

On Allowing a General Form for Unobserved Heterogeneity in the Multiple Discrete-Continuous Probit Model: Formulation and Application to Tourism Travel

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
Phone: 512-471-4535; Fax: 512-475-8744
Email: bhat@mail.utexas.edu

and

King Abdulaziz University, Jeddah 21589, Saudi Arabia

Sebastian Astroza

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

Aarti C. Bhat

The University of Texas at Austin
College of Natural Sciences and Liberal Arts
Email: aartibhat@utexas.edu

*corresponding author

First version: July 2015

Revised version: January 2016

ABSTRACT

This paper proposes a new econometric formulation and an associated estimation method for a finite discrete mixture of normals (FDMN) version of the multiple discrete-continuous probit (MDCP) model. To our knowledge, this is the first such formulation and application of an MDCP model in the econometric literature. Using the New Zealand Domestic Travel Survey data set, the model is applied to analyze individual-level decisions regarding recreational destination locations and the number of trips to each destination. The results provide insights into the demographic and other factors that influence individuals' preferences for different destinations, and show that the FDMN MDCP model is able to identify different segments of the sample, each one of them with different effects of the exogenous variables on destination choice.

Keywords: multiple discrete-continuous models, finite discrete mixture of normals, MACML approach, endogenous segmentation, recreational trips, long distance trips, tourism travel.

1. INTRODUCTION

There are several approaches to understanding the decision process when consumers have to choose an alternative from a set and then determine the amount of the chosen alternative to consume. Classical discrete choice models assume that alternatives are mutually exclusive and only one alternative can be chosen. Alternatively, multiple discrete-continuous (MDC) models expand the decision by allowing consumers to choose multiple alternatives at the same time, along with the continuous dimension of the amount of consumption. MDC models have been applied not only in the case of consumer brand choice and purchase quantity, but also in contexts such as household vehicle type and usage, recreational destination choice and number of trips, land-use type and intensity, and stock portfolio selection choice and investment amounts.

The MDC model that has dominated the recent literature is based on a utility maximization framework that assumes a non-linear (but increasing and continuously differentiable) utility function to accommodate the relationship between the decreasing marginal utility (satiation) and the increasing investment in an alternative. Consumers maximize this utility within their budget constraints. The optimal consumption quantities are obtained by writing the Karush-Kuhn-Tucker (KKT) first-order conditions of the utility function with respect to the investment quantities. A very general utility form for this KKT approach was proposed by Bhat (2008). In Bhat's utility function form, stochasticity is introduced in the baseline preference for each alternative to acknowledge the presence of unobserved 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). The most common distributions used for the kernel stochastic error term (across alternatives) are the generalized extreme value (GEV) distribution (see Bhat, 2008; Pinjari, 2011; Castro *et al.*, 2012) and the multivariate normal distribution (see Kim *et al.*, 2002 and Bhat *et al.*, 2013). The first distribution leads to a closed-form MDC GEV model structure, while the second leads to an MDC probit (MDCP) model structure.

Researchers have also introduced random structures for the coefficients on the exogenous variables (or response coefficients) that allow heterogeneity (across individuals) in the sensitivity to exogenous variables in discrete choice models. There are three possible approaches to introduce randomness in the response coefficients. The first approach uses continuous random structures for the coefficients on the exogenous variables. Within this approach, the most common assumption is that the random response coefficients are realizations from a multivariate

