A New Spatial Multiple Discrete-Continuous Modeling Approach to Land Use Change Analysis

Chandra R. Bhat*
The University of Texas at Austin, Dept of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712
Phone: 512-471-4535; Fax: 512-475-8744; Email: bhat@mail.utexas.edu
and
King Abdulaziz University, Jeddah 21589, Saudi Arabia

Subodh K. Dubey
The University of Texas at Austin, Dept of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712
Phone: 512-471-4535, Fax: 512-475-8744; E-mail: subbits@gmail.com

Mohammad Jobair Bin Alam
King Abdulaziz University, Department of Civil Engineering
P.O. Box 80204, Jeddah, 21589, Kingdom of Saudi Arabia
Phone: +966-2-6402000 (Ext.: 51339), Fax: +966-2-6952179, Email: malam@kau.edu.sa

Waleed H. Khushefati
King Abdulaziz University, Department of Civil Engineering
P.O. Box 80204, Jeddah, 21589, Kingdom of Saudi Arabia
Phone: +966-2-6402000 (Ext.: 51339), Fax: +966-2-6952179, Email: wkhushefati@kau.edu.sa

*corresponding author

ABSTRACT
This paper formulates a multiple discrete-continuous probit (MDCP) land-use model within a spatially explicit economic structural framework for land-use change decisions. The spatial MDCP model is capable of predicting both the type and intensity of urban development patterns over large geographic areas, while also explicitly acknowledging geographic proximity-based spatial dependencies in these patterns. At a methodological level, the paper focuses on specifying and estimating a spatial MDCP model that allows the dependent variable to exist in multiple discrete states with an intensity associated with each discrete state. The formulation also accommodates spatial dependencies, as well as spatial heterogeneity and heteroscedasticity, in the dependent variable, and should be applicable in a wide variety of fields where social and spatial dependencies between decision agents (or observation units) lead to spillover effects in multiple discrete-continuous choices (or states). A simulation exercise is undertaken to evaluate the ability of the proposed maximum approximate composite marginal likelihood (MACML) approach to recover parameters from a cross-sectional spatial MDCP model. The results show that the MACML approach does well in recovering parameters. An empirical demonstration of the approach is undertaken using the city of Austin parcel level land use data.

Keywords: spatial econometrics, multiple discrete-continuous model, random-coefficients, land use analysis, MACML approach.

1. INTRODUCTION

Land-use change models are used in a variety of fields such as planning, urban science, ecological science, climate science, geography, watershed hydrology, environmental science, political science, and transportation to examine future land-use scenarios as well as to evaluate the potential effects of policies directed toward engendering a socially or economically or ecologically desirable pattern of future land-use that minimizes negative externalities. More recently, there has been substantial attention in the scientific literature on biodiversity loss, deforestation consequences, and carbon emissions increases caused by patterns of urban and rural land-use development, and associated climate change impacts (for example, see Lewis et al., 2011). This is not surprising, since one of the most important "habitat" elements characterizing Earth's terrestrial and aquatic ecosystems is the land use pattern (another is climate pattern, which is increasingly becoming closely related to the land use pattern). In this paper, we contribute to the vibrant and interdisciplinary literature on land-use analysis by proposing a new econometric approach to specify and estimate a model of land-use change that is capable of predicting both the type and intensity of urban development patterns over large geographic areas, while also explicitly acknowledging geographic proximity-based spatial dependencies in these patterns. 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 spatial multiple discrete-continuous probit (MDCP) model that allows the dependent variable to exist in multiple discrete states with an intensity associated with each discrete state. The formulation also accommodates spatial heterogeneity and heteroscedasticity in the dependent variable, and should be applicable in a wide variety of fields where social and spatial dependencies between decision agents (or observation units) lead to spillover effects in multiple discrete-continuous choices (or states). At an empirical level, the paper models land-use in multiple discrete states, along with the area invested in each land-use discrete state, within each spatial unit in an entire urban region. The model is a hybrid of three different strands of model types used in the land-use analysis literature.

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

The Econometric Context

In the past decade, there has been increasing interest and attention on recognizing and explicitly accommodating spatial (and social) dependence among decision-makers (or other observation units) in urban and regional modeling, agricultural and natural resource economics, public economics, geography, sociology, political science, and epidemiology. The reader is referred to a special issue of *Regional Science and Urban Economics* entitled "Advances in spatial econometrics" (edited by Arbia and Kelejian, 2010) and another special issue of the *Journal of Regional Science* entitled "Introduction: Whither spatial econometrics?" (edited by Partridge et al., 2012) for a collection of recent papers on spatial dependence, and to Elhorst (2010), Anselin (2010), Ferdous and Bhat (2013) and Brady and Irwin (2011) for overviews of recent developments in the spatial econometrics field. Within the past few years, there has particularly been an explosion in studies that recognize and accommodate spatial dependency in discrete choice models. The typical way this is achieved is by applying spatial lag and spatial error-type structures developed in the context of continuous dependent variables to the linear (latent) propensity variables underlying discrete choice dependent variables (see reviews of this literature in Fleming, 2004; Franzese and Hays, 2008; LeSage and Pace, 2009; Hays et al., 2010; Brady and Irwin, 2011; and Sidharthan and Bhat, 2012). The two dominant techniques, both based on simulation methods, for the estimation of such spatial 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 even moderate estimation sample sizes (see Bhat, 2011 and Franzese et al., 2010). Recently, Bhat and colleagues have suggested a maximum approximate composite marginal likelihood (MACML) inference approach for estimating spatial multinomial probit (MNP) models and a composite marginal likelihood (CML) inference approach for estimating spatial binary/ordered probit models. The MACML approach uses the CML approach, but also makes an additional analytic approximation to evaluate the multivariate normal cumulative distribution (MVNCD) function during estimation. These methods are easy to implement, require no simulation, and involve 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 social/spatial dependence structures.

At the same time that spatial considerations are receiving widespread attention in the discrete choice literature, multiple discrete-continuous (MDC) models have also seen substantial application in different disciplines, including regional science (Kaza et al., 2012), transportation (Bhat, 2005; 2008; Bhat et al., 2013a), time use (Habib and Miller, 2008; Pinjari and Bhat, 2010), marketing and retailing (Kim et al., 2002; Allenby et al., 2010; Satomura et al., 2011), energy economics (Ahn et al., 2008), environmental economics (see von Haefen et al., 2004; Kuriyama et al., 2010), and tourism (LaMondia et al., 2008; Van Nostrand et al., 2013). In MDC situations, consumers may consume multiple alternatives at the same time, with a continuous dimension of the amount of consumption associated with each alternative. To be precise, there are multiple alternatives, with zero consumption of an alternative implying that the alternative is not consumed, and a positive consumption amount implying that the alternative is consumed. Examples of such MDC contexts include land-use type and intensity of land-use over a spatial unit, household vehicle type holdings and usage, consumer brand choice and purchase quantity, and recreational destination location choice and number of trips. While a variety of modeling approaches have been used in the literature to accommodate MDC choice contexts, the one that has dominated the recent literature is based on a utility maximization framework that assumes a non-linear (but increasing and continuously differentiable) utility structure to accommodate decreasing marginal utility (or satiation) with increasing investment in an alternative. Consumers are assumed to maximize this utility subject to a budget constraint. The optimal consumption quantities (including possibly zero investments in some alternatives) are obtained by writing the Karush-Kuhn-Tucker (KKT) first-order conditions of the utility function with respect to the investment quantities. Researchers from many disciplines have used such a KKT approach, and several additively separable and non-linear utility structures have been proposed. Of these, the general utility form proposed by Bhat (2008) subsumes other non-linear utility forms as special cases, and allows a clear interpretation of model parameters. In Bhat's utility function form and other more restrictive utility forms, stochasticity is introduced in the baseline preference for each alternative to acknowledge the presence of unobserved (to the analyst) factors that may impact the utility of each alternative (the baseline preference is the marginal utility of each alternative at the point of zero consumption of the alternative). As in traditional discrete choice models, the

most common distributions used for the kernel stochastic error term (across alternatives) are the generalized extreme value (GEV) distribution (see Bhat, 2008; Pinjari, 2011; Castro et al., 2012) and the multivariate normal distribution (see Kim et al., 2002 and Bhat et al., 2013b). The first distribution leads to a closed-form MDC generalized extreme value (or MDCGEV) model structure, while the second to a MDC probit (or MDCP) model structure. In both these structures, the analyst can further superimpose a mixing random distribution of coefficients in the baseline preference to accommodate unobserved heterogeneity across consumers (or observation units). Assuming a normal mixing error distribution, the use of a GEV kernel error term leads to a mixing of the normal distribution with a GEV kernel (leading to the mixed MDCGEV model or MMDCGEV structure), while the use of a probit kernel leads back to an MDCP model structure (because of the conjugate nature of the multivariate normal distribution in terms of addition). In this paper, we will use the MDCP structure because it allows us to use the MACML inference approach even in the presence of spatial dependence. This is the first such formulation and application of a spatial MDCP model in the econometric literature.[1]

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 drawing elements from three different types of models proposed and applied in the literature.

The first type of models, usually referred to as pattern-based models and developed by geographers and natural scientists, is well suited for land-use modeling over relatively large geographic extents (such as urban regions or entire states or even countries). The unit of analysis in these pattern-based models is typically an aggregated spatial unit (such as a large grid or a traffic analysis zone or a Census tract or a County or a State). One basis for these models originates from the mathematical representations of the discrete state of a cell (a very fine disaggregate unit of space) as a deterministic or probabilistic function of the states of neighboring cells in an earlier time period (see, for example, Wu and Webster, 1998; Clarke et

---

[1] A couple of recent studies using an MDC structure have accommodated spatial effects in the systematic component of utility (see Kaza et al., 2012 and Richards et al., 2012). However, these models are no different from aspatial MDC models from a formulation and estimation standpoint, since the resulting model is the closed-form MDCEV model.

al., 1997; Engelen and White, 2008). In these cellular automata-based models, the analyst hypothesizes the nature of the deterministic or probabilistic updating functions, simulates the states of cells over many "virtual" time periods, and aggregates up the states of the cells at the end to obtain land-use patterns. While such models may be able to "fit" the land-use patterns at the aggregated spatial unit level, the imposed updating functions are not based on actual data. Thus, there is no direct evidence linking the updating mechanism at the cell level to the spatial evolution of land-use patterns at the aggregate spatial unit level. Also, since such models do not use exogenous variables such as sociodemographic characteristics of spatial units, transportation network features, and other environmental features as the basis for explaining land-use, the policy value of these models is limited. An alternative basis of pattern-based models is to use empirical models estimated at the aggregate spatial unit level that relates variables such as distance to urban center, pedoclimatic or biophysical factors of the land in the spatial unit (such as slope, water content, aeration, and elevation), and transportation network and accessibility variables to land-use patterns (see, for example, Landis and Zhang, 1998a, 1998b; Brown et al., 2000; Parker et al., 2003; Brown and Duh, 2004; Robinson and Brown, 2009). Once estimated, these models can be used in a simulation setting to predict land-use patterns in response to different exogenously imposed policy scenarios. Unfortunately, these empirical models have not been formulated in a manner that appropriately recognizes the multiple discrete-continuous nature of land-use patterns in the aggregated spatial units. Further, these models typically do not adequately consider population characteristics of spatial units in explaining land-use patterns within that unit.

The second type of models, usually referred to as process-based models and considered by economists, is based on explicitly modeling landowners' decisions of land-use type choice for their parcels. The most important aspect of these process-based models is that they explicitly consider the human element in land-use modeling; that is, landowner decisions (regarding the type of land-use to invest their parcel in), as influenced by a suite of economic, biophysical, accessibility, and policy variables, are acknowledged as the fundamental drivers of land-use patterns. The emphasis is on using the land-owner as the unit of analysis, rather than a piece of land. To elucidate, landowners are considered as 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 and Towe et al., 2008). The stream of returns from converting a parcel from the current land-use to some other land-use is 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). Such process-based models allow for the analysis of a rich set of policy scenarios, by enabling the modeling of individual-level behavioral changes to exogenously imposed policy scenarios. However, in addition to difficulties associated with incorporating spatial considerations at this micro-level, the data and computing demands can be very high when using process-based models, especially when the analysis is being conducted at the level of entire urban regions or states in a country (see Kaza et al., 2012). Further, individual landowners may not have *carte blanche* authority to develop their land any way they want to, because of the presence of land-use and zoning regulations. Besides, multiple parcels in very close proximity tend to get similarly developed, because multiple parcels can be under the purview of a single decision-making agent such as a county board or a community board (see McMillen and McDonald, 1991; Mayer and Somerville, 2000; Munroe et al., 2005).

The third type of models, referred to as spatial-based models, puts emphasis on spatial dependence among spatial units (in pattern-based models) or among landowners (in process-based models), as caused by diffusion effects, or zoning and land-use regulation effects, or social interaction effects, or observed and unobserved location-related influences (see Jones and Bullen, 1994 and Miller, 1999). Indeed, as expressed by Tobler's (1970) first law of geography, "everything is related to everything else, but close things more so". While some of these proximity-based spatial effects may be accommodated through the appropriate construction of spatial variables (such as accessibility to city centers and market places), there will inevitably be unobserved spatial variables (such as say neighborhood soil quality or attitudes/politics) that will create unobserved dependencies in land-use patterns of proximally located spatial units. Several different spatial formulations have been considered in land-use modeling to accommodate such spatial dependencies, though the two most dominant remain the spatial lag and spatial error formulations. Of these, the spatial lag structure is more appealing.[2] The spatial lag formulation

---

[2] 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

also generates spatial heteroscedasticity. In addition to the spatial lag-based and resulting heteroscedasticity effects just discussed, it is also likely that there is spatial heterogeneity (*i.e.*, differences in relationships between the dependent variable of interest and the independent variables across decision-makers or spatial units in a study region (see, Fotheringham and Brunsdon, 1999; Bhat and Zhao, 2002; Bhat and Guo, 2004). Thus, it behooves the analyst to accommodate local variations (*i.e.*, recognize spatial non-stationarity) in the relationship across a study region rather than settle for a single global relationship.

In the current study, we adopt an aggregate spatial unit of analysis of a quarter-of-a-mile square grid cell to study land-use over an entire urban region of Austin, Texas, with each grid having the "option" of investing (and converting) from one package of land-uses to another alternative package of land-uses. In doing so, some grids can invest entirely in a single land-use. The grid-level land-use is obtained by aggregating underlying parcel-level land-use information. However, we supplement this pattern-based modeling view with a process-based modeling view. Specifically, while the clear linkage between parcels and their human landowners in typical process-based models is admittedly not present, we consider a rich set of population demographics of the citizenry of each aggregate grid to approximate a collective decision-making process for that grid. In addition, the land-use in a grid may also be determined by community or county boards through zoning regulations. Besides, by using a grid size that is not too aggregate, we retain some of the process-based model characteristics of having a connection between the spatial unit of analysis and human decision-makers. But since there is no clear label possible for the decision-maker of a grid, we will use the terminology of the "grid" both as a spatial unit of analysis as well as the decision-maker for the spatial unit of analysis. The hybrid model just discussed is further enhanced by considering all the spatial analysis aspects considered in spatial-based models. Thus, while Kaza et al. (2012) also consider a hybrid land-use model based on Bhat's (2008) MDCEV model, we consider the important spatial issues of dependence and heterogeneity due to unobserved as well as observed factors, as well as the resulting spatial heteroscedasticity, in our modeling approach. We also accommodate a general

---

"hard to defend" position that "space matters in the error process but not in the 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. Of course, the procedure developed here can also be extended to Spatial Durbin and other spatial specifications, but we leave these for future application efforts. The basic concepts we propose here to accommodate spatial dependence in MDCP models are the same regardless of the spatial dependence structure.

covariance matrix for the utilities of grid investments in the land-use categories. In accommodating all these effects, we adopt an MDCP model rather than the MDCEV model, since it is next to impossible to incorporate global spatial issues within the MDCEV structure when dealing with even a moderate number of spatial units.

2. MODELING METHODOLOGY

Model Formulation

We derive the spatial MDCP model in the empirical context of the type and intensity of land-use over a grid, though the same formulation can be used in the many other multiple discrete-continuous contexts identified in the section entitled "The Econometric Context". Also, in the discussion in this section, we will assume that each grid has the potential to invest in all possible land-uses. The case when some grids cannot be developed for specific land-use purposes (say, due to zoning or hazard mitigation restrictions) poses no complications whatsoever, since the only change in such a case is that the dimensionality of the integration in the likelihood contribution changes from one grid to the next. The next section presents the set-up for the aspatial MDCP model in a way that makes it convenient to extend to the spatial MDCP set-up discussed in the subsequent section.

