Incorporating Spatial Dynamics and Temporal Dependency in Land Use Change Models

Raghuprasad Sidharthan
The University of Texas at Austin
Dept of Civil, Architectural and Environmental Engineering
1 University Station C1761, Austin TX 78712-0278
Phone: 512-471-4535, Fax: 512-475-8744
E-mail: raghu@mail.utexas.edu

Chandra R. Bhat*
The University of Texas at Austin
Dept of Civil, Architectural and Environmental Engineering
1 University Station C1761, Austin TX 78712-0278
Phone: 512-471-4535, Fax: 512-475-8744
E-mail: bhat@mail.utexas.edu

*corresponding author

Original version: May 19, 2011
Revised version: October 19, 2011

ABSTRACT

This paper formulates an empirical discrete land-use model within a spatially explicit economic structural framework for land-use change decisions. The underlying framework goes beyond mechanistic fitting models for the spatial process of land use change to more closely link landowner decision behavior to land use patterns. At the same time, the paper explicitly considers spatial "spillover" effects in the decisions of land-owners of proximately located parcels. These "spillover" or peer influences may be due to strategic or collaborative partnerships between land owners, and can be associated with observed variables to the analyst (such as accessibility to city centers and market places) and unobserved variables to the analyst (such as perhaps soil quality and neighborhood attitudes/politics). In addition to spatial spillover effects, it is also likely that there is heterogeneity in the decision-making process of different land owners because of differential responsiveness to various signals relevant to decision-making. This leads to a stationary across-time correlation in land uses for the same spatial unit. The paper accommodates these technical considerations by formulating a random-coefficients spatial lag discrete choice model using a fine resolution for the spatial unit of analysis. Time-varying random effects are also considered to capture the effects of time-varying unobserved factors (for instance, unobserved land owner attitudes regarding specific land uses may shift over time). The model is estimated using Bhat's (2011) maximum approximate composite marginal likelihood (MACML) inference approach. The analysis is undertaken using the City of Austin parcel-level land use database for multiple years (1995, 2000, 2003, and 2006). The estimation results indicate that proximity to highways and other roadways, distance from flood plains, parcel location in the context of existing development, and distance from schools are all important determinants of land-use. As importantly, the results provide very strong evidence of temporal dependency and spatial dynamics in land-use decisions. There is also a suggestion that major highways may not only physically partition regions, but may also act as social barriers for didactic interactions among individuals.

Keywords: spatial econometrics, spatial multipliers, discrete spatial panel, random-coefficients, land use analysis.

1. INTRODUCTION

This paper proposes a new econometric approach to specify and estimate a model of land-use change, based on the now rich theoretical literature on land use conversion decisions made by economic agents to maximize net returns (see Plantinga and Irwin, 2006). As such, the motivations of this paper stem both from a methodological perspective as well as an empirical perspective. At a methodological level, the paper focuses on specifying and estimating a multi-period multinomial probit model, accounting for observation unit-specific inter-temporal dependencies, and a spatial lag structure across observation units. The model also accommodates spatial heterogeneity in the model. The model should be applicable in a wide variety of fields where social and spatial interactions (or didactic interactions) between decision agents lead to spillover effects. The inference methodology used is the maximum approximate composite marginal likelihood (MACML) approach proposed by Bhat (2011), and is strongly motivated by the very difficult computational problems that arise from the use of a Bayesian Markov chain-Monte Carlo (MCMC) or classical maximum simulated likelihood (MSL) inference approaches. At an empirical level, the paper models the discrete indicators for the type of land-use of each spatial unit within a discrete choice model framework. The model brings together the quantitative (but aspatial or highly stylized spatial effects) perspective of land-use analysis that dominates the economic literature with the qualitative (but richer spatial dynamics and heterogeneity) perspective of land-use analysis that is quite prevalent in the ecological literature (see Irwin, 2010 for a discussion of the different perspectives of economists and ecologists in the context of urban land use change analysis). In this manner, the current paper also attempts to develop a stronger linkage between the spatial unit of analysis used in economic models of land-use change and the didactic interactions between land-owners of proximally-spaced spatial units. Thus, the empirical model is closely tied to the underlying theoretical underpinnings of the land-use model.

The next section discusses the econometric context for the current paper, while the subsequent section presents the empirical context.

1.1. The Econometric Context

In the past decade, there has been increasing attention in discrete choice modeling on accommodating spatial dependence across decision agents or observational units to recognize the potential presence of diffusion effects, social interaction effects, or unobserved location-related influences (see Jones and Bullen, 1994, and Miller, 1999). Specifically, spatial lag and spatial error-type structures developed in the context of continuous dependent variables to accommodate spatial dependence (see, for instance, Dubin, 1998, Cho and Rudolph, 2007, Messner and Anselin, 2004, Anselin, 2006, Elhorst, 2010ab, Lee and Yu, 2010) are being considered for discrete choice dependent variables (see reviews of this literature in Franzese *et al*. 2010, Brady and Irwin, 2011, and Bhat *et al*., 2010a). But almost all of this research focuses on binary or ordered response choice variables by applying global spatial structures to the linear (latent) propensity variables underlying the choice variables (for example, see Fleming, 2004, Franzese and Hays, 2008, Franzese *et al*., 2010, and LeSage and Pace, 2009). The two dominant techniques, both based on simulation methods, for the estimation of such spatial binary/ordered discrete models are the frequentist recursive importance sampling (RIS) estimator (which is a generalization of the more familiar Geweke-Hajivassiliou-Keane or GHK simulator; see Beron and Vijverberg, 2004) and the Bayesian Markov Chain Monte Carlo (MCMC)-based estimator (see LeSage and Pace, 2009). However, both of these methods are confronted with multi-dimensional normal integration, and are cumbersome to implement in typical empirical contexts with moderate to large estimation sample sizes (see Bhat, 2011 and Smirnov, 2010).[1]

The RIS and MCMC methods become even more difficult to implement in a spatial unordered multinomial choice context because the likelihood function entails a multidimensional integral of the order of the number of observational units factored up by the number of alternatives minus one (in the case of multi-period data, as in the current paper, the integral dimension gets factored up further by the number of time periods of observation). Thus, it is no surprise that there has been little research on including spatial dependency effects in unordered choice models. However, Bhat (2011) suggested a maximum approximate composite marginal likelihood (MACML) for spatial multinomial probit (MNP) models that is easy to implement, is based on a frequentist likelihood-based approach, and requires no simulation. The MACML estimation of spatial MNP models involves only univariate and bivariate cumulative normal distribution function evaluations, regardless of the number of alternatives or the number of choice occasions per observation unit, or the number of observation units, or the nature of

[1] The reader is referred to Franzese *et al*., 2010 for an excellent discussion of the many challenges that arise with the frequentist RIS or Bayesian MCMC procedures in spatial discrete choice models.

social/spatial dependence structures. In this paper, we use Bhat's MACML inference approach to estimate a spatial MNP model with random coefficients as well as temporal dependence.

There are four precursors of the current research that are worth noting. The recent studies by Carrión-Flores *et al*. (2009) and Smirnov (2010) superimposed a spatial lag structure over a multinomial logit (MNL) model. Carrión-Flores *et al.* estimated the resulting spatial model using a linearized version of Pinkse and Slade's (1998) Generalized Method of Moments (GMM) approach (as proposed by Klier and McMillen, 2008 for the binary choice model), while Smirnov employed a pseudo-maximum likelihood (PML) estimator to obtain model parameters. Smirnov's PML estimator is essentially based on estimating the spatial autoregressive term in the spatial lag model by recognizing the implied heteroscedasticity generated by the spatial correlation, while ignoring the spatial correlation across observational units. The approaches of Carrión-Flores *et al*. and Smirnov simplify inference by avoiding multidimensional integration. However, they are both based on a two-step instrumental variable estimation technique after linearizing around zero interdependence, and so work well only for the case of large estimation sample sizes and weak spatial dependence. Chakir and Parent (2009) estimated a multinomial probit model of land-use change, similar to the empirical focus of the current paper. However, they employed a Bayesian MCMC method, which requires extensive simulation, is time-consuming, is not straightforward to implement, and can create convergence assessment problems.[2] Sener and Bhat (2011) allowed spatial error dependence in a multinomial logit model of choice, but their approach is not applicable to a spatial lag structure. The reader will also note that none of the above studies consider random coefficients to account for spatial heterogeneity and temporal dependence effects.

1.2. The Empirical Context

There are several approaches to studying and modeling land-use change. Irwin and Geoghegan (2001) and Irwin (2010) provide a good taxonomy of these approaches. In the current paper, we derive our empirical discrete choice model based on an economic structural framework for land-

[2] Franzese *et al.* (2010) and LeSage and Pace (2009) point out a mistake in the original MCMC method proposed for spatial probit models by LeSage (2000). Essentially, the earlier LeSage (2000) study provided the false perception that Bayesian MCMC was simpler and faster than frequentist methods, because LeSage inadvertently used a univariate truncated normal distribution in translating the latent variables to the observed variables, while a multivariate truncated normal distribution is needed for this purpose. The net result is that the Bayesian MCMC "parallels the computation intensity of the classical RIS strategy" (Franzese *et al*., 2010).

use change decisions within a spatially explicit framework. This underlying framework goes beyond mechanistic fitting models for the spatial process of land use change to more closely linking landowner decision behavior to land use patterns. At the same time, we explicitly consider spatial dynamics (caused by interdependence among individual landowners) that lead to the land-use decisions of one landowner affecting that of the landowners of proximally located properties. To elucidate, consider landowners as being economic agents who make forward-looking inter-temporal land use decisions based on profit-maximizing behavior regarding the conversion of a parcel of land to some other economically viable land use (for example, see Capozza and Li, 1994). The stream of returns from converting a parcel from the current land-use to some other land-use has to be weighed against the costs entailed in the conversion from the current land-use to some other land-use. The premise then is that the land use at any time will correspond to the land use type with the highest present discounted sum of future net returns (stream of returns minus the cost of conversion). Some of the factors affecting the stream of returns and the cost of conversion (and, therefore, the net returns) will be observed (such as road accessibility, distance from flood plain, and the availability and quality of amenities), while others will not. Thus, the net returns may be considered as a latent variable that includes a systematic component and an unobserved component. In addition, spatial interactions are likely to naturally arise because land owners of proximately located spatial units (say, parcels) are likely to be influenced by each other's perceptions of net returns from a certain land-use type investment. These peer influences may be due to strategic or collaborative partnerships between land owners associated with observed variables to the analyst (such as accessibility to city centers and market places) and unobserved variables to the analyst (such as perhaps soil quality and neighborhood attitudes/politics). Such spatial interactions can be captured by relating the latent continuous "net returns" from each land-use type for a parcel (as perceived by the land owner of that parcel) with the corresponding latent "net returns" from surrounding parcels (as perceived by the land owners of those surrounding parcels) using a spatial lag formulation.[3] But,

[3] Interestingly, many spatial formulations for land-use modeling have considered spatial interactions to be a "nuisance" issue, and have employed a spatial error structure. However, didactic and related interactions between land owners require the use of a spatial lag structure that allows spillover effects, as also suggested by Carrión-Flores *et al*. (2009). Further, more generally, and as emphasized by McMillen (2010), it is much easier to justify a parametric spatial lag structure when accommodating spatial dependence, while the use of a parametric spatial error structure is "troublesome because it requires the researcher to specify the actual structure of the errors". Beck *et al.* (2006) also find theoretical and conceptual issues with the spatial error model and refer to it as being "odd", because the formulation rests on the "hard to defend" position that "space matters in the error process but not in the

in addition to the spatial lag-based interaction effect just discussed, it is also likely that there is heterogeneity in the decision-making process of different land owners because of differential responsiveness to various signals relevant to decision-making. For instance, different land owners may perceive the effects of market place proximity on the net returns differently based on their individual experiences, risk-taking behavior, and even vegetation conservation values. This would then translate to a land owner-specific random coefficients formulation for the "net returns", leading to a stationary across-time correlation in land uses for the same spatial unit. Such land owner-specific random coefficients and resulting temporal correlations of the land-owner's choices across time have been ignored thus far in the literature. Some earlier studies have considered a generic time-stationary random effect (that is, a random coefficient only on the intercept) for each spatial unit in their spatial error formulation, but such a formulation is restrictive relative to the more general random-coefficients spatial lag formulation used here. In addition to such a general time-stationary random-coefficients effect, there may also be time-varying correlation effects for landowners in their assessment of net returns. Such effects may be due to personality characteristics (such as, say risk averseness or risk acceptance behavior) that fade over time or recent personal experiences.