normal distribution.¹ But this can lead to a misspecification if some other non-normal distribution characterizes the taste heterogeneity for one or more coefficients (see Train, 1998; Amador *et al.*, 2005; Train and Sonnier, 2005; Hensher *et al.*, 2005; Fosgerau, 2005; Greene *et al.*, 2006; Balcombe *et al.*, 2009; and Torres *et al.*, 2011). The second approach uses a discrete distribution for the response coefficients. This approach leads to the familiar latent class model with an endogenous segmentation that allocates individuals probabilistically to segments as a function of exogenous variables (see Bhat, 1997; Greene and Hensher, 2003; Train, 2008; Bastin *et al.*, 2010; Cherchi *et al.*, 2009; and Sobhani *et al.*, 2013). The problem with this approach, however, is that homogeneity in response is assumed within each latent class. The third approach uses a hybrid semi-parametric approach that combines a continuous response surface for the coefficients with a latent class approach (see, for example, Campbell *et al.*, 2010; Bujosa *et al.*, 2010; Greene and Hensher, 2013; and Xiong and Mannering, 2013). In this approach, the response coefficients are typically assumed to be realizations of a discrete mixture of multivariate normal distributions. That is, the relationship between the propensity variable and exogenous variables is assumed to belong to one of several latent (discrete) classes. Within each of these classes, the coefficients are drawn from a continuous multivariate normal distribution. The resulting finite discrete mixture of normal (FDMN) model generalizes the heterogeneity form because the normally distributed random parameters approach and the latent class approach consist of special cases—the first approach resulting when there is only one latent class and the

¹ To put things in context within the broader literature on accommodating non-normal coefficients, note that the second latent segmentation approach is equivalent to a non-parametric approach in which all random coefficients are assumed to have the same number of nodal points, with the number of nodal points being equal to the number of latent segments. The nodal points correspond to the segment-specific values (for each coefficient) in the latent segmentation set-up, and the probability masses at these nodal points for each individual correspond to the segment membership probabilities for that individual. This latent segmentation set-up is a restrictive version of a more general non-parametric specification in which the number of nodal points is allowed to vary across coefficients and both the nodal points and probability masses are separately estimated for each coefficient (see, for example, Bastin *et al.*, 2010). However, this general non-parametric approach is seldom used because consistency is achieved only in very large samples and parameter estimates generally have high variance (Mittelhammer and Judge, 2011). On the other hand, a continuous distribution offers substantial efficiency in the number of mixing parameters. In this regard, the paper by Bhat and Sidharthan (2012) is of particular note because it enables a non-normal (skew) continuous distribution to be used. However, their approach still enforces a unimodal distribution for the coefficients. The finite discrete mixture of normals (FDMN) is a good hybrid semi-parametric approach that combines the flexibility of the discrete mixture distribution with the efficiency advantage of the continuous distributions. It is by far the most widely used semi-parametric approach in the statistical and econometric literature because of this good balance between flexibility and efficiency (see Geweke and Keane, 1999, Frühwirth-Schnatter, 2011, and Ferdous *et al.*, 2011).

second resulting when the multivariate normal distribution becomes degenerate within each latent class.

Several earlier studies have included heterogeneity in the sensitivity to exogenous variables in the MDC context. Bhat *et al.* (2013) proposed an estimation approach for the MDCP model that allows taste variation through the inclusion of random parameters. They demonstrated the ability to recover the parameters based on a simulation experiment, using both cross-sectional and panel data, and applied the model to analyze recreational long-distance travel. On the same topic of recreational travel, Kuriyama *et al.* (2010) proposed a latent class KKT model based on the linear expenditure system with translated constant elasticity of substitution utility functions proposed by Hanemman (1978). Sobhani *et al.* (2013) and Wafa *et al.* (2015) use a latent class approach with the MDCEV kernel structure. In Sobhani *et al.* (2013), the authors propose an estimation approach combining the full information maximum likelihood and the expectation maximization approaches. The latent class MDCEV model is applied to study non-workers' daily decisions regarding vehicle type and usage in conjunction with activity type and accompaniment choice decisions. Wafa *et al.* (2015) proposed a latent class MDCEV model to study the spatial transferability of activity-travel models.

In this paper, we propose an FDMN version of the MDCP model. To our knowledge, this is the first such formulation and application of an MDCP model in the econometric literature. We also propose the use of Bhat's (2011) maximum approximate composite marginal likelihood (MACML) inference approach for the estimation. This approach is computationally efficient and does not involve quasi-Monte Carlo simulation techniques of the type proposed in Bhat (2000) and Bhat (2001). The advantage of the MACML approach relative to simulation techniques is that it involves only univariate and bivariate cumulative normal distribution function evaluations in the likelihood function, regardless of the number of alternatives or segments in the latent classification. Using a 2012 New Zealand Domestic Travel Survey data set, the model is applied to analyze individual-level decisions regarding recreational destination locations and the number of trips to each destination. The results provide insights into the demographic and other factors that influence individuals' preferences for different recreational destinations, and show that the FDMN MDCP model is able to identify different discrete segments of the sample, each one of them with different stochastic effects of the exogenous variables on destination choice (and the effects varying across the discrete segments).