The Aspatial MDCP Model

Let q $(q = 1, 2, ..., Q)$ be the index for grids and let k $(k = 1, 2, ..., K)$ be the index for land use types. In the empirical context of this paper, the alternative land use types include (1) residential land-use (including single family, duplexes, three/four-plexes, apartments, condominiums, mobile homes, group quarters, and retirement housing), (2) commercial land-use (including commercial, office, hospitals, government services, educational services, cultural services, and parking), (3) industrial land-use (including manufacturing, warehousing, resource extraction (mining), landfills, and miscellaneous industrial), and (4) undeveloped land-use (including open and undeveloped spaces, preserves, parks, golf courses, and agricultural open spaces). The last among these alternatives serves as an "essential outside good" in that all grid cells inevitably will have at least some of their land area that remains undeveloped.[3]

---

[3] The presence of the "undeveloped" land use category as an outside good ensures that each grid is invested in at least one of the alternatives. This is in the spirit of the Hicksian composite commodity approach in consumer theory

Following Bhat (2008), consider a vector $\boldsymbol{x}_q$ of dimension $K \times 1$ with elements x_{qk} $(x_{qk} > 0 \ \forall \ q)$, where x_{qk} is a specific value of land-use investment in land-use type k in grid q To determine grid q's optimal allocation of its land area E_q among the K alternative land-uses (that is, to obtain the optimal values of x_{qk}), we consider the presence of a baseline utility ψ_{qk} ($\psi_{qk} > 0$) for each land-use type k and grid q that represents the marginal utility at the point of zero parcel area under land use k in grid q. The idea is that the first unit of land in grid q will be assigned to the land-use type k that has the highest baseline utility. But the marginal utility accrued from investing in any land-use type k is assumed to decrease as the grid is invested more and more in the land-use type k. Thus, at some point, an additional unit of investment in land-use type k may provide less marginal value than the first unit of investment in some other land-use type k', at which point there is investment in land-use type k'. Assume this process plays out until all the land area E_q is allocated among the many land-uses. In that case, the marginal utility of the consumed alternatives at the optimal point will be equal, and will be greater than the baseline marginal utility of the non-consumed alternatives. This conceptualization forms the basis for the MDCP model. Formally, consider the following utility-maximizing function subject to the binding land area constraint:

(1)
$$\max U_q(\boldsymbol{x}_q) = \left[\sum_{k=1}^{K-1} \gamma_k \psi_{qk} \ln\left(\frac{x_{qk}}{\gamma_k} + 1 \right) \right] + \psi_{qK} \ln x_{qK}$$

$$s.t. \sum_{k=1}^{K} x_{qk} = E_q \, ,$$

---

in that one replaces all the elementary alternatives that are not of primary interest (for example, the non-built up land-use types in the empirical analysis of the current paper) by a single composite "undeveloped" land use. The analysis proceeds then by considering the composite good as an "outside" good and modeling consumption in this outside good as well as in the more finely categorized "inside" goods representing the group of main interest to the analyst (in this case, the alternatives other than the undeveloped land-use category). This approach is very general, and can be used to study any categorization of land-use types. For example, in some land-use and climate change studies, the amount of area in dense vegetation may be the focus of interest, in which case the area in dense vegetation may be included as an "inside" land use category, while still maintaining other kinds of undeveloped land and perhaps even built-up land uses as an "outside" category. Finally, we should note that the model developed in this paper can be easily modified to the case when there is no outside category, and zero investment is possible in all land-use categories.

where the utility function $U_q(\boldsymbol{x}_q)$ is quasi-concave, increasing and continuously differentiable.[4] γ_k is a parameter associated with land-use type k. The utility function form in Equation (1) allows corner solutions (*i.e.*, zero consumptions) for the land-use alternatives 1 through $K-1$ through the parameters γ_k (as long as $\gamma_k > 0: k = 1, 2, ..., K-1$). On the other hand, the functional form for the final land-use alternative (the undeveloped land-use alternative) ensures that some land in each grid is in an undeveloped state (this can be easily proved given the specific utility structure used; see Bhat, 2008). The parameter γ_k for the first $K-1$ land-use types also contributes to the rate at which the marginal utility of investing in land-use type k decreases with increasing investment x_k in the land-use type, which can be seen by computing the marginal utility of consumption with respect to good k ($k = 1, 2, ..., K-1$):

$$\frac{\partial U_q(\boldsymbol{x}_q)}{\partial x_{qk}} = \psi_{qk} \left(\frac{x_{qk}}{\gamma_k} + 1 \right)^{-1}.$$ As long as $\gamma_k > 0$, the marginal utility decreases with increasing x_{qk},

as required. Further, if γ_k increases, the rate of the marginal utility decrease itself decreases.

To complete the model structure, the baseline utility ψ_{qk}, which has to be non-negative, is parameterized as follows for each alternative:

(2) $\qquad \psi_{qk} = \exp(\tilde{z}_{qk}, \xi_{qk}) = \exp(\boldsymbol{\beta}_q' \tilde{z}_{qk} + \xi_{qk})$ or $\overline{\psi}_{qk}^* = \ln(\psi_{qk}) = \boldsymbol{\beta}_q' \tilde{z}_{qk} + \xi_{qk}$,

where $\tilde{z}_{qk}$ is a D-dimensional vector of attributes that characterizes land-use type k and grid q (including a dummy variable for each alternative except the last outside alternative, to capture intrinsic preferences for each alternative relative to the last alternative), $\boldsymbol{\beta}_q$ is a grid-specific vector of coefficients (of dimension $D \times 1$), and ξ_{qk} captures the idiosyncratic (unobserved) characteristics that impact the baseline utility of land-use type k and grid q. We assume that the error term vector $\boldsymbol{\xi}_q = (\xi_{q1}, \xi_{q2}, ..., \xi_{qK})'$ is multivariate normally distributed for a given grid q: $\boldsymbol{\xi}_q \sim MVN_K(\boldsymbol{0}_K, \breve{\boldsymbol{\Lambda}})$, where $MVN_K(\boldsymbol{0}_K, \breve{\boldsymbol{\Lambda}})$ indicates a K-variate normal distribution with a mean vector of zeroes denoted by $\boldsymbol{0}_K$ and a covariance matrix $\breve{\boldsymbol{\Lambda}}$. Further, to allow heterogeneity in

---

[4] The utility form in Equation (1) is a restrictive form of the more general utility form proposed by Bhat (2008). In our empirical analysis, we started from the more general form and estimated alternative estimable restrictive forms. But the specific form in Equation (1) provided the best representation of utility. Additional details of the origins of the utility form may be found in Bhat (2008).

responsiveness to exogenous variables across grids (*i.e.*, spatial heterogeneity), we consider $\boldsymbol{\beta}_q$ as a realization from a multivariate normal distribution with mean vector $\boldsymbol{b}$ and covariance $\boldsymbol{\Omega} = \mathbf{LL}'$. That is, $\boldsymbol{\beta}_q \sim MVN_D(\boldsymbol{b}, \boldsymbol{\Omega})$. It is not necessary that all elements of $\boldsymbol{\beta}_q$ 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 $\boldsymbol{\beta}_q$ are random. The vectors $\boldsymbol{\beta}_q$ and $\boldsymbol{\xi}_q$ are assumed to be independent of each other. For future reference, we also write $\boldsymbol{\beta}_q = \boldsymbol{b} + \tilde{\boldsymbol{\beta}}_q$, where $\tilde{\boldsymbol{\beta}}_q \sim MVN_D(\mathbf{0}_D, \boldsymbol{\Omega})$.[5]

As in the multinomial probit model, only differences in the logarithm of the baseline utilities matter, not the actual logarithm of the baseline utility values (see Bhat, 2008). Thus, it will be easier to work with the logarithm of the baseline utilities of the first $K-1$ alternatives, and normalize the logarithm of the baseline utility for the last alternative to zero. That is, we write:

$$\overline{\psi}_{qk} = \overline{\psi}_{qk}^* - \overline{\psi}_{qK}^* = \boldsymbol{\beta}_q'(\tilde{\boldsymbol{z}}_{qk} - \tilde{\boldsymbol{z}}_{qK}) + (\xi_{qk} - \xi_{qK})$$

(3)
$$= \boldsymbol{\beta}_q' \boldsymbol{z}_{qk} + \varepsilon_{qk}, \ \boldsymbol{z}_{qk} = \tilde{\boldsymbol{z}}_{qk} - \tilde{\boldsymbol{z}}_{qK}, \varepsilon_{qk} = (\xi_{qk} - \xi_{qK}) \ \forall k \neq K$$

$$\overline{\psi}_{qK} = \overline{\psi}_{qK}^* - \overline{\psi}_{qK}^* = 0 \text{ for } k = K.$$

It should be clear from above that only the covariance matrix, say $\boldsymbol{\Lambda}$ of the error differences $\varepsilon_{qk} = (\xi_{qk} - \xi_{qK})$, is estimable, and not the covariance matrix $\breve{\boldsymbol{\Lambda}}$ of the original error terms. Further, with the formulation as in Equation (1), where the sum of the investments across land-use types (which constitute the dependent variables) is equal to the total land area in the grid, an additional scale normalization needs to be imposed (see Bhat, 2008). A convenient normalization is to set the first element of $\boldsymbol{\Lambda}$ (that is, $\boldsymbol{\Lambda}_{11}$ to one). Further, technically speaking, the fully unrestricted substitution pattern implied by the full covariance matrix for $\boldsymbol{\Lambda}$ comes at the expense of rendering the estimated parameters of the $\boldsymbol{\Lambda}$ matrix completely uninterpretable (see Train, 2009; page 113 for a similar discussion in the case of traditional multinomial probit models). The approach we adopt in this paper to make the parameters behaviorally interpretable is to impose the not-so-implausible structure that, for each grid, the error term of the "outside"

---

[5] Note, however, that the parameters (in the $\boldsymbol{\beta}_q$ vector) on the dummy variables specific to each alternative (except the last) have to be fixed parameters in the cross-section model, since their randomness is already captured in the covariance matrix $\bar{\boldsymbol{\Lambda}}$.

alternative ξ_{qK} is independent of the error terms of the inside alternatives ξ_{qk} ($k = 1, 2, ..., K - 1$).

With this assumption, each covariance matrix element of $\mathbf{\Lambda}$ can then immediately be interpreted as a direct indicator of the extent of variance and covariance in the utilities of the inside alternatives.[6]

The analyst can solve for the optimal consumption allocations corresponding to Equation (1) by forming the Lagrangian and applying the Karush-Kuhn-Tucker (KKT) conditions. The Lagrangian function for the problem, after substituting $\psi_{qk} = \exp(\overline{\psi}_{qk})$ (equal to $\exp(\boldsymbol{\beta}'_q \mathbf{z}_{qk} + \varepsilon_{qk})$ for $k = 1, 2, ..., K - 1$ and equal to $\exp(0) = 1$ for $k = K$) in Equation (1) is:

$$
(4) \qquad L_q = \sum_{k=1}^{K-1} \gamma_k \exp(\boldsymbol{b}' \mathbf{z}_{qk} + \tilde{\boldsymbol{\beta}}'_q \mathbf{z}_{qk} + \varepsilon_{qk}) \ln\left(\frac{x_{qk}}{\gamma_k} + 1\right) + \ln(x_{qK}) - \lambda_q \left[\sum_{k=1}^{K} x_{qk} - E_q\right],
$$

where λ_q is the Lagrangian multiplier associated with the land area constraint (that is, it can be viewed as the marginal utility of total land area). The KKT first-order condition for the "optimal" investment x^*_{qK} in undeveloped land (which is always positive for each grid) implies the following: $\left(x^*_{qK}\right)^{-1} - \lambda_q = 0$; that is, $\lambda_q = \left(x^*_{qK}\right)^{-1}$. The KKT first-order conditions for the optimal land investments for the inside alternatives (the x^*_{qk} values for $k = 1, 2, ..., K - 1$) are given by:

$$
(5) \qquad \exp(\boldsymbol{b}' \mathbf{z}_{qk} + \tilde{\boldsymbol{\beta}}'_q \mathbf{z}_{qk} + \varepsilon_{qk}) \left(\frac{x^*_{qk}}{\gamma_k} + 1\right)^{-1} - \lambda_q = 0 \text{ , if } x^*_{qk} > 0, \ k = 1, 2, ..., K - 1
$$

$$
\exp(\boldsymbol{b}' \mathbf{z}_{qk} + \tilde{\boldsymbol{\beta}}'_q \mathbf{z}_{qk} + \varepsilon_{qk}) \left(\frac{x^*_{qk}}{\gamma_k} + 1\right)^{-1} - \lambda_q < 0, \text{ if } x^*_{qk} = 0, \ k = 1, 2, ..., K - 1.
$$

---

[6] To be precise, assume that the variance of ξ_{qK} is 0.5. Then, to normalize $\mathbf{\Lambda}_{11}$ to one, we should have that the variance of ξ_{q1} is also 0.5. Let the variance of ξ_{qk} ($k = 2, 3, ..., K - 1$) be σ_k^2 and the covariance between ξ_{qk} and $\xi_{qk'}$ ($k = 1, 2, 3, ..., K - 1; k \neq k'$) be $\sigma_{kk'}$. Then, the matrix $\mathbf{\Lambda}$ of the error differences $\varepsilon_{qk} = (\xi_{qk} - \xi_{qK})$ is:

$$
\mathbf{\Lambda} = \begin{bmatrix}
1 & 0.5 + \sigma_{12} & 0.5 + \sigma_{13} & \cdots & 0.5 + \sigma_{1,K-1} \\
0.5 + \sigma_{12} & 0.5 + \sigma_2^2 & 0.5 + \sigma_{23} & \cdots & 0.5 + \sigma_{2,K-1} \\
0.5 + \sigma_{13} & 0.5 + \sigma_{23} & 0.5 + \sigma_3^2 & \cdots & 0.5 + \sigma_{3,K-1} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0.5 + \sigma_{1,K-1} & 0.5 + \sigma_{2,K-1} & 0.5 + \sigma_{3,K-1} & \cdots & 0.5 + \sigma_{K-1}^2
\end{bmatrix}.
$$

Substituting $\lambda_q = \left(x_{qK}^*\right)^{-1}$ into the above Equation, and taking logarithms, we can rewrite the KKT conditions as:

(6)
$$y_{qk}^* = (V_{qk} - V_{qK}) + \widetilde{\varepsilon}_{qk} = 0, \text{ if } x_{qk}^* > 0, \ k = 1, 2, ..., K-1$$

$$y_{qk}^* = (V_{qk} - V_{qK}) + \widetilde{\varepsilon}_{qk} < 0, \text{ if } x_{qk}^* = 0, \ k = 1, 2, ..., K-1,$$

where $V_{qk} = \boldsymbol{b'z_{qk}} - \ln\left(\dfrac{x_{qk}^*}{\gamma_k} + 1\right)$ for $k = 1, 2, ..., K-1$, $V_{qK} = -\ln\left(x_{qK}^*\right)$, and $\widetilde{\varepsilon}_{qk} = \widetilde{\boldsymbol{\beta}}_q' \boldsymbol{z_{qk}} + \varepsilon_{qk}$.

The Spatial MDCP (or SMDCP) Model

We retain all notations from the aspatial model, and begin the formulation of the spatial model from Equation (3), and write the logarithm of the baseline utilities (taken as the difference from the logarithm of the baseline utility of the last alternative) for the alternatives as follows:

(7)
$$\overline{\psi}_{qk} = \boldsymbol{\beta}_q' \boldsymbol{z_{qk}} + \varepsilon_{qk} + \delta_k \sum_{q'} w_{qq'} \overline{\psi}_{q'k}, \text{ for } k = 1, 2, ..., K-1$$

$$\overline{\psi}_{qK} = 0 \text{ for } k = K.$$

The only difference from Equation (3) is the presence of the component $\delta_k \sum_{q'} w_{qq'} \overline{\psi}_{q'k}$ in the logarithm of the baseline utilities for the inside alternatives. This component takes the typical spatial lag specification used extensively in spatial econometrics, and causes the logarithm of the baseline utilities to be spatially interdependent across grids based on the spatial proximity of grids. In particular, $w_{qq'}$ is a distance-based spatial weight corresponding to grids q and q' (with $w_{qq} = 0$ and $\sum_{q'} w_{qq'} = 1$) for every q, and δ_k ($0 < \delta_k < 1$) is the spatial lag autoregressive parameter specific to land-use type k ($k = 1, 2, ..., K-1$).[7]

---

[7] Unlike other spatial econometric studies in the context of traditional unordered discrete choice (such as Sener and Bhat, 2012 and Sidharthan and Bhat, 2012) that do not allow the spatial lag parameter to vary across alternatives, we allow this parameter to vary across alternatives in the current study, as should be obvious from the subscript k in δ_k. This is because of an important nuance. In the current study, the spatial dependence patterns for the first $K–1$ alternatives effectively determine the spatial pattern for the last K^{th} alternative (because of the land-use constraint). We expressly acknowledge this "identification" problem in spatial dependence by not allowing the spatial lag on the Kth alternative (that is using this K^{th} alternative as the base for introducing spatial dependence effects). Of course, this K^{th} alternative is easily identified in the current paper as the "outside" alternative that is always chosen. However, in traditional discrete choice models where only one "inside" alternative can be chosen (and a similar identification problem arises because only utility differences matter), deciding which alternative to use as the base

The reader will note that the spatial specification in Equation (7) allows spatial spillovers for the same land-use type as well as across land-use types. This is because, as in standard discrete choice models where the spatial lag structure is applied to the linear (latent) propensity variables underlying the discrete choice dependent variable (see reviews of this literature in LeSage and Pace, 2009; Hays et al., 2010; Brady and Irwin, 2011; and Sidharthan and Bhat, 2012), the spatial lag structure of Equation (7) is associated with the linear (latent) log(baseline utility) measure underlying the multiple discrete-continuous variable. Thus, even as a variable in a certain grid A increases the investment intensity in a specific land-use type in that grid (because of say a positive effect of the variable on the (log) baseline utility of the specific land-use type), it leads to an increase in investment intensity in the same land-use type of neighboring grids (assuming a positive spatial autoregressive lag coefficient in Equation (7) for the land-use type, as we would generally expect).[8] At the same time, everything else being the same, an increase in the investment intensity in land-use type k in grid A and other neighboring grids will have a negative effect on investment intensity in other land-use types in grid A and the neighboring grids (because the area of each grid is fixed). Thus, the specification does allow spatial spillover effects across land-use types.

We now set out additional notation to write the baseline utility in a compact form. Define the following:

$$\overline{\mathbf{\Psi}}_q = (\overline{\psi}_{q1}, \overline{\psi}_{q2}, ..., \overline{\psi}_{q,K-1})',\ \boldsymbol{\varepsilon}_q = (\varepsilon_{q1}, \varepsilon_{q2}, ..., \varepsilon_{q,K-1})'\ [(K-1) \times 1 \text{ vectors}]$$

$$\overline{\mathbf{\Psi}} = (\overline{\mathbf{\Psi}}_1', \overline{\mathbf{\Psi}}_2', ..., \overline{\mathbf{\Psi}}_Q')', \qquad \boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1', \boldsymbol{\varepsilon}_2', ..., \boldsymbol{\varepsilon}_Q')' \qquad [Q(K-1) \times 1 \text{ vectors}]$$

---

for introducing spatial dependence is not at all clear. Importantly, the determination of the base alternative for spatial dependence effects is not innocuous, since different results would be obtained by using different alternatives as the base (this exchangeability problem has seldom been discussed in the literature). This is the reason that earlier studies of traditional unordered discrete choice models impose the same spatial lag parameter for all alternatives, which resolves the identification problem as well as does not have the problem of non-unique results.