The implementation of the economic land use change framework discussed above is facilitated by the recent public availability of longitudinal and high resolution spatial land-use data (collected using aerial photography, remote-sensing, and real-estate appraisal information), which enables the modeling of land use at a fine spatial level such as a parcel. In particular, the observed land use data for each spatial unit is in the form of categorical data. Also, the choice of land use is mutually exclusive. Thus, the theoretical "net returns" land use change framework leads naturally to an empirical discrete choice model at a very fine level of spatial resolution (see Bockstael, 1996, Carrión-Flores and Irwin, 2004, Chakir and Parent, 2009, and Carrión-Flores *et al.*, 2009). In such a model, the "net returns" concept is replaced by an "instantaneous utility" of each landowner to have a spatial unit in a certain land use type. This utility is a function of exogenous variables and unobserved variables, and the land use observed at a spatial unit corresponds to the one with highest utility. While earlier studies have used such a cross-sectional

---

substantive portion of the model". As they point out, the implication is that if a new independent variable is added to a spatial error model "so that we move it from the error to the substantive portion of the model", the variable magically ceases to have a spatial impact on neighboring observations. Overall, we submit that land-use models should be developed using the spatial lag formulation or its many variants, and by explicitly linking land owner decision behavior to land use patterns.

discrete choice model, no earlier land-use study that we are aware of has considered and applied a discrete choice formulation that simultaneously accommodates the spatial dynamics through a spatial lag structure, spatial heterogeneity through spatial-unit specific random coefficients, time-varying as well as time-stationary unobserved components extracted from multiperiod observations on the same spatial units, as well as a flexible contemporaneous covariance structure across the utilities of the different land use type alternatives.

2. MODELING METHODOLOGY

2.1. Model Formulation

Let the instantaneous utility U_{qti} obtained by the landowner of parcel q (q = 1, 2, …, Q) at time t (t = 1, 2, …, T) with land use i (i = 1, 2, …, I) be a function of a (K × 1)-column vector of exogenous attributes x_{qti} . This utility is spatially interdependent across landowners (due to spillover effects based on spatial proximity of parcels) as well as has a temporally interdependent component (due to unobserved factors specific to each landowner). Thus, we write the utility U_{qti} using a spatial lag structure as follows:

$$U_{qti} = \delta \sum_{q'} w_{qq'} \, U_{q'ti} + \tilde{\alpha}_{qi} + \boldsymbol{\beta}'_q \mathbf{x}_{qti} + \tilde{\varepsilon}_{qti} \tag{1}$$

where $w_{qq'}$ is the usual distance-based spatial weight corresponding to units q and q' (with $w_{qq} = 0$ and $\sum_{q'} w_{qq'} = 1$) for each (and all) q, δ $(0 < \delta < 1)$ is the spatial lag autoregressive parameter, $\tilde{\alpha}_{\mathbf{qi}}$ is a normal random-effect term capturing time-stationary preference effects of the landowner of parcel q for land use i, and $\boldsymbol{\beta}_{\mathbf{q}}$ is a parcel-specific (K×1)-vector of coefficients assumed to be a realization from a multivariate normal distribution with mean vector **b** and covariance $\tilde{\boldsymbol{\Omega}} = \mathbf{LL}'$. It is not necessary that all elements of $\boldsymbol{\beta}_{\mathbf{q}}$ be random; that is, the analyst may specify fixed coefficients on some exogenous variables in the model, though it will be convenient in presentation to assume that all elements of $\boldsymbol{\beta}_{\mathbf{q}}$ are random. For later use, we will write $\boldsymbol{\beta}_{\mathbf{q}} = \mathbf{b} + \breve{\boldsymbol{\beta}}_{\mathbf{q}}$, where $\breve{\boldsymbol{\beta}}_q \sim MVN_K(0, \tilde{\boldsymbol{\Omega}})$ (MVN_K represents the multivariate normal distribution of dimension K). Also, for later use, we will write $\tilde{\alpha}_{\mathbf{qi}} = \tilde{a}_i + \breve{\alpha}_{qi}$, and let the mean and variance-covariance matrix of the vertically stacked (I×1)-vector of random-effect terms

$\tilde{\boldsymbol{\alpha}}_{\boldsymbol{q}}$ $\left[=\left(\tilde{\alpha}_{q1}, \tilde{\alpha}_{q2}, ..., \tilde{\alpha}_{qI}\right)'\right]$ be $\tilde{\mathbf{A}}$ and $\tilde{\boldsymbol{\Lambda}}$, respectively. $\tilde{\varepsilon}_{qti}$ in Equation (1) is a normal error term uncorrelated with $\tilde{\boldsymbol{\beta}}_{\boldsymbol{q}}$ and all $\tilde{\alpha}_{\boldsymbol{qi}}$ terms (i = 1, 2, …, I), and also uncorrelated across observation units q. However, the $\tilde{\varepsilon}_{qti}$ terms may have a covariance (dependency) structure across land uses i (due to unobserved factors at time t that simultaneously increase or simultaneously decrease the utility of certain types of land uses) and also a covariance structure across time to recognize time-varying preference effects of the landowner of parcel q. For the time varying effects, it is reasonable to consider that the dependency effects fade over time, and so we consider a first order autoregressive temporal dependency process: $\tilde{\varepsilon}_{qti} = \rho \tilde{\varepsilon}_{q,t-1,i} + \tilde{\eta}_{qti}$, with ρ $\left(0 < \rho < 1\right)$ being the temporal autoregressive parameter. The error term $\tilde{\eta}_{\boldsymbol{qti}}$ is temporally uncorrelated, but can be correlated across alternatives - $\tilde{\boldsymbol{\eta}}_{\boldsymbol{qt}}$ $\left[=\left(\tilde{\eta}_{qt1}, \tilde{\eta}_{qt2}, ..., \tilde{\eta}_{qtI}\right)'\right] \sim MVN_I(0, \tilde{\boldsymbol{\Psi}})$. As usual, appropriate scale and level normalization must be imposed on $\tilde{\mathbf{A}}$, $\tilde{\boldsymbol{\Lambda}}$ and $\tilde{\boldsymbol{\Psi}}$ for identifiability. Specifically, only utility differentials matter in discrete choice models. Take the utility differentials with respect to the first alternative. Then, only the elements $\alpha_{qi1} = \tilde{\alpha}_{qi} - \tilde{\alpha}_{q1}$ $(i \neq 1)$ and its covariance matrix $\boldsymbol{\Lambda}_1$, and the covariance matrix $\boldsymbol{\Psi}_1$ of $\eta_{qti1} = \tilde{\eta}_{qti} - \tilde{\eta}_{qt1}$ $(i \neq 1)$, are estimable. However, as discussed in Bhat (2011), the MACML inference approach, like the traditional GHK simulator, takes the difference in utilities against the chosen alternative during estimation. Thus, consider that land use m_{qt} exists at parcel q at time t. This implies that values of $\alpha_{qim_{qt}} = \tilde{\alpha}_{qi} - \tilde{\alpha}_{qm_{qt}}$ $(i \neq m_{qt})$, and the covariance matrices $\boldsymbol{\Lambda}_{m_{qt}}$, and $\boldsymbol{\Psi}_{m_{qt}}$ are desired for parcel q at time t. However, though different random effects differentials and different covariance matrices are used for different parcels and different time periods, all of these must originate in the same values of the undifferenced error term vector $\tilde{\mathbf{A}}$ and covariance matrices $\tilde{\boldsymbol{\Lambda}}$ and $\tilde{\boldsymbol{\Psi}}$. To achieve this consistency, we normalize $\tilde{\alpha}_{q1} = 0 \ \forall \ q$. This implies that $\tilde{a}_1 = 0$. Also, we develop $\boldsymbol{\Lambda}$ from $\boldsymbol{\Lambda}_{\boldsymbol{1}}$ by adding an additional row on top and an additional column to the left. All elements of this additional row and additional column are filled with values of zeros. Similarly, we construct $\boldsymbol{\Psi}$ from $\boldsymbol{\Psi}_1$ by adding a row on top and a column to the left. This first row and the first column of the matrix $\tilde{\boldsymbol{\Psi}}$ are also filled with zero values. However, an additional normalization needs to be imposed on $\tilde{\boldsymbol{\Psi}}$ because the scale is also not identified. For this, we normalize the element of $\tilde{\boldsymbol{\Psi}}$ in the second row and second column to the value of one. Note that all these normalizations do not place any restrictions, and a fully general specification is the result. But they are needed for econometric identification.

We now set out notation to write the likelihood function in a compact form. Define the following:

$\mathbf{U}_{qt} = (U_{qt1}, U_{qt2}, ..., U_{qtI})'$, $\tilde{\boldsymbol{\varepsilon}}_{qt} = (\tilde{\varepsilon}_{1t1}, \tilde{\varepsilon}_{qt2}, ..., \tilde{\varepsilon}_{qtI})'$, $\tilde{\boldsymbol{\eta}}_{qt} = (\tilde{\eta}_{qt1}, \tilde{\eta}_{qt2}, ..., \tilde{\eta}_{qTI})'$ ($I \times 1$ vectors),

$\mathbf{U}_{q} = (\mathbf{U}'_{q1}, \mathbf{U}'_{q2}, ...\mathbf{U}'_{qT})'$, $\tilde{\boldsymbol{\varepsilon}}_{q} = (\tilde{\boldsymbol{\varepsilon}}'_{q1}, \tilde{\boldsymbol{\varepsilon}}'_{q2}, ...\tilde{\boldsymbol{\varepsilon}}'_{qT})'$, $\tilde{\boldsymbol{\eta}}_{q} = (\tilde{\boldsymbol{\eta}}'_{q1}, \tilde{\boldsymbol{\eta}}'_{q2}, ...\tilde{\boldsymbol{\eta}}'_{qT})'$ ($TI \times 1$ vectors),

$\mathbf{U} = (\mathbf{U}'_{1}, \mathbf{U}'_{2}, ...\mathbf{U}'_{Q})'$, $\tilde{\boldsymbol{\varepsilon}} = (\tilde{\boldsymbol{\varepsilon}}'_{1}, \tilde{\boldsymbol{\varepsilon}}'_{2}, ...\tilde{\boldsymbol{\varepsilon}}'_{Q})'$, $\tilde{\boldsymbol{\eta}} = (\tilde{\boldsymbol{\eta}}'_{1}, \tilde{\boldsymbol{\eta}}'_{2}, ...\tilde{\boldsymbol{\eta}}'_{Q})'$ ($QTI \times 1$ vectors),

$\breve{\boldsymbol{\alpha}}_{q} = (\breve{\alpha}_{q1}, \breve{\alpha}_{q2}, ..., \breve{\alpha}_{qI})'$ ($\mathrm{I} \times 1$ vector), $\breve{\boldsymbol{\alpha}} = \left[(1_T \otimes \breve{\alpha}_1)', (1_T \otimes \breve{\alpha}_2)', ...(1_T \otimes \breve{\alpha}_Q)'\right]'$ ($QTI \times 1$ vector), $\mathbf{x}_{qt} = (\mathbf{x}_{qt1}, \mathbf{x}_{qt2}, ..., \mathbf{x}_{qtI})'$ ($I \times K$ matrix), $\mathbf{x}_{q} = (\mathbf{x}'_{q1}, \mathbf{x}'_{q2}, ..., \mathbf{x}'_{qT})'$ ($TI \times K$ matrix), $\mathbf{x} = (\mathbf{x}'_{1}, \mathbf{x}'_{2}, ..., \mathbf{x}'_{Q})'$ ($QTI \times K$ matrix), and $\breve{\boldsymbol{\beta}} = \left(\breve{\boldsymbol{\beta}}'_1, \breve{\boldsymbol{\beta}}'_2, ..., \breve{\boldsymbol{\beta}}'_Q\right)'$ ($QK \times 1$ vector) . Let $\mathbf{IDEN}_{\boldsymbol{E}}$ be the identity matrix of size E, $\mathbf{1}_E$ be a column vector of size E with all of its elements taking the value of one, and $\mathbf{1}_{EE}$ be a square matrix of size E with all unit elements. Also, define the following matrices:

$$\mathbf{R} = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} (T \times T \text{ matrix}), \ \tilde{\mathbf{x}} = \begin{bmatrix} \mathbf{x}_1 & 0 & 0 & 0. & \cdots & 0 \\ 0 & \mathbf{x}_2 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{x}_3 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \mathbf{x}_Q \end{bmatrix} (QTI \times QK \text{ matrix}), \quad (2)$$

$\mathbf{S} = \left[\mathbf{IDEN}_{QTI} - \left\{(\delta \mathbf{W} \otimes \mathbf{IDEN}_T) \otimes \mathbf{IDEN}_I\right\}\right]^{-1}$ ($QTI \times QTI$ matrix), $\mathbf{W}$ is the ($Q \times Q$) weight matrix with the weights $w_{qq'}$ as its elements, and $\mathbf{C} = \left[\mathbf{IDEN}_{QTI} - \mathbf{IDEN}_Q \otimes \left(\rho \mathbf{R} \otimes \mathbf{IDEN}_I\right)\right]^{-1} = \mathbf{IDEN}_Q \otimes \left(\left[\mathbf{IDEN}_{TI} - \left(\rho \mathbf{R} \otimes \mathbf{IDEN}_I\right)\right]^{-1}\right)$ ($QTI \times QTI$ matrix). Then, we can write Equation (1) in matrix notation as:

$$\mathbf{U} = \mathbf{S}\left[\left(\mathbf{1}_{QT} \otimes \tilde{\mathbf{A}}\right) + \mathbf{x}\mathbf{b} + \breve{\boldsymbol{\alpha}} + \tilde{\mathbf{x}}\breve{\boldsymbol{\beta}} + \mathbf{C}\tilde{\boldsymbol{\eta}}\right] \quad (3)$$