2. METHODOLOGY

2.1 Model Formulation

Following Bhat (2008), consider a choice scenario where a consumer q ($q = 1, 2, \dots, Q$) belonging to a segment g ($g = 1, 2, \dots, G$) maximizes his/her utility subject to a binding constraint, as shown in Equation (1):

$$\begin{aligned} \max U_q(\mathbf{x}_q) | \text{segment } g &= \sum_{k=1}^K \frac{\gamma_{qgk}}{\alpha_{qgk}} \psi_{qgk} \left(\left(\frac{x_{qk}}{\gamma_{qgk}} + 1 \right)^{\alpha_{qgk}} - 1 \right) \\ \text{s.t. } \sum_{k=1}^K p_{qk} x_{qk} &= E_q, \end{aligned} \quad (1)$$

where the utility function $U_q(\mathbf{x}_q)$, given that consumer q belongs to segment g , is quasi-concave, increasing and continuously differentiable; \mathbf{x}_q is the consumption quantity (vector of dimension $K \times 1$ with elements x_{qk} so that $x_{qk} \geq 0$ for all k ; k is an index for good k), and γ_{qgk} , α_{qgk} , and ψ_{qgk} are parameters associated with good k and consumer q , given that consumer q belongs to segment g .² In the 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 . Assume, for now, that there is no essential outside good, so that corner solutions (zero consumptions) are possible for all goods k (relaxing this assumption is straightforward and simplifies the analysis considerably). The parameter ψ_{qgk} represents the baseline marginal utility for good k , given that the individual q belongs to population segment g (*i.e.*, ψ_{qgk} is the marginal utility of good k at the point of no consumption of good k , given that q belongs to segment g). The parameter γ_{qgk} allows a corner solution for good k and also serves as a translation-based satiation parameter, while α_{qgk} serves as an exponential-based satiation parameter. As discussed in detail in Bhat (2008), only one parameter of the set γ_{qgk} or α_{qgk} will be empirically identified, so the analyst will have to estimate a γ -profile (in which $\alpha_{qgk} \rightarrow 0$) or an α -profile (in which the γ_{qgk} terms are normalized to the value of one). Both these profiles can be estimated, and the one that provides a

² Though we will refer to the alternatives for consumption as “goods” in this section, it is important to note that the consumption alternatives can also refer to consumption of different types of activities or, as in the empirical analysis of the current paper, to “consumption” of different destination regions for leisure trips.

better data fit may be selected. Also, for the γ -profile, we will need $\gamma_{qgk} > 0 \forall k$, and, for the α -profile, we will need $\alpha_{qgk} \leq 1 \forall k$. In the current paper, we will retain the general utility form of Equation (1) to keep the presentation general.

In Equation (2) we introduce observed heterogeneity across individuals within segment g and stochasticity through the baseline marginal utility function ψ_{qgk} :

$$\psi_{qgk} = \exp(\boldsymbol{\beta}'_{qg} \mathbf{z}_{qk}), \quad (2)$$

where \mathbf{z}_{qk} is a D -dimensional vector of attributes that characterizes good k and the consumer q (including a constant for each good except one, to capture intrinsic preferences for each good relative to a base good); $\boldsymbol{\beta}_{qg}$ is a consumer-specific vector of coefficients (of dimension $D \times 1$) that allows unobserved taste variation across all consumers q in segment g and allows different observed responsiveness across all consumers q based on different values of the elements of the vector \mathbf{z}_{qk} . In the current paper, we consider $\boldsymbol{\beta}_{qg}$ to be a realization from a multivariate normal distribution: $\boldsymbol{\beta}_{qg} \sim f_D(\mathbf{b}_g, \boldsymbol{\Omega}_g)$. For future reference, we also write $\boldsymbol{\beta}_{qg} = \mathbf{b}_g + \tilde{\boldsymbol{\beta}}_{qg}$, where $\tilde{\boldsymbol{\beta}}_{qg} \sim f_D(\mathbf{0}_D, \boldsymbol{\Omega}_g)$.