[8] An increase in $\overline{\psi}_{qk}$ leads to an increase in $\psi_{qk} = \exp(\overline{\psi}_{qk})$. Then, as discussed in the previous section of the paper,

$\dfrac{\partial U_q(\boldsymbol{x}_q)}{\partial x_{qk}} = \psi_{qk} \left(\dfrac{x_{qk}}{\gamma_k} + 1 \right)^{-1}$. Thus, as the baseline utility ψ_{qk} increases for a grid q, the marginal utility at any

consumption level x_{qk} (including at $x_{qk} = 0$) also increases. By the formulation, then, the likelihood of grid q being invested in land-use type k increases, as does the amount of land invested in land-use type k if the grid is already invested in the land-use type k. Overall, there is then an increase in investment intensity in land-use type k in grid q because of both these discrete as well as continuous effects. In our spatial MDCP model based on Equation (7), an increase in the baseline utility ψ_{qk} also increases the baseline utility for grid type k for neighboring grids, which increases investment intensity in land-use type k even in these neighboring grids.

$\mathbf{z}_q = (z_{q1}, z_{q2}, ..., z_{q,K-1})'$ $[(K-1) \times D$ matrix], $\mathbf{z} = (z_1', z_2', ..., z_Q')'$ $[Q(K-1) \times D$ matrix], and

$\tilde{\boldsymbol{\beta}} = (\tilde{\boldsymbol{\beta}}_1', \tilde{\boldsymbol{\beta}}_2', ..., \tilde{\boldsymbol{\beta}}_Q')'$ $(QD \times 1$ vector).

Let $\mathbf{IDEN}_K$ be the identity matrix of size K. Also, define the following matrices:

$$(8) \qquad \ddot{\mathbf{z}} = \begin{bmatrix} z_1 & 0 & 0 & \cdots & 0 \\ 0 & z_2 & 0 & \cdots & 0 \\ 0 & 0 & z_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & z_Q \end{bmatrix} \quad [Q(K-1) \times QD \text{ matrix}], \text{ and}$$

$$(9) \qquad \boldsymbol{\delta} = \begin{bmatrix} \delta_1 & 0 & 0 & \cdots & 0 \\ 0 & \delta_2 & 0 & \cdots & 0 \\ 0 & 0 & \delta_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \delta_{K-1} \end{bmatrix} \quad [(K-1) \times (K-1) \text{ matrix}].$$

Let $\tilde{W}$ be the $(Q \times Q)$ weight matrix with weight $w_{qq'}$ as its elements, and let $\mathbf{1}_{QQ}$ be a $(Q \times Q)$ matrix with each element taking the value of one. Next, define $\mathbf{W} = (\mathbf{1}_{QQ} \otimes \boldsymbol{\delta}) .* (\tilde{W} \otimes \mathbf{IDEN}_{K-1})$, where "$\otimes$" is the kronecker product and ".$*$" stands for the element-by-element multiplication of two matrices. Let $\mathbf{S} = [\mathbf{IDEN}_{Q(K-1)} - \mathbf{W}]^{-1}$ $[Q(K-1) \times Q(K-1)$ matrix]. Then, we can write Equation (7) for all alternatives k ($k = 1, 2, ..., K-1$) and all grids $q = 1, 2, ..., Q$ in matrix notation as:

$$(10) \qquad \overline{\boldsymbol{\psi}} = \mathbf{S}(\mathbf{z}b + \ddot{\mathbf{z}}\tilde{\boldsymbol{\beta}} + \varepsilon) = \mathbf{S}\mathbf{z}b + \mathbf{S}(\ddot{\mathbf{z}}\tilde{\boldsymbol{\beta}} + \varepsilon) \quad [Q(K-1) \times 1 \text{ vector}].$$

Let $[.]_e$ indicate the eth element of the column vector $[.]$, and let $d_{qk} = (q-1)K + k$. Equation (10) can be equivalently written as:

$$(11) \qquad \overline{\psi}_{qk} = [\mathbf{S}\mathbf{z}b]_{d_{qk}} + [\mathbf{S}(\ddot{\mathbf{z}}\tilde{\boldsymbol{\beta}} + \varepsilon)]_{d_{qk}}, \ k = 1, 2, ..., K-1.$$

Using the same approach as for the aspatial case, the KKT conditions for the land-use pattern for each grid q take the same form for y_{qk}^* as in Equation (6) with the new definitions of V_{qk} ($k = 1, 2, ..., K-1$), V_{qK}, and $\tilde{\varepsilon}_{qk}$ as follows:

15

(12)
$$V_{qk} = [\mathbf{Szb}]_{d_{qk}} - \ln\left(\frac{x^*_{qk}}{\gamma_k} + 1\right) \text{ for } k = 1, 2, ..., K-1,$$

$$V_{qK} = -\ln(x^*_{qK}), \text{ and } \widetilde{\varepsilon}_{qk} = [\mathbf{S}(\ddot{\mathbf{z}}\widetilde{\beta} + \varepsilon)]_{d_{qk}}.$$

Now, stack the elements y^*_{qk} ($k = 1, 2, ..., K-1$) in the following order:

(13)
$$\mathbf{y}^*_q = (y_{q1}, y_{q2}, ..., y_{q,K-1})', \text{ a } (K-1) \times 1 \text{ vector, and}$$

$$\mathbf{y}^* = (\mathbf{y}_1^{*'}, \mathbf{y}_2^{*'}, ..., \mathbf{y}_Q^{*'})', \text{ a } (Q \times (K-1)) \times 1 \text{ vector.}$$

Define the following additional matrices:

(14)
$$\mathbf{B}_q = (V_{q1} - V_{qK}, V_{q2} - V_{qK}, ..., V_{q,K-1} - V_{qK})' \ [(K-1) \times 1 \text{ vector}],$$

$$\mathbf{B} = (\mathbf{B}_1', \mathbf{B}_2', ..., \mathbf{B}_Q')' \ [Q(K-1) \times 1 \text{ vector}].$$

It is easy to see that $\mathbf{y}^*$ has a mean vector of $\mathbf{B}$. To determine the covariance matrix of $\mathbf{y}^*$, define the following additional matrices:

$$\widetilde{\Lambda} = \mathbf{IDEN}_Q \otimes \Lambda \ [Q(K-1) \times Q(K-1) \text{ matrix }],$$

(15)
$$\widetilde{\Omega} = \ddot{\mathbf{z}}(\mathbf{IDEN}_Q \otimes \Omega)\ddot{\mathbf{z}}' \ [Q(K-1) \times Q(K-1) \text{ matrix }], \text{ and}$$

$$\Sigma = \mathbf{S}[\widetilde{\Lambda} + \widetilde{\Omega}]\mathbf{S}' \ [Q(K-1) \times Q(K-1) \text{ matrix}].$$

Then, we obtain that $\mathbf{y}^* \sim MVN_{Q \times (K-1)}(\mathbf{B}, \Sigma)$.

3. MODEL ESTIMATION

Development of the Maximum Likelihood Estimator

Let $\gamma = (\gamma_1, \gamma_2, ..., \gamma_{K-1})'$. The parameters to estimate in the spatial MDCP model include the γ parameter vector, the $\mathbf{b}$ vector, the elements of the spatial lag parameter matrix δ, and the covariance matrices - Λ and Ω. The data provide information on the vector of exogenous variables $\mathbf{z}_q$ for each grid q, the spatial weight matrix $\widetilde{W}$, and the observed land-use square footage vector across the land-use alternatives in each grid q: $\mathbf{x}^*_q = (x^*_{q1}, x^*_{q2}, ...x^*_{q,K})'$. Note that a specific grid may not be invested in some land-use types, in which case the corresponding x^*_{qk} values take a value of zero. Let θ be the collection of these parameters: $\theta = [\gamma', \mathbf{b}', \text{Vech}(\delta), \text{Vech}(\Lambda), \text{Vech}(\Omega)]$, where $\text{Vech}(\Lambda)$ and $\text{Vech}(\Omega)$ represent the column

vector of upper triangle elements of $\mathbf{\Lambda}$ and $\mathbf{\Omega}$, respectively, and $\text{Vech}(\mathbf{\delta})$ represents the column vector of diagonal element of $\mathbf{\delta}$.

Next, partition the vector $\mathbf{y}^*$ into a sub-vector $\widetilde{\mathbf{y}}_{NC}^*$ of length $L_{NC} \times 1$ ($[0 \le L_{NC} \le Q(K-1)]$) corresponding to the grid and land-use type combinations in which there is no land investment, and another sub-vector $\widetilde{\mathbf{y}}_C^*$ of length $L_C \times 1$ ($[0 \le L_C \le Q(K-1)]$) for the grid and land-use type combinations in which there is land investment ($[L_{NC} + L_C = Q(K-1)]$). In forming the sub-vector $\widetilde{\mathbf{y}}_C^*$, the outside alternative is not included. Let $\widetilde{\mathbf{y}}^* = \left(\left[\widetilde{\mathbf{y}}_{NC}^* \right]', \left[\widetilde{\mathbf{y}}_C^* \right]' \right)'$, which may be obtained from $\mathbf{y}^*$ as $\widetilde{\mathbf{y}}^* = \mathbf{R}\mathbf{y}^*$, where $\mathbf{R}$ is a re-arrangement matrix of dimension $Q(K-1) \times Q(K-1)$ with zeroes and ones. For example, consider the case of three grids and five land-use alternatives. The last alternative is the "undeveloped" land-use state, which is the outside alternative. Among the remaining four alternatives, let grid 1 be invested in alternatives 1 and 4 (not invested in alternatives 2 and 3), let grid 2 be invested in alternatives 2 and 3 (not invested in alternatives 1 and 4), and let grid 3 be invested in alternative 1 (not invested in alternatives 2, 3, and 4). In this case, $L_{NC} = 7$ and $L_C = 5$. Then, the re-arrangement matrix $\mathbf{R}$ is:

$$
(16) \quad \mathbf{R} == \left[\begin{array}{cccc|cccc|cccc}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
\hline
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0
\end{array} \right] = \left[\begin{array}{c} \mathbf{R}_{NC} \\ \hline \mathbf{R}_C \end{array} \right],
$$

where the upper sub-matrix $\mathbf{R}_{NC}$ corresponds to the grid and land-use alternative combinations with no land investment (of dimension $L_{NC} \times Q(K-1)$) and the lower sub-matrix $\mathbf{R}_C$

corresponds to the grid and land-use alternative combinations (excluding the outside alternative for each grid) with positive land investment (of dimension $L_C \times Q(K-1)$). Note also that $\tilde{\boldsymbol{y}}_{NC}^* = \mathbf{R}_{NC}\boldsymbol{y}^*$ and $\tilde{\boldsymbol{y}}_C^* = \mathbf{R}_C\boldsymbol{y}^*$.[9]

Consistent with the above re-arrangement, define $\tilde{\boldsymbol{H}} = \mathbf{R}\boldsymbol{B}$, $\tilde{\boldsymbol{H}}_{NC} = \mathbf{R}_{NC}\boldsymbol{B}$, $\tilde{\boldsymbol{H}}_C = \mathbf{R}_C\boldsymbol{B}$, and $\tilde{\boldsymbol{\Sigma}} = \mathbf{R}\boldsymbol{\Sigma}\mathbf{R}'$. Define $\boldsymbol{x}^* = \left(\left[\boldsymbol{x}_1^*\right]', \left[\boldsymbol{x}_2^*\right]', \ldots \left[\boldsymbol{x}_Q^*\right]' \right)'$. Then, the maximum likelihood function may be obtained as:

$$(17) \qquad L_{ML}(\boldsymbol{\theta}) = \text{Prob}(\boldsymbol{x}^*) = \det(\mathbf{J}) \int_{\boldsymbol{h}_{NC}=-\infty}^{\boldsymbol{0}_{L_{NC}}} f_{Q(K-1)}\left(\mathbf{h}_{NC}, \boldsymbol{0}_{L_C} \mid \tilde{\boldsymbol{H}}, \tilde{\boldsymbol{\Sigma}}\right) d\boldsymbol{h}_{NC},$$

where $\det(\mathbf{J})$ is the determinant of the Jacobian of the transformation from $\boldsymbol{y}^*$ to the consumption quantities $\boldsymbol{x}^*$ (see Bhat, 2008). The matrix $\mathbf{J}$ of dimension $(L_C \times L_C)$ is block-diagonal with each block matrix $\mathbf{J}_q$ of size $(L_{qC} \times L_{qC})$ corresponding to a specific grid q (L_{qC} is the number of inside alternatives in which grid q is invested in). The block diagonality of $\mathbf{J}$ arises because $\dfrac{\partial y_{qk}^*}{\partial x_{q'h}^*} = 0$ for all $q \neq q'$ and $k, h \in \tilde{L}_{qC}$ ($\tilde{L}_{qC}$ is the set of inside alternatives in which grid q is invested, so that L_{qC} is the cardinality of the set $\tilde{L}_{qC}$; for future use, we will also define $\tilde{L}_{q,NC}$ as the set of alternatives in which grid q is not invested in, with $L_{q,NC}$ being the cardinality of the set $\tilde{L}_{q,NC}$). Let $\breve{L}_{qC}$ be the set of all land-use alternatives in which grid q is invested in (that is, those in the set $\tilde{L}_{qC}$ plus the outside alternative K). Using the derivation approach in Bhat (2005) for each block matrix $\mathbf{J}_q$, and due to the block-diagonality of the larger matrix $\mathbf{J}$, we are able to write:

$$(18) \qquad \det(\mathbf{J}) = \prod_{q=1}^{Q} \left[\left\{ \prod_{k \in \breve{L}_{qC}} \frac{1-\alpha_k}{x_{qk}^* + \gamma_k} \right\} \left\{ \sum_{k \in \breve{L}_{qC}} \left(\frac{x_{qk}^* + \gamma_k}{1-\alpha_k} \right) \right\} \right].$$

---

[9] $\mathbf{R}_{NC}$ has as many rows and columns as the number of grid and land-use alternative combinations with no land investment (each column corresponds to an alternative except the Kth alternative). Then, for each row, $\mathbf{R}_{NC}$ has a value of "1" in one of the columns corresponding to a grid- alternative combination that is not invested in (starting from the first alternative that is not invested in for the first grid and working down to the last alternative that is not invested in for the last grid). Each row has strictly one column with a value of "1" and the value of "0" everywhere else. A similar construction is involved in creating the $\mathbf{R}_C$ matrix.

The likelihood function in Equation (17) involves integration of dimension L_{NC}. This is of very high dimensionality in the typical case of sample sizes of 500 grids or more. The lower bound of L_{NC} is equal to zero, corresponding to the case when each grid is invested in each land-use alternative. The upper bound is equal to $(K-1)*Q$, corresponding to the case when each grid is invested in only the undeveloped (outside) land-use alternative state. Of course, in practice, the situation will be somewhere between these two extreme values for L_{NC}, but the value for L_{NC} will be sufficient to render maximization of the likelihood function using traditional simulation methods almost impractical. In particular, existing estimation methods, including the Maximum Simulated Likelihood (MSL) method and the Bayesian Inference method, become cumbersome and encounter convergence problems even for moderately sized Q (Bhat et al., 2010). In this paper, we instead use Bhat's Maximum Approximate Composite Marginal Likelihood (MACML) inference approach for estimation.

The 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, maximizes a surrogate likelihood function that compounds much easier-to-compute, lower-dimensional, marginal likelihoods (see Bhat, 2014 for a recent extensive review of CML methods; Lindsay et al., 2011;

Bhat, 2011; Varin et al., 2011; and Yi et al., 2011 are also useful references). The CML approach, which belongs to the more general class of composite likelihood function approaches (see Lindsay, 1988), may be explained in a simple manner as follows. In the SMDCP model, instead of developing the likelihood function for the entire set of Q observations, as in Equation (17), one may compound (multiply) pairwise probabilities of grid q having the land-use pattern x_q^*, grid q' having the land-use pattern $x_{q'}^*$, grid q'' having the land-use pattern $x_{q''}^*$, and so on. The CML estimator (in this instance, the pairwise CML estimator) is then the one that maximizes the compounded probability of all pairwise events. The properties of the CML estimator may be derived using the theory of estimating equations (see Cox and Reid, 2004; Yi et al., 2011). Specifically, under usual regularity assumptions (Molenberghs and Verbeke, 2005, page 191; Xu and Reid, 2011), 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; for a formal proof, see Bhat, 2014). Parameter restrictions in the CML estimation approach may be tested using the composite likelihood ratio statistic (see Bhat, 2014).

