A New Estimation Approach for the Multiple Discrete-Continuous Probit (MDCP) Choice Model

Chandra R. Bhat*
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712-1172
Tel: 512-471-4535, Fax: 512-475-8744
Email: bhat@mail.utexas.edu
and
King Abdulaziz University, Jeddah 21589, Saudi Arabia

Marisol Castro
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712-1172
Tel: 512-471-4535, Fax: 512-475-8744
Email: m.castro@utexas.edu

Mubassira Khan
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712-1172
Tel: 512-471-4535, Fax: 512-475-874
Email: mubassira@utexas.edu

*corresponding author

Original: July 18, 2012
Revised: April 25, 2013

ABSTRACT

This paper develops a blueprint (complete with matrix notation) to apply Bhat's (2011) Maximum Approximate Composite Marginal Likelihood (MACML) inference approach for the estimation of cross-sectional as well as panel multiple discrete-continuous probit (MDCP) models. A simulation exercise is undertaken to evaluate the ability of the proposed approach to recover parameters from a cross-sectional MDCP model. The results show that the MACML approach does very well in recovering parameters, as well as appears to accurately capture the curvature of the Hessian of the log-likelihood function. The paper also demonstrates the application of the proposed approach through a study of individuals' recreational (*i.e.*, long distance leisure) choice among alternative destination locations and the number of trips to each recreational destination location, using data drawn from the 2004-2005 Michigan statewide household travel survey.

Keywords: Multiple discrete-continuous model, maximum approximate composite marginal likelihood, recreation choice.

1. INTRODUCTION

Consumers often encounter two inter-related decisions at a choice instance -- which alternative(s) to choose for consumption from a set of available alternatives, and the amount to consume of the chosen alternatives. Classical discrete choice models, such as the multinomial logit (MNL) and probit (MNP), allow an analysis of consumer preferences in situations when only one alternative can be chosen for consumption from among a set of available and mutually exclusive alternatives. These models assume that the alternatives are perfect substitutes of one another. However, there are several multiple discrete-continuous (MDC) choice situations where consumers choose to consume multiple alternatives at the same time, along with the continuous dimension of the amount of consumption. Examples of such MDC contexts include, but are not limited to, household vehicle type holdings and usage, airline fleet mix and usage, individuals' choice of recreational destination locations and number of trips to the selected locations, activity type choice and duration spent in different activity types, brand choice and purchase quantity, energy equipment choice and energy consumption, and stock selection and investment amount.

A variety of modeling approaches have been used in the literature to accommodate MDC choice contexts, including (a) the use of a traditional random utility-based (RUM) single discrete choice models by identifying all combinations or bundles of the "elemental" alternatives and treating each bundle as a "composite" alternative, and (b) the use of multivariate probit (logit) methods (see Manchanda *et al*., 1999, Baltas, 2004, Edwards and Allenby, 2003, and Bhat and Srinivasan, 2005). However, the first approach leads to an explosion in the number of composite alternatives as the number of elemental alternatives increases, while the second approach represents more of a statistical stitching of univariate models rather being based on an explicit utility-maximizing framework for multiple discreteness. Besides, it is difficult to incorporate the continuous dimension of consumption quantity in these approaches. Another approach for MDC situations that is rooted firmly in the utility maximization framework assumes a non-linear (but increasing and continuously differentiable) utility structure to accommodate decreasing marginal utility (or satiation) with increasing consumption. Consumers are assumed to maximize this utility subject to a budget constraint. The optimal consumption quantities (including possibly zero consumptions of some alternatives) are obtained by writing the Karush-Kuhn-Tucker (KKT) first-order conditions of the utility function with respect to the consumption quantities. Researchers from many disciplines have used such a KKT approach, and several additively separable and non-linear utility structures have been proposed in the literature (see Hanemann, 1978, Wales and Woodland, 1983, Kim *et al*., 2002, von Haefen and Phaneuf, 2005, Phaneuf and Smith, 2005, Bhat, 2005, 2008, and Kuriyama *et al*., 2011). 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 this 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). Since the baseline preference has to be positive for the overall utility function to be valid, the baseline preference is parameterized as the exponential of a systematic component (capturing the effect of exogenous variables) as well as a stochastic error term. As in traditional discrete choice models, the most common distributions used for the stochastic error term are the multivariate normal (see Kim *et al*., 2002) and generalized extreme value distributions (see Bhat, 2008, Pinjari and Bhat, 2011, Pinjari, 2011). The first distribution leads to an MDC probit (or MDCP) model structure, while the second to a closed-form MDC generalized extreme value (or MDCGEV) model structure (the closed-form MDC extreme value or MDCEV model structure is a special case of the MDCGEV model). In all these cases, the analyst can further superimpose a mixing random distribution structure in the baseline preference to accommodate unobserved taste variations across consumers in the sensitivity to relevant exogenous attributes (such as differential sensitivity due to unobserved factors to travel time and travel cost in a recreation destination choice model). All studies to date in the MDC context that we are aware of have used a normal mixing distribution. The mixing distribution can also be used to accommodate heteroscedasticity and correlations across alternatives (due to generic unobserved preferences) in the MDCEV and MDCGEV model structures.

In the context of a normal mixing error distribution, the use of a GEV kernel structure 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). The domain of integration (to uncondition out the unobserved mixing elements in the consumption probability) in the MMDCGEV structure is the entire multidimensional real space, while the domain of integration in the MDCP structure is a truncated (orthant) space. In both these structures, the multidimensional integration does not have a closed-form solution, and so it is usually undertaken using simulation techniques. The MMDCGEV structure is typically estimated using quasi-Monte Carlo simulations in combination with a quasi-Newton optimization routine in a maximum simulated likelihood (MSL) inference approach (see Bhat, 2001, 2003). The MDCP structure, on the other hand, is typically estimated using the Geweke-Hajivassiliou-Keane (GHK) simulator or the Genz-Bretz (GB) simulator that accommodate the orthant integration domain (see Bhat *et al.*, 2010 for a detailed description of these simulators). Between the MMDCGEV and MDCP structures, the former structure has been the model form of choice in the economics and transportation fields because simulation techniques to evaluate multidimensional integrals are generally easier when the domain is the entire real space rather than orthant spaces. In any case, the consistency, efficiency, and asymptotic normality of these MSL-based simulation estimators is critically predicated on the condition that the number of simulation draws rises faster than the square root of the number of individuals in the estimation sample. Unfortunately, as the number of dimensions of integration increases, the computational cost to ensure good asymptotic estimator properties can be prohibitive and literally infeasible (in the context of the computation resources available, the time available for estimation, and the need for considering a suite of different variable specifications), especially because the accuracy of simulation techniques is known to degrade rapidly at medium-to-high dimensions. The resulting increase in simulation noise can lead to convergence problems during estimation. Also, since the hessian (or second derivatives) needed with the MSL approach to estimate the asymptotic covariance matrix of the estimator is itself estimated on a highly nonlinear and non-smooth second derivatives surface of the log-simulated likelihood function, it can be difficult to accurately compute this covariance matrix (see Craig, 2008 and Bhat *et al.*, 2010). This has implications for statistical inference even if the asymptotic properties of the estimator are well established.[1]

In this paper, we propose the use of Bhat's (2011) Maximum Approximate Composite Marginal Likelihood or MACML inference approach for the estimation of multiple discrete-continuous models. This inference approach is simple, computationally very efficient, and

[1] Bayesian simulation using Markov Chain Monte Carlo (MCMC) techniques (instead of MSL techniques) may also be used for the estimation of MDCGEV and MDCP model structures (for example, see Kim *et al.*, 2002, Fang, 2008, and Brownstone and Fang, 2010). However, these Bayesian techniques also require extensive simulation, are time-consuming, are not straightforward to implement, and create convergence assessment problems as the number of dimensions of integration increases.

simulation-free. While Bhat's original MACML inference proposal was developed for the estimation of multinomial probit models in a traditional discrete choice setting, we show how it also can be gainfully employed for the estimation of MDC models. The proposed MACML approach for MDC models is simple to code and apply using readily available software for likelihood estimation. It also represents a conceptually and pedagogically simpler inference procedure relative to simulation techniques, and involves only univariate and bivariate cumulative normal distribution function evaluations in the likelihood function (in addition to the evaluation of a closed-form multivariate normal density function), regardless of the number of alternatives or the number of choice occasions per individual in a panel setting, or the nature of social/spatial dependence structures imposed. In the MACML inference approach, the MDCP model structure is much easier to estimate because of the conjugate addition property of the multivariate normal distribution, while the MACML estimation of the MMDCGEV structure models requires a normal scale mixture representation for the extreme value error terms, and adds an additional layer of computational effort. Given that the use of a GEV kernel or a multivariate normal (MVN) kernel is simply a matter of convenience, and that the MVN kernel allows a more general covariance structure for the kernel error terms, we will henceforth focus in this paper on the MDCP model structure.

The paper is structured as follows. The next section presents the MACML inference approach for the cross-sectional MDCP model structure, while Section 3 illustrates the approach for the panel MNCP model structures. Section 4 presents details of a simulation effort to examine the ability of the MACML estimator to recover parameters from finite samples in a cross-sectional setting. Section 5 demonstrates an application to study households' leisure travel choice among recreational destination locations and the number of trips to each recreational destination location using data drawn from the 2004-2005 Michigan statewide household travel survey. The final section offers concluding thoughts and directions for further research.[2]

[2] Due to space considerations, we will not discuss the intricate technical details of the MACML inference approach in this paper. This inference approach involves the combination of two basic concepts – the analytic approximation of the multivariate normal cumulative distribution (or MVNCD) function and the use of a composite marginal likelihood (or CML) inference approach. Readers are referred to Bhat (2011) for technical details.

2. CROSS-SECTIONAL MDCP MODEL

2.1 Model Formulation

In the discussion in this section, we will assume that the number of consumer goods in the choice set is the same across all consumers. The case of different numbers of consumer goods per consumer 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 consumer to the next.

Following Bhat (2008), consider a choice scenario where a consumer q (q = 1, 2, …, Q) maximizes his/her utility subject to a binding budget constraint:

$$\max U_q(\boldsymbol{x}_q) = \sum_{k=1}^{K} \frac{\gamma_{qk}}{\alpha_{qk}} \psi_{qk} \left(\left(\frac{x_{qk}}{\gamma_{qk}} + 1 \right)^{\alpha_{qk}} - 1 \right)$$
$$s.t. \quad \sum_{k=1}^{K} p_{qk} x_{qk} = E_q, \tag{1}$$

where the utility function $U_q(\boldsymbol{x}_q)$ is quasi-concave, increasing and continuously differentiable, $\boldsymbol{x}_q \geq 0$ is the consumption quantity (vector of dimension K×1 with elements x_{qk}), and γ_{qk}, α_{qk}, and ψ_{qk} are parameters associated with good k and consumer q. In the linear budget constraint, E_q is the total expenditure (or income) of consumer q, and p_{qk} is the unit price of good k as experienced by consumer q. The utility function form in Equation (1) assumes that there is no essential outside good, so that corner solutions (*i.e.*, zero consumptions) are allowed for all the goods k. This assumption is being made only to streamline the presentation; relaxing this assumption is straightforward and, in fact, simplifies the analysis (see Bhat, 2008).[3] The

[3] The issue of an essential outside good is related to a complete versus incomplete demand system. In a complete demand system, the demands of all goods (that exhaust the consumption space of consumers) are modeled. However, the consideration of complete demand systems can be impractical when studying consumptions in finely defined commodity/service categories. In such situations, it is common to use an incomplete demand system, either in the form of a two stage budgeting approach or in the form of a Hicksian composite commodity approach. In the two stage budgeting approach, the first stage entails allocation between a limited number of broad groups of consumption items, followed by the incomplete demand system allocation within the broad group of interest to elementary commodities/services within that group (the elementary commodities/services in this broad group of primary interest are referred to as “inside” goods, with consumers selecting at least one of these goods for consumption). The plausibility of such a two stage budgeting approach, in general, requires strong homothetic preferences within each broad group and strong separability of preferences (see Menezes *et al*., 2005). In the Hicksian composite commodity approach, one replaces all the elementary alternatives within each broad group that is not of primary interest by a single composite alternative representing the broad group (one needs to assume in this approach that the prices of elementary goods within each broad group of consumption items vary proportionally).

parameter γ_{qk} in Equation (1) allows corner solutions for good *k*, but also serves the role of a satiation parameter. The role of α_{qk} is to capture satiation effects, with smaller value of α_{qk} implying higher satiation for good *k*. ψ_{qk} represents the stochastic baseline marginal utility; that is, it is the marginal utility at the point of zero consumption (see Bhat, 2008 for a detailed discussion).

The utility function in Equation (1) represents a general and flexible functional form under the assumption of additive separable preferences (see Bhat and Pinjari, 2010 for modifications of the utility function to accommodate non-additiveness). It constitutes a valid utility function if $\gamma_{qk} > 0$, $\alpha_{qk} \leq 1$, and $\psi_{qk} > 0$ for all *q* and *k*. Also, as indicated earlier, γ_{qk} and α_{qk} influence satiation, though in quite different ways: γ_{qk} controls satiation by translating consumption quantity, while α_{qk} controls satiation by exponentiating consumption quantity. Empirically speaking, it is difficult to disentangle the effects of γ_{qk} and α_{qk} separately, which leads to serious empirical identification problems and estimation breakdowns when one attempts to estimate both parameters for each good. Thus, Bhat (2008) suggests estimating both a γ-profile (in which $\alpha_{qk} \to 0$ for all goods and all consumers, and the γ_{qk} terms are estimated) and an α-profile (in which the γ_{qk} terms are normalized to the value of one for all goods and consumers, and the α_{qk} terms are estimated), and choose the profile that provides a better statistical fit. However, in this section, we will retain the general utility form of Equation (1) to keep the presentation general. But, for notational simplicity, we will drop the index "*q*" from the γ_{qk} and α_{qk} terms in the rest of this paper. In practice, if a *γ*-profile is used, the parameter γ_{qk} can be allowed to vary across consumers by parameterizing it as an exponential function of relevant consumer-specific variables (and interactions of consumer-specific and alternative attributes). The exponential function ensures that $\gamma_{qk} > 0 \forall$ *q* and *k*. On the other hand, if an *α*-

---

The analysis proceeds then by considering the composite goods as "outside" goods and modeling consumption in these outside goods as well as in the finely categorized "inside" goods representing the consumption group of main interest to the analyst. It is common in practice in this Hicksian approach to include a single outside good with the inside goods. If this composite outside good is not essential, then the consumption formulation is similar to that of a complete demand system. If this composite outside good is essential, then the formulation needs minor revision to accommodate the essential nature of the outside good (see Bhat, 2008).

profile is used, the parameter α_{qk} can be parameterized as one minus the exponential function of relevant consumer-specific attributes (and interactions of consumer-specific and alternative attributes).

To complete the model structure, stochasticity is added by parameterizing the baseline utility as follows:

$$\psi_{qk} = \exp(\boldsymbol{\beta}_q' \boldsymbol{z}_{qk} + \xi_{qk}), \tag{2}$$

where $\boldsymbol{z}_{qk}$ is a D-dimensional vector of attributes that characterize good k and the consumer q (including a dummy variable for each good except one, to capture intrinsic preferences for each good except one good that forms the base), $\boldsymbol{\beta}_q$ is a consumer-specific vector of coefficients (of dimension $D\times1$), and ξ_{qk} captures the idiosyncratic (unobserved) characteristics that impact the baseline utility of good k and consumer q. We assume that the error terms ξ_{qk} are multivariate normally distributed across goods k for a given consumer q: $\boldsymbol{\xi}_q = (\xi_{q1}, \xi_{q2}, ..., \xi_{qK})' \sim MVN_K(\boldsymbol{0}_K, \boldsymbol{\Lambda})$, where $MVN_K(\boldsymbol{0}_K, \boldsymbol{\Lambda})$ indicates a K-variate normal distribution with a mean vector of zeros denoted by $\boldsymbol{0}_K$ and a covariance matrix $\boldsymbol{\Lambda}$. Further, to allow taste variation due to unobserved individual attributes, we consider $\boldsymbol{\beta}_q$ as a realization from a multivariate normal distribution: $\boldsymbol{\beta}_q \sim MVN_D(\boldsymbol{b}, \boldsymbol{\Omega})$. 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 $\tilde{\boldsymbol{\beta}}_q \sim MVN_D(0_D, \boldsymbol{\Omega})$. Note, however, that the parameters (in the $\boldsymbol{\beta}_q$ vector) on the dummy variables specific to each alternative have to be fixed parameters in the cross-section model, since their randomness is already captured in the covariance matrix $\boldsymbol{\Lambda}$.

