Land Use Change: A Spatial Multinomial Choice Analysis

Carmen E. Carrion-Flores

Alfonso Flores-Lagunes

Ledia Guci

UF UNIVERSITY of FLORIDA

CAMP RESOURCES XVI

August 13 - 14, 2009

Asheville Renaissance Hotel

Asheville, North Carolina

Motivation

- Changes in land-use patterns deserve examination
  - Accelerated expansion of urban/exurban areas
    - urban sprawl; loss of cropland and forests; etc.
  - Impacts: economic, social, environmental, global
  - Policy implications
- Analysis of land-use changes requires spatial analysis
  - Challenging in a discrete-choice framework
  - Ignoring spatial dependence leads to inconsistent estimates
  - Recent spatial estimators apply mostly to binary choice

Objectives

- Employ a spatially-explicit econometric model of land-use conversion that reflects the landowner's decision making process
  - Develop a spatial discrete-choice estimator for multiple land uses that is computationally feasible in large samples
- Apply the model to actual data
  - Data from Medina county, Ohio (Cleveland area)

Economic Model of Land-Use Conversion

- Spatially-explicit model of land-use change
  - E.g., Bockstael 1996; McMillen 1989; Irwin 2002; Irwin and Bockstael 2001; Hite et.al., 2003
- Some model assumptions:
  - A risk-neutral, price-taking landowner maximizes the present discounted value of the expected stream of future net land returns
  - Expected returns to conversion are a function of land attributes
  - Only a number of factors that affect conversion costs are observable

Economic Model of Land-Use Conversion

- Land net returns for parcel *i* in use *k:*

$$Y_{ik} = f(X_{ik}) + e_{ik}$$

where X_{ik} denotes observed parcel characteristics

- The landowner will choose to convert parcel *i* to use *l* if :

$$Y_{il}(X_{il}) > \max_{l \neq k}\{Y_{ik}(X_{ik}) - C_{ik}(Z_{ik})\}, \forall k$$

where C_{ik} is the cost of conversion and Z_{ik} denotes parcel characteristics that affect conversion costs

Economic Model of Land-Use Conversion

- The "latent" net returns are given by:

$$Y_{ik}^{*} = f_{ik}(X, Z) + e_{ik}$$

where *X* and *Z* are observed but not *Y**

- Hence, a parcel of land will remain in use *k* if:

$$Y_{ik}^{*} > Y_{il}^{*}, \forall\, k \neq l$$

Methodology

- Spatial autoregressive lag (SAL) model:

$$Y = \rho WY + X\beta + \varepsilon$$

where *W* is a spatial weighting matrix and ρ is the spatial lag parameter

ρ >0 implies clustering

ρ <0 implies dispersion

- Can be applied to land-use change decisions

Methodology

- Continuous Y:
  - Maximum likelihood methods (Anselin,1988)
  - GMM approaches (Kelijian and Prucha, 1998)
- Discrete Y:
  - Proposed estimation procedures mainly limited to the binary case (e.g., Case 1992; McMillen 1992; LeSage 2000; Pinkse and Slade 1998)
  - Many of these procedures become infeasible in large samples due to inversion of large matrices or use of simulation techniques.

Proposed Estimator: Spatial MNL

- Extend Klier and McMillen (2008) binary choice "linearized logit" model to a multinomial setting
- Simple→ linearization reduces estimation to two steps:

    1) A standard multinomial logit model

    2) Two-stage least squares

SMNL Estimator

- Model:

$$Y^* = \rho W Y^* + X\beta + \varepsilon, \quad \varepsilon \sim iid$$

$$Y^* = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1}\varepsilon$$

Covariance matrix:

$$V(e) = [(I - \rho W)'(I - \rho W)]^{-1}$$

- Define $d_{ik}=1$ if choice k is observed, =0 o/w, hence MNL model if errors are logistic
- But there is heteroskedasticity and autocorrelation in the errors unless ρ =0.