The optimal consumption vector \mathbf{x}_q can be solved based on the constrained optimization problem of Equation (1) by forming the Lagrangian function and applying the KKT conditions, conditional on the individual belonging to segment g . The Lagrangian function for the problem is provided in Equation (3):

$$\ell_q \mid q \in \text{segment } g = \sum_{k=1}^K \frac{\gamma_{qgk}}{\alpha_{qgk}} \exp(\mathbf{b}'_g \mathbf{z}_{qk} + \tilde{\boldsymbol{\beta}}'_{qg} \mathbf{z}_{qk}) \left\{ \left(\frac{x_{qk}}{\gamma_{qgk}} + 1 \right)^{\alpha_{qgk}} - 1 \right\} - \lambda_{qg} \left[\sum_{k=1}^K p_{qk} x_{qk} - E_q \right], \quad (3)$$

where λ_{qg} is a segment g -specific Lagrangian multiplier associated with the expenditure constraint. The KKT first-order conditions for the optimal consumption x_{qk}^* , given that consumer q belongs to segment g , are as shown in Equation (4):

$$\begin{aligned} \exp(\mathbf{b}'_g \mathbf{z}_{qk} + \tilde{\boldsymbol{\beta}}'_{qg} \mathbf{z}_{qk}) \left(\frac{x_{qk}^*}{\gamma_{qgk}} + 1 \right)^{\alpha_{qgk}-1} - \lambda_{qg} p_{qk} &= 0, \text{ if } x_{qk}^* > 0, \quad k = 1, 2, \dots, K \\ \exp(\mathbf{b}'_g \mathbf{z}_{qk} + \tilde{\boldsymbol{\beta}}'_{qg} \mathbf{z}_{qk}) \left(\frac{x_{qk}^*}{\gamma_{qgk}} + 1 \right)^{\alpha_{qgk}-1} - \lambda_{qg} p_{qk} &< 0, \text{ if } x_{qk}^* = 0, \quad k = 1, 2, \dots, K. \end{aligned} \quad (4)$$

The optimal demand, conditional on individual q belonging to segment g , satisfies the above conditions and 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. To accommodate this singularity, let m_q be, without loss of generality, the consumed good with the lowest value of k for the q^{th} consumer (note that the consumer must consume at least one good given $E_q > 0$). For this m_q^{th} good, $x_{qm_q}^* > 0$, which implies Equation (5) from the first set of KKT conditions in Equation (4):

$$\lambda_{qg} = \frac{\exp(\mathbf{b}'_g \mathbf{z}_{qm_q} + \tilde{\boldsymbol{\beta}}'_{qg} \mathbf{z}_{qm_q})}{p_{qm_q}} \left(\frac{x_{qm_q}^*}{\gamma_{qgm_q}} + 1 \right)^{\alpha_{qgm_q} - 1}. \quad (5)$$

Substituting back in Equation (4) for the other goods k ($k=1,2,\dots,K; k \neq m_q$), and taking logarithms and simplifying, we may write the KKT conditions as Equation (6):

$$\begin{aligned} y_{qgkm_q}^* &= 0, \text{ if } x_{qk}^* > 0, k=1,2,\dots,K, k \neq m_q, \\ y_{qgkm_q}^* &< 0, \text{ if } x_{qk}^* = 0, k=1,2,\dots,K, k \neq m_q. \end{aligned} \quad (6)$$

where $y_{qgkm_q}^* = y_{qgk} - y_{qgm_q}$; $y_{qgk}^* = V_{qgk} + \tilde{\boldsymbol{\beta}}'_{qg} \mathbf{z}_{qk}$; and $V_{qgk} = \mathbf{b}'_g \mathbf{z}_{qk} + (\alpha_{qgk} - 1) \ln \left(\frac{x_{qk}^*}{\gamma_{qgk}} + 1 \right) - \ln p_{qgk}$.