To write the pairwise CML function, let $L_{qq',NC} = L_{q,NC} + L_{q',NC}$ and $L_{qq',C} = L_{qC} + L_{q'C}$. Define a vector $y_{qq'}^*$ of size $[2(K-1) \times 1]$ as follows:

$$(19) \qquad y_{qq'}^* = \left(\left[y_q^* \right]', \left[y_{q'}^* \right]' \right)'.$$

Let $\Delta_{qq'}$ be a selection matrix of size $2 \times Q$. This matrix has the value of "1" in the top row and the column q, and the value of "1" in the bottom row and column q'. All other cells of this matrix are filled with values of zero. Then, $y_{qq'}^* \sim MVN_{2 \times (K-1)}(B_{qq'}, \Sigma_{qq'})$, where $B_{qq'} = (\Delta_{qq'} \otimes \mathbf{IDEN}_{K-1})B$, and $\Sigma_{qq'} = (\Delta_{qq'} \otimes \mathbf{IDEN}_{K-1})\Sigma(\Delta_{qq'} \otimes \mathbf{IDEN}_{K-1})'$. Next, define the re-arrangement matrices $\mathbf{R}_{qq'}$ (of dimension $2(K-1) \times 2(K-1)$), $\mathbf{R}_{qq',NC}$ (of dimension $L_{qq',NC} \times 2(K-1)$), and $\mathbf{R}_{qq',C}$ (of dimension $L_{qq',C} \times 2(K-1)$) similar to the corresponding re-arrangement matrices defined on the entire sample for the maximum likelihood approach. Also, define $\widetilde{B}_{qq',NC} = \mathbf{R}_{qq',NC}B_{qq'}$, $\widetilde{B}_{qq',C} = \mathbf{R}_{qq',C}B_{qq'}$, and $\widetilde{\Sigma}_{qq'} = \mathbf{R}_{qq'}\Sigma_{qq'}\mathbf{R}'_{qq'} = \begin{bmatrix} \widetilde{\Sigma}_{qq',NC} & \widetilde{\Sigma}'_{qq',NC,C} \\ \widetilde{\Sigma}_{qq',NC,C} & \widetilde{\Sigma}_{qq',C} \end{bmatrix}$,

where $\widetilde{\boldsymbol{\Sigma}}_{qq',NC} = \mathbf{R}_{qq',NC}\boldsymbol{\Sigma}_{qq'}\mathbf{R}'_{qq',NC}$, $\widetilde{\boldsymbol{\Sigma}}_{qq',C} = \mathbf{R}_{qq',C}\boldsymbol{\Sigma}_{qq'}\mathbf{R}'_{qq',C}$, and $\widetilde{\boldsymbol{\Sigma}}_{qq',NC,C} = \mathbf{R}_{qq',NC}\boldsymbol{\Sigma}_{qq'}\mathbf{R}'_{qq',C}$. Let

$$\breve{\boldsymbol{B}}_{qq',NC} = \widetilde{\boldsymbol{B}}_{qq',NC} + \widetilde{\boldsymbol{\Sigma}}'_{qq',NC,C}(\widetilde{\boldsymbol{\Sigma}}_{qq',C})^{-1}(-\widetilde{\boldsymbol{B}}_{qq',C}), \qquad \breve{\boldsymbol{\Sigma}}_{qq',NC} = \widetilde{\boldsymbol{\Sigma}}_{qq',NC} - \widetilde{\boldsymbol{\Sigma}}'_{qq',NC,C}(\widetilde{\boldsymbol{\Sigma}}_{qq',C})^{-1}\widetilde{\boldsymbol{\Sigma}}_{qq',NC,C},$$

$$\widetilde{\boldsymbol{B}}^*_{qq',C} = \boldsymbol{\omega}^{-1}_{\widetilde{\Sigma}_{qq',C}}\left(-\widetilde{\boldsymbol{B}}_{qq',C}\right), \qquad \breve{\boldsymbol{B}}^*_{qq',NC} = \boldsymbol{\omega}^{-1}_{\breve{\Sigma}_{qq',NC}}\left(-\breve{\boldsymbol{B}}_{qq',NC}\right), \qquad \breve{\boldsymbol{\Sigma}}^*_{qq',NC} = \boldsymbol{\omega}^{-1}_{\breve{\Sigma}_{qq',NC}}\breve{\boldsymbol{\Sigma}}_{qq',NC}\boldsymbol{\omega}^{-1}_{\breve{\Sigma}_{qq',NC}}, \qquad \text{and}$$

$$\widetilde{\boldsymbol{\Sigma}}^*_{qq',C} = \boldsymbol{\omega}^{-1}_{\widetilde{\Sigma}_{qq',C}}\widetilde{\boldsymbol{\Sigma}}_{qq',C}\boldsymbol{\omega}^{-1}_{\widetilde{\Sigma}_{qq',C}} \quad \text{where} \quad \boldsymbol{\omega}_{\breve{\Sigma}_{qq',NC}} \text{ is the diagonal matrix of standard deviations of } \breve{\boldsymbol{\Sigma}}_{qq',NC}$$

and $\boldsymbol{\omega}_{\widetilde{\Sigma}_{qq',C}}$ is the diagonal matrix of standard deviations of $\widetilde{\boldsymbol{\Sigma}}_{qq',C}$. Let $\overline{\boldsymbol{\omega}}_{\widetilde{\Sigma}_{qq',C}}$ be the product of the

diagonal elements of $\boldsymbol{\omega}_{\widetilde{\Sigma}_{qq',C}}$, and write the determinant of the Jacobian corresponding to grids q

and q' as $\det(\mathbf{J}_{qq'}) = \prod_{l=q,q'}\left[\left\{\prod_{k\in\bar{L}_{IC}}\frac{1-\alpha_k}{x^*_{lk}+\gamma_k}\right\}\left\{\sum_{k\in\bar{L}_{IC}}\left(\frac{x^*_{lk}+\gamma_k}{1-\alpha_k}\right)\right\}\right]$. Then, using the marginal and

conditional distribution properties of the multivariate normal distribution, the pairwise CML function for the SMDCP model can be written as:

$$\begin{aligned}
(20) \qquad L_{CML}(\boldsymbol{\theta}) &= \text{Prob}(x^*_q, x^*_{q'}) \\
&= \prod_{q=1}^{Q-1}\prod_{q'=q+1}^{Q}\det(\mathbf{J}_{qq'})\times\left(\overline{\boldsymbol{\omega}}_{\widetilde{\Sigma}_{qq',C}}\right)^{-1}\left[\phi_{L_{qq',C}}\left(\widetilde{\boldsymbol{B}}^*_{qq',C},\widetilde{\boldsymbol{\Sigma}}^*_{qq',C}\right)\right]\times\left[\Phi_{L_{qq',NC}}\left(\breve{\boldsymbol{B}}^*_{qq',NC},\breve{\boldsymbol{\Sigma}}^*_{qq',NC}\right)\right].
\end{aligned}$$

The CML function above requires the computation of the multivariate normal cumulative distribution (MVNCD) function that is utmost of dimension $(K-1)*2$ integrals (instead of $(K-1)*Q$ in the full maximum likelihood case). Such integrals may be computed easily using the MVNCD approximation method embedded in the MACML method (the MVNCD function approximates the pairwise probabilities in Equation (20) using only univariate and bivariate cumulative normal distribution functions; see Bhat, 2011 and Bhat and Sidharthan, 2011).

The MACML estimator is obtained by maximizing the logarithm of the function in Equation (20) after evaluating the MVNCD function using the analytic approximation. Since the MACML estimator entails only the computation of univariate and bivariate cumulative normal distribution functions, it is extremely quick to evaluate. The covariance matrix is given by the inverse of Godambe's (1960) sandwich information matrix (see Zhao and Joe, 2005). Bhat (2011) exploits the fading spatial dependence pattern implied by the spatial lag structure (due to the decaying nature of the distance weight matrix, combined with the spatial lag parameter being less than 1) to propose a specific implementation of Heagerty and Lumley's (2000) windows sampling procedure to estimate this sandwich information matrix.

The pairwise CML function of Equation (20) comprises $Q(Q-1)/2$ grid pairs of probability computations. To further accelerate the estimation, one can reduce the number of grid pairs because spatial dependency drops quickly with inter-grid distance. In fact, as demonstrated by Bhat et al. (2010) and Varin and Czado (2010), retaining all pairs not only increases computational costs, but may also reduce estimator efficiency. We examine this issue by creating different distance bands and, for each specified distance band, we consider only those pairings in the CML function that are within the spatial distance band. Then, we develop the asymptotic variance matrix $\mathbf{V}_{CML}(\hat{\boldsymbol{\theta}})$ for each distance band and select the threshold distance value that minimizes the total variance across all parameters as given by $tr[\mathbf{V}_{CML}(\hat{\boldsymbol{\theta}})]$ (*i.e.*, the trace of the matrix $[\mathbf{V}_{CML}(\hat{\boldsymbol{\theta}})]$).

A final issue regarding estimation. The analyst needs to ensure the positive definiteness of the two covariance matrices $\boldsymbol{\Omega}$ and $\boldsymbol{\Lambda}$. Once this is ensured, and as long as $0 < \delta_k < 1 \ \forall \ k$, $\boldsymbol{\Sigma}$ will be positive definite. In our estimation, the positive-definiteness of each of the $\boldsymbol{\Omega}$ and $\boldsymbol{\Lambda}$ 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 δ_k ($k = 1, 2, ..., K-1$) autoregressive terms, we parameterize these as $\delta_k = 1/[1 + \exp(\widetilde{\delta}_k)]$. Once estimated, the $\widetilde{\delta}_k$ estimates can be translated back to estimates of δ_k.

4. SIMULATION STUDY

There are two objectives of this simulation study. The first is to examine the ability of the MACML estimator to recover parameters from finite samples in the spatial MDCP model by generating simulated data sets with known underlying model parameters. The second is to examine the effects of (a) imposing a restrictive independent and identically distributed covariance among the baseline utilities of the alternatives, (b) ignoring spatial heterogeneity, and (c) ignoring spatial dependence.

Experimental Design

We consider a four alternative case in the simulation exercise, as in the empirical analysis of the current study. The last alternative, as in the earlier sections, is assumed to be the outside alternative. Assume three independent variables in the z_{qk} vector in the baseline utility. The values of each of the three independent variables for the alternatives are drawn from a standard univariate normal distribution, and a synthetic sample of 2000 realizations of the exogenous variables is generated, corresponding to Q=2000 grids. The spatial pattern of the grids as well as the total land area of each grid is based on an actual grid configuration obtained from the 2010 land use survey data for Austin, Texas (this is the data set also used in the empirical analysis of this paper in Section 5; the use of an actual grid configuration lends more credibility to the simulation exercise rather than generating a synthetic grid configuration). For the weight matrix $(\widetilde{W})$, we use a continuous inverse of distance specification in the simulation analysis based on the line distance (in kilometers) between the centroids of the grids on the coordinate system.[10] Once generated, the independent variable values, the grid configuration and weights, and the grid total area are held fixed in the rest of the simulation exercise.

The coefficient vector $\boldsymbol{\beta}_q$ is allowed to be random according to a bivariate normal distribution for the first two variables, but set to be fixed for the third variable. The mean vector for $\boldsymbol{\beta}_q$ is assumed to be $\boldsymbol{b} = (0.5, -1, 1)$. The covariance matrix $\boldsymbol{\Omega}$ for the two random coefficients is specified as follows:

$$(21) \qquad \boldsymbol{\Omega} = \begin{bmatrix} 0.81 & 0.54 \\ 0.54 & 1.00 \end{bmatrix} = \mathbf{L}_\Omega \mathbf{L}'_\Omega = \begin{bmatrix} 0.90 & 0.00 \\ 0.60 & 0.80 \end{bmatrix} \begin{bmatrix} 0.90 & 0.60 \\ 0.00 & 0.80 \end{bmatrix}.$$

The Cholesky decomposition of $\boldsymbol{\Omega}$ guarantees the positive definiteness of $\boldsymbol{\Omega}$. In the estimations, the likelihood function is reparameterized in terms of the lower Cholesky factor $\mathbf{L}_\Omega$, and the three associated Cholesky parameters $l_{\Omega 1} = 0.9$, $l_{\Omega 2} = 0.6$, and $l_{\Omega 3} = 0.8$ are estimated.

---

[10] More generally, the spatial weight matrix may be based on a continuous representation of distance (such as the inverse distance, or the inverse of the square of distance, or the inverse of exponential distance) or on a discrete representation of distance (such as a simple contiguity indicator of whether or not two grids are adjacent), or on a hybrid representation of distance (such as the shared boundary length between two contiguous grids). These different representations and functional forms for the weight matrix may be tested in any empirical context, as we undertake in our empirical analysis in Section 5.

Collectively, these three parameters, stacked vertically into a column vector, will be referred to as $\mathbf{l}_\Omega$. Next, the covariance matrix $\mathbf{\Lambda}$ for $\varepsilon_q = (\varepsilon_{q1}, \varepsilon_{q2}, \varepsilon_{q3})'$ is specified as follows (see also Section 3):

$$
\mathbf{\Lambda} = \begin{bmatrix} 1.00 & 0.50+0.20 & 0.50+0.40 \\ 0.50+0.20 & 0.50+0.80 & 0.50+0.31 \\ 0.50+0.40 & 0.50+0.31 & 0.50+0.99 \end{bmatrix} = \begin{bmatrix} 1.00 & 0.70 & 0.90 \\ 0.70 & 1.30 & 0.81 \\ 0.90 & 0.81 & 1.49 \end{bmatrix}
$$

(22)
$$
= \mathbf{L}_\Lambda \mathbf{L}_\Lambda' = \begin{bmatrix} 1.00 & 0.00 & 0.00 \\ 0.70 & 0.90 & 0.00 \\ 0.90 & 0.20 & 0.80 \end{bmatrix}\begin{bmatrix} 1.00 & 0.70 & 0.90 \\ 0.00 & 0.90 & 0.20 \\ 0.00 & 0.00 & 0.80 \end{bmatrix}.
$$

In the above matrix, the first element is normalized (and fixed) to the value of 1. There are five Cholesky matrix elements to be estimated in $\mathbf{L}_\Lambda$ ($l_{\Lambda 1} = 0.7$, $l_{\Lambda 2} = 0.9$, $l_{\Lambda 3} = 0.9$, $l_{\Lambda 4} = 0.2$, and $l_{\Lambda 5} = 0.8$). Collectively, these elements, vertically stacked into a column vector, will be referred to as $\mathbf{l}_\Lambda$.

In the simulations, we set the γ_k parameters for the first three alternatives to the value of one. The three γ_k parameters to be estimated are collected in the vector $\mathbf{l}_\gamma$. Finally, to examine the potential impact of different levels of spatial dependence on the ability of the CML approach to recover model parameters, we consider two sets of values of the spatial autoregressive coefficients corresponding to low dependence ($\delta_1 = 0.1$, $\delta_2 = 0.2$, $\delta_3 = 0.3$) and high dependence ($\delta_1 = 0.6$, $\delta_2 = 0.7$, $\delta_3 = 0.8$).

The set-up above is used to develop the $[Q(K-1)\times 1]$ vector $\mathbf{Szb}$ and the covariance matrix $\mathbf{\Sigma}$ (see the section entitled "The Spatial MDCP (or SMDCP) Model"). Since $\overline{\mathbf{\psi}} \sim MVN_{Q(K-1)}(\mathbf{Szb}, \mathbf{\Sigma})$, a specific realization of the $[Q(K-1)\times 1]$ vector for $\overline{\mathbf{\psi}}$ is drawn from the multivariate normal distribution with mean $\mathbf{Szb}$ and covariance structure $\mathbf{\Sigma}$. Then, using subsets of this $\overline{\mathbf{\psi}}$ vector corresponding to each grid, and the specified γ vector, we generate the investment quantity vector x_q^*, using the forecasting algorithm proposed by Pinjari and Bhat (2011). The above data generation process is undertaken 30 times with different realizations of the $\overline{\mathbf{\psi}}$ vector to generate 30 different data sets each for the low spatial dependence case and the high spatial dependence case. The datasets had a good proportion of grids that are not invested in

one or more of the first three alternatives (by construction, all grids are invested in the fourth alternative). On average, across all the datasets, the percentage of grids solely in one of the first three alternatives was 21.5 percent, in only two of the first three alternatives was 30.0 percent, and in all three was 48.5 percent.

The MACML estimator is applied to each data set to estimate data specific values of $b, l_\Omega, l_\Lambda, l_\gamma$, and δ. A single random permutation is generated for each individual (the random permutation varies across individuals, but is the same across iterations for a given individual) to decompose the multivariate normal cumulative distribution (MVNCD) function into a product sequence of marginal and conditional probabilities (see Section 2.1 of Bhat, 2011).[11] All the $Q(Q-1)/2$ pairings of grids are considered in the MACML estimator. The estimator is applied to each dataset 10 times with different permutations to obtain the approximation error, computed as the standard deviation of estimated parameters among the 10 different estimates on the same data set.

Performance Evaluation

The performance of the MACML inference approach in estimating the parameters of the spatial MDCP model and the corresponding standard errors is evaluated as follows:

(1) Estimate the MACML parameters for each data set and for each of 10 independent sets of permutations. Estimate the standard errors (s.e.) using the Godambe (sandwich) estimator.

(2) For each data set *s*, compute the mean estimate for each model parameter across the 10 random permutations used. Label this as MED, and then take the mean of the MED values across the data sets to obtain **a mean estimate**. Compute the **absolute percentage (finite sample) bias** (APB) of the estimator as:

(23) $$APB = \left| \frac{\text{mean estimate - true value}}{\text{true value}} \right| \times 100.$$

(3) Compute the standard deviation of the MED values across datasets, and label this as the **finite sample standard error or FSSE** (essentially, this is the empirical standard error).

---

[11] Technically, the MVNCD approximation should improve with a higher number of permutations in the MACML approach. However, when we investigated the effect of different numbers of random permutations per individual, we noticed little difference in the estimation results between using a single permutation and higher numbers of permutations, and hence we settled with a single permutation per individual.

(4) For each data set, compute the mean standard error for each model parameter across the 10 draws. Call this MSED, and then take the mean of the MSED values across the 30 data sets and label this as **the asymptotic standard error or ASE** (essentially this is the standard error of the distribution of the estimator as the sample size gets large).

(5) Next, to evaluate the accuracy of the asymptotic standard error formula as computed using the MACML inference approach for the finite sample size used, compute the relative efficiency of the estimator as:

(24) $$\text{Relative efficiency} = \frac{\text{ASE}}{\text{FSSE}}.$$

Relative efficiency values in the range of 0.8-1.2 indicate that the ASE, as computed using the Godambe matrix in the CML method, does provide a good approximation of the FSSE. In general, the relative efficiency values should be less than 1, since we expect the asymptotic standard error to be less than the FSSE. But, because we are using only a limited number of data sets to compute the FSSE, values higher than one can also occur. The more important point is to examine the closeness between the ASE and FSSE, as captured by the 0.8-1.2 ranger for the relative efficiency value.

(6) Compute the standard deviation of the parameter values around the MED parameter value for each data set, and take the mean of this standard deviation value across the data sets; label this as the **approximation error (APERR)**.

Additional Restrictive Model Comparisons with the Proposed Model

The main purpose of the methodology proposed here is to accommodate (1) a full covariance in the baseline utilities (within each grid), (2) spatial heterogeneity (across grids) in the responsiveness to exogenous variables, and (3) spatial dynamics that generate dependency effects in the land-use types and intensities (across grids). To examine the implications of ignoring these considerations when they are actually present, we estimate three additional models on the 30 data sets generated for the high spatial dependence level. The first model, which we will refer to as the spatial IID MDCP (or SIMDCP) model, assumes a restrictive independent and identically distributed specification for the error terms in the baseline utilities; that is, for the $\xi_{qk}(k = 1, 2, ..., K)$ terms. This translates to the following restrictive covariance matrix for the ε_q vector:

$$\Lambda = \begin{bmatrix} 1.00 & 0.50 & 0.50 \\ 0.50 & 1.00 & 0.50 \\ 0.50 & 0.50 & 1.00 \end{bmatrix}$$

(25)

$$= \mathbf{L}_\Lambda \mathbf{L}'_\Lambda = \begin{bmatrix} 1.000 & 0.000 & 0.000 \\ 0.500 & 0.866 & 0.000 \\ 0.500 & 0.287 & 0.816 \end{bmatrix} \begin{bmatrix} 1.000 & 0.500 & 0.500 \\ 0.000 & 0.866 & 0.287 \\ 0.000 & 0.000 & 0.816 \end{bmatrix}.$$

The second model, which we will refer to as the spatial homogenous MDCP (or SHMDCP) model assumes away any spatial heterogeneity; that is, it assumes that all the elements of the covariance matrix Ω (and, therefore, the elements of the l_Ω vector) are identically zero. The third model, which we will refer to as the aspatial MDCP (or simply the MDCP) model, restricts all the spatial autocorrelation coefficients to zero; that is, it assumes that all the δ parameters are identically zero. We compare these three restrictive formulations with the proposed SMDCP, based on the APB measure for all non-constrained parameters and the adjusted composite log-likelihood ratio test (ADCLRT) value (see Pace et al., 2011 and Bhat, 2011 for more details on the ADCLRT statistic, which is the equivalent of the log-likelihood ratio test statistic when a composite marginal likelihood inference approach is used; this statistic has an approximate chi-squared asymptotic distribution).