Let $[.]_e$ indicate the e^{th} element of the column vector [.], and let $d_{qti} = (q-1)TI + (t-1)I + i$. Equation (3) can be equivalently written as:

$$U_{qti} = \left[\mathbf{S}\left\{\left(\mathbf{1}_{QT} \otimes \tilde{\mathbf{A}}\right) + \mathbf{xb}\right\}\right]_{d_{qti}} + \left[\mathbf{S}\left\{\breve{\boldsymbol{\alpha}} + \tilde{\mathbf{x}}\breve{\boldsymbol{\beta}} + \mathbf{C}\tilde{\boldsymbol{\eta}}\right\}\right]_{d_{qti}} \tag{4}$$

Define $V_{qti} = \left[\mathbf{S}\left\{\left(\mathbf{1}_{QT} \otimes \tilde{\mathbf{A}}\right) + \mathbf{xb}\right\}\right]_{d_{qti}}$ and $\varepsilon_{qti} = \left[\mathbf{S}\left\{\breve{\boldsymbol{\alpha}} + \tilde{\mathbf{x}}\breve{\boldsymbol{\beta}} + \mathbf{C}\tilde{\boldsymbol{\eta}}\right\}\right]_{d_{qti}}$. The landowner of parcel q chooses the land use at time t that provides maximum utility. As earlier, let the land use of parcel q at time t be m_{qt}. In the utility differential form, we may write Equation (4) as:

$$y_{qtim_{qt}} = U_{qti} - U_{qtm_{qt}} = H_{qtim_{qt}} + \xi_{qtim_{qt}};\ H_{qtim_{qt}} = V_{qti} - V_{qtm_{qt}} \ \ and\ \ \xi_{qtim_{qt}} = \varepsilon_{qti} - \varepsilon_{qtm_{qt}};\ i \neq m_{qt} \tag{5}$$

Then stack the utility differentials $y_{qtim_{qt}}\ (= U_{qti} - U_{qtm_{qt}}, i \neq m_{qt})$ in the following order: $\mathbf{y}_{qt} = (y_{qt1m_{qt}}, y_{qt2m_{qt}}, ..., y_{qtIm_{qt}})'$, an $(I-1)\times 1$ vector; $\mathbf{y}_{q} = (\mathbf{y}'_{q1}, \mathbf{y}'_{q2}, ..., \mathbf{y}'_{qT})'$, an $\left[(I-1)\times T\right]\times 1$ vector; and $\mathbf{y} = (\mathbf{y}'_1, \mathbf{y}'_2, ..., \mathbf{y}'_Q)'$, an $\left[(I-1)\times T \times Q\right]\times 1$ vector. Correspondingly, let $\mathbf{H}_{qt} = (H_{qt1m_{qt}}, H_{qt2m_{qt}}, ..., H_{qtIm_{qt}})'$, an $(I-1)\times 1$ vector; $\mathbf{H}_{q} = (\mathbf{H}'_{q1}, \mathbf{H}'_{q2}, ..., \mathbf{H}'_{qT})'$, an $\left[(I-1)\times T\right]\times 1$ vector; and $\mathbf{H} = (\mathbf{H}'_1, \mathbf{H}'_2 ..., \mathbf{H}'_Q)'$, an $\left[(I-1)\times T \times Q\right]\times 1$ vector. It is easy to see that $\mathbf{y}$ has a mean vector **H.** To determine the covariance matrix of $\mathbf{y}$, several additional matrix definitions are needed. Define $\boldsymbol{\Lambda} = \mathbf{IDEN}_Q \otimes \left(\mathbf{1}_{TT} \otimes \tilde{\boldsymbol{\Lambda}}\right)$ ($QIT \times QIT$ matrix), $\boldsymbol{\Omega} = \tilde{\mathbf{x}}(\mathbf{I}_Q \otimes \tilde{\boldsymbol{\Omega}})\tilde{\mathbf{x}}'$ ($QTI \times QTI$ matrix), and $\boldsymbol{\Psi} = \mathbf{IDEN}_{QT} \otimes \tilde{\boldsymbol{\Psi}}$ ($QTI \times QTI$ matrix). Let $\tilde{\mathbf{F}} = \mathbf{S}\left[\boldsymbol{\Lambda} + \boldsymbol{\Omega} + \mathbf{C}\boldsymbol{\Psi}\mathbf{C}'\right]\mathbf{S}'$ and define **M** as an $[(I-1)\times T \times Q]\times[I \times T \times Q]$ block diagonal matrix, with each block diagonal having $(I-1)$ rows and I columns corresponding to the t^{th} observation time period on parcel q. This $(I-1)\times I$ matrix for parcel q and observation time period t corresponds to an $(I-1)$ identity matrix with an extra column of -1's added as the m_{qt} $^{\text{th}}$ column. For instance, consider the case of Q = 2, T = 2, and I = 4. Let parcel 1 be observed to be in land-use 2 in time period 1 and in land-use 1 in time period 2, and let parcel 2 be in land-use 3 in time period 1 and in land-use 4 in time period 2. Then **M** takes the form below.

$$\mathbf{M} = \left[\begin{array}{cccc|cccc|cccc|cccc}
1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1
\end{array}\right] \quad (6)$$

Finally, we obtain the multivariate distribution of the utility differentials $\mathbf{y} : \mathbf{y} \sim \boldsymbol{MVN}(\mathbf{B}, \boldsymbol{\Sigma}),$ where $\boldsymbol{\Sigma} = \mathbf{M}\tilde{\mathbf{F}}\mathbf{M}'$. Next, let $\boldsymbol{\theta}$ be the collection of parameters to be estimated: $\boldsymbol{\theta} = \left[\mathbf{b}'; \text{Vech}(\tilde{\boldsymbol{\Omega}}); \tilde{\mathbf{A}}', \text{Vech}(\tilde{\boldsymbol{\Lambda}}), \text{Vech}(\tilde{\boldsymbol{\Psi}}), \delta, \rho\right]'$, where Vech($\tilde{\boldsymbol{\Omega}}$) represents the row vector of upper triangle elements of $\tilde{\boldsymbol{\Omega}}$. Then, the likelihood of the observed sample may be written succinctly as Prob[$\mathbf{y}^* < \mathbf{0}$].

$$L_{ML}(\boldsymbol{\theta}) = \text{Prob}[\mathbf{y}^* < 0] = F_{Q\times T\times(I-1)}(-\mathbf{B}, \boldsymbol{\Sigma}) \quad (7)$$

where $F_{Q\times T\times(I-1)}$ is the multivariate cumulative normal distribution of $Q \times T \times (I-1)$ dimensions. Despite advances in simulation techniques and computational power, the evaluation of such a high dimensional integral is literally infeasible using traditional frequentist and Bayesian simulation techniques. For instance, in frequentist methods, where estimation is typically undertaken using pseudo-Monte Carlo or quasi-Monte Carlo simulation approaches (combined with a quasi-Newton optimization routine in a maximum simulated likelihood (MSL) inference), the computational cost to ensure good asymptotic estimator properties can be prohibitive as the number of dimensions of integration increases (see Bhat *et al.*, 2010b for a detailed discussion of frequentist simulation procedures and problems under high integration dimensionality). Similar problems arise in Bayesian Markov Chain Monte Carlo (MCMC) simulation approaches, which remain cumbersome, require extensive simulation, are time consuming, and pose convergence assessment problems as the number of dimensions increases (see Müller and Czado, 2005, Ver Hoef and Jansen, 2007, and Franzese *et al*., 2010 for discussions).

In a recent paper, Bhat (2011) proposed a maximum approximate composite marginal likelihood (MACML) approach for multinomial probit models, which is used in the current paper. The MACML inference approach is briefly discussed next.

2.2. The Maximum Approximate Composite Marginal Likelihood (MACML) Approach

The MACML approach combines a composite marginal likelihood (CML) estimation approach with an approximation method to evaluate the multivariate standard normal cumulative distribution (MVNCD) function. The composite likelihood approach replaces the likelihood function with a surrogate likelihood function of substantially lower dimensionality, which is then subsequently evaluated using an *analytic approximation* method rather than simulation techniques. This combination of the CML with the specific analytic approximation for the MVNCD function is effective because it involves only univariate and bivariate cumulative normal distribution function evaluations, regardless of the spatial and/or temporal complexity of the model structure. The approach is able to recover parameters and their covariance matrix estimates quite accurately and precisely because of the smooth nature of the first and second derivatives of the approximated analytic log-likelihood function (unlike the non-smooth first and second derivatives that arise in simulation approaches). The MVNCD approximation method is based on linearization with binary variables (see Bhat, 2011).

The MACML approach, similar to the parent CML approach (see Varin *et al*., 2011 for a recent review of CML approaches), represents a conceptually and pedagogically simple simulation-free procedure relative to simulation techniques. The approach may be explained in a simple manner as follows. In the current empirical context, instead of developing the likelihood of the entire sample, consider developing a surrogate likelihood function that is the product of the probability of easily computed marginal events. For instance, one may compound (multiply) pairwise probabilities of parcel q being in land use i at time t *and* being in land use j at time s, of parcel q being in land use i at time t *and* parcel q' being in land use j at time s, and so on and so forth. The CML estimator is then the one that maximizes the compounded probability of all pairwise events (see Varin and Vidoni, 2009, Engle *et al*., 2007, Bhat *et al*., 2010b, and Bhat and Sener, 2009 for earlier applications of the estimator for binary and ordered-response systems).

Alternatively, the analyst can also consider larger subsets of observations, such as triplets or quadruplets or even higher dimensional subsets (see Engler *et al*., 2006 and Caragea and Smith, 2007). However, doing so in the MNP context defeats the purpose of the approach because it leads to high dimensionality of integration, especially when the number of alternatives is high. Besides, it is generally agreed that the pairwise approach is a good balance between statistical and computational efficiency. The properties of the general CML estimator may be derived using the theory of estimating equations (see Cox and Reid, 2004). Specifically, under usual regularity assumptions (Molenberghs and Verbeke, 2005, page 191), the CML estimator is consistent and asymptotically normal distributed (this is because of the unbiasedness of the CML score function, which is a linear combination of proper score functions associated with the marginal event probabilities forming the composite likelihood). The CML function may be written as:

$$L_{CML}(\boldsymbol{\theta}) = \prod_{q=1}^{Q} \prod_{q'=q}^{Q} \prod_{t=1}^{T} \prod_{t'=t}^{T} \text{Prob}\,(C_{qt} = m_{qt}, C_{q't'} = m_{q't'}) \text{ with } q \neq q' \text{ when } t = t', \quad (8)$$

where C_{qt} is an index for the land use in which parcel q is at time t. Each of these pairwise probabilities is of $(I-1)\times 2$ dimensions, which may be computed easily using the MVNCD approximation method embedded in the MACML method (the MVNCD function approximates the pairwise probabilities in Equation (8) using only univariate and bivariate cumulative normal distribution functions; see Bhat, 2011).[4]

The pairwise marginal likelihood function of Equation (8) comprises $QT(QT-1)/2$ pairs of pairwise probability computations. But, in a spatial-temporal case where spatial dependency drops quickly with inter-observation distance, the pairs formed from the closest observations provide much more information than pairs that are very far away. In fact, as demonstrated by Varin and Vidoni (2009), Bhat *et al*. (2010a), and Varin and Czado (2008) in different empirical contexts, retaining all pairs may reduce estimator efficiency. We examine this

[4] It should be noted that while the MVNCD approximation of Equation (8) provides a relatively simple objective function to be maximized with respect to the parameters, the resulting function can theoretically have multiple maxima (note, however, that this is also true of the likelihood function of pretty much every other multinomial discrete choice model except the multinomial logit model). There is no way out of this "multiple maxima" situation, other than for the analyst to test various different starting points and see whether the parameters converge to the same point. We undertook such an analysis with some of the data sets generated as part of testing whether the MACML procedure is able to recover parameters (see next section), and found that the parameters always converged to the same point. Of course, this does not mean that there are no multiple optima, because it is impossible to test the infinite number of possible starting parameter spaces; but our testing does suggest reasonable stability of the maximization procedure.