The above conditions are conditional on individual q belonging to segment g . Within this context, two important identification considerations need to be noted (additional identifications considerations due to multiple segments will be noted later). First, a dummy variable (or constant) corresponding to one of the K goods should not be introduced, since only differences in the y_{qgk}^* terms matter (this is similar to a standard discrete choice model). Similarly, consumer-specific variables that do not vary across goods can be introduced only for $(K-1)$ goods, with the remaining alternative being the base. Let the first alternative be the base for the dummy variable and for consumer-specific variables that do not vary across goods. That is, let $z_{q1constant} = 0$ (and correspondingly, the element in \mathbf{b}_g corresponding to this first alternative's constant is fixed at 0 and the variance element contribution in $\boldsymbol{\Omega}_g$ corresponding to this alternative's constant is also fixed at 0; in addition, all covariance elements in $\boldsymbol{\Omega}_g$ corresponding to this first alternative's constant also are set to zero). Also, let $z_{q1l} = 0$ for all consumer-specific variables l that do not

vary across goods (and correspondingly, the elements in \mathbf{b}_g for these variables for the first alternative are fixed at zero and so are all variances/covariances in $\mathbf{\Omega}_g$ for these variables for the first alternative).

2.2 Consumer Role in a Finite Mixture of Segments

The derivation thus far is based on the notion that consumer q belongs to a single segment g . But now consider the case that consumer q belongs to a finite mixture of segments—that is, the actual assignment of consumer q to a specific segment is not observed, but we are able to attribute different probabilities π_{qg} ($g=1,2,\dots,G$) to consumer q belonging to different

segments. We require that $0 \leq \pi_{qg} \leq 1$, and $\sum_{g=1}^G \pi_{qg} = 1$. To enforce these restrictions, and

following Bhat (1997), we use the logit link function shown in Equation (7):

$$\pi_{qg} = \frac{\exp(\boldsymbol{\mu}'_g \mathbf{w}_q)}{\sum_{h=1}^G \exp(\boldsymbol{\mu}'_h \mathbf{w}_q)}, \quad (7)$$

where \mathbf{w}_q is a vector of individual exogenous variables, and $\boldsymbol{\mu}_1 = \mathbf{0}$ serves as a vector identification condition. This probabilistic assignment to segments is tantamount to using a

mixture of multivariate normal distributions for $\boldsymbol{\beta}_q$: $\lambda(\boldsymbol{\beta}_q = \mathbf{a}) | \mathbf{b}, \mathbf{\Omega} = \sum_{g=1}^G \pi_{qg} f_D(\mathbf{a}; \mathbf{b}_g, \mathbf{\Omega}_g)$,

where $f_D(\cdot; \mathbf{b}_g, \mathbf{\Omega}_g)$ is the multivariate normal density function with mean vector \mathbf{b}_g and covariance matrix $\mathbf{\Omega}_g$. \mathbf{b} is a vector obtained by stacking the \mathbf{b}_g vectors vertically, and $\mathbf{\Omega}$ is the

matrix obtained by block-diagonally stacking the $\mathbf{\Omega}_g$ matrices. Specifically, one may write

$\psi_{qk} = \exp(\boldsymbol{\beta}'_q \mathbf{z}_{qk})$, which, with the mixture of MVN distributions as above for $\boldsymbol{\beta}_q$, leads to the

segment-specific baseline utility functions of the form of Equation (2) with a probabilistic segment assignment π_{qg} . The mixture of normal distributions is a semi-parametric distribution

that relaxes the normal distribution for $\boldsymbol{\beta}_q$ commonly used in typical MDC models, while

allowing the distribution itself to be a function of individual-level attributes through the π_{qg}

terms. The mixture distribution effectively combines the flexibility of the latent class model with

the parsimony of the continuous multivariate normal distribution assumption for $\boldsymbol{\beta}_q$. In

particular, if each individual belongs to a single segment that is known *a priori* (that is, if $\pi_{qg} = 1$ for a specific segment g and zero for all other segments, and if this is known a priori for each individual q) and $\gamma_{qgk} = \gamma_{qk} \forall g$ and $\alpha_{qgk} = \alpha_{qk} \forall g$, the model collapses to a random-coefficient MDCP model (or RC-MDCP in the rest of this paper) as in Bhat *et al.* (2013). On the other hand, if the multivariate normal distribution within each segment becomes degenerate (*i.e.*, $\Omega_g = 0$ for all g), then the model collapses to a latent class MDCP (LC-MDCP) model.

