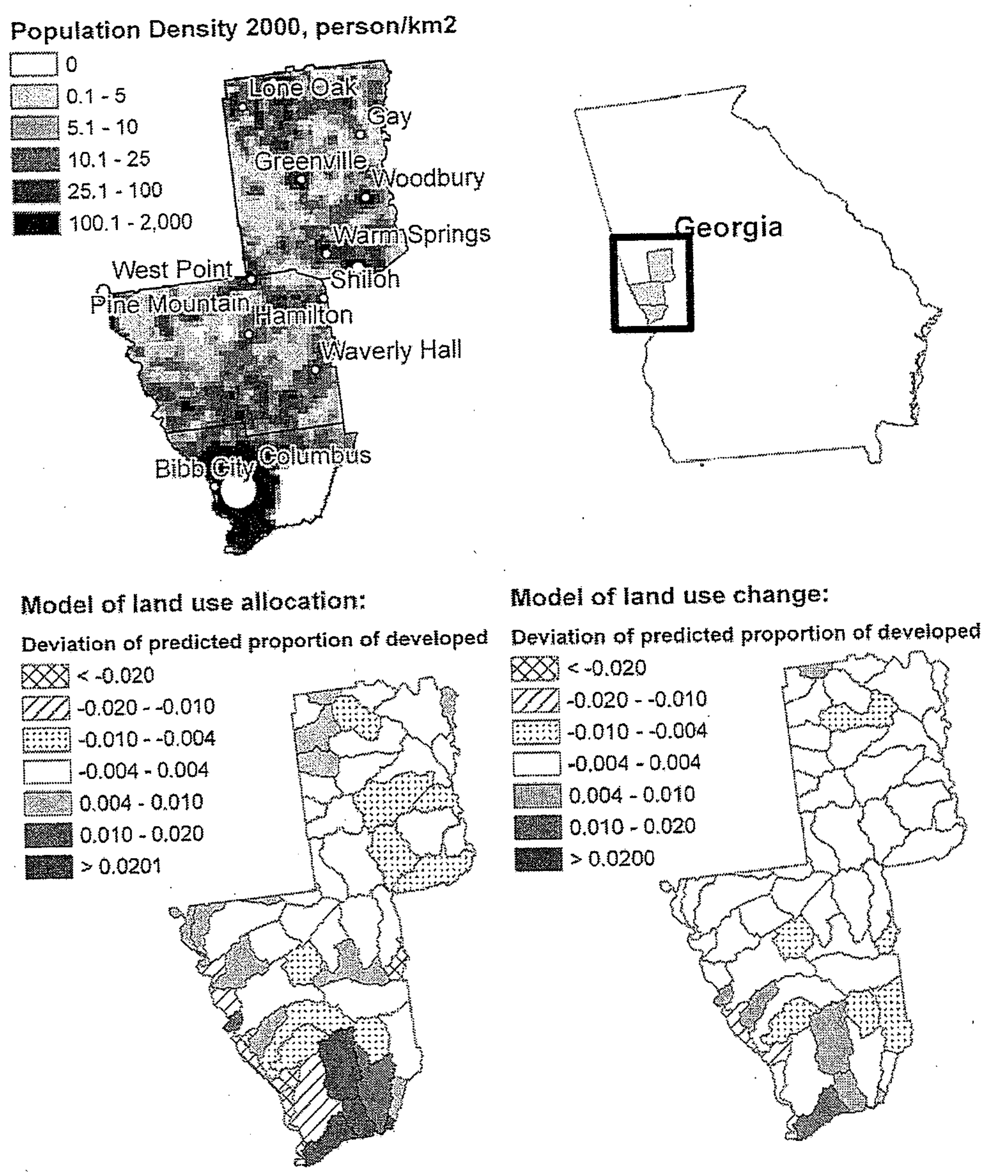

Figure 1. Spatial patterns of population density and deviation of proportions of developed lands predicted by two models from actual values at a watershed level.

There are several limitations of this study primarily due to our desire to make the models as simple as possible for the purpose of comparison. The conditional logit model assumes that the independence of irrelevant alternatives (IIA) property holds. This is a strong assumption. It can be relaxed by applying nested or random parameter logit models. Furthermore, we do not take into account zoning and other regulations, which determines, among other things, possibility and maximum density of development.

REFERENCES

Ahn, S., Plantinga, A., & Alig, R. J. (2000). Predicting future forestland area: A comparison of econometric approaches. *Forest Science, 46*(3), 363-376.

Alig, R. J., & Healy, R. G. (1987). Urban and built-up land area changes in the United States: An empirical investigation of determinants. *Land Economics, 63*(3), 215-226.

Barlowe, R. (1978). *Land Resource Economics, third ed.* Englewood Cliffs, NJ: Prentice-Hall.

Bockstael, N. E. (1996). Modeling economics and ecology: The importance of a spatial perspective. *American Journal of Agricultural Economics, 78*(5), 1168–1180.