issue by creating different distance bands (including the band that includes all pairings) and, for each specific distance band, considering only those unordered pairings in the CML function that are within the distance band. Then, we develop the asymptotic variance matrix $V_{CML}(\hat{\boldsymbol{\theta}})$ for each distance band and select the threshold distance value (say $\tilde{d}_{thresh}$) that minimizes the total variance across all parameters as given by $tr[V_{CML}(\hat{\boldsymbol{\theta}})]$ (*i.e.*, the trace of the matrix $[V_{CML}(\hat{\boldsymbol{\theta}})]$.[5]

The CML estimator of $\boldsymbol{\theta}$ is consistent and asymptotically normal distributed with asymptotic mean $\boldsymbol{\theta}$ and covariance matrix given by the inverse of Godambe's (1960) sandwich information matrix (see Zhao and Joe, 2005):

$$V_{CML}(\hat{\boldsymbol{\theta}}) = [G(\boldsymbol{\theta})]^{-1} = [H(\boldsymbol{\theta})]^{-1} J(\boldsymbol{\theta}) [H(\boldsymbol{\theta})]^{-1}, \text{ where} \quad (9)$$

$$H(\boldsymbol{\theta}) = E\left[-\frac{\partial^2 \log L_{CML}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} \right] \text{ and}$$

$$J(\boldsymbol{\theta}) = E\left[\left(\frac{\partial \log L_{CML}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} \right) \left(\frac{\partial \log L_{CML}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'} \right) \right].$$

The "bread" matrix $H(\boldsymbol{\theta})$ of Equation (9) can be estimated in a straightforward manner using the Hessian of the negative of the MACML likelihood function, evaluated at the MACML estimate $\hat{\boldsymbol{\theta}}$. On the other hand, the "vegetable" matrix $J(\boldsymbol{\theta})$ is not that straightforward to estimate. But the decaying nature of the distance weight matrix can be used to create pseudo-independent subsamples of the data using the windows sampling method proposed by Heagerty and Lumley (2000). Based on this windows sampling method, Bhat (2011) suggests overlaying the spatial region under consideration with a square grid providing a total of D internal and external nodes. Then, select the observational unit closest to each of the D grid nodes to obtain D observational units from the original Q observational units ($\tilde{d} = 1, 2, 3, \ldots, D$). Let $\tilde{\mathbf{C}}$ be a $Q \times D$ matrix with its $\tilde{d}^{th}$ column filled with a $Q \times 1$ vector of 0s and 1s, with a zero value in the q'^{th} row (q' =1, 2, …, Q) if the observational unit q' is not within the specified threshold distance $\tilde{d}_{thresh}$ of unit $\tilde{d}$,

[5] We do not test different time period bands like we do distance bands. This is because, unlike the spatial dependency pattern, the temporal dependency includes a time-invariant component that does not fade over time. Thus, for each observation unit, all time periods need to be considered, unless the time-stationary dependency is negligible, which we generally do not believe will be the case.

and a one otherwise (by construction, $\tilde{\mathbf{C}}_{q'\tilde{d}} = 1$ if $q' = \tilde{d}$). Also, let $\breve{\mathbf{C}} = 1_T \otimes \tilde{\mathbf{C}}$. Then, the columns of $\breve{\mathbf{C}}$ provide pseudo-independent sets of observational units.[6] Let the score matrix corresponding to the pairings in column $\tilde{d}$ of matrix $\breve{\mathbf{C}}$ be $S_{CML,d}(\boldsymbol{\theta})$. Also, Let $N_{\tilde{d}}$ be the sum of the $\tilde{d}^{th}$ column of $\breve{\mathbf{C}}$, and let $\tilde{W}$ be the total number of pairings used in the CML function of Equation (8) (after considering the distance threshold $\tilde{d}_{thresh}$). Then, the $\mathbf{J}$ matrix maybe empirically estimated as:

$$\mathbf{J} = \frac{\tilde{W}}{D}\left[\sum_{d=1}^{D}\left[\frac{1}{N_d}\left([S_{CML,d}(\boldsymbol{\theta})][S_{CML,d}(\boldsymbol{\theta})]'\right)_{\hat{\theta}}\right]\right]. \tag{10}$$

One additional issue regarding estimation. The analyst needs to ensure the positive definiteness of the three covariance matrices, $\tilde{\boldsymbol{\Omega}}, \tilde{\boldsymbol{\Lambda}},$ and $\tilde{\boldsymbol{\Psi}}$. Once this is ensured, and as long as $0 < \rho < 1$ and $0 < \delta < 1$, $\boldsymbol{\Sigma}$ will be positive definite. In our estimation, the positive-definiteness of each of the $\tilde{\boldsymbol{\Omega}}, \tilde{\boldsymbol{\Lambda}},$ and $\tilde{\boldsymbol{\Psi}}$ matrices is guaranteed by writing the logarithm of the pairwise-likelihood in terms of the Cholesky-decomposed elements of these matrices, and maximizing with respect to these elements of the Cholesky factor. Essentially, this procedure entails passing the Cholesky elements as parameters to the optimization routine, constructing the covariance matrix internal to the optimization routine, then computing $\boldsymbol{\Sigma}$, and finally picking off the appropriate elements of the matrix for the pairwise likelihood components. To ensure the constraints on the autoregressive terms ρ and δ, we parameterize these terms as $\rho = 1/[1+\exp(\tilde{\rho})]$ and $\delta = 1/[1+\exp(\tilde{\delta})]$, respectively. Once estimated, the $\tilde{\rho}$ and $\tilde{\delta}$ estimates can be translated back to estimates of ρ and δ.

2.3. Simulation Study

We have undertaken a simple simulation exercise to examine the ability of the MACML estimation approach to recover the parameters in the context of a four-alternative choice situation

[6] As indicated by Bhat (2011), there needs to be a balance here between the number of sets of pairings *D* and the proximity of points. The smaller the value of *D,* the less proximal are the sets of observation units and more likely that the sets of observational pairings will be independent. However, at the same time, the value of *D* needs to be reasonable to obtain a good empirical estimate of *J*, since this empirical estimate is based on averaging the cross-product of the score functions (computed at the convergent parameter values) across the *D* sets of observations.

(I = 4) with four time periods (T = 4) (this scenario matches with the dimensions of the empirical study in this paper). A total of Q = 200 observation units are assumed (the observation units correspond to parcels in the case of the empirical application in the current paper). Three independent variables are used. The details are available in an online supplementary note at http://www.caee.utexas.edu/prof/bhat/ABSTRACTS/LandUse/Supp_Note.pdf. The simulation results illustrate the ability of the MACML method to recover the true parameters remarkably well for the spatial lag unordered response model with temporal autocorrelation. Future studies should more extensively examine the performance of the MACML estimation approach under alternative spatial and temporal dependency patterns, as well as investigate estimator efficiency considerations.

3. APPLICATION

3.1. The Data and the Context

The data used in this paper is from the City of Austin, Texas. Parcel level land use inventory data for the years 1995, 2000, 2003 and 2006 are used. This data is available in the Environmental Systems Research Institute's (ESRI's) shape file format for all the four years for a 2 mile extraterritorial jurisdiction (ETJ) of the City of Austin, covering a total area of 1795 sq km. (693 sq mile). The land use type for each parcel is available at a fine level of detail; however, for the current study, they are aggregated into four mutually exclusive land use categories. These are (1) residential (including single family, duplexes, three/four-plexes, apartments, condominiums, mobile homes, group quarters, and retirement housing), (2) commercial (including commercial, office, hospitals, government services, educational services, cultural services, and parking), (3) industrial (including manufacturing, warehousing, resource extraction (mining), landfills, and miscellaneous industrial), and (4) undeveloped (including open and undeveloped spaces, preserves, parks, golf courses, and agricultural open spaces).

An area measuring 23.5 sq km (9.06 sq miles) in the suburbs of Austin city is selected for the analysis. The interstate highway, IH-35, divides this analysis area into an eastern section with two-thirds of the total area and a western section with the remaining one-third of the area. A part of the eastern section falls within the City of Pflugerville, a suburban Austin city. Mopac (Loop 1), another major expressway in Austin, also runs in the North-South direction about half a mile west of the western boundary of the analysis area. Apart from IH-35 and Mopac, three minor arterials and a major arterial pass through the analysis area. All the explanatory variables were created from the GIS data obtained from the City of Austin, except the flood plains data, which was obtained from the Capital Area Council of Governments (CAPCOG).

For the econometric analysis, the area is divided into 400 square cells each of size 242m×242m. The land use in the parcel at the centroid of each cell is designated as the land use for that grid cell. If the centroidal point falls well within the right-of-way of an arterial roadway or other roadways with high land-use access functionality, the corresponding grid cell is assigned the predominant land use of the adjacent area. However, if the centroidal point of a grid cell falls within the right-of-way of IH-35, which primarily serves the functionality of through movement, we removed the corresponding grid cell from analysis (a total of five grid cells were accordingly removed, leaving a sample of 395 grid cells observed at each of four time points).

In the rest of this paper, and for ease in presentation, we will use the terms "grid cell" and "parcel" interchangeably, though the analysis is technically being conducted at the grid cell level. The explanatory variables for each parcel considered in the model include road access measures (distance to IH-35, distance to Mopac, distance to the nearest non-freeway roadway, and interactions of these variables), location relative to the flood plains, an interaction term of proximity to road access with proximity to the flood plain (distance to nearest road divided by distance to the nearest flood plain), being situated in Pflugerville city, and proximity to schools.[7] To construct distances (all measured in kilometers) from each parcel to the roadways, a road network data in polyline format (obtained from the City of Austin) was overlaid on the analysis area, and the Euclidean distance from the parcel to roadways was calculated. To construct distances from each parcel to the nearest flood plain, the flood plain data in polygon format (obtained from the Capital Area Council of Governments) was overlaid onto the analysis area, and Euclidean distances were computed from each parcel centroid to the nearest floodplain polygon. School data was available as point data, and this was overlaid on the analysis area to obtain the distance from a parcel to the nearest school.

Among the exogenous variables considered, we expect that land-owners of parcels in close proximity to highways will most likely invest their parcels in commercial and industrial land-uses. On the other hand, one can expect parcels located far from highways and roadways to

[7]A floodplain is an area susceptible to flooding. Such areas in the United States are identified by the Federal Emergency Management Agency (FEMA) in its Flood Insurance Rate Maps, which show spatial regions likely to be affected by a 100-year flood (1% chance of a flood of this magnitude during the year).

remain undeveloped, as land-owners are not likely to see much net returns in developing these parcels. Similarly, we can expect parcels in close proximity to flood plains not to be built up. In addition, we consider an interaction effect of distance to the nearest roadway divided by distance to the nearest flood plain. This captures the potential "push-pull" non-linear positive effect (on the propensity of a parcel being undeveloped) of being afar from roadways *and* being proximal to a flood plain. However, the land-owner of a parcel that is distant from roadways may see some "net returns" potential in developing the parcel if the parcel is also far away from the flood plains. Similarly, the land-owner of a parcel that is close to a flood plain may still invest the parcel in some kind of development if the parcel is close to roadways. All of these effects are captured by introducing the "distance to nearest roadway divided by distance to the nearest flood plain" variable. The Pflugerville city dummy variable is introduced to capture the effects of a differential development/tax incentive structure in Pflugerville relative to the remainder of the analysis region. Finally, the proximity to schools is likely to be an incentive to develop the parcel for residential land-use.

Figure 1 shows the analysis area along with the roadways in the region, the boundary of Pflugerville, the locations of the flood plains and the land-use type of each grid point for the year 1995 (see the legend for land-use type at the bottom right of the figure). Figure 2 is the corresponding figure for the year 2000. Several observations may be made just from a visual scan of the figures. First, there is a clustering of parcels in industrial and commercial land-uses immediately adjacent to IH-35. Second, there are more parcels in an undeveloped state as one goes eastwards, away from Mopac. Third, parcels close to the floodplains indeed are more likely to be in an undeveloped land-use state. Fourth, the share of parcels within Pflugerville city in commercial land-use appears higher than in other areas of the analysis region. Fifth, while residences are not necessarily closely clustered around each school, there is a tendency to have quite a few residential parcels within a reasonable range of schools (this visual scan suggests the need to test distance bands from parcels to schools rather than a simple continuous representation of distance from school). Sixth, there is clear evidence of parcels with the same land-use in close proximity, reinforcing the notion of spatial dynamics at play. This effect is particularly obvious when looking at each of the eastern and western sections of the area (as delineated by IH-35) individually. Seventh, one can see how the clustering effect of similar land-uses manifests itself in the change from 1995 to 2000. Specifically, it can be clearly observed from Figures 1 and 2 that many parcels in an undeveloped state in 1995 are in a developed state in 2000. Most of these conversions are to residential land-use, though there also is a clear surge in industrial land-use in 2000. As can be noted, there is a distinct clustering pattern in parcels that change from an undeveloped state to each of the residential and industrial land-use types.