For the comparisons, we use a single replication per data set (the replication is the same one for the SMDCP model and all the restrictive models; that is, we use a single permutation per individual that varies across individuals but is held fixed across the SMDCP and other models). We do so rather than run 10 replications for each of the SMDCP and the more restrictive models because, as we will present in the next section, the approximation error in the parameters is negligible for any given data set. The ADCLRT statistic needs to be computed for each data set separately, and compared with the chi-squared table value with the appropriate degrees of freedom. In this paper, we identify the number of times (corresponding to the 30 model runs, one run for each of the 30 data sets) that the ADCLRT value rejects the SIMDCP, SHMDCP, and MDCP models in favor of the SMDCP model.

Simulation Results

      Recoverability of Parameters in the SMDCP Model. Tables 1a and 1b present the results for the first part of the simulation exercise focusing on parameter recoverability. Table 1a corresponds to the low spatial dependence case, while Table 1b corresponds to the high spatial dependence case. As indicated earlier, there are four alternatives $(K = 4)$, leading to up to a six $[= (K - 1) * 2]$ dimensional integral in the CML function.

      The parameter estimate results in Tables 1a and 1b indicate that the MACML method does very well in recovering the parameters, as can be observed by comparing the mean estimates of the parameters with the true values. The absolute percentage bias (APB) is no more than 5 percent for any parameter (see column titled "Absolute Percentage Bias") in the low dependence case, with an overall mean value of 1.90 percent across all parameters, as indicated at the bottom of the table (see the row labeled "overall mean value across parameters"). The APB values are somewhat higher for the high dependence case (Table 1b), with an overall mean value of 3.4 percent across all parameters. This is not surprising, since the high dependence case generates high interdependence between grids, and leads to a much more non-linear surface of the CML function over which to optimize. Between the low and high dependence cases, there are no substantial differences in recovery ability for the mean value of coefficients on the exogenous variables in the baseline utility function and the covariance matrix of these coefficients (*i.e.*, the **b** vector elements and the $\boldsymbol{l}_\Omega$ vector elements). However, there are differences in the recovery ability for the other parameters, with the APB values being relatively high for the $\boldsymbol{\gamma}$ vector values in the high dependence case. This result is a reflection of somewhat greater difficulty in pinning the satiation parameters in the presence of spatial autoregressive parameters. As indicated earlier, a higher spatial autoregressive parameter causes the logarithm of the baseline utilities to be more spatially interdependent across grids, thus increasing the non-linearity in the baseline utility function (see Equation 11). At the same time, the satiation parameters (captured by the elements of the $\boldsymbol{\gamma}$ vector) generate non-linearity in the overall utility function (see Equation 1). Both these non-linearities come together in the probability expression through their combined presence in the V_{qk} component in Equation (12), which manifests itself in the mean of the distribution over which there is an integration in the CML function of Equation (20). Thus, as one form of non-

linearity (*i.e.*, generated by the spatial autoregressive parameter vector $\boldsymbol{\delta}$) increases, it becomes difficult to estimate the $\boldsymbol{\gamma}$ parameter vector contributing to another form of non-linearity.

The APB values for the parameters of the Cholesky decomposition of the covariance matrix associated with the error term (*i.e.*, the $\boldsymbol{l}_\Lambda$ values) are generally small and lower than 5 percent. The highest APB is for the $l_{\Lambda 4}$ parameter, though this could also be attributed to the low true value of this parameter (which inflates the absolute percentage bias value). Finally, the recovery of spatial autoregressive parameters is exceptionally good in the low dependence case, with all three parameters having an associated APB value of zero (simply a remarkable happenstance). The parameters are also quite well recovered in the high dependence case, though not as well as in the low dependence case, a result of the additional non-linearity generated in the CML function.

The standard error estimates of the parameters indicate good empirical efficiency of the MACML estimator. Across all parameters, the finite sample standard error (FSSE) is a mere 2 percent of the mean parameter estimate in the low spatial dependence case and 4.8 percent of the mean parameter estimate in the high spatial dependence case, indicating very good empirical efficiency of the MACML estimator for the SMDCP model. Another observation from the finite sample standard error estimates is that these estimates (as a percentage of the mean estimates) are generally higher for the $\boldsymbol{\gamma}$ parameters relative to the other parameters, specifically for the high spatial dependence case, reinforcing the finding earlier that the $\boldsymbol{\gamma}$ parameters are more difficult to recover than other parameters.

The finite sample standard errors and the asymptotic standard errors obtained using the Godambe matrix in the CML method are also close, with the relative efficiency value between 0.8-1.2 for most of the parameters in both the low dependence and high dependence cases. The efficiency values are outside the 0.8-1.2 range for some elements of the mean vector of $\boldsymbol{\beta}_q$, some of the Cholesky decomposition elements of the error covariance matrix (*i.e.*, the elements of the $\boldsymbol{l}_\Lambda$ vector) and some of the spatial autoregressive parameters (*i.e.*, the elements of the $\boldsymbol{\delta}$ vector). However, even this is rather deceptive, since the values are simply an artifact of the low values of the finite sample error and asymptotic standard error for these parameters. In particular, the absolute differences in the finite sample error and asymptotic standard error are quite small even for these parameters. For instance, the relative efficiency values for the elements of the $\boldsymbol{\delta}$

vector in the low dependence case are consistently low, but the absolute difference between the finite sample error and asymptotic standard error ranges from 0.001 to 0.003. Overall, the average relative efficiency across all the parameters is 0.92 for the low dependency case and 1.03 for the high dependency case, indicating that the asymptotic formula is performing well in estimating the finite sample standard error. Further, as for the FSSE values, the ASE estimate from the MACML procedure, on average across all parameters, is only 2.24 percent of the mean estimate in the low dependence case and 4.8 percent of the mean estimate in the high dependence case, indicating good efficiency of the MACML procedure even using the ASE estimate for the FSSE.

Finally, the last column of Tables 1a and 1b present the approximation error (APERR) for each of the parameters, because of the use of different permutations. These entries indicate that the APERR is of the order of 0.005 or less, across both the low dependence and high spatial dependence cases. More importantly, the approximation error (as a percentage of the FSSE or the ASE), averaged across all the parameters, is of the order of 7.5 percent of the sampling error for the low dependence case and of the order of 3.5 percent of the sampling error for the high dependence case. This is clear evidence that even a single permutation (per observation) of the MACML estimator provides adequate precision, in the sense that the convergent values are about the same for a given data set regardless of the permutation used for the decomposition of the multivariate probability expression within the MACML approach. This is indeed a remarkable result.

Comparison between the SMDCP Model and more Restrictive Models. In this section, we compare the performance of the SMDCP model formulation with the more restrictive formulations, when the data generated actually conforms to the SMDCP (see the section entitled "Additional Restrictive Model Comparisons with the Proposed Model"). This provides a sense of the biases that may accrue from using a restrictive specification.

The comparison of the SMDCP model with the spatial IID MDCP (or SIMDCP) model tests the restriction that there is no covariance in the unobserved determinants of the baseline utilities of the land-use types within each grid. The column panel entitled "SIMDCP model" in Table 2 presents the mean estimates and the APB values for the elements of the l_{Ω}, γ, and δ vector (the elements in the l_{Λ} vector are all constrained as shown in Equation (25)). As can be

observed, the mean estimates are, in general, not as close to the true parameters as in the SMDCP model. This is particularly noticeable for the γ vector elements, with the APB for the parameter γ_3 being as high as 42 percent. The overall APB across all parameters is 9.64 percent relative to 4.05 percent in the SMDCP model, clearly indicating the biases that occur if one assumes a restrictive independent and identically distributed specification for the error terms in the baseline utilities when the data does not conform to such a specification (note that the overall APB of 4.05 percent for the SMDCP model is computed based on the 30 datasets and the same single set of replications as for the SIMDCP model; also, the overall APB value for the SMDCP model in this comparison is computed excluding the l_Λ vector elements that are constrained in the SIMDCP model). The superiority of the SMDCP model is further reinforced by the ADCLRT test with five degrees of freedom (corresponding to the five elements in the l_Λ vector). The table chi-squared value with five degrees of freedom is 11.07 at the 95 percent confidence level, and the ADCLRT test value between the SMDCP and SIMDCP models exceeds this value for each of the 30 data sets. Thus, the ADCLRT clearly rejects the SIMDCP model in favor of the SMDCP model.

The comparison of the SMDCP model with the spatial homogenous MDCP (or SHMDCP) model tests the restriction that there is no randomness across grids in the response to exogenous covariates; that is, the restriction that all elements of the covariance matrix $\boldsymbol{\Omega}$ (and, therefore, the elements of the l_Ω vector) are identically zero. The column panel entitled "SHMDCP model" in Table 2 presents the mean estimates and the APB values for the elements of the $\boldsymbol{b}$, l_Λ, $\boldsymbol{\gamma}$, and $\boldsymbol{\delta}$ vectors. As should be obvious, the APB values are very large across the board, with an average APB value of over 32 percent relative to 3.52 percent for the SMDCP models for the elements of the $\boldsymbol{b}$, l_Λ, $\boldsymbol{\gamma}$, and $\boldsymbol{\delta}$ vectors. Clearly, there are large biases in the parameters when heterogeneity is ignored, as also reflected in the fact that the ADCLRT test value between the SHMDCP and SMDCP models is higher than the table chi-squared value with three degrees of freedom for each of the 30 datasets at even beyond the 99 percent confidence level (the appropriate table chi-squared value is 11.34). Thus, ignoring spatial heterogeneity can lead to serious model misspecification and inferences.

Finally, the comparison of the SMDCP model with the MDCP model tests the restriction that there are no spatial interdependence effects at play; that is, that all the elements of the $\boldsymbol{\delta}$

31

vector are identically zero. Again, the mean APB of 13.53 percent in the MDCP model is higher than the mean APB of 3.61 percent for the non- δ parameters in the SMDCP model. The APB for the γ vector elements are again the highest, with that for the γ_3 parameter being 76 percent. In particular, there is a substantial underestimation in the γ parameters. The table chi-squared value with three degrees of freedom is 11.34 at the 99 percent confidence level, and the ADCLRT test value between the SMDCP and MDCP models exceeds this value for each of the 30 data sets in our simulation.

Overall, the simulation results show that, irrespective of the magnitude of spatial and temporal dependences, the MACML estimator recovers the parameters of the proposed spatial MDCP very well. The MACML estimator also seems to be quite efficient based on the low FSSE estimates. Further, the asymptotic standard error formula estimates the FSSE quite well, and the approximation error due to the use of the analytic approximation is very small. Additionally, the results clearly highlight the bias in estimates if error covariance, or spatial heterogeneity, or spatial dependence is ignored when both are actually present. An interesting suggestion from our simulation study is that ignoring spatial heterogeneity is of much more serious consequence than ignoring error covariance effects or spatial lag dynamics. Further theoretical and empirical exploration of this finding is left for future work.

5. APPLICATION DEMONSTRATION

In this paper, we demonstrate the application of the proposed SMDCP model by analyzing land-use patterns in Austin, Texas.

The Data and the Context

The data used in this study is drawn from parcel-level land-use inventory data for the year 2010, as obtained from the City of Austin, TX. This data is available in the Environmental Systems Research Institute's (ESRI's) shape file format. The land use type for each parcel is available at a fine level of detail; however, for the current study, the land use types are aggregated into the following four mutually exclusive land use categories: (1) commercial land-use (including commercial, office, hospitals, government services, educational services, cultural services, and parking), (2) industrial land-use (including manufacturing, warehousing, resource extraction (mining), landfills, and miscellaneous industrial uses), (3) residential land-use (including single

family, duplexes, three/four-plexes, apartments, condominiums, mobile homes, group quarters, and retirement housing), and (4) undeveloped land-use (including open and undeveloped spaces, preserves, parks, golf courses, and agricultural open spaces). The last among these alternatives serves as an "essential outside good" in that all grid cells inevitably will have at least some of their land area that remains undeveloped

For the current analysis, an area measuring 377.98 km^2 (145.94 mi^2) covering the central business district and important surrounding areas is considered. As shown in Figure 1a, there are two major highways (Interstate Highway 35 and Loop-1 MoPac) in the study area, running roughly parallel to each other and from the northeast to southwest. In addition, several other major thoroughfares in the Austin area are also represented, including Ben White Blvd (State Highway 71) that forms the southern boundary of the study region, US-290, US-183 that runs diagonally from the northwest to the southeast at the north end of the study area and then directly south at the south end of the study area, Loop 360, and FM-2222. Several major arterials also go through the study area, including Lamar Blvd (roughly parallel to IH-35 and MoPac, and between these two highways), Parmer Lane (toward the north), Cesar Chavez Street (just south of the downtown area), Martin Luther King Jr (MLK) Blvd (just north of the downtown area), Congress Avenue, and Dessau Road. In the rest of this paper, we will use the label "major thoroughfares" to refer to both the major thoroughfares as well as the major arterials identified above. Also, we will consider the Austin Central Business District (CBD) zone as the "square" bounded by Lamar Blvd to the west, MLK Blvd to the north, IH-35 to the east, and Cesar Chavez to the south (see Figure 1a).

The study area is divided into 2383 square grids, each of size 0.25 mi × 0.25 mi (an area of 0.0625 square miles). The use of grids as the unit of analysis is not new, and has been adopted earlier in, amongst many others, Kline (2003) and Kaza et al. (2012). Each grid can be in multiple land-uses, and we obtain the grid-level square footage in each of the four land-use types (that, together, constitute the multiple-continuous dependent variable in our model system) by aggregating the underlying parcel-level land-use inventory data obtained from the City of Austin. If there are no parcels at all in a specific land-use type in a grid, then the square footage in that land-use type in the grid is equal to zero (corner "solution" in the terminology of Section 2). On the other hand, if there are one or more parcels of a specific land-use type in a grid, the square footage in that land-use type in the grid is positive, and has a continuous intensity component.

Thus, the multiple discrete-continuous dependent variable vector for a specific grid q may be $\boldsymbol{x}_q^* = (0,.0000,0.0500,0.0000,0.0125)'$, which indicates that the grid is invested in the second and fourth land-use types, but not the first and third land-use types (note also that the sum of the elements equals 0.0625 square miles). All the explanatory variables for the analysis are created through appropriate aggregation to the grid-level, using geographic information system (GIS) data obtained from the City of Austin (except for the floodplains data, which were obtained from the Capital Area Council of Governments).[12]

The explanatory variables include (1) road access measures (distance to MoPac, distance to IH-35, distance to US-183, and distance to other nearest major thoroughfares), (2) distance to nearest school, (3) distance to the nearest hospital, (4) fraction of grid area that is under a floodplain, (5) an interaction term of proximity to road access with proximity to the floodplain (distance to nearest road divided by distance to the nearest floodplain), (6) average elevation of the grid, and (7) whether the grid is in the Austin CBD zone or not. To construct distances (in miles) from each grid 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 grid centroid to the roadways was calculated. School and hospital data were available as point data, which were overlaid on the analysis area to obtain the distance from a grid centroid to the nearest school and hospital. To calculate the amount of area under a floodplain for each grid, the floodplain polygon shape file was intersected with the grid structure and the intersected area was obtained as the area under a floodplain for the corresponding grid. To construct distances from each parcel to the nearest floodplain, the floodplain 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 grid centroid to the nearest floodplain polygon. To calculate the average elevation of a grid, all the contour lines passing through the grid were recorded, and their average was assigned as the average elevation of the grid.[13] Further, we classified each grid into whether or not it was a "high elevation" grid. In particular, grids with an average elevation greater than 0.10 miles were classified as high elevation grids. We tested both forms of the average elevation explanatory variable (*i.e.*, as a continuous variable and categorical variable) in

---

[12]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 percent chance of a flood of this magnitude during the year).
[13] A contour line is a line joining the points of equal elevation above a given level, such as mean sea level.

the utility specification and retain the specification which provides better data fit. The assignment of a grid as belonging to the CBD zone or not was based on whether any part of the grid was in the CBD zone.

Among the exogenous variables, we expect grids located in close proximity to the major highways and thoroughfares to be more likely to be developed. On the other hand, we expect grids located far from highways and thoroughfares to remain undeveloped, as there is no incentive (less or no net returns) to convert the area into commercial, industrial or residential land use. This should also be obvious from Figures 1b through 1d that map developmental patterns to the location of roadways in the study area. The clustering of commercial development in and around the major highways and thoroughfares is very obvious in Figure 1b. For the industrial land-use in Figure 1c, one can once again notice the clustering of industrial land-use in and around US-183, especially in the vicinity of US-183 and MoPac at the north end of the study area, around US-183 between Dessau Road and US-290 at the east end, and in the neighborhood of US-183 at the southeast end of the study area. The clustering of industrial land-use around US-183 is not surprising, since this thoroughfare is an important diagonal conduit in Austin that passes close to the airport at the southeast end, and has good connectivity to IH-35, while also being away from the downtown area of Austin that is expensive for the large amounts of land needed for industrial purposes. The pattern of residential development is less obvious from the figure, because of the dense development of residences all through the study area.