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 Equation (2) in Equation (1) is:

$$\mathscr{L}_q = \sum_{k=1}^{K} \frac{\gamma_k}{\alpha_k} \exp(\boldsymbol{b}' \boldsymbol{z}_{qk} + \tilde{\boldsymbol{\beta}}_q' \boldsymbol{z}_{qk} + \xi_{qk}) \left\{ \left(\frac{x_{qk}}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\} - \lambda_q \left[\sum_{k=1}^{K} p_{qk} x_{qk} - E_q \right], \tag{3}$$

where λ_q is the Lagrangian multiplier associated with the expenditure constraint (that is, it can be viewed as the marginal utility of total expenditure or income). The KKT first-order conditions for the optimal consumption allocations (the x^*_{qk} values) are given by:

$$\begin{aligned}
&\exp(\boldsymbol{b}'\boldsymbol{z}_{qk} + \tilde{\boldsymbol{\beta}}'_q\boldsymbol{z}_{qk} + \xi_{qk})\left(\frac{x^*_{qk}}{\gamma_k}+1\right)^{\alpha_k-1} - \lambda_q\, p_{qk} = 0 \text{, if } x^*_{qk} > 0 \text{, } k = 1,2,...,K \\
&\exp(\boldsymbol{b}'\boldsymbol{z}_{qk} + \tilde{\boldsymbol{\beta}}'_q\boldsymbol{z}_{qk} + \xi_{qk})\left(\frac{x^*_{qk}}{\gamma_k}+1\right)^{\alpha_k-1} - \lambda_q\, p_{qk} < 0 \text{, if } x^*_{qk} = 0 \text{, } k = 1,2,...,K.
\end{aligned} \tag{4}$$

The optimal demand satisfies the conditions above plus the budget constraint $\sum_{k=1}^{K} p_{qk}x^*_{qk} = E_q$. The budget constraint implies that only K–1 of the x^*_{qk} values need to be estimated, since the quantity consumed of any one good is automatically determined from the quantities consumed of all the other goods. To accommodate this constraint, let m_q be the consumed good with the lowest value of k for the q^{th} consumer. For instance, if the choice set has seven goods $(K = 7)$ and the consumer q chooses goods 2, 3 and 5, then $m_q = 2$. The order in which the goods are organized does not affect the model formulation or estimation, since the definition of m_q only serves as a reference to compare marginal utilities (note also that the consumer q should choose at least one good given that $E_q > 0$). For the good m_q, the Lagrangian multiplier may then be written as:

$$\lambda_q = \frac{\exp(\boldsymbol{b}'\boldsymbol{z}_{qm_q} + \tilde{\boldsymbol{\beta}}'_q\boldsymbol{z}_{qm_q} + \xi_{qm_q})}{p_{qm_q}}\left(\frac{x^*_{qm_q}}{\gamma_{m_q}}+1\right)^{\alpha_{m_q}-1}. \tag{5}$$

Substituting for λ_q from above into Equation (4) for the other goods k ($k = 1,2,...,K$; $k \neq m_q$), and taking logarithms, we can rewrite the KKT conditions as:

$$\begin{aligned}
&V_{qk} + \tilde{\boldsymbol{\beta}}'_q\boldsymbol{z}_{qk} + \xi_{qk} = V_{qm_q} + \tilde{\boldsymbol{\beta}}'_q\boldsymbol{z}_{qm_q} + \xi_{qm_q} \text{, if } x^*_{qk} > 0 \text{, } k = 1,2,...,K \text{, } k \neq m_q \\
&V_{qk} + \tilde{\boldsymbol{\beta}}'_q\boldsymbol{z}_{qk} + \xi_{qk} < V_{qm_q} + \tilde{\boldsymbol{\beta}}'_q\boldsymbol{z}_{qm_q} + \xi_{qm_q} \text{, if } x^*_{qk} = 0 \text{, } k = 1,2,...,K \text{, } k \neq m_q,
\end{aligned} \tag{6}$$

where $V_{qk} = \boldsymbol{b}'\boldsymbol{z}_{qk} + (\alpha_k - 1)\ln\left(\frac{x^*_{qk}}{\gamma_k}+1\right) - \ln p_{qk}$. Letting $y_{qk} = V_{qk} + \tilde{\boldsymbol{\beta}}'_q\,\boldsymbol{z}_{qk} + \xi_{qk}$, and $y^*_{qkm_q} = y_{qk} - y_{qm_q}$, the KKT conditions in Equation (6) are equivalent to:

$$y^*_{qkm_q} = 0, \text{ if } x^*_{qk} > 0, \; k = 1,2,...,K, \; k \neq m_q$$
$$y^*_{qkm_q} < 0, \text{ if } x^*_{qk} = 0, \; k = 1,2,...,K, \; k \neq m_q. \tag{7}$$

Three important identification issues need to be noted here because the KKT conditions above are based on differences, as reflected in the $y^*_{qkm_q}$ terms. First, a constant cannot be identified in the $\boldsymbol{b}'\boldsymbol{z}_{qk}$ term for one of the K goods. Similarly, consumer-specific variables that do not vary across goods can be introduced for K–1 goods, with the remaining good being the base. Second, only the covariance matrix of the error differences is estimable. Taking the difference with respect to the first good, only the elements of the covariance matrix $\boldsymbol{\Lambda}_1$ of $\varepsilon_{qk1} = \xi_{qk} - \xi_{q1}$, $k \neq 1$ are estimable. However, the KKT conditions take the difference against the first consumed good m_q by consumer q. Thus, in translating the KKT conditions to the consumption probability for consumer q, the covariance matrix $\boldsymbol{\Lambda}_{m_q}$ is desired. Since m_q will vary across consumers q, $\boldsymbol{\Lambda}_{m_q}$ will also vary across consumers. But all the $\boldsymbol{\Lambda}_{m_q}$ matrices must originate in the same covariance matrix $\boldsymbol{\Lambda}$ for the original error term vector ξ_q. To achieve this consistency, $\boldsymbol{\Lambda}$ is constructed from $\boldsymbol{\Lambda}_1$ by adding an additional row on top and an additional column to the left. All elements of this additional row and column are filled with values of zeros. $\boldsymbol{\Lambda}_{m_q}$ may then be obtained appropriately for each consumer q based on the same $\boldsymbol{\Lambda}$ matrix. Third, an additional scale normalization needs to be imposed on $\boldsymbol{\Lambda}$ if there is no price variation across goods for each consumer q (*i.e.*, if $p_{qk} = \tilde{p}_q \; \forall k$ and $\forall q$). For instance, one can normalize the element of $\boldsymbol{\Lambda}$ in the second row and second column to the value of one. But, if there is some price variation across goods for even a subset of consumers, there is no need for this scale normalization and all the $K(K–1)/2$ parameters of the full covariance matrix of $\boldsymbol{\Lambda}_1$ are estimable (see Bhat, 2008 for a discussion of this scale normalization issue).

2.2 Model Estimation

The parameters to estimate include the α_k parameters (for an α-profile), the γ_k parameters (for a γ-profile), the $\boldsymbol{b}$ vector, and the elements of the covariance matrices $\boldsymbol{\Omega}$ and $\boldsymbol{\Lambda}$. In the rest of this section, we will use the following key notation: $f_G(.;\boldsymbol{\mu},\boldsymbol{\Sigma})$ for the multivariate normal density function of dimension G with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, $\boldsymbol{\omega}_{\Sigma}$ for the diagonal matrix of standard deviations of $\boldsymbol{\Sigma}$ (with its r^{th} element being $\omega_{\Sigma,r}$), $\phi_G(.;\boldsymbol{\Sigma}^*)$ for the multivariate standard normal density function of dimension G and correlation matrix $\boldsymbol{\Sigma}^*$, such that $\boldsymbol{\Sigma}^* = \boldsymbol{\omega}_{\Sigma}^{-1}\boldsymbol{\Sigma}\boldsymbol{\omega}_{\Sigma}^{-1}$, $F_G(.;\boldsymbol{\mu},\boldsymbol{\Sigma})$ for the multivariate normal cumulative distribution function of dimension G with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, and $\Phi_G(.;\boldsymbol{\Sigma}^*)$ for the multivariate standard normal cumulative distribution function of dimension G and correlation matrix $\boldsymbol{\Sigma}^*$.

To develop the likelihood function, define $\mathbf{M}_q$ as an identity matrix of size K–1 with an extra column of "–1" values added at the m_q^{th} column. Also, stack y_{qk}, V_{qk} , and ξ_{qk} into K×1 vectors: $\boldsymbol{y}_q = (y_{q1}, y_{q2},..., y_{qK})'$, $\boldsymbol{V}_q = (V_{q1}, V_{q2},..., V_{qK})'$, and $\boldsymbol{\xi}_q = (\xi_{q1}, \xi_{q2},..., \xi_{qK})'$, respectively, and let $\mathbf{z}_q = (\boldsymbol{z}_{q1}, \boldsymbol{z}_{q2},..., \boldsymbol{z}_{qK})'$ be a K×D matrix of variable attributes. Then, we may write, in matrix notation, $\boldsymbol{y}_q = \boldsymbol{V}_q + \mathbf{z}_q\tilde{\boldsymbol{\beta}}_q + \boldsymbol{\xi}_q$ and $\boldsymbol{y}_q^* = \mathbf{M}_q\,\boldsymbol{y}_q \sim MVN_{K-1}(\boldsymbol{H}_q, \boldsymbol{\Psi}_q)$, where $\boldsymbol{H}_q = \mathbf{M}_q\boldsymbol{V}_q$ and $\boldsymbol{\Psi}_q = \mathbf{M}_q(\mathbf{z}_q\boldsymbol{\Omega}\mathbf{z}_q' + \boldsymbol{\Lambda})\mathbf{M}_q'$. Next, partition the vector $\boldsymbol{y}_q^*$ into a sub-vector $\tilde{\boldsymbol{y}}_{q,NC}^*$ of length $L_{q,NC}\times 1$ ($0 \le L_{q,NC} \le K-1$) for the non-consumed goods, and another sub-vector $\tilde{\boldsymbol{y}}_{q,C}^*$ of length $L_{q,C}\times 1$ ($0 \le L_{q,C} \le K-1$) for the consumed goods ($L_{q,NC} + L_{q,C} = K-1$). Let $\tilde{\boldsymbol{y}}_q^* = \left(\left[\tilde{\boldsymbol{y}}_{q,NC}^*\right]', \left[\tilde{\boldsymbol{y}}_{q,C}^*\right]'\right)'$, which may be obtained from $\boldsymbol{y}_q^*$ as $\tilde{\boldsymbol{y}}_q^* = \mathbf{R}_q\,\boldsymbol{y}_q^*$, where $\mathbf{R}_q$ is a re-arrangement matrix of dimension (K–1)×(K–1) with zeros and ones. For example, consider a consumer q who chooses among five goods (K=5), and selects goods 2, 3, and 5 for consumption. Thus, $m_q = 2$, $L_{q,NC} = 2$ (corresponding to the non-consumed goods 1 and 4), and $L_{q,C} = 2$ (corresponding to the consumed goods 3 and 5, with good 2 serving as the base good needed to take utility differentials). Then, the re-arrangement matrix $\mathbf{R}_q$ (for goods 1, 3, 4, and 5) is:

$$\mathbf{R}_q = \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{c} \mathbf{R}_{q,NC} \\ \hline \mathbf{R}_{q,C} \end{array}\right], \tag{8}$$

where the upper sub-matrix $\mathbf{R}_{q,NC}$ corresponds to the non-consumed goods (of dimension $L_{q,NC} \times (K-1)$) and the lower sub-matrix $\mathbf{R}_{q,C}$ corresponds to the consumed goods (of dimension $L_{q,C} \times (K-1)$). Note also that $\tilde{\boldsymbol{y}}^*_{q,NC} = \mathbf{R}_{q,NC}\, \boldsymbol{y}^*_q$ and $\tilde{\boldsymbol{y}}^*_{q,C} = \mathbf{R}_{q,C}\, \boldsymbol{y}^*_q$. $\mathbf{R}_{q,NC}$ has as many rows as the number of non-consumed alternatives and as many columns as the number of alternatives minus one (each column corresponds to an alternative, except the m_q^{th} alternative). Then, for each row, $\mathbf{R}_{q,NC}$ has a value of "1" in one of the columns corresponding to an alternative that is not consumed, and the value of "0" everywhere else. A similar construction is involved in creating the $\mathbf{R}_{q,C}$ matrix.

Consistent with the above re-arrangement, define $\tilde{\boldsymbol{H}}_q = \mathbf{R}_q \boldsymbol{H}_q$, $\tilde{\boldsymbol{H}}_{q,NC} = \mathbf{R}_{q,NC} \boldsymbol{H}_q$, $\tilde{\boldsymbol{H}}_{q,C} = \mathbf{R}_{q,C} \boldsymbol{H}_q$, and $\tilde{\boldsymbol{\Psi}}_q = \mathbf{R}_q \boldsymbol{\Psi}_q \mathbf{R}'_q = \begin{bmatrix} \tilde{\boldsymbol{\Psi}}_{q,NC} & \tilde{\boldsymbol{\Psi}}'_{q,NC,C} \\ \tilde{\boldsymbol{\Psi}}_{q,NC,C} & \tilde{\boldsymbol{\Psi}}_{q,C} \end{bmatrix}$, where $\tilde{\boldsymbol{\Psi}}_{q,NC} = \mathbf{R}_{q,NC} \boldsymbol{\Psi}_q \mathbf{R}'_{q,NC}$, $\tilde{\boldsymbol{\Psi}}_{q,C} = \mathbf{R}_{q,C} \boldsymbol{\Psi}_q \mathbf{R}'_{q,C}$, and $\tilde{\boldsymbol{\Psi}}_{q,NC,C} = \mathbf{R}_{q,NC} \boldsymbol{\Psi}_q \mathbf{R}'_{q,C}$. Then, the likelihood function corresponding to the consumption quantity vector $\boldsymbol{x}^*_q$ for consumer q may be obtained from the KKT conditions in Equation (7) as:

$$L_q = \det(\mathbf{J}_q) \int_{\boldsymbol{h}_{q,NC}=-\infty}^{0} f_{K-1}\left(\boldsymbol{h}_{q,NC}, \boldsymbol{0}_{L_{q,C}} \mid \tilde{\boldsymbol{H}}_q, \tilde{\boldsymbol{\Psi}}_q\right) \boldsymbol{dh}_{q,NC}, \tag{9}$$

where $\det(\mathbf{J}_q)$ is the determinant of the Jacobian of the transformation from $\boldsymbol{y}^*_q$ to the consumption quantities $\boldsymbol{x}^*_q$ (see Bhat, 2008):

$$\det(\mathbf{J}_q) = \left\{ \prod_{k \in \mathcal{C}_q} \frac{1-\alpha_k}{x^*_{qk} + \gamma_k} \right\} \left\{ \sum_{k \in \mathcal{C}_q} \left(\frac{x^*_{qk} + \gamma_k}{1-\alpha_k} \right) \left(\frac{p_{qk}}{p_{qm_q}} \right) \right\}, \tag{10}$$

where $\mathcal{C}_q$ is the set of goods consumed by consumer q (including good m_q).