Table 1 shows the percentage shares of parcels in each of the four land use types at each of the four years of analysis. A high share of the parcels is either in residential or undeveloped land-uses, with the commercial and industrial land-uses representing about 20% of the total share. Another observation from this table, also visible from Figures 1 and 2, is the boom in residential land development that occurred between 1995 and 2000 (and the reduction in the share of undeveloped land during the same period). This boom is consistent with ground reality in the Austin region (Glaeser *et al.*, 2006). Historically speaking, Austin, like other cities in Texas, has had relatively weak land use zoning policies. Thus, the economic prosperity of the late 90s (and into the first year of the new millennium) led to substantial and relatively uncontrolled development in the Austin area, resulting in the emergence of several low density residential enclaves at the fringes of the main city (such as the area considered in this paper). Of course, this growth tapered off and came to a literal standstill after 2001 (see also Table 1), attributable to the economic recession that began around March 2001 rather than to any land-use regulations.[8]

3.2. Variable Specification and Spatial Weight Matrix Formulation

Many different variable specifications, functional forms, and variable interactions were considered to determine the final model specification. The roadway access variables (distance from IH-35, Mopac, and other arterials) as well as the distance to the closest flood plain polygon were considered both in linear and non-linear forms (such as the logarithm of distance, the square of distance, and spline variables that allow piece-wise linear effects of distance on the utilities). In addition, we also considered dummy variables for different ranges of distance for these variables (for instance, parcel is within 200 meters of IH-35, parcel is within 300 meters of

[8] To be sure, Austin has had comprehensive development plans since 1928, including the Austin Tomorrow Comprehensive Plan (ATCP) adopted by the City Council in the late 1970s. The ATCP was not acted upon due to lack of consensus and shifts in the Austin City Council make-up over the years, furthered by the non-involvement of the Austin development community. The ATCP was resurrected in 2008 with interim updates. In the meantime, the Austin City also established an explicit and streamlined land use planning process with significant public participation to develop a future land use map (FLUM), which provides the framework for zoning regulations (see City of Austin, 2008). But these issues are not relevant for the period of analysis in the current paper.

IH-35, *etc.*). The Pflugerville city location dummy variable was introduced as a switch variable taking the value of ‘1’ for parcels within the City of Pflugerville and ‘0’ otherwise. The proximity to school effect was considered similar to the other continuous variables, and included alternative functional forms of distance from the nearest school as well as dummy variables for different ranges of distance from school (such as parcel is located within 300 meters of a school and within 1 kilometer of a school). In addition to the variables just discussed, we also included a “1995 dummy variable” to capture the rather substantial temporal shifts in shares among the land-use categories between this first year and the subsequent years (see Table 1). Further, various interactions of the continuous and the categorical variables were also considered whenever adequate observations were available to test such interaction effects.

The final model specification was obtained after extensive explorations and testing, and based on statistical fit, intuitiveness, parsimony considerations, and the preliminary insights offered by the visual scan of Figure 1 (as discussed in the previous Section). Specifically, in terms of statistical fit, we used the adjusted composite likelihood ratio test (AD*CLRT*) statistic (see Pace *et al*., 2011 and Bhat, 2011) to compare nested models and the composite likelihood information criterion (CLIC) introduced by Varin and Vidoni (2005) to test non-nested models. Table 2 provides the descriptive statistics of the independent variables in the final model specification. We also examined alternative specifications for the construction of the spatial weights, including inverse distance and the inverse of the square of distance, the inverse of exponential distance, a simple contiguity indicator, and a contiguity weight but based on shared boundary length rather than a simple indicator. Further, based on the insights from the visual scan of Figure 1, we decided to test two spatial variants for the weight specification. The first was to develop the weight matrix between any two parcels over the entire analysis region. The second was to assume no spatial dependency between parcels on the western and eastern sections of the analysis area (as determined by IH-35), but assuming dependence between parcels within each of the two sections. That is, if two parcels are located on the same side of IH-35, then the spatial weight for the pair is non-zero based on the weight matrix; otherwise, the spatial weight for the pair is assigned a value of zero. At the end of this extensive testing, which was undertaken using all pairwise interactions in the CML function, the best weight specification involved the inverse of the square of distance specification with spatial dependency confined to parcels within each of the western and eastern sections of the analysis area (with no dependence between parcels lying on opposite sides of I-35). This selection from among the many non-nested weight specifications was undertaken using the composite likelihood information criterion (CLIC) introduced by Varin and Vidoni (2005), which takes the following form:

$$CLIC = \log L_{CML}(\hat{\boldsymbol{\theta}}) - tr\left[\hat{\mathbf{J}}(\hat{\boldsymbol{\theta}})\hat{\mathbf{H}}(\hat{\boldsymbol{\theta}})^{-1}\right] \tag{11}$$

where $\hat{\boldsymbol{\theta}}$ represents the estimated model parameter vector, and $\hat{\mathbf{J}}(\hat{\boldsymbol{\theta}})$ and $\hat{\mathbf{H}}(\hat{\boldsymbol{\theta}})$ are the estimated "vegetable" and "bread" matrices as discussed in Equations (9) and (10), respectively. The model that provides a higher value of CLIC is preferred. For instance, Table 3 provides the values of the log-composite likelihood at convergence $\log L_{CML}(\hat{\boldsymbol{\theta}})$, the trace value in the CLIC statistic ($tr(\mathbf{J}(\hat{\boldsymbol{\theta}})\mathbf{H}(\hat{\boldsymbol{\theta}})^{-1})$), and the CLIC statistic value for the model that constructs the weight matrix over the entire region (Full region-based weight matrix model) and the model that constructs the weight matrix over each of the eastern and western sections of the analysis region (Partitioned region-based weight matrix model), with the preferred inverse of the square of distance as the basis for the weight matrix. As can be observed from the CLIC statistic column, the partitioned region-based spatial weight matrix model is superior in representing spatial effects in the current empirical context, indicating the lack of didactic interactions between land-owners of parcels on either side of IH-35. This result emphasizes the social separation that can be caused by a physical barrier such as a freeway.[9]

Finally, using the preferred combination of the variable specification, and the partitioned weight matrix with the inverse of the square of distance as the separation measure, we undertook an efficiency analysis to determine the optimal distance band for including pairwise interactions in the CML function, based on minimizing the trace of the variance-covariance matrix given by $tr[\mathbf{V}_{CML}(\hat{\boldsymbol{\theta}})]$ (see Section 2.2). The $tr[\mathbf{V}_{CML}(\hat{\boldsymbol{\theta}})]$ value was the lowest for a distance band of 400 meters (other distance bands considered included 800, 1200 and 7000 meters, the last one representing the case of including all pairs of parcel-year observations in the CML function).

[9] One of the reviewers encouraged us to develop a more rigorous confidence level-based procedure to test this social barrier hypothesis, and suggested a bootstrapping of the CLIC statistic. To do so, we generate 100 data sets using the estimated values for the partitioned region-based weight matrix model (PRWM) model. Then, for each data set, we estimate the full region-based weight matrix model (FRWM) and the PRWM model, and subsequently obtain the corresponding CLIC statistics (say, CLIC-FRWM and CLIC-PRWM). In 75% of the bootstrap-generated data sets, we obtained CLIC-PRWM>CLIC-FRWM, providing confidence that the "social barrier" finding is not simply an artifact of sampling. We would like to thank the referee for suggesting that we pursue such an effort.

Thus, all subsequent results for models including spatial dependency are based on the 400 meters distance band.

The next section discusses the results of the following two models in more detail: (1) the multinomial probit model with no temporal and spatial dependencies or the MNP model (in the notation of Section 2.1, this model imposes the restrictions that $\tilde{\mathbf{\Omega}} = \mathbf{LL}'$ is a $K \times K$-matrix of zero values, $\tilde{\mathbf{\Lambda}}$ is an $I \times I$-matrix of zeros, $\rho = 0$, and $\delta = 0$), and (2) the multinomial probit model with temporal and spatial dependencies (MNPTS). In both of these models, we could not reject the null hypothesis that, after accommodating the exogenous variables, the covariance matrix $\tilde{\mathbf{\Psi}}$ had the structure below:

$$\tilde{\mathbf{\Psi}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0.5 & 0.5 \\ 0 & 0.5 & 1 & 0.5 \\ 0 & 0.5 & 0.5 & 1 \end{bmatrix}, \tag{12}$$

which is equivalent to the specification that the intrinsic utility preferences are independent and identically distributed across the four alternatives (with the scale normalized to 0.5). However, note that the MNPTS model does incorporate both dependence and heteroscedasticity across the overall utilities of the alternatives because of the random coefficients on the exogenous variables. Finally, in the MNPTS model, we could not reject the hypothesis that the covariance matrix $\tilde{\mathbf{\Lambda}}$ had the following form (the utilities are arranged in the following order of land-use type: residential, commercial, industrial, and undeveloped):

$$\tilde{\mathbf{\Lambda}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \Lambda_{com} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \Lambda_{und} \end{bmatrix}, \tag{13}$$

The covariance matrix above indicates that there are no time-stationary random effects in the utilities for the residential and industrial land-uses. More intuitively speaking, land-owners are likely to have intrinsic (unobserved and randomly distributed) time-invariant utility “biases” (or preferences) for commercial and undeveloped land-use types, but not for residential and industrial land-use types. This also implies that the utilities for commercial and undeveloped land-use types are correlated across time due to time-invariant land-owner preferences.

3.3. Model Estimation Results

The results of the MNP and the MNPTS models are presented in Table 4. We first discuss the effects of variables on the utilities of alternatives (Section 3.3.1), next the temporal and spatial effects (Section 3.3.2), then the model fit comparisons (Section 3.3.3), and finally the variable magnitude effects (Section 3.3.4). A '-' entry in a cell of Table 4 indicates that the corresponding "row" variable did not have a statistically significant effect on the utility of the corresponding "column" land-use category.

3.3.1. Variable Effects on Utility of Alternatives

The estimated coefficients of the two models in Table 4 are not directly comparable, since the scales of the error terms in the utilities are different. But the mean coefficient estimates are the same in sign in both models. All the results are consistent with the hypotheses in Section 3.1. The constant terms do not have any substantive interpretations, and simply represent adjustments in the utilities of alternatives after accommodating the other variables in the model. The presence of standard deviations on the constants for the commercial and undeveloped land-uses (in the MNPTS model) indicates time-invariant preference heterogeneity across landowners in the utilities for these land-uses, as discussed earlier. Parcels located proximal to IH-35 are more likely to be invested in commercial and industrial land-uses, though the functional form of proximity to IH-35 in the utilities of these two land-uses takes different forms. For commercial land-use, the proximity to IH-35 enters as a distance band of 350 meters from IH-35, which is consistent with the clustering of commercial parcels close to IH-35 in Figure 1. However, for industrial land-use, the linear form of distance to IH-35 (interacted with distance to nearest non-freeway road) enters the utility function, again consistent with the relative scatter of industrial parcels around IH-35. More generally, industrial facilities (and therefore their land-owners) gain from proximity to freeways. At the same time, zoning setback guidelines can preclude owners of parcels that are immediately adjacent to freeways from investing their land in industrial use (which is why the distance band specification did not come out statistically significant for the industrial land-use alternative). Also, land-owners of parcels close to other major roads can benefit from placing their land in industrial use because of improved transportation accessibility. These behaviors are captured by the negative coefficient for the industry land use category on the interaction variable of the distance to IH-35 and the distance to the nearest non-freeway road.

The results in Table 4 also indicate that parcels farther away from Mopac are more likely to be in an undeveloped state. Mopac is a major expressway connecting the analysis area to the Austin Central Business District (CBD), so it is not surprising that land-owners of parcels located closer to Mopac are more likely to develop their parcels, while land-owners of parcels far away from Mopac may not see the value in developing their land (see Carrión-Flores and Irwin, 2004 and Chakir and Parent, 2009, who also discuss how proximity and access to central metropolitan areas and major roadways can impact land-use decisions). The "push-pull" non-linear effect of distance to the nearest road and distance to the nearest flood plain is clear from the positive coefficient on the ratio of these two variables. Parcels situated within Pflugerville city, according to the MNP model, provide high "net returns" (relative to parcels outside Pflugerville) if invested in residential or commercial land-uses (particularly the latter) rather than being undeveloped or invested in industrial land-use. However, according to the MNPTS model, on average, parcels within Pflugerville are less likely (relative to parcels outside Pflugerville) to be in residential land-use than being undeveloped or in industrial use. However, there is substantial heterogeneity in this effect, as can be observed from the large estimated standard deviation of the random coefficient on this "Parcel lies within Pflugerville City" variable for the residential land-use alternative. The mean and standard deviation effects on the variable indicate that, for 47.4% of the land-owners of the parcels in the City of Pflugerville, the utility of investing in residential land-use is higher than the utility of leaving the land undeveloped or investing in industrial land-use; for the remaining 52.6% of land-owners of parcels in the City of Pflugerville, the reverse situation holds. Such heterogeneity is a natural result of the tension between the urban amenities (access to retail places and public services such as hospitals) on the one hand that may increase the demand for residential development in already dense residential areas, and the urban "disamenities" (such as traffic congestion effects and air quality problems) on the other hand that may decrease demand for residential development in already dense residential neighborhoods (see Anas *et al.*, 1998; Carrión-Flores and Irwin, 2004 and Irwin and Bockstael, 2002), But, consistent with the MNP model, the MNPTS model also shows a higher propensity of parcels within Pflugerville City to be invested in commercial land-use than invested in industrial land-use or left undeveloped (see Carrión-Flores *et al.*, 2009). Also, as expected, the proximity to schools is likely to be an incentive to develop the parcel for residential land-use (see Li and Liu, 2007). Finally, the dummy variable for 1995 shows the lower share of parcels in residential land-use and the higher share of parcels in undeveloped land-use in 1995 relative to the other years, as highlighted earlier in Section 3.1.