The proximity to schools is likely to be associated with commercial and residential land-use development (see Li and Liu, 2007), while proximity to hospitals is likely to be an incentive for development of a grid into commercial land use. This latter effect may be attributable to the need for commercial outlets such as eateries and shopping places for hospital employees, and patients and guests. Additionally, we expect grids with a high percentage of area under a floodplain to be less likely to be built up. Further, we consider an interaction effect of distance to the nearest roadway divided by distance to the nearest floodplain. This captures the potential "push-pull" nonlinear positive effect generated by the interaction of distance from roadways and distance to a floodplain. Thus, grids that are distant from roadways may still be developed, - if the grids are also far away from floodplains. Similarly, grids that are close to a floodplain may still be developed, - if the grids are close to roadways. Such "push-pull" effects would be reflected by a negative sign on the "distance to nearest roadway divided by distance to the

nearest floodplain" variable specific to the commercial, industrial, and residential alternatives. Next, grids at higher elevations are more likely (than those at lower elevation) to be in an undeveloped state or in residential land-use than in commercial and industrial land-uses, while we expect grids in the CBD zone to be primarily invested in commercial land-use (this is also reflected in Figures 1b through 1d). The CBD zone in Austin also has several running areas, parks, nature preserves, and picnic trails to promote outdoor activities and active living, and these are reflected in a higher intensity of land-use in the undeveloped state (relative to residential and industrial land-use) in the CBD zone (though this is not discernible from Figure 1 because we have not shown the undeveloped land-use state in the study area). Finally, there is a distinct clustering pattern in the land-use development for each of the three land-use types, suggesting the importance of considering spatial dependency effects.

Table 3a provides descriptive statistics on the dependent variable relating to the investment in each land-use type. Note again that there is both a discrete element as well as a continuous element to the dependent variable. If a grid is not invested in a specific land-use type, the area in that land-use type is zero. If a grid is invested in a specific land-use type, the area in that land-use type is positive and has a continuous component corresponding to the square miles invested in that land-use type. Thus, the dependent variable vector in the model, which corresponds to the areas in each land-use type for each grid, has elements that take a value of zero or, if positive, a continuous value. The second column of Table 3a indicates the number (percentage) of grids invested in each land-use type. As expected, all grids have some land area that is undeveloped, as reflected in the last row of the first numeric column. Among the other land-use types, the highest percentage (82 percent) of grids is invested in residential land-use, while only a small percentage (24 percent) of grids is invested in industrial land-use (equivalently, 18 percent of grids have zero land-use invested in residential land-use and 76 percent of grids have zero land-use invested in industrial land-use). A little more than half of all grids are invested in commercial land-use. The third column of the table provides the intensity of investment among those grids invested in each land-use type. The column indicates that, on average, a grid that is invested in residential land-use has a larger area invested in this land-use than the area investment in commercial or industrial land-uses among grids invested in these other land-use types (note that the sum across rows for this column exceeds 0.0625 square miles because the averages for each land-use type are taken only across those grids that have some

36

positive investment in the land-use type). The last two columns in Table 3a provide information on the fraction of grids that are solely in undeveloped land, and the fraction of grids invested in multiple "inside" land-use types. In particular, the last row of these two columns indicates that only 8 percent of all grids are solely in an undeveloped state, while 92 percent of grids are invested in at least one of commercial, industrial, or residential land-uses in addition to having some land in an undeveloped state. The other rows indicate, for example, that only 8 percent of grids that are invested in the commercial land-use (in addition to having some land in an undeveloped state) are not also invested in any of the other two land-uses of industrial and residential land-uses, while 92 percent of such grids are invested in at least one of industrial or residential land-uses. The results also show, consistent with Figures 1b through 1d, that a grid is more likely to be invested solely in residential land-use than to be invested solely in commercial or industrial land-uses. Overall, these last two columns very clearly indicate that grids are likely to be invested in multiple land-use types at once, strongly supporting the use of a multiple discrete-continuous model for grid-level land-use modeling.

Model Specification

Several weight matrix specifications were considered in our empirical analysis to characterize the nature of the dynamics of the spatial lag dependence. These included (1) a contiguity specification that generates spatial dependence based on whether or not two grids are contiguous, (2) another contiguity specification but based on shared boundary length, (3) the inverse of a continuous distance specification where the distance is measured as the Euclidean distance (crow fly distance) from the centroids of each grid, (4) the inverse of the square of the continuous distance specification, and (5) the inverse of the root of the continuous distance specification. For the last three continuous distance-based specifications, we also explored alternative distance bands to select the pairs of observations for inclusion in the composite marginal likelihood (CML) estimation. As indicated earlier, this distance band determination may be based on minimizing the trace of the variance matrix of parameters given by $tr[\mathbf{V}_{CML}(\hat{\boldsymbol{\theta}})]$. Our results did not show substantial variations in the trace value for different distance bands (regardless of the specific continuous functional form used to represent the distance separation and the variable specification used), though the best estimator efficiency was obtained at about 0.25 miles for all the three continuous distance specifications formulations and all variable specifications we

attempted. Further, the results indicated that for all variable specifications, the best spatial weight matrix specification was consistently the inverse of the continuous distance specification with the 0.25 mile distance band. This determination was based on the composite likelihood information criterion (CLIC) statistic, which may be used to compare the data fit of non-nested formulations (see Varin and Vidoni, 2005). This CLIC statistic takes the form shown below:

$$(26) \qquad \text{CLIC} = \log L_{CML}(\hat{\mathbf{\theta}}) - tr\left[\hat{\mathbf{J}}(\hat{\mathbf{\theta}})\hat{\mathbf{H}}(\hat{\mathbf{\theta}})^{-1}\right],$$

where $\hat{\mathbf{\theta}}$ is the estimated model parameter vector, and $\hat{\mathbf{J}}(\hat{\mathbf{\theta}})$ and $\hat{\mathbf{H}}(\hat{\mathbf{\theta}})$ are the "vegetable" and "bread" matrices used in the estimation of the asymptotic variance matrix $\mathbf{V}_{CML}(\hat{\mathbf{\theta}})$ (see Bhat, 2011 for details of how these matrices may be estimated in a spatial context). In the current context, the weight specification that provides the highest value of the CLIC statistic is preferred over the other competing weight specifications. Of all the weight matrix specifications that were considered here, the best specifications and the corresponding CLIC statistics are presented in Table 3b. The results in the table clearly show the superiority of the inverse of the continuous distance specification over other weight matrix specifications. Thus, all subsequent results in this paper correspond to the inverse distance weight specification with a 0.25 mile distance band.

Concurrent with the weight matrix specification, we also explored several different variable specifications and functional forms of the variables, including linear and non-linear functional forms for continuous variables (such as the logarithm of distance, the square of distance, and spline variables that allow piece-wise linear effects of distance from grid centroid to roadways). In addition, we also considered dummy variables for different ranges of distance for these variables (for instance, grid is within 2 miles of IH-35). Further, various interactions of the many variables were also considered whenever adequate observations were available to test such interaction effects. The final specification was based on intuitive, data fit, and statistical significance considerations. Interestingly, all the distance variables were best reflected in linear continuous form. On the other hand, the average elevation variable was best reflected in categorical form. Table 3c provides descriptive statistics of the independent variables in the final specification. The results of the final specification themselves are discussed in the next section.

Model Estimation and Results

Table 4a presents the results of the proposed SMDCP model. The entries provide the estimate of each "row" exogenous variable on the "column" baseline utility, along with the corresponding t-statistic of the estimate. Note that we have only three land-use alternatives listed in Table 4a because the "undeveloped" land-use alternative constitutes the base alternative. Also, in the case of a random coefficient on a specific variable, we provide both a mean estimate with the corresponding t-statistic as well as an estimate of the standard deviation of the distribution of the parameter with its corresponding t-statistic. In this regard, we attempted a (normally distributed) random coefficients specification for the variables through a general specification of the Ω matrix. However, only the variance parameters corresponding to the constant specific to commercial land-use turned out to be statistically significant. Further, we could not reject the null hypothesis that the off-diagonal (covariance) elements of the Ω matrix corresponding to these random coefficients were all zero.

The first row of variables in Table 4a corresponds to the alternative specific constants for each land-use alternative. These constant terms do not have any substantive interpretations, and simply represent adjustments to the baseline utilities of alternatives after accommodating the other variables in the model. The presence of a statistically significant standard deviation for the commercial land use constant indicates that there is unobserved heterogeneity in grid investments in commercial land-use, attributable perhaps to such unobserved factors as zoning regulations and community perceptions regarding commercial development. In the following sections, we discuss the effects of the non-constant variables on the baseline utilities.

Variable Effects on the Baseline Utility of Alternatives. The results in Table 4a show that grids in the proximity of MoPac are more likely to be invested in commercial and residential land uses and less likely to be invested in industrial land use relative to being in an undeveloped state (technically, but equivalently, the results show that as distance from MoPac increases, grids are more likely to be invested in industrial land use and less so in commercial and residential land uses (relative to being undeveloped). This is not surprising, since MoPac connects the north of Austin directly to the CBD area of Austin. On the other hand, grids that are close to IH-35 have, on average, a higher propensity of being invested in commercial and industrial land use than residential land use. IH-35 is the major interstate highway passing through Austin, and

39

serves as a main north-south conduit for both passenger and truck movement in the region and beyond. The low mean propensity for residential land use in the proximity of IH-35, may be attributable to a desire to live somewhat away from the traffic congestion and noise pollution problems that accompany living in and around IH-35, though the results indicate substantial heterogeneity in this effect. Also, as discussed in the section entitled "The Data and the Context", the majority of industrial areas are located in the proximity of US-183, which is captured in the model through the highly statistically significant negative coefficient on the baseline utility of the industrial land use alternative. As expected, the propensity to invest under commercial and industrial land uses decreases with increase in distance from thoroughfares (though there is substantial heterogeneity in this effect for the industrial land-use alternative). This is a reflection of the importance of ease of access to thoroughfares for commercial and industrial businesses (see Carrión-Flores and Irwin, 2004 and Chakir and Parent, 2009 who also discuss how proximity to major roadways can impact land use decisions). On the other hand, the propensity to invest in residential land use increases with increase in distance to thoroughfares, perhaps because households would rather keep some distance from high traffic activity areas.

Other results of the effects of variables on the baseline utilities are consistent with the hypotheses in the section entitled "The Data and the Context", though there are substantial heterogeneity effects across grids in the push-pull influence of distance to thoroughfares divided by distance to floodplains, and the high elevation dummy variable effect specific to commercial land-use. As expected, the baseline utility for the commercial and undeveloped land-uses is higher than for the industrial and residential land-uses for grids contained within the CBD zone.

Finally, in Table 4a, the satiation parameter estimates indicate that, when there is investment in each of the commercial, industrial, and residential land uses, the residential land use investment intensity in a grid (in terms of square miles) tends to exceed that of the commercial and industrial land use intensities (note that the γ_k parameter corresponding to residential land use is much higher than the corresponding parameters for the commercial and industrial land uses). This result is also consistent with the higher mean acreage (per grid) in residential land-use than the mean acreage in commercial and industrial land uses.

Variance-Covariance Parameters. The estimated variance-covariance structure among the baseline utilities (that is, the Λ matrix) is presented in Table 4b. The variance term (*i.e.*, the

diagonal element) indicates a high variance in the baseline utility of industrial land-use. There is also a significant and high covariance (implied correlation of 0.62) between the baseline utilities for commercial and industrial land use, indicating the presence of common unobserved grid-specific factors that increase (or decrease) the propensity of a grid to be invested in these two land-uses. A similar positive and significant covariance (implied correlation of 0.26) exists in the investments in residential and commercial land-uses, though there is little covariance (implied correlation of 0.08) in the industrial and residential baseline utilities.

Spatial Dependency Parameters. The results indicate the presence of spatial dependence in land use development decisions. Specifically, the estimated spatial autoregressive coefficient δ_k is 0.300 for the commercial land-use alternative (t-statistic of 2.36), 0.623 for the industrial land-use alternative (t-statistic of 2.09), and 0.477 for the residential land-use alternative (t-statistic of 4.95). These estimates strongly support the hypothesis of the presence of spatial dependency effects in the baseline utilities of proximally located spatial grids. That is, there is strong evidence of dyadic dependence between proximally located grids.

Model Selection and Statistical Fit. The statistically significant spatial autoregressive parameters are evidence of the presence of spatial dependency (across grids) in the baseline utility for each alternative. Another way to examine the role of spatial dependence is to compare the data fit of the proposed SMDCP model with an aspatial MDCP model that ignores spatial dependence. This can be undertaken using the adjusted composite likelihood ratio test (ADCLRT). In Particular, the composite log-likelihood value for the SMDCP model is -76250.00 (43 parameters estimated) and for the MDCP model is -76320.00 (40 parameters estimated). The ADCLRT statistic is 34.72, which is much higher than the corresponding critical chi-squared value with three degrees of freedom at any reasonable level of significance. This result clearly indicates the superiority of the SMDCP model relative to the MDCP model.

Aggregate Elasticity Effects

The estimated parameter coefficients in Table 4a provide a sense of the direction of variable effects on the baseline utilities of different land use types. However, these estimated parameters do not directly provide the magnitude of the impact of variables on the acreage investment in

each land use category. To characterize the magnitude and direction of variable effects on acreage in each land use category, 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 across all the grids. We achieve this by computing the share of each land use category in a grid and aggregating these shares across grids for each land use category. Note that the shares incorporate both the discrete and continuous elements, because the share of a specific land-use type in a grid can decrease because that land-use type is not invested in anymore or because of a decrease in the intensity of investment in that land-use type even if it is still invested in. Similarly, the share of a specific land-use type in a grid can increase because that land-use type is invested in (and was not before) and now has a positive investment intensity, or because of an increase in the intensity of investment in that land-use type if that land-use type is already invested in.

Specifically, for the SMDCP model, we use the following procedure. We first draw a realization of the vectors b and δ from their sampling distributions, and construct a realization of the $[Q(K-1)\times 1]$ vector Szb. Next, because we allow random parameters, we draw a realization from the sampling distribution of Ω (which is the covariance matrix of $\tilde{\beta}_q$; see Section 2.3). Subsequently, we also draw a realization from the sampling distribution of Λ (which is the covariance matrix of ε_q). Using these, and the expression in Equation (15), we obtain a realization of the covariance matrix Σ. Then, we draw a realization from the $Q(K-1)\times 1$ vector $\overline{\psi}$ using the realization of Szb and the realization of Σ (note that $\overline{\psi} \sim MVN_{Q(K-1)}(Szb, \Sigma)$). Using the realization of $\overline{\psi}$, we use the Pinjari and Bhat (2011) forecasting algorithm to predict the share of each land use category for each grid (based on $\overline{\psi}$ and the realizations of the satiation parameters from their respective sampling distributions). The expected aggregate share of each land-use alternative is obtained by averaging the corresponding predicted land use share for that alternative across all grids. With the preliminaries above, we compute an aggregate "elasticity" effect for each variable. For dummy variables, the procedure is as follows: (1) set the value of the dummy variable to zero for all the grids in the sample and compute the expected share of each land use category, (2) set the value of the dummy variable to one for all the grids in the sample and compute the expected share of each land use category, and (3) compute the effective percentage change in the expected share of each land use category across all grids in the sample

by taking the difference between the expected share obtained in step (2) and step (1) and dividing by the result from step (1). For continuous variables, we increase the value of the variable by 25 percent for each grid and compute the percentage change in the expected shares.

The procedure in the previous paragraph provides the aggregate elasticity effects of variables for one set of realizations from the sampling distributions of the parameters (and accommodates random coefficients too). In the current paper, we take 10,000 sets of realizations, and compute a mean aggregate elasticity effect for each variable and each land-use alternative as the average of the aggregate elasticity effects across the 10,000 realizations. The standard deviation of these values provides the standard error of the aggregate elasticity effects. For convenience, we will label the mean aggregate elasticity effects simply as the elasticity effects in the presentation below.

A similar, but easier to implement, procedure is used to obtain the elasticity effects and the standard errors of these effects from the MDCP model. The elasticity effects and their standard errors for the MDCP and SMDCP models are presented in Table 4c. The first entry in Table 4c indicates that, in general, a grid located 25 percent farther away from MoPac (than another otherwise observationally identical grid) would be about 4.92 percent less invested in commercial land use than the grid closer to MoPac. Other entries may be similarly interpreted.[14] The last sub-column within each land use alternative column provides the *P* value for the difference in elasticity estimates between the MDCP and SMDCP 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 MDCP and SMDCP models are in the same direction for most variables, and are consistent with the discussions in the previous section. However, the elasticity effects from the SMDCP model are generally higher in magnitude than those from the MDCP model, a consequence of the spillover effects in the SMDCP model that causes a spatial multiplier effect. Specifically, a change in a variable for one grid (say grid A) does not only directly influence the baseline utilities for this grid, but affects the baseline utility of neighboring grids as well due to the positive spatial autocorrelation parameters (this is captured by **S** matrix in Equation 10). The MDCP model ignores such spatial dependencies and assumes that a change in a variable at one grid impacts only the land use at that grid. The differences in the elasticity

---

[14] Very technically speaking, if all grids were located 25 percent farther away from MoPac than their current locations, the aggregate share of commercial land-use in the entire study region would decrease by 4.92 percent. But the more informal interpretation provided in the text is easier to comprehend.

effects between the MDCP and SMDCP models are statistically significant for many variables and land-use type alternatives. In addition, for a couple of variables, the elasticity effects from the MDCP and SMDCP models are even in the opposite directions. For example, the MDCP model predicts an increase (by 5.64 percent) in industrial land use share due to a 25 percent increase in the distance between a grid and the nearest thoroughfare, while the SMDCP model predicts, more in tune with expectations, a decrease (by 7.84 percent) in industrial land use share for the same situation. So, overall, there are differences in the elasticity predictions between the MDCP and SMDCP models, both in terms of magnitude as well as direction of effect. Combined with the improved data fit offered by the SMDCP model, the elasticity effects highlight the importance of considering spatial dependency rather than *a priori* ignoring this dependency.

6. CONCLUSIONS

This paper formulates a spatial multiple discrete-continuous probit (SMDCP) model that should be applicable in a wide variety of fields where social and spatial dependencies lead to spillover effects in multiple discrete-continuous choices (or states). The paper also accommodates spatial heterogeneity in response to exogenous covariates and heteroscedasticity in the dependent variable. The resulting model formulation becomes too cumbersome to be estimated using existing estimation methods, including the frequentist recursive importance sampling (RIS) estimator and the Bayesian Markov Chain Monte Carlo (MCMC) estimator. Instead, we propose the use of a maximum approximate composite marginal likelihood (MACML) estimation method for the proposed SMDCP model. The MACML method is easy to implement, requires no simulation, and involves only univariate/bivariate cumulative normal distribution evaluations.

The paper undertakes a simulation exercise to evaluate the ability of the MACML approach to recover model parameters. The simulation results show that, irrespective of the magnitude of spatial and temporal dependences, the MACML estimator recovers the parameters of the model very well. The MACML estimator also seems to be quite efficient, and the approximation error due to the use of the analytic approximation is very small. Additionally, the simulation study demonstrates that ignoring error covariance across the baseline utility of alternatives within spatial units, or spatial heterogeneity, or spatial dependence, when present but ignored, will introduce substantial bias in model parameters.