Using the marginal and conditional distribution properties of the multivariate normal distribution, the above likelihood function can be written as:

$$\begin{aligned} L_q &= \det(\mathbf{J}_q) \times f_{L_{q,C}}(\boldsymbol{0}_{L_{q,C}}; \tilde{\boldsymbol{H}}_{q,C}, \tilde{\boldsymbol{\Psi}}_{q,C}) \times F_{L_{q,NC}}(\boldsymbol{0}_{L_{q,NC}}; \breve{\boldsymbol{H}}_{q,NC}, \breve{\boldsymbol{\Psi}}_{q,NC}) \\ &= \det(\mathbf{J}_q) \times \left(\prod_{g=1}^{L_{q,C}} \omega_{\tilde{\boldsymbol{\Psi}}_{q,C},g} \right)^{-1} \left(\phi_{L_{q,C}}(\boldsymbol{\omega}^{-1}_{\tilde{\boldsymbol{\Psi}}_{q,C}}(-\tilde{\boldsymbol{H}}_{q,C}), \tilde{\boldsymbol{\Psi}}^*_{q,C}) \right) \times \Phi_{L_{q,NC}}(\boldsymbol{\omega}^{-1}_{\breve{\boldsymbol{\Psi}}_{q,NC}}(-\breve{\boldsymbol{H}}_{q,NC}), \breve{\boldsymbol{\Psi}}^*_{q,NC}), \end{aligned} \tag{11}$$

where $\breve{\boldsymbol{H}}_{q,NC} = \tilde{\boldsymbol{H}}_{q,NC} + \tilde{\boldsymbol{\Psi}}_{q,NC,C}(\tilde{\boldsymbol{\Psi}}_{q,C})^{-1}(-\tilde{\boldsymbol{H}}_{q,C})$, $\breve{\boldsymbol{\Psi}}_{q,NC} = \tilde{\boldsymbol{\Psi}}_{q,NC} - \tilde{\boldsymbol{\Psi}}_{q,NC,C}(\tilde{\boldsymbol{\Psi}}_{q,C})^{-1}\tilde{\boldsymbol{\Psi}}'_{q,NC,C}$, $\tilde{\boldsymbol{\Psi}}^{*}_{q,C} = \boldsymbol{\omega}^{-1}_{\tilde{\boldsymbol{\Psi}}_{q,C}}\tilde{\boldsymbol{\Psi}}_{q,C}\boldsymbol{\omega}^{-1}_{\tilde{\boldsymbol{\Psi}}_{q,C}}$, and $\breve{\boldsymbol{\Psi}}^{*}_{q,NC} = \boldsymbol{\omega}^{-1}_{\breve{\boldsymbol{\Psi}}_{q,NC}}\breve{\boldsymbol{\Psi}}_{q,NC}\boldsymbol{\omega}^{-1}_{\breve{\boldsymbol{\Psi}}_{q,NC}}$.

The multivariate normal cumulative distribution (MVNCD) function in Equation (11) is of dimension $L_{q,NC}$, which can have a dimensionality of up to (K–1). As indicated in Section 1, typical simulation-based methods to approximate this MVNCD function can get inaccurate and time-consuming as K increases. An alternative is to use the maximum approximate composite marginal likelihood (MACML) approach (Bhat, 2011), in which the multiple integrals are evaluated using a fast analytic approximation method. The MACML estimator is based solely on univariate and bivariate cumulative normal distribution evaluations, regardless of the dimensionality of integration, which considerably reduces computation time compared to other simulation techniques to evaluate multidimensional integrals (see Bhat and Sidharthan, 2011 for an extended simulation analysis of the ability of the MACML method to recover parameters). As we mentioned before, the MACML approach was proposed to estimate mixed multinomial probit models (MNP), but can be extended to other modeling frameworks that result in MVNCD function evaluations, such as the proposed MDCP modeling framework. A brief description of the MACML approach is discussed in the Appendix and the code for the MACML estimation of the MDCP model is available at http://www.caee.utexas.edu/prof/bhat/FULL_CODES.htm.

There is one very important issue that still needs to be dealt with. This concerns the positive definiteness of covariance matrices. The positive-definiteness of $\tilde{\boldsymbol{\Psi}}_q$ in the likelihood function can be ensured by using a Cholesky-decomposition of the matrices $\boldsymbol{\Omega}$ and $\boldsymbol{\Lambda}$, and estimating these Cholesky-decomposed parameters. Note that, to obtain the Cholesky factor for $\boldsymbol{\Lambda}$, we first obtain the Cholesky factor for $\boldsymbol{\Lambda}_1$, and then add a column of zeros as the first column and a row of zeros as the first row to the Cholesky factor $\boldsymbol{\Lambda}_1$.

3. PANEL MDCP MODEL

3.1 Model Formulation

In this section we consider the case of panel data or repeated observations. We will assume that the number of consumer goods and choice occasions are the same across all consumers. Extension to the case of varying number of consumer goods or choice occasions per individual is straightforward. Using the notation of Section 2.1, consider the following utility maximization process with t $(t = 1,2,...,T)$ denoting the choice occasion (or time period):

$$\max U_{qt}(\boldsymbol{x}_{qt}) = \sum_{k=1}^{K} \frac{\gamma_{qtk}}{\alpha_{qtk}} \psi_{qtk} \left(\left(\frac{x_{qtk}}{\gamma_{qtk}} + 1 \right)^{\alpha_{qtk}} - 1 \right)$$

$$s.t. \quad \sum_{k=1}^{K} p_{qtk} x_{qtk} = E_{qt} , \tag{12}$$

In this formulation, the subscript t in the utility parameters implies that the parameters can change over time. However, we will drop the index q and t from the γ_{qtk} and α_{qtk} terms for notational simplicity.

The baseline utility ψ_{qtk} for the q^{th} consumer $(q = 1,2,...,Q)$ at choice occasion t $(t = 1,2,...,T)$ for the k^{th} good ($k = 1,2,...,K$) is parameterized as follows:

$$\psi_{qtk} = \exp(\boldsymbol{\beta}'_q \boldsymbol{z}_{qtk} + \xi_{qtk}) , \tag{13}$$

where $\boldsymbol{z}_{qtk}$ is a D×1 column-vector of exogenous attributes that characterizes good k at choice occasion t for consumer q (including a dummy variable for each good to capture time-invariant intrinsic preference effects of consumer q for good k relative to one of the goods that serves as the base) and $\boldsymbol{\beta}_q$ is the corresponding D×1 column vector of consumer-specific coefficients. $\boldsymbol{\beta}_q$ is assumed to be a realization from a multivariate normal distribution: $\boldsymbol{\beta}_q = \mathrm{b} + \widetilde{\boldsymbol{\beta}}_q$, where $\widetilde{\boldsymbol{\beta}}_q \sim MVN_D(\boldsymbol{0}_D, \boldsymbol{\Omega})$. ξ_{qtk} in Equation (13) is a normal error term uncorrelated with $\boldsymbol{\beta}_q$ and also uncorrelated across consumers. However, the terms ξ_{qtk} may have a covariance structure across goods k (to capture dependencies in the baseline preference of goods due to unobserved factors) and/or across time t (to recognize the time-varying preferences of consumer q). For the latter, we assume a parsimonious first order autoregressive process: $\xi_{qtk} = \delta \xi_{qt-1k} + \eta_{qtk}$, where δ is the autoregressive parameter, $|\delta| < 1$. The η_{qtk} terms are uncorrelated over time [$\mathrm{cov}(\eta_{qtk}, \eta_{qt'k}) = 0$, $\forall q$, $\forall k$, $\forall t, t'$, $t \neq t'$] and contemporaneously correlated across goods. That is,

$\boldsymbol{\eta}_{qt} = (\eta_{qt1}, \eta_{qt2}, ..., \eta_{qtK})' \sim MVN_K(\boldsymbol{0}_K, \boldsymbol{\Lambda})$.[4] The identification considerations for $\boldsymbol{\Lambda}$ are the same as in the cross-sectional case.

Following the procedure of Section 2.1, one obtains the following KKT conditions for consumer q at choice occasion t:

$$\begin{aligned} &y^*_{qtkm_{qt}} = 0, \text{ if } x^*_{qtk} > 0, \; k = 1,...,K, \; k \neq m_{qt} \\ &y^*_{qtkm_{qt}} < 0, \text{ if } x^*_{qtk} = 0, \; k = 1,...,K, \; k \neq m_{qt}, \end{aligned} \tag{14}$$

where m_{qt} is the consumed good with the lowest value of k for the q^{th} consumer at the t^{th} choice occasion, $y^*_{qtkm_{qt}} = y_{qtk} - y_{qtm_{qt}}$ $(k \neq m_{qt})$, $y_{qtk} = V_{qtk} + \tilde{\boldsymbol{\beta}}'_q \boldsymbol{z}_{qtk} + \xi_{qtk}$, and $V_{qtk} = \boldsymbol{b}'\boldsymbol{z}_{qtk} + (\alpha_k - 1) \ln\left(\frac{x^*_{qtk}}{\gamma_k} + 1\right) - \ln p_{qtk}$.

3.2 Model Estimation

The parameters to estimate are the α_k parameters or the γ_k parameters (depending on the profile used for the utility function), the $\boldsymbol{b}$ vector, the δ scalar, and the elements of the covariance matrices $\boldsymbol{\Omega}$ and $\boldsymbol{\Lambda}$. To develop the likelihood function, define $\mathbf{M}_q$ as a $T(K–1)\times TK$ block diagonal matrix, with each block diagonal having K–1 rows and K columns corresponding to the q^{th} consumer in the t^{th} choice instance. This $(K–1)\times K$ matrix for consumer q and time period t corresponds to a K–1 identity matrix with an extra column of "–1" values added to the m^{th}_{qt} column. Also, stack y_{qtk}, V_{qtk} and ξ_{qtk} into the $K\times 1$ vectors $\boldsymbol{y}_{qt} = (y_{qt1}, y_{qt2}, ..., y_{qtK})'$, $\boldsymbol{V}_{qt} = (V_{qt1}, V_{qt2}, ..., V_{qtK})'$ and $\boldsymbol{\xi}_{qt} = (\xi_{qt1}, \xi_{qt2}, ..., \xi_{qtK})'$, respectively. Then define the $TK\times 1$ vectors $\boldsymbol{y}_q = (\boldsymbol{y}'_{q1}, \boldsymbol{y}'_{q2}, ..., \boldsymbol{y}'_{qT})'$, $\boldsymbol{V}_q = (\boldsymbol{V}'_{q1}, \boldsymbol{V}'_{q2}, ..., \boldsymbol{V}'_{qT})'$, and $\boldsymbol{\xi}_q = (\boldsymbol{\xi}'_{q1}, \boldsymbol{\xi}'_{q2}, ..., \boldsymbol{\xi}'_{qT})'$. The variable matrix is written as $\mathbf{z}_q = (\mathbf{z}_{q1}, \mathbf{z}_{q2}, ..., \mathbf{z}_{qT})'$ (matrix of dimension $TK\times D$), where $\mathbf{z}_{qt} = (\boldsymbol{z}_{qt1}, \boldsymbol{z}_{qt2}, ..., \boldsymbol{z}_{qtK})$ (matrix of dimension $D\times K$). Further, let $\mathbf{I}_T$ be the identity matrix of dimension T and let $\boldsymbol{1}_T$ be a column vector of size T with all elements taking the value of one. Now, define the $T\times T$ matrix $\mathbf{A}$ as an identity matrix of size $(T–1)$ with an extra first row and an extra last column of zeros:

[4] Unlike in the cross-sectional case, random coefficients can be estimated in the panel case on the parameters (in the $\boldsymbol{\beta}_q$ vector) on the dummy variables specific to each alternative (except one that serves as the base). This is because we can disentangle consumer-specific intrinsic preferences from choice instance-specific intrinsic preferences based on the repeated choices made by the same consumer.

$$\mathbf{A}=\begin{bmatrix} 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 \end{bmatrix}. \tag{15}$$

Using this matrix and the previous definitions, we can write $\boldsymbol{y}_q$ compactly as:

$$\boldsymbol{y}_q = \boldsymbol{V}_q + \mathbf{z}_q \widetilde{\boldsymbol{\beta}}_q + \mathbf{S}\boldsymbol{\eta}_q, \tag{16}$$

where $\mathbf{S} = \left(\mathbf{I}_{TK} - \delta\,\mathbf{A}\otimes\mathbf{I}_K\right)^{-1}$ is a matrix of dimension $TK \times TK$. Next, for each consumer q, we can write the vector of differences $\boldsymbol{y}_q^*$ as:

$$\boldsymbol{y}_q^* = \mathbf{M}_q\,\boldsymbol{y}_q = \mathbf{M}_q\boldsymbol{V}_q + \left(\mathbf{M}_q\mathbf{z}_q\right)\widetilde{\boldsymbol{\beta}}_q + \mathbf{M}_q\mathbf{S}\boldsymbol{\eta}_q. \tag{17}$$

The above vector of dimension $T(K-1)\times 1$ follows a multivariate normal distribution $\boldsymbol{y}_q^* \sim MVN_{T(K-1)}(\boldsymbol{H}_q, \boldsymbol{\Psi}_q)$, where $\boldsymbol{H}_q = \mathbf{M}_q\boldsymbol{V}_q$ and $\boldsymbol{\Psi}_q$ is the $T(K-1)\times T(K-1)$ covariance matrix defined as $\boldsymbol{\Psi}_q = \mathbf{M}_q\left[\mathbf{z}_q\boldsymbol{\Omega}\,\mathbf{z}_q' + \mathbf{S}\left(\mathbf{I}_T\otimes\boldsymbol{\Lambda}\right)\mathbf{S}'\right]\mathbf{M}_q'$.

As earlier, create a rearrangement matrix $\mathbf{R}_q$ to reorganize the $\boldsymbol{y}_q^*$ vector such that the elements of $\boldsymbol{y}_q^*$ corresponding to the non-consumed goods (across all choice occasions of the consumer) appear first, in order from the first time period to the last. For each consumer q and choice occasion t, let $L_{qt,NC}$ ($0 \le L_{qt,NC} \le K-1$) be the number of non-consumed goods, and let $L_{qt,C}$ ($0 \le L_{qt,C} \le K-1$) be the number of consumed goods, excluding good m_{qt}. Also, let $L_{q,NC} = \sum_{t=1}^{T} L_{qt,NC}$ and $L_{q,C} = \sum_{t=1}^{T} L_{qt,C}$. For example, consider a consumer q with two choice occasions $(T=2)$ and five goods $(K=5)$. In the first choice occasion, the consumer chooses goods 2, 3, and 5 and, in the second choice occasion, the consumer selects goods 1 and 5. Thus, in the first choice occasion $m_{q1}=2$, $L_{q1,NC}=2$ (corresponding to the non-consumed goods 1 and 4), and $L_{q1,C}=2$ (corresponding to the consumed goods 3 and 5). In the second choice occasion, $m_{q2}=1$, $L_{q2,NC}=3$ (non-consumed goods 2, 3, and 4), and $L_{q2,C}=1$ (consumed good 5). Then, $L_{q,NC}=5$ and $L_{q,C}=3$. In this case, the rearrangement matrix $\mathbf{R}_q$ is:

$$\mathbf{R}_q = \left[\begin{array}{cccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{c} \mathbf{R}_{q,NC} \\ \hline \mathbf{R}_{q,C} \end{array}\right], \tag{18}$$

where the upper sub-matrix $\mathbf{R}_{q,NC}$ (of dimension $L_{q,NC} \times T(K-1)$) corresponds to the non-consumed goods and the lower sub-matrix $\mathbf{R}_{q,C}$ (of dimension $L_{q,C} \times T(K-1)$) corresponds to the consumed goods (excluding the good m_{qt} for each time period). Then, the re-arranged vector $\tilde{\boldsymbol{y}}_q^* = \mathbf{R}_q \, \boldsymbol{y}_q^*$, and the corresponding sub-vectors of non-consumed and consumed goods are $\tilde{\boldsymbol{y}}_{q,NC}^* = \mathbf{R}_{q,NC} \, \boldsymbol{y}_q^*$ and $\tilde{\boldsymbol{y}}_{q,C}^* = \mathbf{R}_{q,C} \, \boldsymbol{y}_q^*$, respectively. Consistent with this rearrangement, reorganize $\boldsymbol{H}_q$ and $\boldsymbol{\Psi}_q$ into $\tilde{\boldsymbol{H}}_q = \mathbf{R}_q \, \boldsymbol{H}_q$ and $\tilde{\boldsymbol{\Psi}}_q = \mathbf{R}_q \boldsymbol{\Psi}_q \mathbf{R}'_q$, and generate the corresponding sub-vectors and sub-matrices of non-consumed and consumed goods, as in the cross-sectional case.