3.3.2. Temporal and Spatial Dependency Effects

Temporal dependency (across years) is introduced in our model in the utilities of each alternative for the same land-owner through time-invariant utility preferences and sensitivities to variables (as captured by the random coefficients specification on the constants and the "parcel lies within Pflugerville City" dummy variable in Table 4), as well as through the time-varying autoregressive error correlation structure to represent land-owner characteristics that may fade over time (as captured by the autoregressive coefficient ρ). As already indicated in the earlier section, the results show the presence of time-invariant dependency in the utilities for the same land-owner. In addition, Table 4 shows a statistically significant and moderate-level autoregressive coefficient of 0.367, indicating the presence of land-owner specific unobserved factors (such as risk averseness or risk acceptance for specific land-use types) that change over time (due to recent events or experiences, or due to lifecycle-related changes). Ignoring these time-varying effects will, in general, lead to inconsistent estimates (due to ignoring the heteroscedasticity generated by these time-varying effects) as well as inefficient estimates (due to ignoring the dependence across the land-use choice occasions of individuals).

The spatial autoregressive parameter in the spatial lag formulation, δ, also turns out to be highly statistically significant with a value of 0.449. This is evidence of the presence of spatial spillover effects caused by didactic interactions between land-owners of proximately located spatial units. These peer influences are due to strategic or collaborative partnerships between land owners associated with observed and unobserved variables to the analyst, supporting and reinforcing our hypothesis of a spatial lag formulation to capture spatial dependency in land-use modeling. However, note that this spatial dependence is confined to each of the eastern and western sections of the analysis region (as defined by IH-35), and does not extend to parcels across the two sections. In other words, IH-35 appears to act not simply as a physical barrier, but also as a barrier to peer interactions and influences.

3.3.3. Model Selection and Statistical Fit

The MNPTS model is clearly superior to the MNP model, as observed from the statistically significant random coefficients, autoregressive temporal dependence parameter, and the spatial lag parameter. Another way to demonstrate the data fit superiority of the MNPTS model over the MNP model is through the adjusted composite likelihood ratio test (ADCLRT) test. The composite log- likelihood value for the MNP model is -53249.32 (12 parameters estimated) and for the MNPTS model is -51669.8 (17 parameters estimated). The two models may be tested using the adjusted composite likelihood ratio test (AD*CLRT*) statistic (see Pace *et al*., 2011 and Bhat, 2011). This statistic has a chi-square asymptotic distribution with 5 degree of freedom. The statistic is about 4737, which is higher than the corresponding critical chi-squared value with five degree of freedom at any reasonable level of significance. This demonstrates very strong evidence of temporal dependence and spatial dynamics at play in land-use decisions.

3.3.4 Aggregate Elasticity Effects

The estimated parameter coefficients in Table 4 provide a sense of the direction of variable effects on the utilities of different land use types. However, these estimated parameters do not directly provide the magnitude of the impact of variables on the probabilities of each land-use category (this is an issue seldom considered in the spatial literature, with many papers simply presenting the parameter results and stopping there). To characterize the magnitude and direction of variable effects on the probabilities, we compute the aggregate-level "elasticity effects" of variables. Specifically, we examine the effects of variables on the expected share of each land-use alternative for the year 2006, given the exogenous variable characteristics of all the 395 parcels. We achieve this by computing the marginal probability of each parcel being in each land-use and aggregating these probabilities across parcels for each land-use category. The computation of the marginal probability of each parcel being in each land-use is relatively straightforward for the MNP model, so we will focus on the procedure for computing the marginal probabilities from the MNPTS model.

For the MNPTS model, we write the utility function of land-use i for the land-owner of parcel q as follows (note that the index 't' does not appear, since we are focusing on a specific year (2006)):

$$U_{qi} = \delta \sum_{q'} w_{qq'} \; U_{q'i} + \tilde{\alpha}_{qi} + \boldsymbol{\beta}'_q \mathbf{x}_{qi} + \tilde{\eta}_{qi} \; ; \; \tilde{\alpha}_{qi} = \tilde{a}_i + \breve{\alpha}_{qi} \, , \; \boldsymbol{\beta}_q = \mathbf{b} + \tilde{\boldsymbol{\beta}}_q \, , \tag{14}$$

where the notation is similar to Section 2.1. Next define the following (for ease in presentation, we maintain the same notations as in Section 2.1 for the re-defined vectors and matrices):

$\mathbf{U}_q = (\mathbf{U}'_{q1}, \mathbf{U}'_{q2}, ..., \mathbf{U}'_{qI})'$ and $\tilde{\boldsymbol{\eta}}_q = (\tilde{\boldsymbol{\eta}}'_{q1}, \tilde{\boldsymbol{\eta}}'_{q2}, ..., \tilde{\boldsymbol{\eta}}'_{qI})'$ ($I \times 1$ vectors),

$\mathbf{U} = (\mathbf{U}'_1, \mathbf{U}'_2, ..., \mathbf{U}'_Q)'$ and $\tilde{\boldsymbol{\eta}} = (\tilde{\boldsymbol{\eta}}'_1, \tilde{\boldsymbol{\eta}}'_2, ..., \tilde{\boldsymbol{\eta}}'_Q)'$ ($QI \times 1$ vectors),

$\breve{\boldsymbol{\alpha}}_q = (\breve{\alpha}_{q1}, \breve{\alpha}_{q2}, ..., \breve{\alpha}_{qI})'$ (I×1 vector), $\breve{\boldsymbol{\alpha}} = \left[(\breve{\alpha}_1)', (\breve{\alpha}_2)', ..., (\breve{\alpha}_Q)'\right]'$ ($QI \times 1$ vector),

$$\mathbf{x}_q = (\mathbf{x}_{q1}, \mathbf{x}_{q2}, ..., \mathbf{x}_{qI})' \; (I \times K \text{ matrix}) \, , \mathbf{x} = (\mathbf{x}'_1, \mathbf{x}'_2, ..., \mathbf{x}'_Q)' \; (QI \times K \text{ matrix}), \text{ and} \tag{15}$$

$$\tilde{\mathbf{x}} = \begin{bmatrix} x_1 & 0 & 0 & 0. \cdots & 0 \\ 0 & x_2 & 0 & 0 \cdots & 0 \\ 0 & 0 & x_3 & 0 \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \; \ldots & \vdots \\ 0 & 0 & 0 & 0 \; \ldots & x_Q \end{bmatrix} \; (QI \times QK \text{ matrix}) \text{, and}$$

$$\mathbf{S} = \left[\boldsymbol{IDEN_{QI}} - (\delta \boldsymbol{W} \otimes \boldsymbol{IDEN_I})\right]^{-1} \; (QI \times QI \text{ matrix}), \tag{16}$$

Then, using other notations as in Section 2.1, we may write the following counterpart of Equation (3) for the year 2006:

$$\mathbf{U} = \mathbf{S}\left[\left(\mathbf{1}_Q \otimes \tilde{\mathbf{A}}\right) + \mathbf{x}\mathbf{b} + \breve{\boldsymbol{\alpha}} + \tilde{\mathbf{x}}\breve{\boldsymbol{\beta}} + \tilde{\boldsymbol{\eta}}\right], \tag{17}$$

We simulate the above QI×1-vector $\mathbf{U}$ thousand times using the estimated values of $\delta, \tilde{\mathbf{A}}, \mathbf{b}$, and by randomly drawing 1000 times from the appropriate normal distributions for $\breve{\boldsymbol{\alpha}}, \breve{\boldsymbol{\beta}},$ and $\tilde{\boldsymbol{\eta}}$. Next, we compare the utilities across alternatives for each parcel q for each of the 1000 draws, assign the chosen alternative for each draw, and take the predicted share of each alternative across the 1000 draws to estimate the probability of each parcel being in each land-use alternative. The aggregate share (across parcels) of each land-use type is obtained by aggregating the parcel-level probabilities of each land-use category.

The elasticity computed is a measure of the aggregate percentage change in the aggregate share of each land-use alternative due to a change in an exogenous variable. We also compute the standard errors of the elasticity effects by using 200 bootstrap draws from the sampling distributions of the estimated parameters.[10] For dummy variables, the value of the variable is changed to one for the subsample of intersections for which the variable takes a value of zero, and to zero for the subsample of parcels for which the variable takes a value of one. We then add the shifts in expected aggregate shares in the two subsamples after reversing the sign of the shifts in the second subsample, and compute the effective percentage change in the expected shares across all parcels in the sample due to a change in the dummy variable from 0 to 1. For continuous variables, we increase the value of the variable by 25% for each parcel and compute the percentage change in the expected shares.

The elasticity effects and their standard errors are computed for the MNP model and the MNPTS model, and are presented in Table 5. The effects (and their standard errors in parenthesis) are presented for the six scenarios listed in the table. The first entry in the table indicates that, on average, a parcel that is within 350 meters from IH-35 is about 35.1% less likely to be in residential land-use relative to a parcel that is beyond 350 meters of IH-35. Similarly, the entry in the first column and second row suggests that a parcel that is 25% farther away from IH-35 than another parcel is about 1.2% more likely to be in residential land-use than the closer-to-IH35 parcel. Other entries may be similarly interpreted. The last sub-column within each alternative column provides the p-value for the difference in elasticity estimates from the MNP and MNPTS models. A '-' in this column implies that the difference is not statistically significant even at the 0.2 level of significance.

The elasticity effects of both the MNP and MNPTS models are in the same direction for all variables, and are consistent with the discussions in the previous section. However, it is clear than the elasticity effects from the MNPTS model are generally higher in magnitude than those from the MNP model, a consequence of the "spillover" effects in the MNPTS model that causes a spatial multiplier effect. Specifically, a change in a variable for one parcel influences the utilities of the land-use alternatives of other parcels, which then have a "circular" influence back on the utilities of the land-use alternatives for the parcel for which a variable has been changed. This "circular" influence is reinforcing because of the positive spatial lag parameter, which implies the spatial multiplier effect (this spatial multiplier effect is captured by the $\mathbf{S}$ matrix in

[10] For ease in computation, we however fix the spatial lag parameter δ in the bootstrapping, so that we do not have to compute the matrix **S** for each bootstrap draw (the matrix **S** entails a high-dimensional matrix inversion).

Equation (17)). The MNP model ignores the presence of such spatial multiplier effects, and assumes that a change in a variable at one parcel impacts only the land-use at that parcel.

The difference in the elasticity effects between the MNP and MNPTS models are, for the most part, statistically significant. Thus, the higher MNPTS-predicted positive effects of a parcel being within 350 meters of IH-35 (rather than being beyond 350 meters of IH-35) on the probabilities of the parcel being in commercial land-use, and the higher MNPTS-predicted negative effect of a parcel being within 350 meters of IH-35 (rather than being beyond 350 meters of IH-35) on the probability of the parcel being in non-commercial land-uses, are all highly statistically significant. Similarly, the differential effects (between the MNP and MNPTS models) of the continuous distance from IH-35 (second variable in Table 5) on the probabilities of the residential and industrial land-uses are highly statistically significant, while the differential effects on the probabilities of commercial land-use are also quite statistically significant. Other differences and their p-values may be similarly extracted from Table 5. The one variable for which there is no statistically significant difference in the MNP and MNPTS elasticity effects is for the Pflugerville City variable (see the last but one row of the table). For this variable, while the elasticity effects are indeed higher from the MNPTS model, the heterogeneity in the utility for the residential land-use type leads to a tempering of the effects on the utilities of other alternatives, which counteracts the spatial multiplier effect. The heterogeneity also leads to higher standard errors for the elasticity estimates. In combination, the tempered effects on elasticities and the higher standard errors lead to less statistically significant differences. But, overall, there are statistically significant differences in elasticity predictions between the MNP and MNPTS models, highlighting the predictive differences between the two models and, in general, the under-estimations of the magnitudes of variable effects from the MNP model.