The model system proposed in the current paper is applied in a demonstration exercise to examine urban land development intensity levels using grid-level data from Austin, Texas. The empirical results provide important insights regarding land-use investment in multiple types of land-uses simultaneously. The results also indicate the superiority, in terms of data fit, of the SMDCP model relative to its restrictive variants. Future efforts need to continue to undertake simulation experiments to evaluate the performance of the MACML approach for estimating models with spatial dependence, and should also focus on harnessing the potential of the proposed SMDCP model for analyzing multiple discrete-continuous contexts in a wide variety of disciplines. Further, future efforts should consider using a more disaggregate classification system for the land-use categories and should include square-footage prices by land-use type as covariates within the baseline utility function. Doing the latter would enable the computation of valuation measures such as, for example, calculating the marginal willingness to pay for being one mile closer to MoPac, or the implied willingness to pay to reside one mile closer to a school. To obtain such valuations, one would follow the prediction procedure of Bhat and Pinjari (2011) to compute the difference, at the individual grid level, in the expected value of the optimal utility between the after situation (moving the grid closer to MoPac or closer to a school) and the before situation. The difference, when normalized by the marginal utility of income (or the negative of the marginal effect of price on overall utility), provides the marginal willingness to pay measures at the individual grid level. This can be aggregated up across grids to obtain an aggregate willingness to pay measure. If square-footage prices are included, it would also be helpful for compensating valuations related to policies such as zoning and greenbelt preservation (i.e., undeveloped land preservation). Here, one would again predict the expected value of optimal utility after imposing constraints (for example, a grid cannot be developed) and the expected value of optimal utility without any constraints, take the difference between the two, and divide by the marginal utility of income.

REFERENCES

Ahn, Jiwoon, Gicheol Jeong, and Yeonbae Kim. 2008. "A Forecast of Household Ownership and Use of Alternative Fuel Vehicles: A Multiple Discrete-Continuous Choice Approach," *Energy Economics*, 30(5), 2091–2104.

Allenby, Greg M., Mark J. Garratt, and Peter E. Rossi. 2010. "A Model for Trade-up and Change in Considered Brands," *Marketing Science* 29(1), 40–56.

Anselin, Luc. 2010. "Thirty Years of Spatial Econometrics," *Papers in Regional Science*, 89(1), 3–25.

Arbia, Giuseppe and Harry Kelejian. 2010. "Advances in Spatial Econometrics," *Regional Science and Urban Economics*, 40(5), 253–366.

Beck, Nathaniel, Kristian Skrede Gleditsch, and Kyle Beardsley. 2006. "Space is More than Geography: Using Spatial Econometrics in the Study of Political Economy," *International Studies Quarterly*, 50(1), 27–44.

Beron, Kurt J. and Wim P.M. Vijverberg. 2004. "Probit in a Spatial Context: A Monte Carlo Analysis," in Luc Anselin, Raymond Florax, and Sergio J. Rey (eds.), *Advances in Spatial Econometrics: Methodology, Tools and Applications*, Berlin: Springer-Verlag, pp. 169–196.

Bhat, Chandra R. 2005. "A Multiple Discrete-Continuous Extreme Value Model: Formulation and Application to Discretionary Time-use Decisions," *Transportation Research Part B*, 39(8), 679–707.

------. 2008. "The Multiple Discrete-Continuous Extreme Value (MDCEV) Model: Role of Utility Function Parameters, Identification Considerations, and Model Extensions," *Transportation Research Part B*, 42(3), 274–303.

------. 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.

------. 2014. "The Composite Marginal Likelihood (CML) Inference Approach with Applications to Discrete and Mixed Dependent Variable Models," *Foundations and Trends in Econometrics*, 7(1), 1–117.

Bhat, Chandra R. and Jessica Guo. 2004. "A Mixed Spatially Correlated Logit Model: Formulation and Application to Residential Choice Modeling," *Transportation Research Part B*, 38(2), 147–168.

Bhat, Chandra R. and Huimin Zhao. 2002. "The Spatial Analysis of Activity Stop Generation," *Transportation Research Part B*, 36(6), 557–575.

Bhat, Chandra R. and Raghuprasad Sidharthan. 2011. "A Simulation Evaluation of the Maximum Approximate Composite Marginal Likelihood (MACML) Estimator for Mixed Multinomial Probit Models," *Transportation Research Part B*, 45(7), 940–953.

Bhat, Chandra R., Marisol Castro, and Mubassira Khan. 2013a. "A New Estimation Approach for the Multiple Discrete-Continuous Probit (MDCP) Choice Model," *Transportation Research Part B*, 55, 1–22.

Bhat, Chandra R., Marisol Castro, and Abdul Rawoof Pinjari. 2013b. "Allowing for Non-Additively Separable and Flexible Utility Forms in Multiple Discrete-Continuous Models," Technical paper, Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin.

Bhat, Chandra R., Cristiano Varin, and Nazneen Ferdous. 2010. "A Comparison of the Maximum Simulated Likelihood and Composite Marginal Likelihood Estimation Approaches in the Context of the Multivariate Ordered-response Model," in W. Greene and R.C. Hill (eds.), *Advances in Econometrics: Maximum Simulated Likelihood Methods and Applications*, Bingley, UK: Emerald Group Publishing Limited, pp. 65–106.

Brady, Michael and Elena Irwin. 2011. "Accounting for Spatial Effects in Economic Models of Land Use: Recent Developments and Challenges Ahead," *Environmental and Resource Economics*, 48(3), 487–509.

Brown, Daniel G. and Jiunn-Der Duh. 2004. "Spatial Simulation for Translating from Land Use to Land Cover," *International Journal of Geographical Information Science*, 18(1), 35–60.

Brown, Daniel G., Bryan C. Pijanowski, and Jiunn-Der Duh. 2000. "Modeling the Relationships Between Land Use and Land Cover on Private Lands in the Upper Midwest, USA," *Journal of Environmental Management*, 59(4), 247–263.

Capozza, Dennis and Yuming Li. 1994. "The Intensity and Timing of Investment: The Case of Land," *American Economic Review*, 84(4), 889–904.

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

Castro, Marisol, Chandra R. Bhat, Ram M. Pendyala, and Sergio R. Jara-Diaz. 2012. "Accommodating Multiple Constraints in the Multiple Discrete-Continuous Extreme Value (MDCEV) Choice Model," *Transportation Research Part B*, 46(6), 729–743.

Chakir, Raja and Olivier Parent. 2009. "Determinants of Land Use Changes: A Spatial Multinomial Probit Approach," *Papers in Regional Science*, 88(2), 327–44.

Clarke, Keith C., Stacy Hoppen, and Leonard Gaydos. 1997. "A Self-Modifying Cellular Automaton Model of Historical Urbanization in the San Francisco Bay Area," *Environment and Planning B: Planning and Design*, 24(2), 247–261.

Cox, David R. and Nancy Reid. 2004. "A Note on Pseudolikelihood Constructed from Marginal Densities," *Biometrika*, 91(3), 729–737.

Elhorst, J. Paul. 2010. "Spatial Panel Data Models," in Manfred M. Fischer and Arthur Getis (eds.), *Handbook of Applied Spatial Analysis:* Software Tools, Methods and Applications, Berlin: Springer-Verlag, pp. 377–407.

Engelen, Guy and Roger White. 2008. "Validating and Calibrating Integrated Cellular Automata based Models of Land Use Change," *The Dynamics of Complex Urban Systems*, 185–211.

Ferdous, Nazneen and Chandra R. Bhat. 2013. "A Spatial Panel Ordered-Response Model with Application to the Analysis of Urban Land-Use Development Intensity Patterns," *Journal of Geographical Systems*, 15(1), 1–29.

Fleming, Mark M. 2004. "Techniques for Estimating Spatially Dependent Discrete Choice Models," in Luc Anselin, Raymond Florax, and Sergio J. Rey (eds.), *Advances in Spatial Econometrics: Methodology, Tools and Applications*, Berlin: Springer-Verlag, pp. 145–168.

Fotheringham, A. Stewart, and Chris Brunsdon. 1999. "Local Forms of Spatial Analysis," *Geographical Analysis*, 31(4), 340–358.

Franzese Jr, Robert J. and Jude C. Hays. 2008. "Empirical Models of Spatial Interdependence," in Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier (eds), *Oxford Handbook of Political Methodology*, New York: Oxford University Press, pp. 570–604.

Franzese Jr, Robert J., Jude C. Hays, and Lena M. Schaeffer. 2010. "Spatial, Temporal, and Spatiotemporal Autoregressive Probit Models of Binary Outcomes: Estimation, Interpretation, and Presentation," *APSA 2010 Annual Meeting Paper*, August.

Godambe, Vidyadhar P. 1960. "An Optimum Property of Regular Maximum Likelihood Estimation," *Annals of Mathematical Statistics*, 31(4), 1208–1211.

Habib, Khandker M.N. and Eric J. Miller. 2008. "Modelling Daily Activity Program Generation Considering Within-day and Day-to-day Dynamics in Activity-travel Behavior," *Transportation*, 35(4), 467–484.

Hays, Jude C., Aya Kachi, and Robert J. Franzese Jr. 2010. "A Spatial Model Incorporating Dynamic, Endogenous Network Interdependence: A Political Science Application," *Statistical Methodology*, 7(3), 406–428.

Heagerty, Patrick J. and Thomas Lumley. 2000. "Window Subsampling of Estimating Functions with Application to Regression Models," *Journal of the American Statistical Association*, 95(449), 197–211.

Irwin, Elena 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, Elena G. and Jacqueline Geoghegan. 2001. "Theory, Data, Methods: Developing Spatially Explicit Economic Models of Land Use Change," *Agriculture, Ecosystems & Environment*, 85(1-3), 7–24.

Jones, Kelvyn and Nina Bullen. 1994. "Contextual Models of House Prices: A Comparison of Fixed and Random-Coefficient Models Developed by Expansion," *Economic Geography*, 70(3), 252–272.

Kaza, Nikhil, Charles Towe, and Xin Ye. 2012. "A Hybrid Land Conversion Model Incorporating Multiple End Uses," *Agricultural and Resource Economics Review*, 40(3), 341–359.

Kim, Jaehwan, Greg M. Allenby, and Peter E. Rossi. 2002. "Modeling Consumer Demand for Variety," *Marketing Science*, 21(3), 229–250.

Kline, Jeffrey D. 2003. "Characterizing Land Use Change in Multidisciplinary Landscape-Level Analyses," *Agricultural and Resource Economics Review*, 32(1), 103–115.

Kuriyama, Koichi, W. Michael Hanemann, and James R. Hilger. 2010. "A Latent Segmentation Approach to a Kuhn–Tucker Model: An Application to Recreation Demand," *Journal of Environmental Economics and Management*, 60(3), 209–220.

LaMondia, Jeffrey J., Chandra R. Bhat, and David A. Hensher. 2008. "An Annual Time Use Model for Domestic Vacation Travel," *Journal of Choice Modelling*, 1(1), 70–97.

Landis, John and Ming Zhang. 1998a. "The Second Generation of the California Urban Futures Model. Part 1: Model Logic and Theory," *Environment and Planning B: Planning and Design*, 25(5) 657–666.

------. 1998b. "The Second Generation of the California Urban Futures Model. Part 2: Specification and Calibration Results of the Land-Use Change Submodel," *Environment and Planning B: Planning and Design*, 25(6) 795–824.

LeSage, James P. and Robert K. Pace. 2009. *Introduction to Spatial Econometrics*. Boca Raton, FL: Chapman & Hall/CRC, Taylor & Francis Group.

Lewis, Simon L., Paulo M. Brando, Oliver L. Phillips, O.L., Geertje M. F. van der Heijden, and Daniel Nepstad. 2011. "The 2010 Amazon Drought," *Science*, 331(6017), 554.

Li, Xia and Xiaoping Liu. 2007. "Defining Agents' Behaviors to Simulate Complex Residential Development using Multicriteria Evaluation," *Journal of Environmental Management*, 85(4), 1063–1075.

Lindsay, Bruce G. 1988. "Composite Likelihood Methods," *Contemporary Mathematics*, 80, 221–239.

Lindsay, Bruce G., Grace Y. Yi, and Jianping Sun. 2011. "Issues and Strategies in the Selection of Composite Likelihoods," *Statistica Sinica*, 21(1), 71–105.

Mayer, Christopher J. and C. Tsuriel Somerville. 2000. "Land Use Regulation and New Construction," *Regional Science and Urban Economics*, 30(6), 639–662.

McMillen, Daniel P. 2010. "Issues in Spatial Analysis," *Journal of Regional Science*, 50(1), 119–141.

McMillen, Daniel P. and John F. McDonald. 1991. "A Simultaneous Equations Model of Zoning and Land Values," *Regional Science and Urban Economics*, 21(1), 55–72.

Miller, Harvey J. 1999. "Potential Contributions of Spatial Analysis to Geographic Information Systems for Transportation (GIS-T)," *Geographical Analysis*, 31(4), 373–399.

Molenberghs, Geert and Geert Verbeke. 2005. *Models for Discrete Longitudinal Data*. Springer Series in Statistics. New York: Springer Science + Business Media, Inc.

Munroe, Darla K., Cynthia Croissant, and Abigail M. York. 2005. "Land Use Policy and Landscape Fragmentation in an Urbanizing Region: Assessing the Impact of Zoning," *Applied Geography*, 25(2), 121–141.

Pace, Luigi, Alessandra Salvan, and Nicola Sartori. 2011. "Adjusting Composite Likelihood Ratio Statistics," *Statistica Sinica*, 21(1), 129–148.

Parker, Dawn C., Steven M. Manson, Marco A. Janssen, Matthew J. Hoffmann, and Peter Deadman. 2003. "Multi-Agent Systems for the Simulation of Land-Use and Land-Cover Change: A Review," *Annals of the Association of American Geographers*, 93(2), 314–337.

Partridge, Mark D., Marlon Boarnet, Steven Brakman, and Gianmarco Ottaviano. 2012. "Introduction: Whither Spatial Econometrics?," *Journal of Regional Science*, 52(2), 167–171.

Pinjari, Abdul Rawoof. 2011. "Generalized Extreme Value (GEV)-Based Error Structures for Multiple Discrete-Continuous Choice Models," *Transportation Research Part B*, 45(3), 474–489.

Pinjari, Abdul Rawoof and Chandra R. Bhat. 2010. "A Multiple Discrete-Continuous Nested Extreme Value (MDCNEV) Model: Formulation and Application to Non-worker Activity Time-use and Timing Behavior on Weekdays," *Transportation Research Part B*, 44(4), 562–583.

------. 2011. "Computationally Efficient Forecasting Procedures for Kuhn-Tucker Consumer Demand Model System: Application to Residential Energy Consumption Analysis," Technical paper, Department of Civil and Environmental Engineering, The University of South Florida.

Richards, Timothy J., Miguel I. Gomez, and Geoffrey Pofahl. 2012. "A Multiple-Discrete/Continuous Model of Price Promotion," *Journal of Retailing*, 88(2), 206–225.

Robinson, Derek T. and Daniel G. Brown. 2009. "Evaluating the Effects of Land-Use Development Policies on Ex-Urban Forest Cover: An Integrated Agent-Based GIS Approach," *International Journal of Geographical Information Science*, 23(9), 1211–1232.

Satomura, Takuya, Jaehwan Kim, and Greg M. Allenby. 2011. "Multiple-Constraint Choice Models with Corner and Interior Solutions," *Marketing Science*, 30(3), 481–490.

Sener, Ipek N. and Chandra R. Bhat. 2012. "Flexible Spatial Dependence Structures for Unordered Multinomial Choice Models: Formulation and Application to Teenagers' Activity Participation," *Transportation*, 39(3), 657–683.

Sidharthan, Raghuprasad and Chandra R. Bhat. 2012. "Incorporating Spatial Dynamics and Temporal Dependency in Land Use Change Models," *Geographical Analysis*, 44(4), 321–349.

Tobler, Waldo R. 1970. "A Computer Movie Simulating Urban Growth in the Detroit Region," *Economic Geography*, 46(2), 234–240.

Towe, Charles A., Cynthia J. Nickerson, and Nancy Bockstael. 2008. "An Empirical Examination of the Timing of Land Conversions in the Presence of Farmland Preservation Programs," *American Journal of Agricultural Economics*, 90(3), 613–626.

Train, Kenneth, 2009. *Discrete Choice Methods with Simulation*, 2nd ed. Cambridge: Cambridge University Press.

Van Nostrand, Caleb, Vijayaraghavan Sivaraman, and Abdul Rawoof Pinjari. 2013. "Analysis of Long-Distance Vacation Travel Demand in the United States: A Multiple Discrete-Continuous Choice Framework," *Transportation*, 40(1), 151–171.

Varin, Cristiano and Claudia Czado. 2010. "A Mixed Autoregressive Probit Model for Ordinal Longitudinal Data," *Biostatistics*, 11(1), 127–138.

Varin, Cristiano and Paolo Vidoni. 2005. "A Note on Composite Likelihood Inference and Model Selection," *Biometrika*, 92(3), 519–528.

Varin, Cristiano, Nancy Reid, and David Firth. 2011. "An Overview of Composite Likelihood Methods," *Statistica Sinica*, 21(1), 5–42.

von Haefen, Roger H., Daniel J. Phaneuf and George R. Parsons. 2004. "Estimation and Welfare Analysis with Large Demand Systems," *Journal of Business & Economic Statistics*, 22(2), 194–205.

Wu, Fulong and Christopher J. Webster. 1998. "Simulation of Land Development Through the Integration of Cellular Automata and Multicriteria Evaluation," *Environment and Planning B: Planning and Design*, 25(1), 103–126.

Xu, Ximing and Nancy Reid. 2011. "On the Robustness of Maximum Composite Likelihood Estimate," *Journal of Statistical Planning and Inference*, 141(9), 3047–3054.

Yi, Grace Y., Leilei Zeng, and Richard J. Cook. 2011. "A Robust Pairwise Likelihood Method for Incomplete Longitudinal Binary Data Arising in Clusters," *The Canadian Journal of Statistics*, 39(1), 34–51.

Zhao, Yinshan and Harry Joe. 2005. "Composite Likelihood Estimation in Multivariate Data Analysis," *The Canadian Journal of Statistics*, 33(3), 335–356.