The likelihood function contribution of consumer q is:

$$L_q = P(\boldsymbol{x}_q^*) = \det(\mathbf{J}_q) \int_{\boldsymbol{v}_{q,NC}=-\infty}^{0} f_{T(K-1)}\left(\boldsymbol{v}_{q,NC}, \boldsymbol{0}_{L_{q,C}} \mid \tilde{\boldsymbol{H}}_q, \tilde{\boldsymbol{\Psi}}_q\right) \boldsymbol{dv}_{q,NC}, \tag{19}$$

where $\mathbf{J}_q$ is the block diagonal Jacobian matrix (dimension $(L_{q,C}+T) \times (L_{q,C}+T)$) with each block diagonal matrix (of size $(L_{qt,C}+1) \times (L_{qt,C}+1)$) corresponding to a specific choice occasion t of consumer q. The block diagonality arises because $\partial y_{qtkm_{qt}}^* \big/ \partial x_{q't'k'}^* = 0$ for all $t \neq t'$ and for all k and k'. Due to the block diagonal nature of $\mathbf{J}_q$ and using Bhat's (2008) derivation, the determinant of $\mathbf{J}_q$ is:

$$\det(\mathbf{J}_q) = \prod_{t=1}^{T} \left[\left\{ \prod_{k \in \mathcal{C}_{qt}} \frac{1-\alpha_k}{x_{qtk}^* + \gamma_k} \right\} \left\{ \sum_{k \in \mathcal{C}_{qt}} \left(\frac{x_{qtk}^* + \gamma_k}{1-\alpha_k} \right) \left(\frac{p_{qtk}}{p_{qtm_{qt}}} \right) \right\} \right]. \tag{20}$$

where the subset $\mathcal{C}_{qt}$ contains the goods consumed by consumer q at time occasion t, including good m_{qt}.

Then, using the same notation as in the cross-sectional case, the likelihood function for consumer q is equivalent to:

$$L_q = \det(\mathbf{J}_q)\times\left(\prod_{g=1}^{L_{q,C}}\omega_{\tilde{\boldsymbol{\Psi}}_{q,C},g}\right)^{-1}\left(\phi_{L_{q,C}}(\boldsymbol{\omega}^{-1}_{\tilde{\boldsymbol{\Psi}}_{q,C}}(-\tilde{\boldsymbol{H}}_{q,C});\tilde{\boldsymbol{\Psi}}^{*}_{q,C})\right)\times\Phi_{L_{q,NC}}(\boldsymbol{\omega}^{-1}_{\breve{\boldsymbol{\Psi}}_{q,NC}}(-\breve{\boldsymbol{H}}_{q,NC});\breve{\boldsymbol{\Psi}}^{*}_{q,NC}) \quad (21)$$

where $\breve{\boldsymbol{H}}_{q,NC}=\tilde{\boldsymbol{H}}_{q,NC}+\tilde{\boldsymbol{\Psi}}_{q,NC,C}(\tilde{\boldsymbol{\Psi}}_{q,C})^{-1}(-\tilde{\boldsymbol{H}}_{q,C})$, $\quad\breve{\boldsymbol{\Psi}}_{q,NC}=\tilde{\boldsymbol{\Psi}}_{q,NC}-\tilde{\boldsymbol{\Psi}}_{q,NC,C}(\tilde{\boldsymbol{\Psi}}_{q,C})^{-1}\tilde{\boldsymbol{\Psi}}'_{q,NC,C}$, $\tilde{\boldsymbol{\Psi}}^{*}_{q,C}=\boldsymbol{\omega}^{-1}_{\tilde{\boldsymbol{\Psi}}_{q,C}}\tilde{\boldsymbol{\Psi}}_{q,C}\boldsymbol{\omega}^{-1}_{\tilde{\boldsymbol{\Psi}}_{q,C}}$, and $\breve{\boldsymbol{\Psi}}^{*}_{q,NC}=\boldsymbol{\omega}^{-1}_{\breve{\boldsymbol{\Psi}}_{q,NC}}\breve{\boldsymbol{\Psi}}_{q,NC}\boldsymbol{\omega}^{-1}_{\breve{\boldsymbol{\Psi}}_{q,NC}}$.

In the likelihood function of Equation (21), $L_{q,\mathrm{NC}}$ can be large, taking a value as high as $T(K-1)$. In Bhat's MACML approach, one maximizes a surrogate likelihood function, labeled as the composite marginal likelihood (CML) function, to obtain parameters (see Section 2.2 of Bhat, 2011 and the Appendix). Here, we suggest the use of a pairwise likelihood function for choice occasions t and t' given by:

$$\begin{aligned}
L_q^{CML} &= \prod_{t=1}^{T-1}\prod_{t'=t+1}^{T}P(\boldsymbol{x}^*_{qt},\boldsymbol{x}^*_{qt'}) = \prod_{t=1}^{T-1}\prod_{t'=t+1}^{T}\left\{\det(\mathbf{J}_{qt'})\int_{\boldsymbol{v}_{qtt',NC}=-\infty}^{0} f_{2(K-1)}\left(\boldsymbol{v}_{qtt',NC},\boldsymbol{0}_{L_{qtt',C}}\mid\tilde{H}_{qtt'},\tilde{\boldsymbol{\Psi}}_{qtt'}\right)d\boldsymbol{v}_{qtt',NC}\right\} \\
&= \prod_{t=1}^{T-1}\prod_{t'=t+1}^{T}\left\{\det(\mathbf{J}_{qtt'})\times\left(\prod_{g=1}^{L_{qtt',C}}\omega_{\tilde{\boldsymbol{\Psi}}_{qtt',C},g}\right)^{-1}\phi_{L_{qtt',C}}(\boldsymbol{\omega}^{-1}_{\tilde{\boldsymbol{\Psi}}_{qtt',C}}(-\tilde{H}_{qtt',C});\tilde{\boldsymbol{\Psi}}^{*}_{qtt',C})\right. \\
&\qquad\left.\times\Phi_{L_{qtt',NC}}(\boldsymbol{\omega}^{-1}_{\breve{\boldsymbol{\Psi}}_{qtt',NC}}(-\breve{H}_{qtt',NC});\breve{\boldsymbol{\Psi}}^{*}_{qtt',NC})\right\}
\end{aligned} \quad (22)$$

where all the notations are similar to earlier, but confined to the t^{th} and $t'^{\,th}$ choice occasions.

4. SIMULATION EVALUATION

The simulation exercises undertaken in this section examine the ability of the MACML estimator to recover parameters from finite samples in a cross-sectional MDCP model by generating simulated data sets with known underlying model parameters. To examine the robustness of the MACML approach to different dimensionalities of integration, we consider both a five-alternative case as well as a ten-alternative case.

4.1 Experimental Design

In each of the five- and ten-alternative case, we consider five independent variables in the $\boldsymbol{z}_{qk}$ vector in the baseline utility. The values of each of the five independent variables for the alternatives are drawn from a standard univariate normal distribution. In particular, a synthetic sample of 5000 realizations of the exogenous variables is generated corresponding to Q=5,000 consumers. Additionally, we generate budget amounts E_q $(q = 1,2,...,Q)$ from a univariate normal distribution with mean 150, and truncated between the values of 100 and 200 (the prices of all goods are fixed at the value of one across all consumers). Once generated, the independent variable values and the total budget are held fixed in the rest of the simulation exercise.

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

$$\boldsymbol{\Omega} = \begin{bmatrix} 0.81 & 0.54 & 0.72 \\ 0.54 & 1.00 & 0.80 \\ 0.72 & 0.80 & 0.89 \end{bmatrix} = \mathbf{L}_{\boldsymbol{\Omega}}\mathbf{L}'_{\boldsymbol{\Omega}} = \begin{bmatrix} 0.90 & 0.00 & 0.00 \\ 0.60 & 0.80 & 0.00 \\ 0.80 & 0.40 & 0.30 \end{bmatrix} \begin{bmatrix} 0.90 & 0.60 & 0.80 \\ 0.00 & 0.80 & 0.40 \\ 0.00 & 0.00 & 0.30 \end{bmatrix} \tag{23}$$

As indicated earlier, the positive definiteness of $\boldsymbol{\Omega}$ is ensured in the estimations by reparameterizing the likelihood function in terms of the lower Cholesky factor $\mathbf{L}_{\boldsymbol{\Omega}}$, and estimating the six associated Cholesky matrix parameters. For future reference and presentation, we will label these six Cholesky parameters as $l_{\Omega 1} = 0.9$, $l_{\Omega 2} = 0.6$, $l_{\Omega 3} = 0.8$, $l_{\Omega 4} = 0.8$, $l_{\Omega 5} = 0.5$, and $l_{\Omega 6} = 0.3$. We will also refer to these parameters collectively as l_{Ω}.

Next, values for the error terms ξ_{qk} are generated for the case of five alternatives by specifying the following 4×4 positive definite covariance matrix $\boldsymbol{\Lambda}_1^5$ for the differenced error terms ε_{qk1} (the superscript on $\boldsymbol{\Lambda}_1$ stands for the 5 alternatives case):

$$\begin{aligned} \boldsymbol{\Lambda}_1^5 &= \begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 1.21 & 0.00 & 0.00 \\ 0.00 & 0.00 & 1.00 & 0.60 \\ 0.00 & 0.00 & 0.60 & 1.00 \end{bmatrix} \\ &= \mathbf{L}_{\boldsymbol{\Lambda}_1^5}\mathbf{L}'_{\boldsymbol{\Lambda}_1^5} = \begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 1.10 & 0.00 & 0.00 \\ 0.00 & 0.00 & 1.00 & 0.00 \\ 0.00 & 0.00 & 0.60 & 0.80 \end{bmatrix} \begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 1.10 & 0.00 & 0.00 \\ 0.00 & 0.00 & 1.00 & 0.60 \\ 0.00 & 0.00 & 0.00 & 0.80 \end{bmatrix}. \end{aligned} \tag{24}$$

In the above matrix, the first element is normalized to the value of 1 because we do not allow price variation in the simulation experiments. There are four Cholesky matrix elements to be estimated in the matrix above ($l_{\mathbf{\Lambda}1} = 1.1$, $l_{\mathbf{\Lambda}2} = 1.0$, $l_{\mathbf{\Lambda}3} = 0.6$, and $l_{\mathbf{\Lambda}4} = 0.8$). The corresponding implied covariance matrix $\mathbf{\Lambda}^5$ for the original error terms ξ_{qk} is then as follows:

$$\mathbf{\Lambda}^5 = \begin{bmatrix} 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 1.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 1.21 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 & 0.60 \\ 0.00 & 0.00 & 0.00 & 0.60 & 1.00 \end{bmatrix}$$
$$= \mathbf{L}_{\mathbf{\Lambda}^5}\mathbf{L}'_{\mathbf{\Lambda}^5} = \begin{bmatrix} 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 1.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 1.10 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.60 & 0.80 \end{bmatrix} \begin{bmatrix} 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 1.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 1.10 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 & 0.60 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.80 \end{bmatrix} \quad (25)$$
[5]

The error terms ξ_{qk} are generated for the case of ten alternatives, similar to the case with five alternatives, by specifying the following 9×9 positive definite covariance matrix $\mathbf{\Lambda}_1^{10}$ for the differenced error terms ε_{qk1} :

$$\mathbf{\Lambda}_1^{10} = \begin{bmatrix} \mathbf{\Lambda}_1^5 & 0 & 0 \\ 0 & \mathbf{\Lambda}_1^5 & 0 \\ 0 & 0 & 1.21 \end{bmatrix} = \begin{bmatrix} \mathbf{L}_{\mathbf{\Lambda}_1^5} & 0 & 0 \\ 0 & \mathbf{L}_{\mathbf{\Lambda}_1^5} & 0 \\ 0 & 0 & 1.1 \end{bmatrix} \begin{bmatrix} \mathbf{L}_{\mathbf{\Lambda}_1^5} & 0 & 0 \\ 0 & \mathbf{L}_{\mathbf{\Lambda}_1^5} & 0 \\ 0 & 0 & 1.1 \end{bmatrix} \quad (26)$$

The parameters corresponding to $\mathbf{\Lambda}_1^{10}$ to be estimated in the 10-alternative case are the same four Cholesky parameters as in the five-alternative case plus two additional parameters: $l_{\mathbf{\Lambda}5} = 1.0$ and $l_{\mathbf{\Lambda}6} = 1.1$. $l_{\mathbf{\Lambda}5}$ corresponds to the square root of the variance of $\varepsilon_{q61} = \xi_{q6} - \xi_{q1}$, which need not be fixed to 1 as we did for $\varepsilon_{q21} = \xi_{q2} - \xi_{q1}$, and $l_{\mathbf{\Lambda}6}$ is the square root of the variance of

[5] Note that while we can specify a full covariance matrix for $\mathbf{\Lambda}_1$ (except for the first element that has to be normalized due to no price variation), we impose a more restrictive structure to keep the parameters to be estimated in our simulation experiments to a reasonable number. This will also generally need to be done in real-world applications (especially as the number of alternatives increases) through behavioral structures on $\mathbf{\Lambda}$ that seem appropriate to the application context. This is needed not simply to contain the number of parameters to be estimated, but also to interpret the estimated covariance matrix parameters (see Train, 2009; page 113 for a similar discussion in the case of traditional multinomial probit models).

$\varepsilon_{q10,1} = \xi_{q10} - \xi_{q1}$. The Cholesky factors corresponding to the error terms ξ_{qk} will be referred to collectively as l_{Λ} in the rest of this paper.

The baseline utility is next computed for each consumer and alternative using Equation (2). In the simulations, we use a γ-profile, and set all the γ parameters to the value of one. Then, for each of the five-alternative and ten-alternative cases, we generate the consumption quantity vector $\boldsymbol{x}_q^*$ for each individual, using the forecasting algorithm proposed by Pinjari and Bhat (2011). The above data generation process is undertaken 20 times with different realizations of the $\boldsymbol{\beta_q}$ vector and the error term ξ_{qk} to generate 20 different data sets each for the five-alternative and the ten-alternative case.

The MACML estimator is applied to each data set to estimate data specific values of $\boldsymbol{b}$, l_{Ω}, l_{Λ}, and γ. 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).[6] The MACML estimator is applied to each dataset 10 times with different permutations to acknowledge that different permutations will lead to different parameter estimates and standard error estimates of parameters.

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

(1) Estimate the MACML parameters for each data set *s* and for each of 10 independent sets of permutations for computing the approximation for the likelihood function contribution of each individual. 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:

[6] 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.

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

(3) Compute the standard deviation for each model parameter across the data sets and across the 10 random permutations for each data set, and label this as the **finite sample standard error or FSEE** (essentially, this is the empirical standard error).

(4) For each data set *s*, compute the median s.e. for each model parameter across the 10 draws. Call this MSED, and then take the mean of the MSED values across the 20 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 APB associated with the ASE of the estimator as:

$$APBASE = \left| \frac{\text{ASE - FSEE}}{\text{FSEE}} \right| \times 100$$

4.2 Simulation Results

Tables 1a and 1b provide the results for the five-alternative case (leading to up to four dimensional integration) and for the ten-alternative case (leading to up to nine dimensional integration), respectively. The tables provide the true value of the parameters (second column), followed by the parameter estimate results and the sampling standard error estimate results.