The elasticity effects from the continuous variables (such as the continuous distance to IH-35) are not directly comparable to those from the dummy variables (such as whether or not the parcel is within 350 meters of IH-35). However, the results identify closeness to IH-35 (whether within a 350 meters band of IH-35 or not), Pflugerville location, and proximity to schools as the dominant variables impacting the land-use type of a given parcel.

4. CONCLUSION

This paper has proposed a new econometric approach to specify and estimate a model of land-use change, based on the now rich theoretical literature on land use conversion decisions made by economic agents to maximize net returns. At a methodological level, the paper has formulated and estimated a multi-period multinomial probit model, accounting for time-varying and time-stationary inter-temporal dependencies as well as a spatial lag structure across observation units. The model also accommodates spatial heterogeneity. The inference methodology used is the maximum approximate composite marginal likelihood (MACML) approach. The paper has modeled the land-use type of parcel-level spatial units in an area north of the City of Austin in Texas. In doing so, the emphasis has been on better linking the quantitative (but aspatial or highly stylized spatial effects) perspective for land-use analysis that dominates the economic literature with the qualitative (but richer spatial dynamics and heterogeneity) perspective for land-use analysis that is quite prevalent in the ecological literature. The empirical results indicate the presence of statistically significant time-invariant and time-varying land-owner-specific unobserved factors as well as the presence of spatial spillover effects caused by didactic interactions between land-owners of proximately located spatial units. Ignoring these dependencies and dynamics will, in general, lead to inconsistent and inefficient estimates of parameter effects. This is highlighted by computing the elasticity effects of variables, which indicates that the model that accommodates temporal dependencies and spatial dynamics predicts magnitude effects that are statistically significantly different from the model that ignores these effects. Important determinants of land-use type include proximity to highways and other roadways, distance from flood plains, parcel location in the context of existing development, and distance from schools. The results also suggest that major transportation roadways can act not only as physical separators of land areas, but also as a barrier to peer interactions and influences.

To conclude, the model structure and inference approach proposed in this paper should be applicable in a wide variety of fields where social and spatial interactions (or didactic interactions) between decision-makers lead to spatial multiplier and spillover effects in the choices of the decision-makers. Of course, as always, there are several directions for future research, including a more rigorous theoretical and simulation-based evaluation of estimator efficiency related to the specification of the composite marginal likelihood function, consideration of more flexible forms of spatial modeling that combine spatial lag and spatial error formulations, and the incorporation of additional parcel-level, pedo-climatic, and regional-level externalities.

ACKNOWLEDGEMENTS

The authors acknowledge the helpful comments of Daniel Griffith and three anonymous reviewers on an earlier version of the paper. The authors also acknowledge support from the Sustainable Cities Doctoral Research Initiative at the Center for Sustainable Development at The University of Texas at Austin. The authors are grateful to Lisa Macias for her help in typesetting and formatting this document.

REFERENCES

Anas, A., Arnott, R., Small, K.A. (1998) Urban spatial structure. *Journal of Economic Literature* 36(3), 1426-1464.

Anselin, L. (2006) Spatial econometrics. In: Mills T., Patterson K. (eds), *Palgrave Handbook of Econometrics: Volume 1, Econometric Theory*, Palgrave Macmillan Basingstoke, 901-969.

Beck, N., Gleditsch, K.S., Beardsley, K. (2006) Space is more than geography: using spatial econometrics in the study of political economy. *International Studies Quarterly* 50(1), 27-44.

Beron, K.J., Vijverberg, W.P.M. (2004) Probit in a spatial context: a Monte Carlo analysis. In: Anselin, L., Florax, R.J.G.M., Rey, S.J. (eds.), *Advances in Spatial Econometrics: Methodology, Tools and Applications*, Springer-Verlag, Berlin, 169-196.

Bhat, C.R. (2011) The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models. *Transportation Research Part B* 45(7), 923-939.

Bhat, C.R., Sener, I.N. (2009) A copula-based closed-form binary logit choice model for accommodating spatial correlation across observational units. *Journal of Geographical Systems* 11(3), 243-272.

Bhat, C.R., Sener, I.N., Eluru, N. (2010a) A flexible spatially dependent discrete choice model: formulation and application to teenagers' weekday recreational activity participation. *Transportation Research Part B* 44(8-9), 903-921.

Bhat, C.R., Varin, C., Ferdous, N. (2010b). A comparison of the maximum simulated likelihood and composite marginal likelihood estimation approaches in the context of the multivariate ordered response model. In: Greene, W., Hill, R.C. (eds.), *Advances in Econometrics: Maximum Simulated Likelihood Methods and Applications*, Emerald Group Publishing Limited, 65-106.

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

Brady, M., Irwin E. (2011) Accounting or spatial effects in economic models of land use: recent developments and challenges ahead. *Environmental Resource Economics* 48(3), 487-509.

Capozza, D.R., Li, Y. (1994) The intensity and timing of investment: the case of land. *American Economic Review* 84(4), 889-904.

Caragea, P.C., Smith, R.L. (2007) Asymptotic properties of computationally efficient alternative estimators for a class of multivariate normal models. *Journal of Multivariate Analysis* 98(7), 1417- 1440.

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.

Carrión-Flores, C.E., Flores-Lagunes, A., Guci, L. (2009) Land use change: a spatial multinomial choice analysis. Presented at the Agricultural and Applied Economics Association 2009 Annual Meeting, July 26-28, 2009, Milwaukee, WI.

Chakir, R., Parent, O. (2009) Determinants of land use changes: a spatial multinomial probit approach. *Papers in Regional Science* 88(2), 327–344

Cho, W.T., Rudolph, T. (2007) Emanating political participation: untangling the spatial structure behind participation. *British Journal of Political Science* 38(2), 273-289.

City of Austin (2008) Neighborhood Planning: Guide to Zoning, http://www.ci.austin.tx.us/planning/neighborhood/downloads/zoning_guide.pdf

Cox, D.R., Reid, N., (2004) A note on pseudolikelihood constructed from marginal densities. *Biometrika* 91(3), 729-737.

Dubin, R.A. (1998) Spatial autocorrelation: a primer. *Journal of Housing Economics* 7(4), 304-327.

Elhorst, J.P. (2010a) Spatial panel data models. In: Fischer, M.M., Getis, A. (eds.), *Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications*, Springer, Berlin, 377-408.

Elhorst, J.P. (2010b) Applied spatial econometrics: raising the bar. *Spatial Economic Analysis* 5(1), 9-28.

Engle, R.F., Shephard, N., Sheppard, K. (2007) Fitting and testing vast dimensional time-varying covariance models. Finance Working Papers, FIN-07-046, Stern School of Business, New York University.

Engler, D.A., Mohapatra, G., Louis, D.N., Betensky, R.A. (2006) A pseudolikelihood approach for simultaneous analysis of array comparative genomic hybridizations. *Biostatistics* 7(3), 399-421.

Fleming, M.M. (2004) Techniques for estimating spatially dependent discrete choice models. In: Anselin, L., Florax, R.J.G.M., Rey, S.J. (eds.), *Advances in Spatial Econometrics: Methodology, Tools and Applications*, Springer-Verlag, Berlin, 145-168.

Franzese, R.J., Hays, J.C. (2008) Empirical models of spatial interdependence. In: Box-Steffensmeier, J.M., Brady, H.E., Collier, D., (eds.), *The Oxford Handbook of Political Methodology*, Oxford University Press, Oxford, 570-604.

Franzese, R.J., Hays, J.C., Schaffer, L.M. (2010) Spatial, temporal, and spatiotemporal autoregressive probit models of binary outcomes: estimation, interpretation, and presentation. APSA 2010 Annual Meeting Paper. Available at: http://ssrn.com/abstract=1643867

Glaeser, E.L., Gyourko, J., Saks, R.E. (2006) Urban growth and housing supply. *Journal of Economic Geography* 6(1), 71-89.

Godambe, V.P. (1960) An optimum property of regular maximum likelihood estimation. *The Annals of Mathematical Statistics* 31(4), 1208-1211.

Heagerty, P., Lumley, T. (2000) Window subsampling of estimating functions with application to regression models. *Journal of the American Statistical Association* 95(449), 197-211.

Irwin, E.G. (2010) New directions for urban economic models of land use change: incorporating spatial dynamics and heterogeneity. *Journal of Regional Science* 50(1), 65-91.

Irwin, E.G., Bockstael, N.E. (2002) Interacting agents, spatial externalities, and the endogenous evolution of residential land use pattern *Journal of Economic Geography* 2(1), 31-54

Irwin, E.G., Geoghegan, J. (2001) Theory, data, methods: developing spatially explicit economic models of land use change. *Agriculture, Ecosystems & Environment* 85(1-3), 7-24.

Jones, K., Bullen, N. (1994) Contextual models of house prices: a comparison of fixed- and random-coefficient models developed by expansion. *Economic Geography* 70(3), 252-272.

Klier, T., McMillen, D.P. (2008) Clustering of auto supplier plants in the U.S.: GMM spatial logit for large samples. *ASA Journal of Business & Economic Statistics* 26(4), 460-471.

Lee, L-F., Yu, J. (2010) Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics* 154(2), 165-185

LeSage, J.P. (2000) Bayesian estimation of limited dependent variable spatial autoregressive models. *Geographical Analysis* 32(1), 19-35.

LeSage, J.P., Pace, R.K. (2009) *Introduction to Spatial Econometrics*. Chapman & Hall/CRC, Taylor & Francis Group, Boca Raton.

Li, X., Liu, X.P. (2007) Defining agents' behaviors to simulate complex residential development using multicriteria evaluation. *Journal of Environmental Management* 85(4), 1063-1075.

McMillen, D.P. (2010) Issues in spatial analysis, *Journal of Regional Science*, 50(1), 119-141.

Messner, S.F., Anselin, L. (2004) Spatial analyses of homicide with areal data. In: Goodchild M., Janelle D. (eds), *Spatially Integrated Social Science*, Oxford University Press, New York, 127-144.

Miller, H.J. (1999) Potential contributions of spatial analysis to geographic information systems for transportation (GIS-T). *Geographical Analysis* 31(4), 373-399.

Molenberghs, G., Verbeke, G. (2005) *Models for Discrete Longitudinal Data*. Springer Series in Statistics, Springer Science + Business Media, Inc., New York.

Müller, G., Czado, C. (2005) An autoregressive ordered probit model with application to high frequency financial data. *Journal of Computational and Graphical Statistics* 14(2), 320-338.

Pace, L., Salvan, A., Sartori, N. (2011) Adjusting composite likelihood ratio statistics. *Statistica Sinica* 21(1), 129-148.

Pinkse, J., Slade, M.E. (1998) Contracting in space: an application of spatial statistics to discrete-choice models. *Journal of Econometrics* 85(1), 125-154.

Plantinga, A.J., Irwin, E.G. (2006) Overview of empirical methods. In: Bell, K.P., Boyle, K.J., Rubin, J. (eds), *Economics of Rural Land-Use Change*, Aldershot, Ashgate Publishing, 113-134.

Sener, I.N., and C.R. Bhat (2011) Flexible spatial dependence structures for unordered multinomial choice models: Formulation and application to teenagers' activity participation. *Transportation*, forthcoming.

Smirnov, O.A. (2010) Spatial econometrics approach to integration of behavioral biases in travel demand analysis. *Transportation Research Record* 2157, 1-10.

Varin, C., Czado, C. (2008) A mixed probit model for the analysis of pain severity diaries. Available at: http://www-m4.ma.tum.de/Papers/Czado/varin_czado_pain_diaries.pdf.

Varin, C., Vidoni, P. (2005) A note on composite likelihood inference and model selection. *Biometrika* 92(3), 519-528.

Varin, C., Vidoni, P. (2009) Pairwise likelihood inference for general state space models. *Econometric Reviews* 28(1-3), 170-185.

Varin, C., Reid, N., Firth, D. (2011) An overview of composite likelihood methods. *Statistica Sinica* 21(1), 5-42.

Ver Hoef, J.M., Jansen, J.K. (2007) Space–time zero-inflated count models of harbor seals. *Environmetrics* 18(7), 697-712.

Zhao, Y., Joe, H. (2005) Composite likelihood estimation in multivariate data analysis. *The Canadian Journal of Statistics* 33(3), 335-356.