Carrión–Flores, C. & Irwin, E. G. (2004). Determinants of Residential Land Use Conversion and Sprawl at the Rural–Urban Fringe. *American Journal of Agricultural Economics, 86*(4), 889–904.

Chomitz, K. M. & Gray, D. A. (1996). Roads, Land Use, and Deforestation: A Spatial Model Applied to Belize. *World Bank Economic Review, 10*(3), 487–512.

GSDI (Georgia Spatial Data Clearinghouse). (2006). Internet site: https://gis1.state.ga.us/. (Accessed October 14, 2006).

Hardie, I. W., & Parks, P. J. (1997). Land use with heterogeneous land quality: An application of an area-based model. *American Journal of Agricultural Economics*, *79*(2), 299–310.

Hauser, J. R. (1978). Testing the accuracy, usefulness, and significance of probabilistic choice models: An information–theoretic approach. *Operations Research*, *26*(3), 406–421.

Hoover, E. M. (1971). *An Introduction to Regional Economics*. New York, NY: Knopf.

Homer, C., Huang, C., Yang, L., Wylie, B., & Coan, M. (2004). Development of a 2001 national Land-cover database for the United States. *Photogrammetric Engineering and Remote Sensing*, 70, 829-840.

Irwin, E. G., & Bockstael, N. E. (2002). Interacting Agents, Spatial Externalities, and the Endogenous Evolution of Residential Land Use Patterns. *Journal of Economic Geography*, *2*(1), 31–54.

Kline, J. D., Moses, A., & Alig, R. J. (2001). Integrating urbanization into landscape-level ecological assessments. *Ecosystems, 4*(1), 3-18.

Lubowski, R. N., Plantinga, A. J., & Stavins, R.N. (2006). Land-use change and carbon sinks: Econometric estimation of the carbon sequestration supply function. *Journal of Environmental Economics and Management, 51*(2), 135–152.

McFadden, D. (1973). Conditional logit analysis of quantitative choice models. In P. Zarembka (Ed.) *Frontiers of Econometrics*. New York, NY: Academic Press.

Nagubadi, V. R., & Zhang, D. (2005). Determinants of timberland use by ownership and forest type in Alabama and Georgia. *Journal of Agricultural and Applied Economics*, *37*(1), 173–186.

Natural Resources Conservation Service. (2000). *1997 National Resources Inventory*. Revised December 2000. Washington, D.C.: Natural Resources Conservation Service, U.S. Department of Agriculture.

Nelson, G., De Pinto, A., Harris, V., & Stone, S. (2005). Land use and road improvements: A spatial perspective. *International Regional Science Review, 27*(3), 297–325.

Parks, P. J., & Murray, B. C. (1994). Land attributes and land allocation: Nonindustrial forest use in the Pacific Northwest. *Forest Science, 40*(3), 558-575.

Plantinga, A. J., Buongiorno, J., & Alig, R. J. (1990). Determinants of changes in non-industrial private timberland ownership in the United States. *Journal of World Forest Resource Management, 5,* 29-46.

Polyakov, M., & Zhang, D. (2008). Property tax policy and land use change. *Land Economics, 8*(3), 396–408.

Ricardo, D. (1817). *The Principles of Political Economy and Taxation*. London: John Murray.

SAS Institute, Inc. *SAS/ETS User's Guide, Version 9.1.* Cary, NC, 2004.

Stavins, R. N., & Jaffe, A. B. (1990). Unintended impacts of public investments on private decisions: The depletion of forested wetlands. *American Economic Review, 80,* 337-352.

Train, K. E. (2003). *Discrete Choice Methods with Simulation*. Cambridge, UK: Cambridge University Press.

Turner, M. G., Wear, D. N. & Flamm, R. O. 1996. Land ownership and land–cover change in the Southern Appalachian Highlands and the Olympic Peninsula. *Ecological Applications, 6*, 1150–1172.

US Census Bureau. (2006). *TIGER/Line*. Internet site: http://www.census.gov/geo/www/tiger/. (Accessed October 20, 2006).

U.S. Census Bureau. (2002). *Census of Population and Housing, 2000* [United States]: Summary File 3, Georgia [Computer file]. Washington, DC: U.S. Census Bureau [producer]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

U.S. Geological Survey (USGS). (1999). *Georgia Land Cover Data Set. Edition 1.* Sioux Falls, SD: U.S. Geological Survey ftp://edcftp.cr.usgs.gov/pub/data/ landcover/states/georgia_FGDC.txt

U.S. Geological Survey (USGS). (2003). *National Land Cover Database Zone 54 Land Cover Layer. Edition 1.0.* Sioux Falls, SD: Geological Survey. http://www.mrlc.gov

von Thünen, J. H. (1826). *Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie*, Hamburg, Germany: Perthes. English translation: *The Isolated State*, Oxford, UK: Pergammon Press.

Zhang, D. & Nagubadi, R. V. (2005). Timberland Use in the Southern United States. *Forest Policy and Economics*, 7(3), 721–731.