LIST OF FIGURES

LIST OF TABLES

FIGURE 1a: Highways, thoroughfares, and CBD location in the analysis area

FIGURE 1b: Commercial land-use distribution
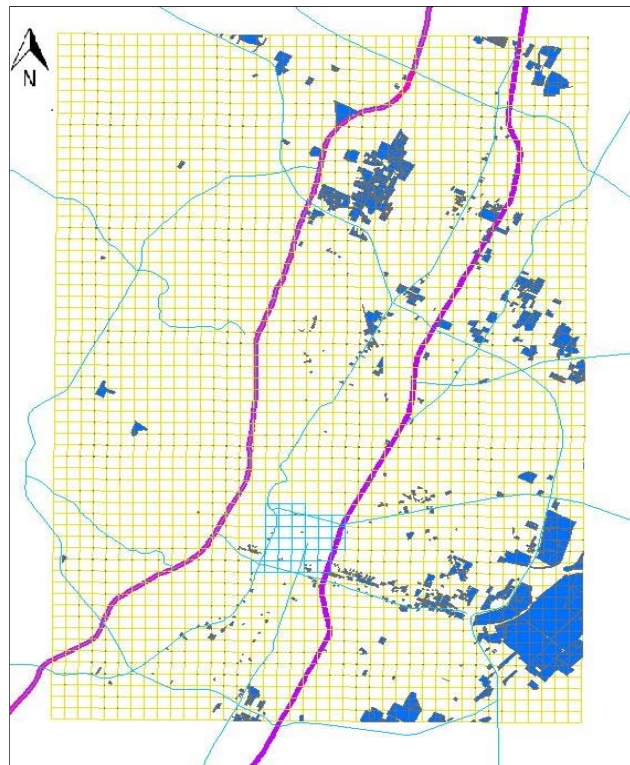
FIGURE 1c: Industrial land-use distribution

FIGURE 1d: Residential land-use distribution

TABLE 1a: Simulation results for the four-alternative case with 30 datasets for low spatial dependency (based on a total of 30×10 runs/dataset=300 runs)

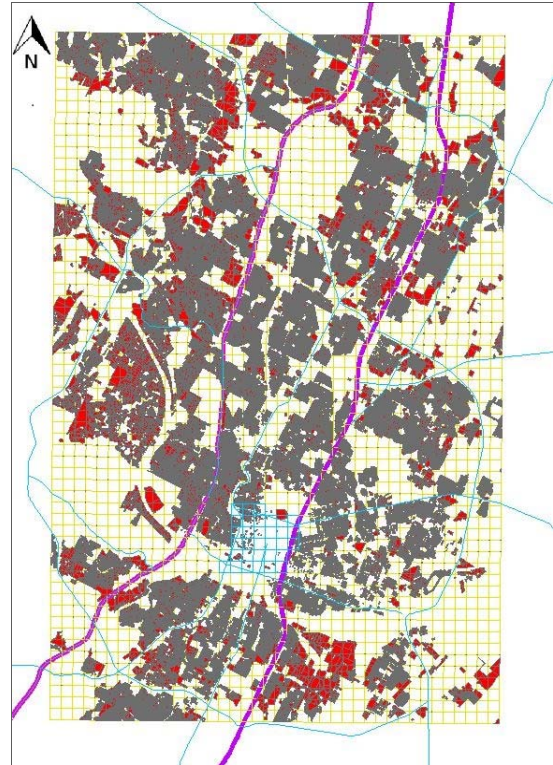| Parameter	True Value	Parameter Estimates			Standard Error Estimates			
		Mean Est.	Abs. Bias	Absolute Percentage Bias (APB)	Finite Sample St. Err. (FSSE)	Asymptotic St. Err. (ASE)	Relative Efficiency	Approximation Error (APERR)
b_1	0.5	0.48	0.02	4.00	0.024	0.030	1.25	0.001722
b_2	-1.0	-1.02	0.02	2.00	0.029	0.028	0.97	0.001781
b_3	1.0	0.99	0.01	1.00	0.023	0.024	1.04	0.001225
$l_{\Omega 1}$	0.9	0.86	0.04	4.44	0.024	0.021	0.88	0.002232
$l_{\Omega 2}$	0.6	0.58	0.02	3.33	0.024	0.029	1.21	0.001310
$l_{\Omega 3}$	0.8	0.78	0.02	2.50	0.028	0.031	1.11	0.001480
γ_1	1.0	0.98	0.02	2.00	0.038	0.038	1.00	0.003031
γ_2	1.0	0.97	0.03	3.00	0.048	0.039	0.82	0.003029
γ_3	1.0	0.96	0.04	4.00	0.049	0.042	0.86	0.003965
$l_{\Lambda 1}$	0.7	0.70	0.00	0.00	0.025	0.019	0.76	0.001797
$l_{\Lambda 2}$	0.9	0.91	0.01	1.11	0.023	0.016	0.70	0.001309
$l_{\Lambda 3}$	0.9	0.90	0.00	0.00	0.021	0.018	0.86	0.002493
$l_{\Lambda 4}$	0.2	0.21	0.01	5.00	0.014	0.016	1.14	0.002852
$l_{\Lambda 5}$	0.8	0.80	0.00	0.00	0.016	0.012	0.75	0.002362
δ_1	0.1	0.10	0.00	0.00	0.005	0.004	0.80	0.000065
δ_2	0.2	0.20	0.00	0.00	0.008	0.006	0.75	0.000175
δ_3	0.3	0.30	0.00	0.00	0.011	0.008	0.73	0.000324
Overall mean value across parameters		0.01	1.90		0.024	0.022	0.92	0.001832

TABLE 1b: Simulation results for the four-alternative case with 30 datasets for high spatial dependency (based on a total of 30×10 runs/dataset=300 runs)

Parameter	True Value	Parameter Estimates			Standard Error Estimates			
		Mean Est.	Abs. Bias	Absolute Percentage Bias (APB)	Finite Sample St. Err. (FSSE)	Asymptotic St. Err. (ASE)	Relative Efficiency	Approximation Error (APERR)
b_1	0.5	0.48	0.02	4.00	0.041	0.052	1.27	0.000943
b_2	-1.0	-1.04	0.04	4.00	0.038	0.047	1.24	0.000792
b_3	1.0	0.98	0.02	2.00	0.022	0.028	1.27	0.000704
$l_{\Omega 1}$	0.9	0.87	0.03	3.33	0.019	0.023	1.21	0.000866
$l_{\Omega 2}$	0.6	0.58	0.02	3.33	0.053	0.047	0.89	0.001881
$l_{\Omega 3}$	0.8	0.80	0.00	0.00	0.041	0.046	1.12	0.001093
γ_1	1.0	0.94	0.06	6.00	0.081	0.082	1.01	0.002657
γ_2	1.0	0.96	0.04	4.00	0.085	0.081	0.95	0.001008
γ_3	1.0	0.89	0.11	11.00	0.070	0.054	0.77	0.000640
$l_{\Lambda 1}$	0.7	0.71	0.01	1.43	0.017	0.017	1.00	0.001736
$l_{\Lambda 2}$	0.9	0.90	0.00	0.00	0.009	0.012	1.33	0.002966
$l_{\Lambda 3}$	0.9	0.89	0.01	1.11	0.020	0.018	0.90	0.002270
$l_{\Lambda 4}$	0.2	0.19	0.01	5.00	0.037	0.029	0.78	0.002260
$l_{\Lambda 5}$	0.8	0.83	0.03	3.75	0.019	0.015	0.79	0.001317
δ_1	0.6	0.60	0.00	0.00	0.048	0.037	0.77	0.000842
δ_2	0.7	0.69	0.01	1.43	0.109	0.105	0.96	0.001897
δ_3	0.8	0.74	0.06	7.50	0.110	0.129	1.17	0.005074
Overall mean value across parameters			0.03	3.40	0.048	0.049	1.03	0.001703

TABLE 2: Effects of ignoring error covariance, spatial heterogeneity and spatial autocorrelation when present (for the high spatial dependence case)

Parameters	True Value	SIMDCP[*]		SHMDCP[+]		MDCP[#]	
		Mean Est.	Absolute Percentage Bias (APB)	Mean Est.	Absolute Percentage Bias (APB)	Mean Est.	Absolute Percentage Bias (APB)
b_1	0.5	0.42	16.00	0.36	28.00	0.48	4.00
b_2	-1.0	-1.07	7.00	-1.02	2.00	-1.01	1.00
b_3	1.0	0.98	2.00	0.88	12.00	1.01	1.00
$l_{\Omega 1}$	0.9	0.89	1.11	—[a]	—	0.89	1.11
$l_{\Omega 2}$	0.6	0.63	5.00	—	—	0.57	5.00
$l_{\Omega 3}$	0.8	0.79	1.25	—	—	0.82	2.50
γ_1	1.0	0.85	15.00	0.73	27.00	0.66	34.00
γ_2	1.0	0.81	19.00	0.67	33.00	0.49	51.00
γ_3	1.0	0.58	42.00	0.26	74.00	0.24	76.00
$l_{\Lambda 1}$	0.7	—	—	0.85	21.43	0.69	1.43
$l_{\Lambda 2}$	0.9	—	—	1.25	38.89	0.91	1.11
$l_{\Lambda 3}$	0.9	—	—	0.99	10.00	0.90	0.00
$l_{\Lambda 4}$	0.2	—	—	0.32	60.00	0.21	5.00
$l_{\Lambda 5}$	0.8	—	—	1.20	50.00	0.85	6.25
δ_1	0.6	0.58	3.33	0.96	60.00	—	—
δ_2	0.7	0.71	1.43	0.80	14.29	—	—
δ_3	0.8	0.78	2.50	0.64	20.00	—	—
Overall mean value across parameters		0.09	9.64	0.24	32.19	0.13	13.53
Mean composite log-likelihood value at convergence		-123728.0236		-127060.8099		-124231.3780	
Number of times the adjusted composite likelihood ratio test (ADCLRT) statistic favors the SMDCP model[b]		All thirty times when compared with $\chi^2_{5,0.95} = 11.07$ value (mean ADCLRT statistic is 26.31)		All thirty times when compared with $\chi^2_{3,0.99} = 11.34$ value (mean ADCLRT statistic is 53.95)		All thirty times when compared with $\chi^2_{3,0.99} = 11.34$ value (mean ADCLRT statistic is 27.47)	

[*]SIMDCP: Spatial IID MDCP.
[+]SHMDCP[*]: Spatial homogeneous MDCP.
[#]MDCP: Aspatial MDCP.
[a] A "—"entry in a cell indicates that the corresponding parameter is not estimated and is fixed to the value mentioned in the section entitled "Additional Restrictive Model Comparisons with the Proposed Model".
[b] The mean composite log-likelihood value for the SMDCP model at converged parameters is -122377.2998.

TABLE 3a: Descriptive statistics of land-use type investments (the dependent variable) in the study area

| Land-use type | Total number (percent) of grids invested in land-use type[a] | Mean land-use area invested (sq mi) | Number of grids (percent of total number) invested…. | |
			only in land-use type and the undeveloped land-use state	in other (inside) land-use types too
Commercial	1304 (55)	0.0136	103 (8)	1201 (92)
Industrial	579 (24)	0.0134	52 (9)	527 (91)
Residential	1953 (82)	0.0267	744 (38)	1209 (62)
Undeveloped	2383 (100)	0.0283	197 (8)	2186 (92)

[a] Percentages across rows in the column do not sum to 100 percent because each grid can be invested in multiple land-use types simultaneously.

TABLE 3b: Model selection based on the weight matrix specification

| | Weight Matrix Specification | | | | |
	Contiguous grid	Shared boundary length	Inverse of continuous distance (0.25 mile distance band)	Inverse of the square of continuous distance (0.25 mile distance band)	Inverse of the root of continuous distance (0.25 mile distance band)
Log-composite likelihood at convergence	-76320.00	-149000.00	-76250.00	-78370.00	-76290.00
Trace Value	628.20	3347.00	530.00	561.40	547.50
CLIC statistics	-76948.20	-152347.00	-76780.00	-78931.40	-76837.50

TABLE 3c: Descriptive statistics of the independent variables used in the model

Variable	Minimum	Maximum	Mean	Standard Deviation
Distance to MoPac (mi)	0.002	7.782	2.52	1.70
Distance to IH-35 (mi)	0.001	8.393	2.64	1.87
Distance to US-183 (mi)	0.004	7.892	2.55	1.88
Distance to nearest thoroughfare (mi)	0.001	2.183	0.50	0.35
Distance to nearest school (mi)	0.005	4.465	1.55	0.96
Distance to nearest hospital (mi)	0.001	2.623	0.65	0.44
Area under floodplain (fraction)	0.000	1.000	0.09	0.16
Distance to nearest thoroughfare / Distance to nearest floodplain	0.000	10.041	1.80	1.52
Average elevation (mi)	0.079	0.187	0.12	0.02
High elevation indicator variable	0.000	1.000	0.82	0.38
CBD indicator variable	0.000	1.000	0.02	0.12

TABLE 4a: Estimation results (t-statistics in parenthesis)

Variables	Spatial Multiple Discrete Continuous Probit (SMDCP) Model		
	Commercial	**Industrial**	**Residential**
Alternative specific constant	-0.488 (-1.15)	1.283 (2.37)	-1.715 (-1.79)
Standard deviation	0.442 (4.49)	—[*]	—
Distance to MoPac (miles)	-0.069 (-4.51)	0.169 (3.03)	-0.063 (-5.47)
Distance to IH-35 (miles)	-0.115 (-3.52)	-0.383 (-5.35)	0.039 (4.15)
Standard deviation	—	—	0.118 (4.42)
Distance to US-183 (miles)	—	-0.323 (-7.95)	—
Distance to nearest thoroughfare (miles)	-0.325 (-2.27)	-1.900 (-3.83)	0.251 (2.888)
Standard deviation	—	2.883 (6.45)	—
Distance to School (miles)	-0.216 (-3.49)	0.536 (3.33)	-0.455 (-10.51)
Distance to Hospital (miles)	-0.255 (-7.11)	0.224 (3.44)	0.027 (1.58)
Fraction of grid area under floodplain	-0.015 (-8.92)	-0.022 (-5.41)	-0.010 (-9.70)
Distance to nearest thoroughfare /Distance to floodplain	-0.358 (-8.88)	-0.372 (-2.98)	0.090 (4.13)
Standard deviation	0.246 (2.15)	0.416 (2.13)	0.165 (6.42)
High elevation indicator variable	-0.265 (-4.51)	-1.429 (-7.74)	0.180 (3.50)
Standard deviation	0.989 (6.57)	—	—
CBD indicator variable	—	-1.079 (-2.55)	-0.776 (-6.84)
Satiation parameter	8.873 (19.01)	3.502 (10.56)	44.939 (14.47)
Spatial lag parameter	0.300 (2.36)	0.623 (2.09)	0.477 (4.95)

[*] A "—"entry in the table indicate that the variable is not statistically significant.

TABLE 4b: Variance-covariance matrix estimates (t-statistics in parenthesis)

Land Use	Commercial	Industrial	Residential
Commercial	1.000 (fixed)	1.445 (4.33)	0.204 (2.30)
Industrial		5.375 (3.22)	0.138 (2.66)
Residential			0.596 (4.94)

TABLE 4c: Aggregate level elasticity effects of the MDCP and SMDCP models (standard-error in parenthesis) based on 1000 runs

Scenario	Commercial			Industrial			Residential			Undeveloped		
	MDCP	SMDCP	P^+	MDCP	SMDCP	P^+	MDCP	SMDCP	P^+	MDCP	SMDCP	P^+
A 25% increase in distance to MoPac	-4.92 (1.82)	-9.87 (1.68)	0.02	6.95 (1.28)	19.86 (5.24)	0.01	-5.91 (2.09)	-8.34 (2.36)	—	-0.11 (0.38)	0.08 (0.32)	—
A 25% increase in distance to IH35	-6.80 (4.17)	-0.15 (5.47)	0.17	-7.26 (1.37)	-15.40 (5.23)	0.07	3.79 (2.76)	13.02 (2.79)	0.01	0.65 (0.58)	0.14 (1.10)	—
A 25% increase in distance to US-183	2.86 (1.15)	6.13 (3.21)	0.17	-5.78 (1.75)	-16.47 (4.11)	0.01	2.40 (0.83)	3.29 (1.81)	—	0.42 (0.32)	0.01 (0.59)	—
A 25% increase in distance to nearest thoroughfare	-2.15 (0.38)	-3.09 (5.44)	—	5.64 (1.19)	-7.84 (6.80)	0.03	4.37 (1.03)	9.05 (1.29)	0.00	1.03 (0.86)	0.07 (0.64)	—
A 25% increase in distance to nearest school	-1.69 (1.04)	-9.57 (1.58)	0.00	5.67 (2.16)	8.61 (2.15)	0.17	-3.64 (1.68)	-15.12 (3.24)	0.00	0.30 (0.23)	0.34 (0.61)	—
A 25% increase in distance to nearest hospital	-9.19 (1.84)	-11.30 (1.80)	0.20	9.68 (3.14)	23.45 (5.44)	0.01	4.64 (3.22)	0.79 (0.89)	0.13	-0.56 (0.47)	-0.05 (0.21)	—
A 25% increase in fraction of grid area under floodplain	-2.73 (1.06)	-3.20 (1.11)	—	-1.79 (0.93)	-2.76 (1.16)	—	-1.13 (0.61)	-5.66 (1.83)	0.01	0.61 (0.07)	0.57 (0.82)	—
A 25% increase in distance to nearest thoroughfare and a 25% decrease in distance to floodplain	-4.31 (1.43)	-6.81 (3.50)	—	-1.56 (0.49)	-6.05 (3.32)	0.09	6.27 (0.99)	12.41 (1.35)	0.00	-0.34 (0.86)	-1.34 (1.11)	—
A switch of the grid location from lower elevation to higher elevation	32.44 (4.46)	46.93 (21.34)	—	-49.26 (2.65)	-43.21 (13.04)	—	41.77 (3.78)	41.80 (22.78)	—	0.36 (0.52)	0.39 (1.73)	—
A switch of the grid location from non CBD zone to CBD zone	13.40 (2.78)	28.16 (22.78)	—	-53.39 (3.55)	-70.84 (18.45)	0.18	-42.91 (3.09)	-47.88 (27.85)	—	1.56 (1.27)	7.06 (2.97)	0.04

*A "—" implies that the difference is not statistically significant even at the 0.2 level of significance.
+P value of the difference.