4.2.1 Five-Alternative Case

The results in Table 1a indicate that the MACML method does extremely well in recovering the parameters, as can be observed by comparing the mean estimate of the parameters with the true values (see the column titled ‘‘parameter estimates”). In fact, the absolute percentage bias (APB) is not higher than 4% for any parameter, with an overall mean value of 0.96% across all parameters, as indicated in the bottom of the table (see the row labeled ‘‘overall mean value across parameters’’ and the column titled ‘‘absolute percentage bias”).[7] The APB values are generally somewhat smaller for the parameters of the Cholesky decomposition of the covariance

[7] The APB values may not match up exactly to the true and estimated values of the parameters presented in the table. This is because of rounding in the estimated values. The same is the case later when computing the APBASE values from the finite sample standard error and asymptotic standard error values.

matrix associated with the error terms (*i.e.*, the l_{Λ} values) than for the other parameters. Also, there is more variation in the APB values among the parameters of the Cholesky decomposition of the covariance matrix associated with the random coefficients (*i.e.*, the l_{Ω} values) than among other parameters. This is not surprising, because the covariance matrix of the random coefficients appears in the most non-linear fashion in the likelihood function of Equation (9) through the overall covariance matrix $\mathbf{\Psi}_q$ ($\mathbf{\Psi}_q = \mathbf{M}_q(\mathbf{z}_q\mathbf{\Omega}\mathbf{z}'_q + \mathbf{\Lambda})\mathbf{M}'_q$; see Section 2.2), leading to somewhat more difficulty in accurately recovering the l_{Ω} parameters. The APB value is particularly high for the $l_{\Omega 5}$ and $l_{\Omega 6}$ parameters, though this could also be attributed to the low true values of these two parameters (which inflates the absolute percentage bias computations).

The standard error estimates of the parameters indicate good empirical efficiency of the MACML estimator. Across all parameters, the finite sample standard error (FSEE) is about 5.3% of the mean estimate, while the corresponding figure for the asymptotic standard error (ASE) is about 5.5%. This result indicates that, for the current experimental setting and sample size, the asymptotic standard error is providing a good estimate of the true finite sample error. The last column of Table 1a presents the absolute percentage bias associated with the ASE estimator (APBASE). Across all parameters, the mean APBASE value is about 9.3% (see last row). The APBASE values for the mean parameters of the random coefficients within the $\boldsymbol{\beta}_q$ vector (*i.e.*, the b_1, b_2, and b_3 parameters) are markedly higher than the APBASE values for the fixed coefficients within the $\boldsymbol{\beta}_q$ vector (*i.e.*, the b_4 and b_5 parameters). This is to be expected because of the multivariate normal distribution underlying the first three coefficients in the $\boldsymbol{\beta}_q$ parameter vector rather than a degenerate distribution for the final two parameters. The APBASE values of l_{Λ} is the highest among all parameters (15.7% on average). However, the associated FSEE value is also the lowest for these l_{Λ} parameters relative to other sets of parameters (average value of FSEE of 0.013 compared to corresponding value of 0.030 across all parameters). The low values of FSEE for the l_{Λ} translates to an inflation in the APBASE values. But the net difference between the ASE and the FSEE values even for the parameter with the highest APBASE (which is $l_{\Lambda 4}$) is only 0.002, a mere 0.26% of the mean estimate of $l_{\Lambda 4}$.

4.2.2 Ten-Alternative Case

The results for this case are presented in Table 1b and, as in the five-alternative case, indicate that the MACML method performs well in recovering the true parameter values. The maximum value of APB across all the parameters is 9%, with an overall mean value of 1.4%. These results suggest that increasing the number of alternatives does not substantially affect the ability of the MACML method to recover parameters (in the current exercise, the difference in APB between the five-alternative and ten-alternative cases is only 0.4%). As in the five-alternative case, with the exception of $l_{\Lambda 6}$ that has the highest APB of 8.97%, the Cholesky decomposition of the covariance matrix associated with the error terms (*i.e.*, the l_{Λ} values) are lower than for other parameters. In contrast, the satiation parameters (*i.e.*, the γ values) consistently present an APB value of more than 1%. This result is a reflection of somewhat greater difficulty in pinning the satiation parameters as the number of alternatives increases, especially since the satiation parameter governs the non-linearity in the utility function. However, even these APB values are all well below 2.5%.

The asymptotic standard error estimates in Table 1b again indicate good efficiency of the MACML estimator, with the asymptotic standard error across all parameters being only about 3% of the mean value of the parameters (3.27% for the FSEE and 3.47% for the ASE). The mean APBASE value across all parameters is 13.7%, slightly higher than in the five-alternative case. It is interesting to note the high APBASE value (64.85%) for b_5, especially because this parameter is a fixed parameter. But this also is because of the low finite standard error value of 0.007 for the parameter, which inflates the 0.005 absolute difference into the 64.85% APBASE value. Indeed, the discrepancy of 0.005 constitutes but 1% of the mean estimate of b_5.

5. ILLUSTRATIVE APPLICATION DEMONSTRATION TO RECREATIONAL TRAVEL DEMAND PATTERNS

5.1 Background

Long distance leisure travel is an important and well embedded element of American households' lifestyle.[8] In 2010, three out of four long distance domestic trips, which constitute

[8] Long-distance travel is usually defined to include travel with a one-way length that exceeds 100 miles (see LaMondia *et al.*, 2008). Note also that, while a trip in an urban context refers to one-way travel, a trip in the long distance travel context usually to refer to the entire round travel to a primary destination and back (this is what

about 1.5 billion person trips annually in the United States, were taken for leisure purposes (U.S. Travel Association (USTA), 2011).[9] The expenditure on long distance leisure travel (which we will refer to as recreational travel in the rest of this paper) has been estimated to generate $82 billion in tax revenue and to have supported 5.2 million jobs in 2010 (USTA, 2011). Indeed, the state of the economy and fuel prices does not seem to have tempered the amount of recreational travel, which actually saw a steady rise from 1.40 billion person trips in 2002 to 1.47 billion person trips in 2005 to the 1.5 billion person trips in 2010 (Holecek and White, 2007, USTA, 2010). Several reasons have been provided to explain this increase in recreation travel, including a sheer "size" effect related to the growth in US population, an increase in paid leave time, enhanced personal control over the travel experience, and marketing efforts to showcase cultural and natural heritage sites (see Alegre and Pou, 2006 and Siegel, 2011).

Even as the total volume of recreation travel has been increasing, so has the share of these trips undertaken close to home in the form of day trips to recreation and entertainment venues (see White, 2011). That is, there has been a shift from the traditional long period vacations undertaken during holidays or over the summer to short period recreation travel built around the work weeks. This shift in recreation travel patterns is a result of multiple considerations, including difficulties in coordinating long vacation getaways due to multiple working individuals in the household, and an increase in the rich and diverse opportunities for recreation offered in every state of the US through programs such as the National Scenic Byways Program. The net result has been a shrinkage in the geographic footprint of recreational travel as well as a significant increase in the mode share of personal auto-based recreation trips (see USTA, 2011).

The substantial and increasing amount of auto-based recreation travel over short distances, in turn, has important transportation air quality planning and tourism implications. From a transportation air quality planning standpoint, the predominantly auto-based recreation travel adds to intra-city urban traffic, and can lead to traffic congestion on the urban transportation network on holidays and weekends (see Jun, 2010 and Liu and Sharma, 2006).

---

would be referred to as a "tour" in an urban context). Leisure travel may be defined as "all journeys that do not fall clearly into the other well-established categories of commuting, business, education, escort, and sometimes other personal business and shopping" (Anable, 2002).

[9] The U.S. Travel Association defines a "person-trip" as one person on a trip away from home overnight in paid accommodations, or on a day or overnight trip to places 50 miles or more, one-way, from home.

Such congestion contributes to lost productivity, lost recreation time, and increased greenhouse gas and mobile source emissions. Understanding recreational travel flow patterns, therefore, can help in planning and implementing transportation control policies to reduce the negative externalities of such travel. From a tourism standpoint, a good understanding of recreational travel patterns helps provide insights into positioning and targeting strategies of services and attractions. States and communities have a vested interest in doing so, because tourism can generate much needed jobs and revenue for the economy.

To be sure, the study of recreational travel demand has received substantial attention both within and outside the transportation domain, with an emphasis on understanding individuals' recreational demand patterns in general (see, for example, LaMondia *et al*., 2010, Hailu and Gao, 2012, Humphreys and Ruseski, 2006, Vaaraa and Materoa, 2011, and Majumdar and Zhang, 2011) and destination choice patterns in particular (the focus of the current analysis). In the context of destination choice, many studies in transportation and other fields have used traditional random utility maximization models to analyze an individual's choice of visiting one destination among a set of available destinations on a single choice instance (see, for example, Hilger and Hanemann, 2006, Pozsgay and Bhat, 2002, Carson *et al*., 2009, Boeri *et al*., 2012, and Siderelis *et al*., 2011). These studies characterize destination locations based on their recreational offerings, facility costs and infrastructure, and travel characteristics. However, such models are unable to accommodate the demand for recreational trips over an extended time horizon, where the decision context shifts from the choice of a single destination to the choice of potentially multiple destinations (along with a count of the number of times each destination may be visited). As a result, the recreation demand field has seen the increasing use of an MDC modeling framework accommodating for unobserved taste heterogeneity across consumers (see, for example, Kuriyama *et al*., 2010, 2011, Van Nostrand *et al*., 2013, von Haefen, 2007, Whitehead *et al*., 2010). However, all of these papers adopt an identical and independent extreme value distribution for the kernel error terms.[10]

[10]As in the MDC recreational demand studies just listed above, we too focus on the count of the number of times each recreational destination is visited. Thus, the "continuous" quantity used is actually a count, as opposed to a truly continuous quantity measure as required by the theoretical model. However, a study by von Haefen and Phaneuf (2003) suggests that treating the integer count of trips as a continuous entity (within an MDC framework) do not lead to substantial biases. Nevertheless, this result can be context-specific, and methods that explicitly formulate a multiple discrete-count model are desirable and are being pursued. Also, in the current study, we do not examine the duration (*i.e*., number of days) of each recreational trip.

In this study, we demonstrate the application of the proposed MACML approach for the flexible MDCP model by analyzing Michigan residents' recreational travel demand and destination choice patterns for sites within Michigan. Covering about 57,022 square miles, the state of Michigan is recognized for its natural diversity, its lakes and streams, forests, beaches, ski areas and its variety of other recreational offerings. The recreational travel spending within Michigan is over $10 billion per year, which contributes more than 73% of the total visitor travel spending in the State of Michigan (Costa, 2009). Tourism in Michigan is a $17.5 billion industry, employs 200,000 people, and contributes to the economies of all 83 Michigan counties (Michigan Tourism Industry Planning Council (MTIPC), 2007). While the tourism industry in Michigan draws visitors from all over the country and the world, it is dominated by regional tourism, drawing about 70% of its revenues from Michigan residents (MTIPC, 2007). To examine this group of travelers, the Michigan Department of Transportation included a long distance retrospective travel survey component in the 2004-2005 Michigan statewide household travel survey (MSHTS), which is used as the primary data source for this study.

5.2 Data Description

5.2.1 Database

Three data sources are used in this research. The first, as just indicated, is the 2004-2005 MSHTS conducted between March 2004 and February 2005, which included a component that asked survey respondents to provide information on all trips 100 miles or longer one-way (to a primary destination for that trip) made up to a year prior to the survey date. The elicited information included the residence city of the respondent, the name of the primary destination city that the respondent visited on each trip, the primary reason for each trip, and the primary type of transportation used to reach the destination. In addition to the long distance trip information, the survey also obtained individual and household socio-demographic information. The second data source is a network level of service file that provided information on personal auto travel time and travel distance between each pair of the 484 origin cities within the State of Michigan. The third data source is a disaggregate spatial land-cover characteristics data obtained from the Geographic Data Library of the Michigan Department of Natural Resources.[11] An

[11] This data has land-cover characteristics available at the level of spatial pockets 900 square meters (0.2224 acres) in size. Each of these spatial pockets is classified into one of seven land-cover categories: urban, agricultural, bare-

elaborate geographic information system based procedure was used in our research to extract information from this third data source to obtain the total land area and acreage at a county level in seven disaggregate land-cover categories (including urban area, agricultural area, bare-land area, forest area, open area, wetland area, and water area).

5.2.2 Data Assembly

The final sample was assembled in a number of steps. First, from the long distance travel survey dataset, only those records corresponding to adults (age ≥ 16 years) with a primary purpose of recreation (vacation, sightseeing and other leisure activity pursuits) were selected. Further, the recreation travel patterns of adults from the same household had substantial overlaps with one another, and so we selected only one individual from each household (the survey was individual-focused, and provided no information on accompanying members on a recreation trip, but our analysis of recreation patterns of adults within the same household showed clear overlap). The individual selected from each household was the one who made the most number of recreation trips during the one-year recall period. Further, only long distance trips to primary destinations within the State of Michigan, and undertaken by the personal auto mode were considered. Second, the destination cities in Michigan were mapped to one of 83 counties, which were then themselves mapped into one of six aggregate destination zones considered in the current analysis (see Figure 1): (1) South-East Lower Peninsula (SELP), (2) South-West Lower Peninsula (SWLP), (3) North-East Lower Peninsula (NELP), (4) North-West Lower Peninsula (NWLP), (5) East Upper Peninsula (EUP), and (6) West Upper Peninsula (WUP). As the names suggest, the first four destination zones are located in the Lower Peninsula of Michigan, and the remaining two destination zones are located in the Upper Peninsula of Michigan. This classification into the six destination zones is the same as that used by the Michigan Economic Development Corporation in its tourism promotion and information campaigns, and is based on the geographic locations and recreational opportunities offered by the six different regions. Third, the total number of trips by each individual to each of the six destination zones was obtained by appropriate aggregation. Fourth, we obtained the total yearly in-state recreational travel budget for each individual as the sum of the individual's trips to the six destination zones

---

land, forest, open, wetland, and water, based on rules developed by the Michigan Department of Natural Resources (see Pacific Meridian Resources, 2001).

identified above.[12] Fifth, the city-to-city personal auto travel time and travel distance data were converted to corresponding residence city-to-destination zone data, by identifying a centroidal city for each of the six destination zones (these centroidal cities were at the center of multiple cities and attracted a majority of the travelers within the respective destination zones). Sixth, the county-based land-cover data were translated to a destination zone-based land-cover data by suitable aggregation over counties within each destination zone. Finally, individual and household socio-demographic information collected in the survey were appended to the long distance travel records.

5.2.3 Summary Statistics on the Choice of Destination Zones

The final sample included 1659 Michigan residents who reported a total of 6620 one-way recreational trips to one of the six Michigan destination zones in the twelve-month period prior to the survey. Of these travelers, 86.4% visited one destination zone, 12.2% visited two destination zones, and 1.5% visited three or more destination zones during the twelve month period of recall. On average, respondents visited slightly more than one (1.15) of the destination zones and made an average of almost four trips. Table 2 presents descriptive statistics by destination zone. The second column of the table provides information on the number (and percentage) of individuals who visited each destination at least once (the percentages add up to more than 100% across the rows of the column because some individuals visit multiple destination zones). The NWLP destination draws the highest percentage of individuals for a visit, while the WUP destination draws the lowest percentage. The NELP, SELP, and the EUP destinations also are popular. The third column presents statistics on the number of visits among those who visited each destination zone. The NELP and NWLP have, in the overall, the most loyal following (see the high mean values in these rows for the "number of trips among those

[12] Our procedure does not consider those Michigan residents who did not report any in-state recreational travel during the one year period preceding the survey. In the terminology of Section 2.1, our empirical analysis corresponds to an incomplete demand system in the form of a two stage budgeting approach. The first stage may be viewed as the allocation of a total leisure travel budget to general non-recreation travel (*i.e.*, non-long distance leisure travel, such as to local recreational spots, to local social or cultural events, or for local or long distance visiting) and recreation travel (*i.e.*, long-distance leisure travel, which is the focus of the current paper). The second stage corresponds to the allocation of any positive count of recreation travel (as determined in the first step) among the six destination zones. In doing so, we are invoking an assumption of strong separability of preferences between non-recreation and recreation travel, which is not unreasonable given that the drivers of these two types of travel are very different (see, for example, LaMondia and Bhat, 2012). Also, note that the budget constraint in this empirical demonstration is the total number of trips, rather than the total expenditure across all trips. This is because the Michigan survey did not have detailed expenditure data on trips.

who visited each destination"). Overall, the results suggest a relatively high baseline preference of Michigan travelers for the north lower peninsula. The particularly attractive year-round water-based activities offered in this part of the state, coupled with its appealing natural diversity and moderate travel cost appear to be reasons for this high baseline preference (as we will note in our empirical estimation results).