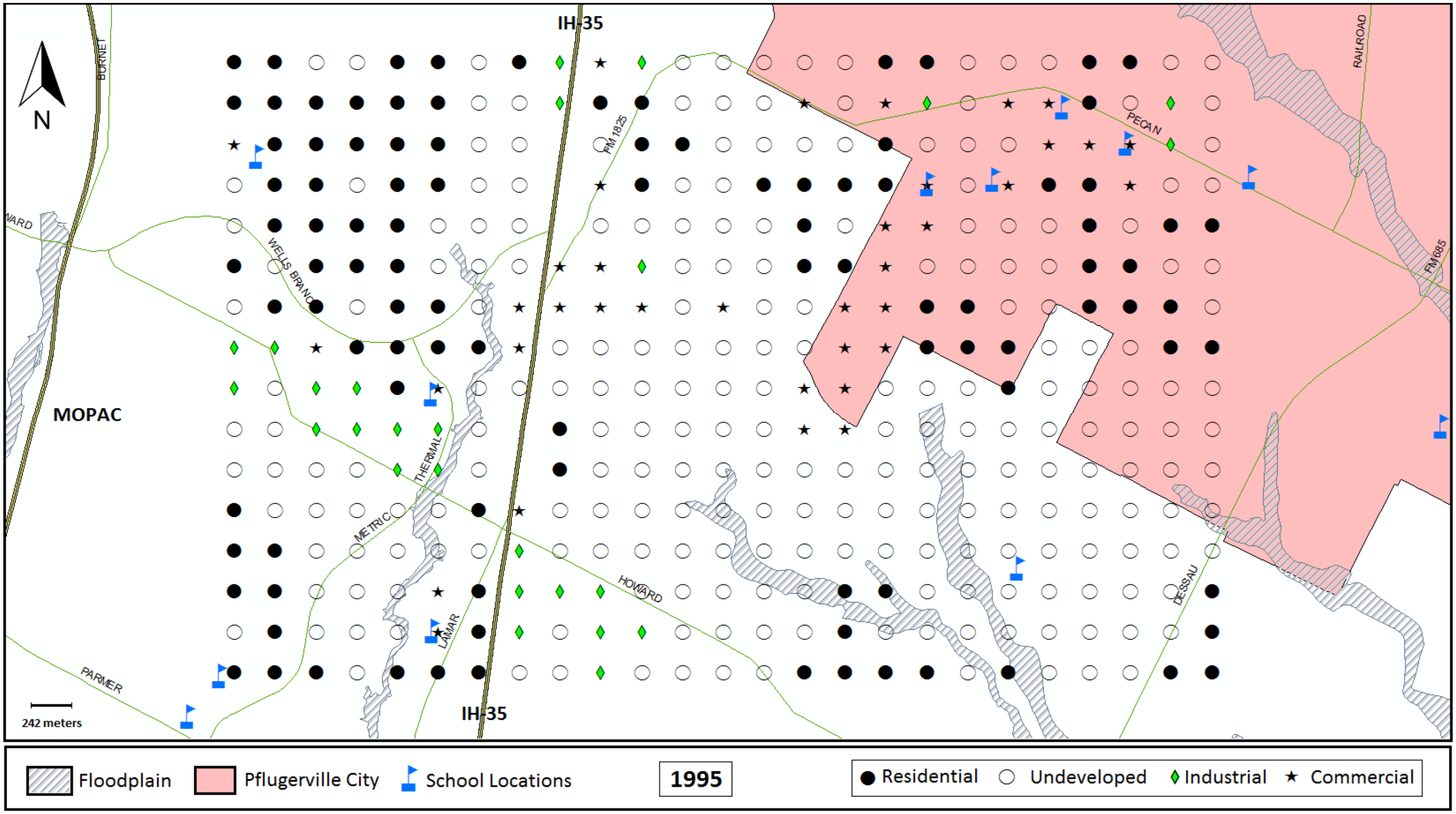

Figure 1. The Analysis Area for the year 1995

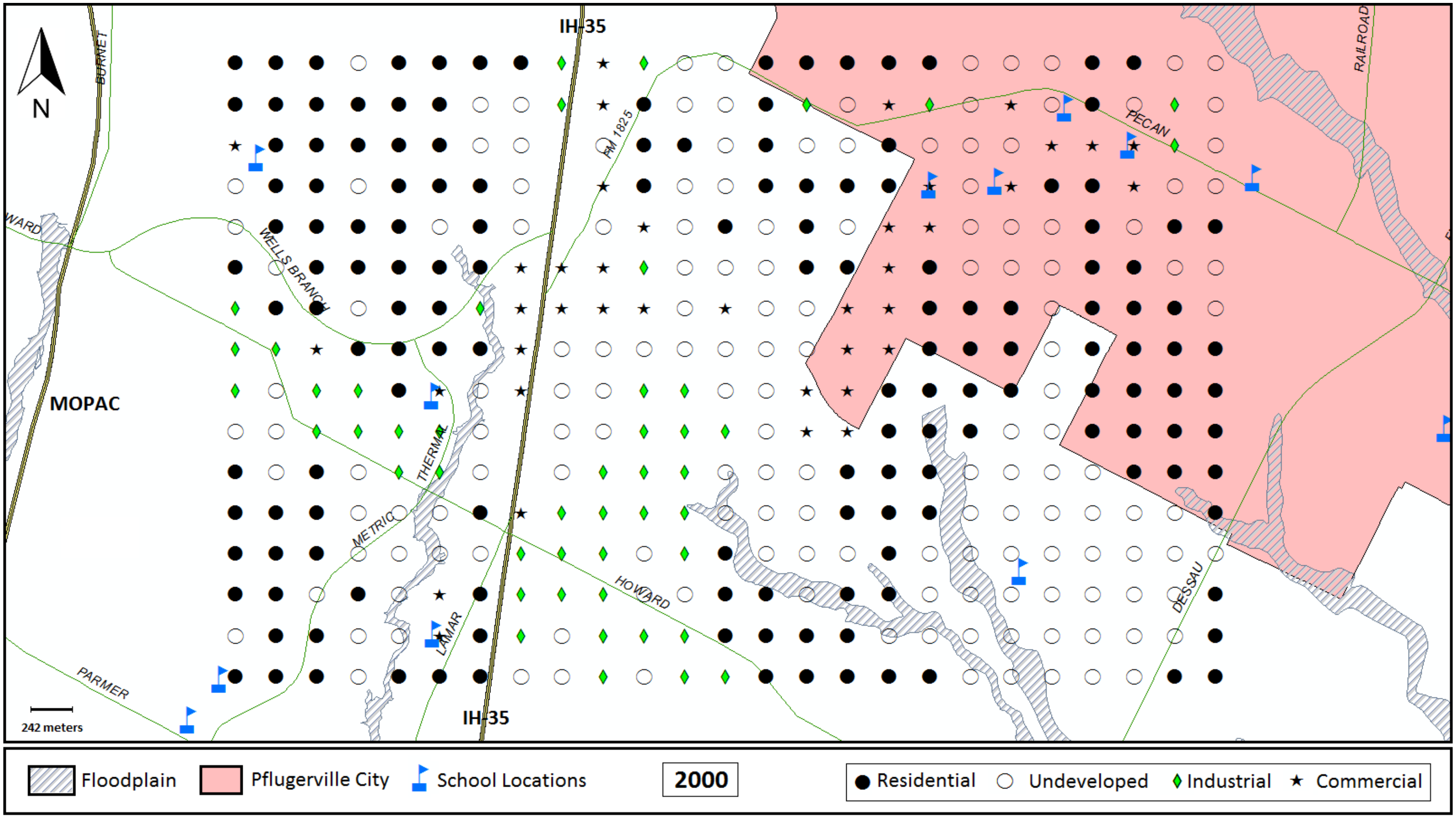

Figure 2. The Analysis Area for the year 2000

Table 1. Percentage of Land by Land Use Types

Land use Type	1995	2000	2003	2006
Residential	25.80%	38.70%	39.70%	39.50%
Undeveloped/Open Area	58.20%	39.50%	42.50%	39.70%
Commercial	9.40%	9.90%	9.90%	13.40%
Industrial	6.60%	11.90%	7.80%	7.30%

Table 2. Descriptive Statistics of the Independent Variables used in the Model

Variable	Min	Max	Mean	Std. Deviation
Distance to IH-35 less than 350 meters	0.0000	1.0000	0.1038	0.3051
Distance to IH-35	0.0440	4.2664	1.7379	1.1556
Distance to nearest non-freeway road (other than IH-35 and Mopac)	0.0004	1.7308	0.5397	0.4254
Distance to IH-35 * Distance to nearest non-freeway road	0.0600	4.3456	1.1479	1.0903
Distance to Mopac Freeway	0.8150	7.154	3.9816	1.7321
Distance to the nearest Roadway	0.0004	1.7308	0.5191	0.4262
Distance to the nearest Flood Plain	0.0012	1.6492	0.6257	0.4065
Distance to the nearest Road / Distance to the nearest flood plain	0.0006	11.5384	1.4367	2.0227
Parcel lies within Pflugerville City	0.0000	1.0000	0.2278	0.4196
Within one kilometer of a school	0.0000	1.0000	0.6127	0.4873

Table 3. Model Selection

Statistic	Full region-based weight matrix model	Partitioned region-based weight matrix model
Log-composite likelihood at convergence	-2568369	-2567958
Trace value	171	136
CLIC statistic	-2568540	-2568094

Table 4. Estimation Results (t-statistics in parenthesis)

Variables	Standard multinomial probit (MNP) model				MNP Spatial lag model with temporal Panel and spatial effects (MNPTS) model autocorrelation			
	Residential	Commercial	Industrial	Undeveloped	Residential	Commercial	Industrial	Undeveloped
Constant	-	-0.864 (-15.47)	-0.133 (-2.31)	-0.269 (-3.38)	-	-1.869 (-3.90)	0.471 (5.64)	-0.597 (-2.20)
Standard deviation	-	-	-	-	-	2.353 (1.25)	-	2.403 (1.46)
Distance to IH-35 less than 350 meters	-	1.090 (12.97)	-	-	-	3.207 (7.26)	-	-
Distance to IH-35 * Distance to nearest non-freeway road	-	-	-0.596 (-8.92)	-	-	-	-0.880 (-5.19)	-
Distance to Mopac Freeway	-	-	-	0.089 (4.33)	-	-	-	0.137 (2.41)
Distance to the nearest Road / Distance to the nearest flood plain	-	-	-	0.101 (7.66)	-	-	-	0.240 (5.38)
Parcel lies within Pflugerville City	0.185 (1.87)	0.899 (9.02)	-	-	-0.338 (-1.26)	2.000 (4.52)	-	-
Standard deviation					5.231 (1.434)	-	-	-
Within one kilometer of a school	0.268 (4.56)	-	-	-	0.657 (7.47)	-	-	-
(t=1995) time dummy	-0.145 (-1.21)	-	-	0.349 (3.06)	-0.11 (-1.16)	-	-	0.624 (4.99)
Temporal autocorrelation - ρ	Implicitly restricted to zero				0.367 (3.98)			
Spatial lag - δ	Implicitly restricted to zero				0.449 (5.39)			

Table 5. Aggregate-Level Elasticity Effects of the MNP and MNPTS Models (standard error in parenthesis)

Scenario	Residential			Commercial			Industrial			Undeveloped		
	MNP	MNPTS	p[†]	MNP	MNPTS	p	MNP	MNPTS	p	MNP	MNPTS	p
A change from the parcel being farther than 350 meters from IH-35 to within 350 meters from IH-35	-35.1 (3.0)	-67.3 (6.5)	0.000	382.5 (45.9)	806.2 (129.1)	0.002	-37.8 (3.2)	-75.2 (7.6)	0.000	-32.7 (2.7)	-54.6 (5.7)	0.001
A 25% increase in the distance to IH-35, but only for those parcels farther than 350 meters from IH-35,	1.2 (0.1)	3.4 (0.4)	0.000	0.9 (0.2)	1.8 (0.5)	0.075	-11.2 (0.6)	-22.1 (2.2)	0.000	1.1 (0.1)	1.0 (0.2)	-*
A 25% increase in the distance to the nearest flood plain	1.9 (0.3)	3.0 (0.5)	0.088	1.4 (0.3)	2.2 (0.5)	0.126	1.2 (0.2)	1.8 (0.5)	-	-2.5 (0.4)	-3.8 (0.6)	0.064
A 25% increase in distance to the nearest road and a 25% decrease in the distance to the nearest flood plain	-6.6 (0.7)	-9.8 (1.4)	0.033	-5.0 (0.7)	-6.7 (1.3)	-	-4.0 (0.5)	-6.3 (1.4)	0.147	8.5 (0.9)	12.3 (1.7)	0.042
A switch of the parcel location from Pflugerville to outside Pflugerville	-5.1 (9.5)	-15.3 (11.1)	-	253.6 (41.7)	380.4 (143.0)	-	-41.8 (7.7)	-24.3 (19.4)	-	-29.1 (6.2)	-34.9 (7.6)	-
A switch of the parcel location from being farther than one kilometer from the closest school to being closer than one kilometer from the closest school	34.7 (9.4)	89.7 (17.3)	0.005	-17.5 (3.5)	-22.2 (3.5)	-	-19.3 (4.1)	-64.5 (5.4)	0.000	-14.8 (3.1)	-21.1 (2.9)	0.143

† p value of the difference

*A '-' implies that the difference is not statistically significant even at the 0.2 level of significance