SMNL Estimator

- The probability that individual i chooses alternative k:

$$P_{ik} = P(d_{ik} = 1|X_i) = \frac{\exp\,(X_i^{**}\beta_k)}{\sum_k \exp\,(X_i^{**}\beta_k)}$$

where:

$$X_i^{**} = (I - \rho W)^{-1}\left[\frac{X_i}{\sigma_i}\right],\ \sigma_i = \sqrt{[(I - \rho W)'(I - \rho W)]^{-1}}$$

- Define the generalized MNL residuals:

$$u_{ik} = d_{ik} - P_{ik}$$

SMNL Estimator

- Idea: Use PS (1998) GMM estimator:

$$min_{\{\beta,\rho\}}\ u'ZMZ'u$$

  - Z matrix of instruments, M positive definite matrix
  - If $M=(Z'Z)^{-1}$ → non-linear TSLS
- Gradient terms: $G_{\beta ik} = P_{ik}(1 - P_{ik})X_i^{**}$

$$G_{\rho i} = P_{ik}(1 - P_{ik})\left[H_i - \frac{X_i^{**}\beta}{\sigma_i^2}\,\Lambda_{ii}\right]$$

where

$$H = (I - \rho W)^{-1}WX^{**},$$

$$\Lambda = (I - \rho W)^{-1}W(I - \rho W)^{-1}(I - \rho W)^{-1}$$

Insight by Klier and McMillen (2008)

- Linearization around $\Gamma_0 = (\beta_0, \rho_0)'$ possible since $G_{\rho i}$ is non- zero when $\rho=0$

- It simplifies the expressions of the gradients as:

$$X_i^{**} = X_i \quad \text{and} \quad \Lambda = W$$

- Linearization avoids inversion of large matrices, making this estimator feasible in large samples

SMNL Estimator

- Step 1: Estimate the model by standard multinomial logit to get estimated β's. Calculate $\hat{u}_{ik}$ and the gradient terms ($\hat{G}_{\beta_{ik}}$ and $\hat{G}_{\rho_{ik}}$).

- Step 2: Regress each gradient term on Z to get the fitted values $\hat{\hat{G}}_{\beta_{ik}}$ and $\hat{\hat{G}}_{\rho_{ik}}$. Regress $(\hat{u}_{ik} + \hat{G}'_{\beta_{ik}} \hat{\beta}_k^{MNL})$ on the fitted values of the gradient terms.

  The coefficients are the estimated values of β and ρ.

Z are the "KP" instruments [X WX W^2X W^3X …]

Monte Carlo Simulation

- Model:

$$Y = \rho WY + X\beta_k + \varepsilon, \quad k = 4$$

Dependent variable:

$$d_{ik} = 1 \text{ for } k = l \text{ if } \sum_{k=0}^{l-1} P_{ik} < u < \sum_{k=0}^{l} P_{ik}, \; P_{i0} = 0; \; u \sim U(0,1)$$

Independent variable: $X \sim U(-1,1)$

Weighting matrix (FL-S, '09): $W_{ij} = 1/(dist)_{ij}^2$

Instruments: $Z = \{X \; WX \; W^2X \; W^3X\}$

Other parameters: N=320 obs, M=1000 reps

$0<\rho<0.9$, $\beta_1=0$ and $\beta_2,\beta_3,\beta_4=1$

Simulation Results

Table 1. Simulation Results for a Sample of 320 Observations

	Standard Multinomial Logit			Spatial Multinomial Logit			
ρ	β1	β2	β3	β1	β2	β3	ρ
0							
Bias	-0.001	0.003	0.005	-0.001	0.003	0.005	-0.016
RMSE	0.080	0.084	0.080	0.080	0.085	0.080	0.248
0.1							
Bias	-0.002	0.002	0.004	-0.002	0.003	0.004	-0.077
RMSE	0.081	0.083	0.079	0.081	0.084	0.079	0.253
0.2							
Bias	-0.003	0.001	0.003	-0.004	0.002	0.003	-0.137
RMSE	0.081	0.084	0.080	0.080	0.084	0.080	0.266
0.3							
Bias	-0.006	-0.002	0.001	-0.006	-0.001	0.001	-0.188
RMSE	0.081	0.083	0.080	0.081	0.083	0.080	0.288
0.4							
Bias	-0.009	-0.007	-0.003	-0.009	-0.006	-0.003	-0.235
RMSE	0.080	0.083	0.079	0.080	0.083	0.079	0.315
0.5							
Bias	-0.019	-0.013	-0.011	-0.018	-0.012	-0.011	-0.261
RMSE	0.080	0.084	0.079	0.080	0.084	0.080	0.342