5.2.4 Utility Form and Exogenous Variable Specification

In the empirical context under study, we estimated both a γ-profile as well as an α-profile (see Section 2.1). Between these, the γ-profile consistently provided a much better data fit than the α-profile for a variety of different exogenous variable specifications, and so is the one used in the empirical analysis of the current paper. The exogenous variables considered in the recreation MDCP model, and their construction, are discussed in turn in the following paragraphs.

The travel cost variable is specified as a function of the respondent's reported household income, the estimated cost of vehicle operation ($0.149/mile), and the travel time and distance between the respondent's residence city and the centroidal city of each destination location. To calculate the travel cost, we follow the standard approach of valuing travel time at a fixed proportion of one-half of the wage rate (see Hanemann *et al*., 2004 for a detailed discussion). The household income is divided by the total number of adult individuals in the household to estimate the individual's wage rate. Specifically, the travel cost is computed as:

Cost (in $) = 2 * (one-way travel distance in miles * 0.149 + one-way travel time in hours * (0.5 * hourly wage)).

The destination zone-based land-cover data by themselves do not provide adequate variation to estimate parameters (because there are only six destination zones, and the land-cover data values for these destination zones do not change across individuals in the sample). But we capture land-cover effects by interacting the land-cover in each destination zone with the travel time from each individual's residence city to the centroidal city of each destination zone. To do so, for each combination of individual q in the sample, land-cover categories i (i = urban, agricultural, bare-land, forest, open, wetland, and water), and destination zone k, we compute an accessibility measure of the Hansen-type (Fotheringham, 1983) as $AC_{qik} = LC_{ik} / TT_{qk}$, where LC_{ik} is the area (in acres) in land-cover category i in destination zone k, and TT_{qk} is the travel time (in hours) from individual q's residence city to the centroid of destination zone k. The accessibility measures (in acres/hour) proxy the opportunities for recreational participation specific to each land-use category in a destination zone normalized by a measure of impedance (travel time) for individual q to reach those opportunities. A positive coefficient on an accessibility measure, say the one corresponding to “water” land-cover, implies that individuals are attracted toward proximal destination zones with substantial water bodies.

In addition to travel cost and destination zone accessibility variables, several household attributes (such as presence of children less than 16 years, number of cars, and number of workers) are interacted with travel cost and the accessibility variables.

Table 3 provides descriptive statistics of the cost-related and land-cover data for each of the six destination zones (we present only these destination zone statistics, rather than the statistics for accessibility measures and household attributes, to keep the presentation concise). Not surprisingly, Table 3 shows that the travel impedance measures (travel time, travel distance, and travel cost) are the highest for the EUP and the WUP destination zones that are well to the north of much of the resident population (see Figure 1), and the lowest for the SWLP and SELP destination zones. The NELP and NWLP destination zones have the highest percentage of land-cover in water that should make these regions particularly attractive as destination zones.

5.3 Empirical Results

The estimation results of the Mixed MDCP model are presented in Table 4. The effects of travel cost and the travel cost variable interacted with the low household income dummy variable need to be considered together. As expected, the effect of travel cost is, on average, negative, though there is also a large standard error for this effect. The combination of the mean and the standard error estimates on the travel cost coefficient indicates that travel cost (as constructed in the current empirical exercise) is valued negatively by about 89% of individuals from households earning \$30,000 or more in annual income, while about 11% of individuals in this income bracket prefer destination zones with higher cost (perhaps because of the recreational experience of travel time itself). Individuals from households with low household income (< \$30,000 per year) particularly prefer destinations zones that are less expensive to travel to relative to individuals from households with high household income (≥ \$30,000 per year), as reflected in the negative coefficient on the interaction of travel cost with the low household income dummy variable. In fact, close to 96% of individuals in this low income bracket prefer lower travel costs, as opposed to 89% in the non-low income bracket.

The urban land-cover accessibility effect reflects a positive disposition toward destination zones with high urban land-cover accessibility, particularly for individuals from households with no children less than 16 years of age (this latter effect is rather small in magnitude and only marginally significant). While there is some unobserved variation across individuals in the effect of the urban land-cover accessibility variable, the combination of the mean and standard error estimates show that the overall effect of the variable remains positive for almost all individuals. The covariance estimate (not shown in Table 4) between the travel cost and urban accessibility random coefficients was 0.044 (t-statistic of 2.17), suggesting that individuals who are less sensitive (more sensitive) to travel costs also prefer (dislike) urban destination zones. That is, individuals who prefer recreation based on man-made urban settings (amusement parks or leisure shopping complexes) appear not to mind spending additional time to get to their destinations, while those who prefer natural and pristine settings are the ones who would rather travel to close destinations to pursue their recreational interests. The effects of the other accessibility measures are intuitive. Destination zones with a relatively high water area and in close proximity are particularly preferred for recreational getaways, perhaps because of the natural beauty coupled with diverse outdoor recreational opportunities around water (such as kayaking, fishing, canoeing, and swimming). Finally, there is a general low baseline preference for destination zones with high land-covers of wetland and open areas.

The baseline constants in the model do not have any substantive interpretations, because of the presence of continuous and ordinal variables in the model. They simply adjust the baseline preferences to fit the data after controlling for the presence of other included variables in the model. The values and standard errors (in parenthesis) of the baseline constants (with the SELP alternative being the base category) are as follows: SWLP: -0.333 (0.09), NELP: -0.322 (0.08), NWLP: 0.407 (0.08), EUP: 0.022 (0.13), and WUP: -0.216 (0.19). Similarly, given the baseline utility function, the satiation parameters fit the number of visits to each destination zone (and also allow corner solutions or no visits to each destination zone). These values are as follows: SELP: 12.471 (2.82), SWLP: 10.322 (1.86), NELP: 84.13 (46.5), NWLP: 47.33 (15.92), EUP: 20.34 (5.99), and WUP: 98.59 (0.19). The satiation values are clearly different from the value of ∞, implying the presence of distinct satiation effects.

The log-likelihood value at convergence of the final model (with an unrestricted error covariance model up to identifiability limits; see Section 2.1) is -3726.2. As indicated earlier, allowing such a general covariance matrix (with 14 parameters) enables an unrestricted substitution pattern, but also does not provide any interpretable information regarding variances and correlations in the baseline utilities of the alternatives. So, we also tested many other restrictive covariance patterns, but none came even close to the data fit offered by the general covariance structure. For instance, the model with a diagonal covariance matrix with unequal variances for the error terms in the baseline utility (with one of the variances being normalized to 0.5 for identification) returned a log-likelihood value of -3787.8. The log-likelihood ratio value for testing this model with our general covariance structure is 123.2, which far exceeds the corresponding chi-squared table value with nine degrees of freedom at even the 0.0001 level of significance. Similar results were obtained with other (identifiable) restrictive structures, such as allowing separate error components for each of the south lower, north lower and upper regions of the state, as well as another separate error component for the east part of the state, in addition to unequal variances (this model returned a log-likelihood value at convergence of -3755.0 with a total of 9 covariance parameters, which is again soundly rejected by our general covariance model). Thus, we retained the general covariance structure in this paper.[13] As a baseline, the log-likelihood value at convergence for the model with only the baseline constants and the satiation parameters, and with an independent and identically distributed error structure for the error terms of the alternatives, is -4120.6. The likelihood ratio test for testing the presence of exogenous variable effects, random coefficients, and a general covariance structure is 788.8, which is substantially larger than the critical chi-squared value with 23 degrees of freedom at any reasonable level of significance. This clearly indicates the value of the model estimated in this paper to predict individuals' recreational destination travel demand.

The results from models such as the one in this paper may be used for compensating value computations and welfare analysis, since they are derived explicitly from utility

[13] We do not present the general covariance matrix results here to conserve on space, and also because the covariance elements do not provide any substantive insights (the results are, however, available from the authors, as are the results of all the more restrictive covariance specifications). Also, it is important to note that we are able to allow a general covariance matrix in our demonstration empirical exercise because the number of alternatives is only six. As the number of alternatives increases, the number of parameters in a general covariance matrix will explode, making it necessarily to impose some *a priori* structure.

maximizing principles. However, the intent in this paper is primarily to demonstrate the application of our proposed MACML approach to estimate the MDCP model.

6. CONCLUSIONS

The current paper develops a blueprint (complete with matrix notation to code and implement the method in software programs) to apply Bhat's (2011) Maximum Approximate Composite Marginal Likelihood (or MACML) inference approach for the estimation of cross-sectional as well as panel multiple discrete-continuous (MDC) models. The MACML inference approach is simple, computationally efficient, simulation-free, and relatively easy to code and apply using readily available software for likelihood estimation. It involves only univariate and bivariate cumulative normal distribution function evaluations in the likelihood function (in addition to the evaluation of a closed-form multivariate normal density function). In the MACML approach, it is much easier to estimate a MDC probit (MDCP) model with normally distributed unobserved heterogeneity effects than normally mixed versions of the MDC generalized extreme value models. This is because of the conjugate addition property of the multivariate normal distribution to addition, which is exploited by the MACML inference approach.

A simulation exercise is undertaken to evaluate the ability of the proposed approach to recover parameters from a cross-sectional MDCP model. Two cases are considered: (1) a five-alternative case with five exogenous variables and (2) a ten-alternative case with five exogenous variables. For both cases, the coefficients of the first three exogenous variables are assumed to be randomly distributed according to a trivariate normal distribution, while the coefficients on the last two variables are fixed. The Cholesky matrix associated with the differenced error terms are estimated for the simulated data. The results show that our proposed approach does very well in recovering the parameters in the MDCP model. In addition, the Hessian of the likelihood function also appears to be computed accurately, as evidenced by the closeness of the asymptotic standard error estimates with the finite sample standard errors.

The paper demonstrates the application of the proposed approach through a study of individuals' recreational (*i.e.*, long distance leisure) choice among alternative destination locations and the number of trips to each recreational destination location, using data drawn from the 2004-2005 Michigan statewide household travel survey. The MDCP model estimates provide insights into the influence of travel cost, destination land-cover specific accessibility factors, and interactions of these with individual sociodemographics. The results indicate statistically significant observed and unobserved heterogeneity in response to travel cost as well as land-cover accessibility measures. The estimated flexible covariance matrix rejects the often-invoked independently and identically distributed error structure for the kernel error terms in the baseline utilities of the alternatives.

Overall, it is remarkable that the MACML approach is able to accurately recover the parameters of the cross-sectional MDCP model, as well as the standard errors of the parameter estimates. Of course, continued exploration of the performance of the MACML inference approach and other alternative approaches is needed through simulation exercises with alternative covariance structures, different numbers of alternatives (such as 15, 20, and more), and different sample sizes to assess parameter recoverability and estimator efficiency in finite sample sizes. Also, future studies should examine the ability of the MACML approach to recover parameters in a panel model. But we hope that the proposed MACML procedure for MDCP models will spawn empirical research into behaviorally rich model specifications within the MDC choice modeling context.

ACKNOWLEDGEMENTS

Karen Faussett of the Michigan Department of Transportation in Lansing assisted with the MHTS data used in the analysis. The authors are grateful to Lisa Macias for her help in formatting this document. Discussions with Prof. Jeff LaMondia provided useful insights. Two referees provided valuable comments on an earlier version of this paper.

REFERENCES

Alegre, J., Pou, L., 2006. An analysis of the microeconomic determinants of travel frequency. Department of Applied Economics, Universitat de les Illes Balears.

Anable, J., 2002. Picnics, pets, and pleasant places: the distinguishing characteristics of leisure travel demand. In Black, W.R., Nijkamp, P. (Eds.), *Social Change and Sustainable Transport*, 181-190, Indiana University Press, Bloomington, IN.

Baltas, G., 2004. A model for multiple brand choice. *European Journal of Operational Research* 154(1), 144-149.

Bhat, C.R., 2001. Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research Part B* 35(7), 677-693.

Bhat, C.R., 2003. Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. *Transportation Research Part B* 37(9), 837-855.

Bhat, C.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.

Bhat, C.R., 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.

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

Bhat, C.R., Pinjari, A.R., 2010. The generalized multiple discrete-continuous extreme value (GMDCEV) model: allowing for non-additively separable and flexible utility forms. Technical paper, Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin.

Bhat, C.R., Sidharthan, R., 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, C.R., Srinivasan, S., 2005. A multidimensional mixed ordered-response model for analyzing weekend activity participation. *Transportation Research Part B* 39(3), 255-278.

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

Boeri, M., Longo, A., Doherty, E., Hynes, S., 2012. Site choices in recreational demand: a matter of utility maximization or regret minimization? *Journal of Environmental Economics and Policy* 1(1), 32-47.

Brownstone, D., Fang, H.A., 2010. A vehicle ownership and utilization choice model with endogenous residential density, University of California Transportation Center, UC Berkeley.

Carson, R.T., Hanemann, M., Wegge, T.C., 2009. A nested logit model of recreational fishing demand in Alaska. *Marine Resource Economics* 24(2), 101-129.

Costa, B., 2009. Michigan 2009 visitor profile. Prepared for the Michigan Economic Development Corporation.

Craig, P., 2008. A new reconstruction of multivariate normal orthant probabilities. *Journal of the Royal Statistical Society: Series B* 70(1), 227-243.

Edwards, Y.D., Allenby, G.M., 2003. Multivariate analysis of multiple response data. *Journal of Marketing Research* 40(3), 321-334.

Fang, H.A., 2008. A discrete-continuous model of households' vehicle choice and usage, with an application to the effects of residential density. *Transportation Research Part B* 42(9), 736–758.

Fotheringham, A.S., 1983. Some theoretical aspects of destination choice and their relevance to production-constrained gravity models. *Environment and Planning A* 15(8), 1121-1132.

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

Hailu, A., Gao, L., 2012. Recreational trip timing and duration prediction: a research note. Working paper no. 1002, School of Agriculture and Resource Economics, University of Western Australia.

Hanemann, W.M., 1978. A methodological and empirical study of the recreation benefits from water quality improvement. Ph.D. dissertation, Department of Economics, Harvard University.

Hanemann, W.M., Pendleton, L., Mohn, C., Hilger, J., Kuriyama, K., Layton, D., Busch, C., Vasquez, F., 2004. Using revealed preference models to estimate the effect of coastal water quality on beach choice in Southern California. University of California at Berkeley, report to the U.S. National Oceanic and Atmospheric Administration.

Hilger, J., Hanemann, W.M., 2006. Heterogeneous preferences for water quality: a finite mixture model of beach recreation in Southern California. California Sea Grant College Program, Research Completion Reports, University of California, San Diego.

Holecek, D., White, R., 2007. Record-high gas prices won't negatively affect travelers, MSU Researcher says. Michigan State University Tourism Center News Release, May 17, 2007.

Humphreys, B.R., Ruseski, J.E., 2006. Economic determinants of participation in physical activity and sport. Working paper no. 0613, International Association of Sports Economists.

Jun, J., 2010. Understanding the variability of speed distributions under mixed traffic conditions caused by holiday traffic. *Transportation Research Part C* 18(4), 599-610.

Kim, J., Allenby, G.M., Rossi, P.E., 2002. Modeling consumer demand for variety. *Marketing Science* 21(3), 229-250.

Kuriyama, K., Hanemann, W.M., Hilger, J.R., 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.

Kuriyama, K., Shoji, Y., Tsuge, T., 2011. Estimating value of mortality risk reduction using the Kuhn-Tucker model: an application to recreation demand. Working Paper, Graduate School of Agriculture, Kyoto University.

LaMondia, J.J., Bhat, C.R., 2012. A conceptual and methodological framework of leisure activity loyalty accommodating the travel context. *Transportation* 39(2), 321-349.

LaMondia, J.J., Bhat, C.R., Hensher, D.A., 2008. An annual time use model for domestic vacation travel. *Journal of Choice Modelling* 1(1), 70-97.

LaMondia, J.J., Snell, T., Bhat, C.R., 2010. Traveler behavior and values analysis in the context of vacation destination and travel mode choices: European Union case study. *Transportation Research Record* 2156, 140-149.

Liu, Z., Sharma, S., 2006. Statistical investigations of statutory holiday effects on traffic volumes. *Transportation Research Record* 1945, 40–48.

Majumdar, S., Zhang Y., 2011. Consumptive nature-based recreation in the United States: welfare measures from truncated and untruncated count data models. *International Journal of Ecological Economics and Statistics* 20(W11), 32-50.

Manchanda, P., Ansari, A., Gupta, S., 1999. The "shopping basket": a model for multicategory purchase incidence decisions. *Marketing Science* 18(2), 95-114.

Menezes, C.F., Wang, X.H., Bigelow, J.P., 2005. Duality and consumption decisions under income and price risk. *Journal of Mathematical Economics* 41(3), 387-405.

Michigan Tourism Industry Planning Council, 2007. 2007-2011 Michigan Tourism Industry Strategic Plan.

Pacific Meridian Resources, 2001. Integrated forest monitoring assessment and prescription (IFMAP), review of remote sensing technologies for the IFMAP project. Prepared for Michigan Department of Natural Resources.

Phaneuf, D.J., Smith, V.K., 2005. Recreation demand models. In Mäler, K-G., Vincent, J.R., (Eds.) *Handbook of Environmental Economics*, Volume 2, 671-762, Elsevier B.V., Amsterdam, The Netherlands.

Pinjari, A.R., 2011. Generalized extreme value (GEV)-based error structures for multiple discrete-continuous choice models. *Transportation Research Part B* 45(3), 474-489.

Pinjari, A.R., Bhat, C.R., 2011. An efficient forecasting procedure for Kuhn-Tucker consumer demand model systems: application to residential energy consumption analysis. Technical paper, Department of Civil and Environmental Engineering, University of South Florida.

Pozsgay, M.A., Bhat, C.R., 2002. Destination choice modeling for home-based recreational trips: analysis and implications for land-use, transportation, and air quality planning. *Transportation Research Record* 1777, 47-54.

Siderelis, C., Moore, R.L., Lee, J., 2011. A mixed logit model of visitors' National Park choices, *Society and Natural Resources* 24(8), 799-813.

Siegel, B., 2011. The power of destination marketing: "Pure Michigan" and "Philadelphia: with love" case studies. Longwoods International.

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

U.S. Travel Association, 2011. U.S. Travel answer sheet.

U.S. Travel Association, 2010. Travel and tourism works for America 2010-2011.

Vaaraa, M., Materoa, J., 2011. Modeling daily outdoor recreation participation and time-use as a two-stage choice process: a Finnish example. *Leisure Sciences: An Interdisciplinary Journal* 33(4), 269-289.

Van Nostrand, C., Sivaraman, V., Pinjari, A.R., 2013. Analysis of long-distance vacation travel demand in the United States: a multiple discrete-continuous choice framework. *Transportation* 40(1), 151-171.

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

von Haefen, R.H., Phaneuf, D.J., 2003. Estimating preferences for outdoor recreation: a comparison of continuous and count data demand system frameworks, *Journal of Environmental Economics and Management* 45, 612-630.

von Haefen, R.H., 2007. Empirical strategies for incorporating weak complementarity into consumer demand models. *Journal of Environmental Economics and Management* 54(1), 15-31.

von Haefen, R.H., Phaneuf, D.J., 2005. Kuhn-Tucker demand system approaches to nonmarket valuation. In Scarpa, R., Alberini, A.A. (Eds.), *Applications of Simulation Methods in Environmental and Resource Economics*, 135-158, Springer, Dordrecht, The Netherlands.

Wales, T.J., Woodland, A.D., 1983. Estimation of consumer demand systems with binding non-negativity constraints. *Journal of Econometrics* 21(3), 263-85.

White, R., 2011. Is the staycation trend a real phenomenon? White paper, White Hutchinson Leisure & Learning Group, Kansas City, MO, January. Available at: http://www.whitehutchinson.com/leisure/articles/Staycation.shtml.

Whitehead, J.C., Phaneuf, D.J., Dumas, C.F., Herstine, J., Hill, J., Buerger, B., 2010. Convergent validity of revealed and stated recreation behavior with quality change: a comparison of multiple and single site demands. *Environmental and Resource Economics* 45(1), 91-112.

Appendix: The basics of the MACML approach

There are two fundamental concepts in the MACML approach to estimate MDCP models. The first is an approximation method to evaluate the multivariate standard normal cumulative distribution (MVNCD) function. The second is the composite marginal likelihood (CML) approach to estimation. For cross-sectional MDCP models, only the MVNCD approximation is involved. In panel MDCP models, both the MVNCD approximation as well as the CML approach are involved. The discussion below is drawn from Bhat (2011), and provided in this paper for completeness following a recommendation by one of the reviewers of the paper.

2.1. Multivariate Standard Normal Cumulative Distribution (MVNCD) Function

In the MACML inference approach, an *analytic approximation* method is used to evaluate the MVNCD function. Unlike Monte-Carlo simulation approaches, even two to three decimal places of accuracy in the analytic approximation is generally adequate to accurately and precisely recover the parameters and their covariance matrix estimates because of the smooth nature of the first and second derivatives of the approximated analytic log-likelihood function. The analytic approximation used is based on decomposition of the MVNCD function into a product of conditional probabilities. To describe the approximation, let $(W_1, W_2, W_3, ..., W_I)$ be a multivariate normally distributed random vector with zero means, variances of 1, and a correlation matrix $\mathbf{\Sigma}$. Then, interest centers on approximating the following orthant probability:

$$\Pr(\boldsymbol{W} < \boldsymbol{w}) = \Pr(W_1 < w_1,\ W_2 < w_2,\ W_3 < w_3, ...,\ W_I < w_I)\,. \tag{A.1}$$

The above joint probability may be written as the product of a bivariate marginal probability and univariate conditional probabilities as follows ($I \geq 3$):

$$\begin{aligned}\Pr(\boldsymbol{W} < \boldsymbol{w}) &= \Pr(W_1 < w_1,\ W_2 < w_2) \times \\ &\prod_{i=3}^{I} \Pr(W_i < w_i \mid W_1 < w_1,\ W_2 < w_2,\ W_3 < w_3, ...,\ W_{i-1} < w_{i-1})\,.\end{aligned} \tag{A.2}$$

Next, define the binary indicator $\tilde{I}_i$ that takes the value 1 if $W_i < w_i$ and zero otherwise. Then $E(\tilde{I}_i) = \Phi(w_i)$, where $\Phi(.)$ is the univariate normal standard cumulative distribution function. Also, we may write the following:

$$\begin{aligned}&\mathrm{Cov}(\tilde{I}_i, \tilde{I}_j) = E(\tilde{I}_i \tilde{I}_j) - E(\tilde{I}_i)E(\tilde{I}_j) = \Phi_2(w_i, w_j, \rho_{ij}) - \Phi(w_i)\Phi(w_j),\ i \neq j \\ &\mathrm{Cov}(\tilde{I}_i, \tilde{I}_i) = Var(\tilde{I}_i) = \Phi(w_i) - \Phi^2(w_i) \\ &\qquad\qquad\qquad\qquad = \Phi(w_i)[1 - \Phi(w_i)]\,,\end{aligned} \tag{A.3}$$

where ρ_{ij} is the ij^{th} element of the correlation matrix $\mathbf{\Sigma}$. With the above preliminaries, consider the following conditional probability:

$$\begin{aligned}&\Pr(W_i < w_i \mid W_1 < w_1,\ W_2 < w_2,\ W_3 < w_3, ...,\ W_{i-1} < w_{i-1}) \\ &= E(\tilde{I}_i \mid \tilde{I}_1 = 1,\ \tilde{I}_2 = 1,\ \tilde{I}_3 = 1, ...,\ \tilde{I}_{i-1} = 1)\,.\end{aligned} \tag{A.4}$$

The right side of the expression may be approximated by a linear regression model, with $\tilde{I}_i$ being the "dependent" random variable and $\tilde{\boldsymbol{I}}_{<i} = (\tilde{I}_1, \tilde{I}_2, ...\tilde{I}_{i-1})$ being the independent random variable vector. In deviation form, the linear regression for approximating Equation (A.4) may be written as:

$$\tilde{I}_i - E(\tilde{I}_i) = \boldsymbol{\alpha}'[\tilde{\boldsymbol{I}}_{<i} - E(\tilde{\boldsymbol{I}}_{<i})] + \tilde{\eta}\ , \tag{A.5}$$

where $\boldsymbol{\alpha}$ is the least squares coefficient vector and $\tilde{\eta}$ is a mean zero random term. In this form, the usual least squares estimate of $\boldsymbol{\alpha}$ is given by:

$$\hat{\boldsymbol{\alpha}} = \boldsymbol{\Omega}_{<i}^{-1} \cdot \boldsymbol{\Omega}_{i,<i}\ \text{, where} \tag{A.6}$$

$$\boldsymbol{\Omega}_{<i} = \mathrm{Cov}(\boldsymbol{I}_{<i}, \boldsymbol{I}_{<i}) = \begin{bmatrix} \mathrm{Cov}(\tilde{I}_1, \tilde{I}_1) & \mathrm{Cov}(\tilde{I}_1, \tilde{I}_2) & \mathrm{Cov}(\tilde{I}_1, \tilde{I}_3) & \cdots & \mathrm{Cov}(\tilde{I}_1, \tilde{I}_{i-1}) \\ \mathrm{Cov}(\tilde{I}_2, \tilde{I}_1) & \mathrm{Cov}(\tilde{I}_2, \tilde{I}_2) & \mathrm{Cov}(\tilde{I}_2, \tilde{I}_3) & \cdots & \mathrm{Cov}(\tilde{I}_2, \tilde{I}_{i-1}) \\ \mathrm{Cov}(\tilde{I}_3, \tilde{I}_1) & \mathrm{Cov}(I_3, I_2) & \mathrm{Cov}(\tilde{I}_3, \tilde{I}_3) & \cdots & \mathrm{Cov}(\tilde{I}_3, \tilde{I}_{i-1}) \\ \vdots & & & & \\ \mathrm{Cov}(\tilde{I}_{i-1}, \tilde{I}_1) & \mathrm{Cov}(\tilde{I}_{i-1}, \tilde{I}_2) & \mathrm{Cov}(\tilde{I}_{i-1}, \tilde{I}_3) & \cdots & \mathrm{Cov}(\tilde{I}_{i-1}, \tilde{I}_{i-1}) \end{bmatrix}\text{, and} \tag{A.7}$$

$$\boldsymbol{\Omega}_{i,<i} = \mathrm{Cov}(\boldsymbol{I}_{<i}, \boldsymbol{I}_i) = \begin{bmatrix} \mathrm{Cov}(\tilde{I}_i, \tilde{I}_1) \\ \mathrm{Cov}(\tilde{I}_i, \tilde{I}_2) \\ \mathrm{Cov}(\tilde{I}_i, \tilde{I}_3) \\ \vdots \\ \mathrm{Cov}(\tilde{I}_i, \tilde{I}_{i-1}) \end{bmatrix}.$$

Finally, putting the estimate of $\hat{\boldsymbol{\alpha}}$ back in Equation (A.5), and predicting the expected value of $\tilde{I}_i$ conditional on $\tilde{\boldsymbol{I}}_{<i} = \mathbf{1}$ (*i.e.*, $\tilde{I}_1 = 1,\ \tilde{I}_2 = 1,\ \tilde{I}_{i-1} = 1$), we get the following approximation for Equation (A.4):

$$\begin{aligned} &\Pr(W_i < w_i \mid W_1 < w_1,\ W_2 < w_2, ...,\ W_{i-1} < w_{i-1}) \approx \\ &\Phi(w_i) + (\boldsymbol{\Omega}_{<i}^{-1} \cdot \boldsymbol{\Omega}_{i,<i})'(1 - \Phi(w_1),\ 1 - \Phi(w_2) ... 1 - \Phi(w_{i-1}))' \end{aligned} \tag{A.8}$$

This conditional probability approximation can be plugged into Equation (A.2) to approximate the multivariate orthant probability in Equation (A.1). The resulting expression for the multivariate orthant probability comprises only univariate and bivariate standard normal cumulative distribution functions.

One remaining issue is that the decomposition of Equation (A.1) into conditional probabilities in Equation (A.2) is not unique. Further, different permutations (*i.e.*, orderings of the elements of the random vector $\boldsymbol{W} = (W_1, W_2, W_3, ..., W_I))$ for the decomposition into the conditional probability expression of Equation (A.2) will lead, in general, to different approximations. In the case when the approximation is used for model estimation (where the integrand in each individual's log-likelihood contribution is a parameterized function of the $\boldsymbol{\beta}$

and $\mathbf{\Sigma}$ parameters), even a single permutation of the $\boldsymbol{W}$ vector per choice occasion should typically suffice (though the single permutation must vary across choice occasions).

2.2. The Composite Marginal Likelihood (CML) Estimator

The composite marginal likelihood (CML) estimation approach is a relatively simple approach that can be used when the full likelihood function is practically infeasible to evaluate due to underlying complex dependencies.

The CML approach, which belongs to the more general class of composite likelihood function approaches, is based on maximizing a surrogate likelihood function that compounds much easier-to-compute, lower-dimensional, marginal likelihoods (see Varin *et al*., 2011 for recent reviews of the CML method). The CML approach works as follows. Assume that the data originate from a parametric underlying model based on a $D \times 1$ vector random variable $\boldsymbol{Y}$ with density function $f(\boldsymbol{y},\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is an unknown $\widetilde{K}$-dimensional parameter vector. Suppose that $f(\boldsymbol{y},\boldsymbol{\theta})$ is difficult or near infeasible to evaluate in reasonable time with the computational resources at hand, so that the corresponding likelihood function from a sampled (observed) vector for $\boldsymbol{Y}$ (say $\boldsymbol{m} = (m_1, m_2, m_3, \ldots m_D)'$) given by $L(\boldsymbol{\theta}; \boldsymbol{m}) = f(\boldsymbol{m},\boldsymbol{\theta})$ is difficult. However, suppose evaluating the likelihood functions of a set of $\widetilde{E}$ observed marginal events (each observed marginal event being a subset of the observed joint event $\boldsymbol{m}$) is easy and/or computationally expedient. Let these observed marginal events be characterized by ($A_1(\boldsymbol{m}), A_2(\boldsymbol{m}), \ldots, A_{\widetilde{E}}(\boldsymbol{m})$). For instance, $A_1(\boldsymbol{m})$ may represent the marginal event that the observed values in the sample for the first two elements of the vector $\boldsymbol{Y}$ are $(m_1, m_2)'$, $A_2(\boldsymbol{m})$ may represent the marginal event that the observed values for the first and third elements of the vector $\boldsymbol{Y}$ are $(m_1, m_3)'$, and so on. Let each event $A_e(\boldsymbol{m})$ be associated with a likelihood object $L_e(\boldsymbol{\theta}; \boldsymbol{m}) = L[\boldsymbol{\theta}; A_e(\boldsymbol{m})]$, which is based on a lower-dimensional marginal joint density function corresponding to the original high-dimensional joint density of $\boldsymbol{Y}$. Then, the general form of the composite marginal likelihood function is as follows:

$$L_{CML}(\boldsymbol{\theta}, \boldsymbol{m}) = \prod_{e=1}^{\widetilde{E}} \left[L_e(\boldsymbol{\theta}; \boldsymbol{m})\right]^{\omega_e} = \prod_{e=1}^{\widetilde{E}} \left[L(\boldsymbol{\theta}; A_e(\boldsymbol{m})\right]^{\omega_e}, \tag{9}$$

where ω_e is a power weight to be chosen based on efficiency considerations. If these power weights are the same across events, they may be dropped. The CML estimator is the one that maximizes the above function (or equivalently, its logarithmic transformation). The CML class of estimators subsumes the usual ordinary full-information likelihood estimator as a special case.