Simulation Results

Table 1. Simulation Results for a Sample of 320 Observations

	Standard Multinomial Logit			Spatial Multinomial Logit			
ρ	β1	β2	β3	β1	β2	β3	ρ
0.6							
Bias	-0.029	-0.024	-0.023	-0.028	-0.022	-0.022	-0.278
RMSE	0.082	0.085	0.080	0.082	0.086	0.080	0.362
0.7							
Bias	-0.050	-0.044	-0.044	-0.049	-0.042	-0.043	-0.258
RMSE	0.083	0.086	0.080	0.082	0.086	0.080	0.383
0.8							
Bias	-0.097	-0.093	-0.092	-0.095	-0.090	-0.090	-0.162
RMSE	0.089	0.092	0.086	0.089	0.092	0.086	0.396
0.9							
Bias	-0.241	-0.237	-0.239	-0.238	-0.233	-0.236	0.211
RMSE	0.132	0.136	0.135	0.131	0.135	0.134	0.699

Notes: Results based on 1,000 replications. See text for the specification of *W* in the SMNL model. The slope's true value is 1 in all cases.

Data

- Parcel-level data from Medina County in Cleveland, Ohio
- Data set: 1990 land use, major roads, soil type, location boundaries, socio-economic data.
- To estimate the model:

  Y: choice of land use (ag, com, res, ind)

  X: proximity to city center, distance to the nearest city, population density, housing density, proportion of surrounding land-use, and min. lot size zoning
- Six W's are used, specified as inverse Euclidean distance varying friction parameter (1 or 2) and cut-off distances (400, 800, 1600 meters).

Estimation Results

Table 2. Multinomial Logit (MNL) and Spatial Multinomial Logit (SMNL) Estimated Coefficients of Land Use Change Model

VARIABLE		MNL Results Est.	MNL Results St. Err.	SMNL Results W_400_f=1 Est.	W_400_f=1 St. Err.	W_800_f=1 Est.	W_800_f=1 St. Err.	W_1600_f=1 Est.	W_1600_f=1 St. Err.	W_400_f=2 Est.	W_400_f=2 St. Err.	W_800_f=2 Est.	W_800_f=2 St. Err.	W_1600_f=2 Est.	W_1600_f=2 St. Err.
Acres:															
	βag	0.081	0.009	0.090	0.004	0.108	0.004	0.125	0.003	0.090	0.004	0.109	0.004	0.126	0.003
	βres	-0.253	0.014	-0.645	0.072	-0.809	0.058	-0.334	0.041	-0.653	0.072	-0.823	0.057	-0.359	0.040
	βcom	0.068	0.009	0.068	0.005	0.066	0.004	0.061	0.004	0.068	0.005	0.066	0.004	0.060	0.004
Totdiscle:															
(x10,000)	βag	-0.329	0.083	-0.320	0.024	-0.314	0.024	-0.304	0.022	-0.321	0.024	-0.314	0.024	-0.307	0.022
	βres	-0.355	0.085	-0.402	0.035	-0.414	0.034	-0.345	0.032	-0.405	0.035	-0.419	0.034	-0.345	0.031
	βcom	-0.381	0.092	-0.386	0.065	-0.383	0.063	-0.371	0.059	-0.386	0.065	-0.381	0.063	-0.371	0.059
Disttonear:															
(x10,000)	βag	0.437	0.163	0.413	0.061	0.348	0.060	0.230	0.056	0.410	0.061	0.345	0.060	0.224	0.056
	βres	0.294	0.171	0.420	0.089	0.467	0.086	0.612	0.080	0.426	0.089	0.484	0.086	0.661	0.079
	βcom	1.053	0.193	1.102	0.168	1.013	0.164	1.010	0.153	1.094	0.168	0.999	0.163	0.994	0.152
Agarea:															
	βag	1.730	0.807	1.643	0.311	1.452	0.303	1.224	0.283	1.645	0.311	1.447	0.302	1.241	0.282
	βres	1.744	0.835	3.292	0.465	4.054	0.435	3.226	0.392	3.331	0.465	4.093	0.434	3.210	0.390
	βcom	2.139	1.013	2.199	0.958	2.167	0.931	1.912	0.868	2.187	0.958	2.178	0.929	1.961	0.865
Reside:															
	βag	-1.894	0.929	-1.790	0.373	-1.592	0.363	-1.156	0.339	-1.785	0.373	-1.584	0.362	-1.099	0.337
	βres	4.239	0.955	6.281	0.650	6.776	0.591	2.720	0.519	6.320	0.650	6.818	0.588	2.786	0.515
	βcom	0.363	1.142	0.517	1.040	0.611	1.011	0.500	0.945	0.527	1.039	0.683	1.010	0.603	0.942
Commarea:															
	βag	-0.007	1.499	-0.089	0.530	-0.070	0.515	0.314	0.481	-0.113	0.530	-0.047	0.514	0.413	0.479
	βres	3.726	1.515	4.938	0.667	5.156	0.638	2.734	0.589	4.925	0.667	5.137	0.637	2.809	0.586
	βcom	4.545	1.688	4.596	1.191	4.396	1.156	4.094	1.077	4.595	1.191	4.413	1.154	4.130	1.072