The properties of the general CML estimator may be derived using the theory of estimating equations. Under usual regularity conditions (these are the usual conditions needed for likelihood objects to ensure that the logarithm of the CML function can be maximized by solving the corresponding score equations), the maximization of the logarithm of the CML function in Equation (9) is achieved by solving the composite score equations given by $s_{CML}(\boldsymbol{\theta}, \boldsymbol{m}) = \nabla \log L_{CML}(\boldsymbol{\theta}, \boldsymbol{m}) = \sum_{e=1}^{\widetilde{E}} \omega_e s_e(\boldsymbol{\theta}, \boldsymbol{m}) = \mathbf{0}$, where $s_e(\boldsymbol{\theta}, \boldsymbol{m}) = \nabla \log L_e(\boldsymbol{\theta}; \boldsymbol{m})$. Since these equations are linear combinations of valid likelihood score functions associated with the event probabilities forming the composite log-likelihood function, they immediately satisfy the requirement of being unbiased. Further, if q independent observations on the vector $\boldsymbol{Y}$ are available (say $\boldsymbol{m^1}, \boldsymbol{m^2}, \boldsymbol{m^3}, \ldots, \boldsymbol{m^Q}$), as would be the case when there are several individuals q (q = 1, 2, 3,…, Q) with panel data or repeated choice data, then, in the asymptotic scenario that $Q \to \infty$ with D fixed, a central limit theorem and a first-order Taylor series expansion can be applied in the usual way (see, for example, Godambe, 1960) to the resulting mean composite score function $\left(= \frac{1}{Q} \sum_{q=1}^{Q} s_{CML,q}(\boldsymbol{\theta}, \boldsymbol{m^q}) \right)$ to obtain consistency and asymptotic normality of the CML estimator:

$$\sqrt{Q}(\hat{\boldsymbol{\theta}}_{CML} - \boldsymbol{\theta}) \xrightarrow{d} N_{\tilde{K}}\left[(0, \boldsymbol{G^{-1}}(\boldsymbol{\theta})\right], \tag{13}$$

where $\boldsymbol{G}(\boldsymbol{\theta})$ is the Godambe information matrix defined as $\boldsymbol{H}(\boldsymbol{\theta})[\boldsymbol{J}(\boldsymbol{\theta})]^{-1}[\boldsymbol{H}(\boldsymbol{\theta})]$. $\boldsymbol{H}(\boldsymbol{\theta})$ and $\boldsymbol{J}(\boldsymbol{\theta})$ take the following form:

$$\boldsymbol{H}(\boldsymbol{\theta}) = E\left[-\frac{\partial^2 \log L_{CML}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right] \text{ and } \boldsymbol{J}(\boldsymbol{\theta}) = E\left[\left(\frac{\partial \log L_{CML}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \left(\frac{\partial \log L_{CML}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \right) \right]. \tag{14}$$

These may be estimated in a straightforward manner at the CML estimate $\hat{\boldsymbol{\theta}}_{\boldsymbol{CML}}$ as follows:

$$\hat{\boldsymbol{H}}(\hat{\boldsymbol{\theta}}) = -\left[\sum_{q=1}^{Q} \frac{\partial^2 \log L_{CML,q}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right]_{\hat{\boldsymbol{\theta}}_{CML}}, \text{ and}$$

$$\hat{\boldsymbol{J}}(\hat{\boldsymbol{\theta}}) = \sum_{q=1}^{Q} \left[\left(\frac{\partial \log L_{CML,q}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \left(\frac{\partial \log L_{CML,q}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \right) \right]_{\hat{\boldsymbol{\theta}}_{\boldsymbol{CML}}}.$$

LIST OF FIGURES

LIST OF TABLES

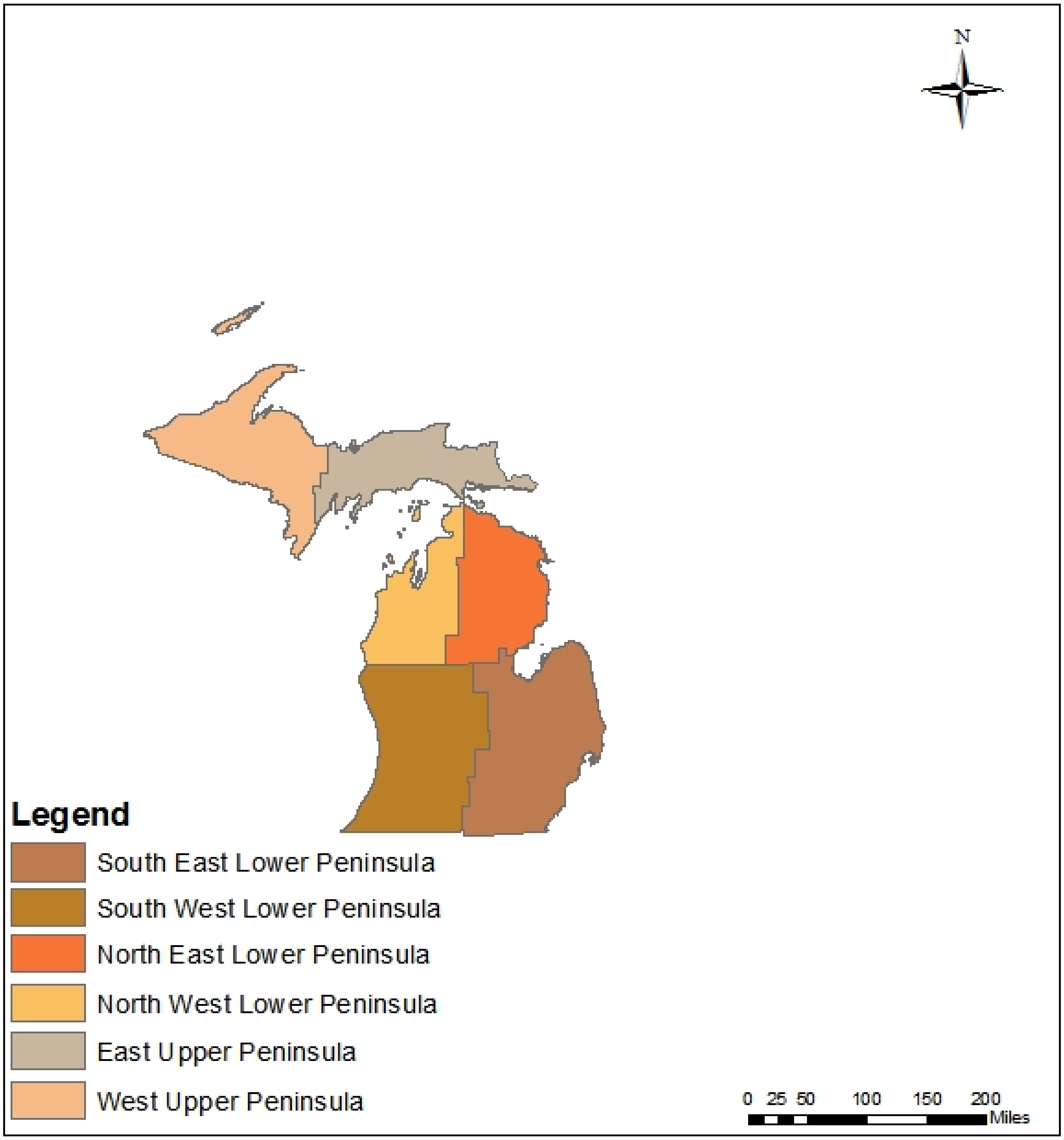

Figure 1: Destination Zones in the Empirical Analysis

Table 1: MDCP Model Estimation Results for the Simulated Data

Table 1a: Simulation results for the five-alternative case

Parameter	True value	Parameter estimates		Standard error estimates		
		Mean estimate	Absolute percentage bias (APB)	Finite sample standard error (FSE)	Asymptotic standard error (ASE)	Absolute percentage bias asymptotic standard error (APBASE)
Mean values of the $\boldsymbol{\beta}_q$ *vector (**b**)*						
b_1	0.500	0.494	1.133 %	0.021	0.019	12.332 %
b_2	-1.000	-0.987	1.279 %	0.021	0.025	20.169 %
b_3	1.000	1.007	0.659 %	0.022	0.025	11.225 %
b_4	-1.000	-0.997	0.299 %	0.013	0.013	1.833 %
b_5	-0.500	-0.505	0.934 %	0.012	0.012	3.051 %
Cholesky parameters characterizing the covariance matrix of the $\boldsymbol{\beta}_q$ *vector* ($l_{\boldsymbol{\Omega}}$)						
$l_{\boldsymbol{\Omega}1}$	0.900	0.898	0.192 %	0.019	0.017	6.142 %
$l_{\boldsymbol{\Omega}2}$	0.600	0.605	0.839 %	0.032	0.035	7.831 %
$l_{\boldsymbol{\Omega}3}$	0.800	0.794	0.733 %	0.032	0.033	5.798 %
$l_{\boldsymbol{\Omega}4}$	0.800	0.791	1.186 %	0.034	0.032	5.181 %
$l_{\boldsymbol{\Omega}5}$	0.400	0.415	3.794 %	0.045	0.049	10.282 %
$l_{\boldsymbol{\Omega}6}$	0.300	0.291	3.127 %	0.105	0.116	10.913 %
Cholesky parameters characterizing the covariance matrix of the $\boldsymbol{\xi}_q$ *vector* ($l_{\boldsymbol{\Lambda}}$)						
$l_{\boldsymbol{\Lambda}1}$	1.100	1.095	0.487 %	0.017	0.019	15.793 %
$l_{\boldsymbol{\Lambda}2}$	1.000	0.995	0.484 %	0.012	0.013	7.849 %
$l_{\boldsymbol{\Lambda}3}$	0.600	0.596	0.713 %	0.018	0.016	10.679 %
$l_{\boldsymbol{\Lambda}4}$	0.800	0.797	0.400 %	0.007	0.009	28.653 %
Satiation parameters ($\boldsymbol{\gamma}$)						
γ_1	1.000	1.002	0.224 %	0.036	0.036	0.405 %
γ_2	1.000	1.007	0.689 %	0.044	0.039	11.310 %
γ_3	1.000	1.016	1.598 %	0.038	0.040	6.161 %
γ_4	1.000	1.004	0.374 %	0.041	0.037	8.996 %
γ_5	1.000	1.001	0.051 %	0.037	0.037	0.881 %
Overall mean value across parameters		0.566	0.960 %	0.030	0.031	9.274 %

Table 1b: Simulation results for the ten-alternative case

Parameter	True value	Parameter estimates		Standard error estimates		
		Mean estimate	Absolute percentage bias (APB)	Finite sample standard error (FSE)	Asymptotic standard error (ASE)	Absolute percentage bias asymptotic standard error (APBASE)
Mean values of the $\boldsymbol{\beta}_q$ vector ($\boldsymbol{b}$)						
b_1	0.500	0.494	1.218 %	0.017	0.016	5.201 %
b_2	-1.000	-1.006	0.646 %	0.028	0.025	10.658 %
b_3	1.000	0.986	1.450 %	0.028	0.025	10.565 %
b_4	-1.000	-1.000	0.048 %	0.013	0.013	1.303 %
b_5	-0.500	-0.502	0.414 %	0.007	0.012	64.850 %
Cholesky parameters characterizing the covariance matrix of the $\boldsymbol{\beta}_q$ vector ($l_{\boldsymbol{\Omega}}$)						
$l_{\boldsymbol{\Omega}1}$	0.900	0.900	0.049 %	0.017	0.015	13.255 %
$l_{\boldsymbol{\Omega}2}$	0.600	0.602	0.342 %	0.030	0.031	2.412 %
$l_{\boldsymbol{\Omega}3}$	0.800	0.800	0.028 %	0.035	0.032	7.280 %
$l_{\boldsymbol{\Omega}4}$	0.800	0.810	1.244 %	0.028	0.029	2.310 %
$l_{\boldsymbol{\Omega}5}$	0.400	0.401	0.288 %	0.036	0.046	27.111 %
$l_{\boldsymbol{\Omega}6}$	0.300	0.284	5.275 %	0.075	0.093	24.057 %
Cholesky parameters characterizing the covariance matrix of the $\boldsymbol{\xi}_q$ vector ($l_{\boldsymbol{\Lambda}}$)						
$l_{\boldsymbol{\Lambda}1}$	1.100	1.099	0.099 %	0.011	0.011	0.330 %
$l_{\boldsymbol{\Lambda}2}$	1.000	1.004	0.376 %	0.008	0.009	9.005 %
$l_{\boldsymbol{\Lambda}3}$	0.600	0.605	0.870 %	0.013	0.010	20.751 %
$l_{\boldsymbol{\Lambda}4}$	0.800	0.796	0.444 %	0.006	0.007	15.982 %
$l_{\boldsymbol{\Lambda}5}$	1.000	1.002	0.173 %	0.011	0.011	4.323 %
$l_{\boldsymbol{\Lambda}6}$	1.100	1.199	8.970 %	0.025	0.030	23.001 %
Satiation parameters ($\boldsymbol{\gamma}$)						
γ_1	1.000	1.012	1.222 %	0.028	0.022	19.877 %
γ_2	1.000	1.010	1.049 %	0.032	0.030	6.650 %
γ_3	1.000	1.021	2.071 %	0.028	0.032	12.803 %
γ_4	1.000	1.016	1.630 %	0.034	0.026	23.893 %
γ_5	1.000	1.018	1.822 %	0.029	0.026	9.422 %
γ_6	1.000	1.013	1.298 %	0.028	0.028	0.117 %
γ_7	1.000	1.018	1.786 %	0.036	0.028	23.741 %
γ_8	1.000	1.015	1.520 %	0.023	0.027	17.473 %
γ_9	1.000	1.016	1.572 %	0.029	0.026	11.032 %
γ_{10}	1.000	1.020	1.974 %	0.032	0.031	2.225 %
Overall mean value across parameters		0.690	1.403 %	0.025	0.026	13.690

Table 2: Recreational Travel Destination Choice and Number of Trips

Destination Zone	Total number (%) of individuals visiting each destination	Number of trips among those who visit each destination			
		Mean	Min.	Max.	Std. Dev.
South-East Lower Peninsula (SELP)	353 (21.3%)	3.05	1	50	4.89
South-West Lower Peninsula (SWLP)	253 (15.3%)	2.86	1	45	4.39
North-East Lower Peninsula (NELP)	366 (22.1%)	4.16	1	40	5.78
North-West Lower Peninsula (NWLP)	445 (26.8%)	4.17	1	60	6.67
East Upper Peninsula (EUP)	337 (20.3%)	2.73	1	40	3.94
West Upper Peninsula (WUP)	158 (9.5%)	3.29	1	32	5.10

Table 3: Destination Zone Characteristics

	South-East Lower Peninsula (SELP)	South-West Lower Peninsula (SWLP)	North-East Lower Peninsula (NELP)	North-West Lower Peninsula (NWLP)	East Upper Peninsula (EUP)	West Upper Peninsula (WUP)
Level of Service Variables (Std. Dev.)						
Travel Time (hours)	2.7 (2.1)	2.9 (2.2)	3.2 (1.2)	3.3 (1.4)	5.1 (1.4)	7.2 (2.3)
Travel Distance (miles)	153.5 (120.0)	162.2 (118.4)	185.7 (65.7)	184.2 (74.1)	290.4 (89.7)	396.1 (134.9)
Cost ($)	97.4 (78.6)	105.7 (82.3)	123.4 (60.6)	124.9 (66.6)	194.9 (94.6)	272.4 (139.7)
Land cover percentage						
Urban	10.8	6.9	2.6	2.9	1.5	2.2
Water	1.3	1.7	3.6	3.8	2.8	3.0
Open Land	9.4	9.6	16.0	17.1	7.0	5.4
Wetland	6.0	5.9	8.0	3.7	19.5	5.1
Agricultural	49.3	47.2	7.4	13.7	3.7	2.8
Sparsely Vegetated	0.2	0.4	0.4	0.7	0.6	0.6
Forest	23.0	28.3	62.0	58.1	64.9	80.9

Table 4: MDCP Model Estimation Results

Variable	Mean		Standard Deviation	
	Estimate	t-stat	Estimate	t-stat
Travel Cost ($/10) and interactions				
Travel cost	-0.850	-8.39	0.706	3.74
Travel cost interacted with low income household (<$30,000 per year)	-0.398	-3.72	-	-
Land cover accessibility measure specific to				
Urban ($/6*10^4$)	0.384	8.59	0.108	2.79
Urban ($/6*10^4$) interacted with presence of children < 16 years	-0.027	-1.43	-	-
Water ($/6*10^4$)	0.764	4.07	-	-
Wetland ($/6*10^4$)	-0.305	-8.16	-	-
Open land ($/6*10^4$)	-0.316	-5.81	-	-
Log-Likelihood at Convergence	-3726.30			