Estimation Results

Table 2. Multinomial Logit (MNL) and Spatial Multinomial Logit (SMNL) Estimated Coefficients of Land Use Change Model

		MNL Results		SMNL Results											
				W_400_f=1		W_800_f=1		W_1600_f=1		W_400_f=2		W_800_f=2		W_1600_f=2	
VARIABLE		Est.	St. Err.	Est.	St. Err.	Est.	St. Err.	Est.	St. Err.	Est.	St. Err.	Est.	St. Err.	Est.	St. Err.
Acres:															
	βag	0.081	0.009	0.090	0.004	0.108	0.004	0.125	0.003	0.090	0.004	0.109	0.004	0.126	0.003
	βres	-0.253	0.014	-0.645	0.072	-0.809	0.058	-0.334	0.041	-0.653	0.072	-0.823	0.057	-0.359	0.040
	βcom	0.068	0.009	0.068	0.005	0.066	0.004	0.061	0.004	0.068	0.005	0.066	0.004	0.060	0.004
Totdiscle:															
(x10,000)	βag	-0.329	0.083	-0.320	0.024	-0.314	0.024	-0.304	0.022	-0.321	0.024	-0.314	0.024	-0.307	0.022
	βres	-0.355	0.085	-0.402	0.035	-0.414	0.034	-0.345	0.032	-0.405	0.035	-0.419	0.034	-0.345	0.031
	βcom	-0.381	0.092	-0.386	0.065	-0.383	0.063	-0.371	0.059	-0.386	0.065	-0.381	0.063	-0.371	0.059
Disttonear:															
(x10,000)	βag	0.437	0.163	0.413	0.061	0.348	0.060	0.230	0.056	0.410	0.061	0.345	0.060	0.224	0.056
	βres	0.294	0.171	0.420	0.089	0.467	0.086	0.612	0.080	0.426	0.089	0.484	0.086	0.661	0.079
	βcom	1.053	0.193	1.102	0.168	1.013	0.164	1.010	0.153	1.094	0.168	0.999	0.163	0.994	0.152
Agarea:															
	βag	1.730	0.807	1.643	0.311	1.452	0.303	1.224	0.283	1.645	0.311	1.447	0.302	1.241	0.282
	βres	1.744	0.835	3.292	0.465	4.054	0.435	3.226	0.392	3.331	0.465	4.093	0.434	3.210	0.390
	βcom	2.139	1.013	2.199	0.958	2.167	0.931	1.912	0.868	2.187	0.958	2.178	0.929	1.961	0.865
Reside:															
	βag	-1.894	0.929	-1.790	0.373	-1.592	0.363	-1.156	0.339	-1.785	0.373	-1.584	0.362	-1.099	0.337
	βres	4.239	0.955	6.281	0.650	6.776	0.591	2.720	0.519	6.320	0.650	6.818	0.588	2.786	0.515
	βcom	0.363	1.142	0.517	1.040	0.611	1.011	0.500	0.945	0.527	1.039	0.683	1.010	0.603	0.942
Commarea:															
	βag	-0.007	1.499	-0.089	0.530	-0.070	0.515	0.314	0.481	-0.113	0.530	-0.047	0.514	0.413	0.479
	βres	3.726	1.515	4.938	0.667	5.156	0.638	2.734	0.589	4.925	0.667	5.137	0.637	2.809	0.586
	βcom	4.545	1.688	4.596	1.191	4.396	1.156	4.094	1.077	4.595	1.191	4.413	1.154	4.130	1.072

Estimation Results

Table 2. Multinomial Logit (MNL) and Spatial Multinomial Logit (SMNL) Estimated Coefficients of Land Use Change Model

	MNL Results		SMNL Results											
			W_400_f=1		W_800_f=1		W_1600_f=1		W_400_f=2		W_800_f=2		W_1600_f=2	
VARIABLE	Est.	St. Err.	Est.	St. Err.	Est.	St. Err.	Est.	St. Err.	Est.	St. Err.	Est.	St. Err.	Est.	St. Err.
Popdens:														
βag	0.006	0.001	0.006	0.001	0.006	0.001	0.006	0.001	0.006	0.001	0.006	0.001	0.006	0.001
βres	0.003	0.001	0.002	0.001	0.002	0.001	0.004	0.001	0.002	0.001	0.002	0.001	0.004	0.001
βcom	0.011	0.002	0.012	0.003	0.011	0.003	0.010	0.003	0.012	0.003	0.011	0.003	0.011	0.003
Housedens:														
βag	-0.018	0.003	-0.017	0.002	-0.017	0.002	-0.018	0.002	-0.017	0.002	-0.017	0.002	-0.018	0.002
βres	-0.011	0.003	-0.011	0.002	-0.012	0.002	-0.014	0.001	-0.011	0.002	-0.012	0.002	-0.013	0.001
βcom	-0.034	0.005	-0.035	0.010	-0.034	0.009	-0.032	0.009	-0.035	0.010	-0.034	0.009	-0.032	0.009
Percpinc:														
(x10,000) βag	-0.352	0.415	-0.361	0.158	-0.329	0.153	-0.276	0.143	-0.361	0.158	-0.323	0.153	-0.283	0.143
βres	0.036	0.426	0.254	0.212	0.262	0.204	-0.284	0.190	0.246	0.212	0.233	0.204	-0.285	0.189
βcom	-1.304	0.495	-1.397	0.433	-1.339	0.422	-1.229	0.394	-1.389	0.433	-1.337	0.421	-1.248	0.392
Largelot:														
βag	0.650	0.236	0.583	0.087	0.439	0.085	0.223	0.079	0.582	0.087	0.429	0.085	0.220	0.079
βres	0.144	0.247	0.579	0.156	0.816	0.148	0.495	0.134	0.601	0.156	0.834	0.148	0.511	0.134
βcom	0.828	0.266	0.832	0.179	0.804	0.174	0.761	0.162	0.829	0.179	0.800	0.174	0.759	0.161
ρ	--	--	0.145	0.009	0.213	0.008	0.406	0.008	0.145	0.009	0.214	0.008	0.401	0.008

Notes: Sample size is 9,760 parcels. All models include indicater variables for the township in which the parcel resides. The columns for the SMNL estimator correspond to different specifications of W that vary the cut-off distance (400, 800, and 1600) and the friction parameter (f=1 or 2). See text for details.

Results Discussion

- Industrial land use becomes less attractive as distance from Cleveland increases
- Local markets and population density are important determinants of ag, com, and res land uses

- Minimum lot size policy affects all land uses but relatively less industrial
- The estimates of the spatial lag parameter indicate the presence of spatial spillover effects

Conclusions

- Spatial dependence is important when analyzing land use decisions
- The proposed SMNL estimator has the following advantages:
  - Easy to estimate and feasible in large sample
  - Wide applicability in analyzing economic decisions
- Future work:
  - A more complete exploration of finite sample properties of SMNL
  - A more comprehensive analysis of the land-use change process
  - Other applications