The use of latent classes, as in the current paper, requires labeling restrictions for identifiability. In particular, the parameter space includes $G!$ subspaces, each associated with a different way of labeling the mixture components. To prevent the interchange of the mixture components, we impose the labeling restriction that the constants specific to the second alternative are increasing across the segments, *i.e.*: $b_{11} < b_{21} < b_{31} < \dots < b_{G1}$ (b_{11} refers to coefficient on the dummy variable for the second alternative in the first segment, b_{21} refers to the coefficient on the dummy variable for the second alternative in the second segment, and so on until b_{G1} refers to the coefficient on the dummy variable for the second alternative in the G th segment).³ To implement the labeling restriction, we parameterize the b_{g1} values as follows: $b_{g1} = b_{g-1,1} + \exp(\kappa_g)$ for $g=2, \dots, G$. Such a labeling restriction is needed because the same model specification (and likelihood function value) results simply by interchanging the sequence in which the segments are numbered. Technically, therefore, multiple sets of parameters (corresponding to a swap of segment values) result in the same likelihood function, creating an identification problem. This identification problem is resolved through the imposition of the labeling restriction above so that the segments become non-interchangeable.⁴ Finally, an additional scale normalization needs to be imposed on Ω_g for one of the g segments 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 variance of the second alternative's constant in the first segment

³ As clearly indicated earlier, the constant coefficients for the first alternative are set to zero in every segment g ($g = 1, 2, \dots, G$) for identification.

⁴ Of course, the labeling restriction discussed above, which we use in the simulation experiments in the current paper, is only one of several possible restrictions to identify the model. Thus, in our empirical analysis, where we do not use constants in the baseline utilities of the alternatives (for reasons discussed in the empirical section), we implement another version of the labeling restriction by requiring that the constant in the first segment in the segment membership model has a maintained value of zero, and there is a descending order of magnitudes of the constants for the other segments in the segment membership model.

($g=1$) to the value of one. But, if there is price variation across even a subset of goods for a subset of consumers, there is no need for this additional scale normalization (see Bhat, 2008).

2.3 Model Estimation

If a γ -profile is used, the parameter γ_{qgk} may be parameterized as $\exp(\tilde{\theta}'_g \mathbf{a}_{qk})$, where \mathbf{a}_{qk} is a vector of explanatory variables and $\tilde{\theta}'_g$ is a corresponding vector of parameters. On the other hand, if an α -profile is used, the parameter α_{qgk} may be parameterized as $1 - \exp(\tilde{\theta}'_g \tilde{\mathbf{a}}_{qk})$ (to maintain the restriction that $\alpha_{qgk} < 1$) or as $\frac{1}{1 + \exp(-\tilde{\theta}'_g \tilde{\mathbf{a}}_{qk})}$ (to maintain the stronger restrictions that $0 < \alpha_{qgk} < 1$; this stronger restriction often helps create stability in estimation).

Let $\theta = (\mathbf{b}'_1, \mathbf{b}'_2, \dots, \mathbf{b}'_G; \bar{\Omega}'_1, \bar{\Omega}'_2, \dots, \bar{\Omega}'_G; \boldsymbol{\mu}'_1, \boldsymbol{\mu}'_2, \dots, \boldsymbol{\mu}'_G; \tilde{\theta}'_1, \tilde{\theta}'_2, \dots, \tilde{\theta}'_G)'$ if a γ -profile is estimated and $\theta = (\mathbf{b}'_1, \mathbf{b}'_2, \dots, \mathbf{b}'_G; \bar{\Omega}'_1, \bar{\Omega}'_2, \dots, \bar{\Omega}'_G; \boldsymbol{\mu}'_1, \boldsymbol{\mu}'_2, \dots, \boldsymbol{\mu}'_G; \tilde{\theta}'_1, \tilde{\theta}'_2, \dots, \tilde{\theta}'_G)'$ if an α -profile is estimated, with $\bar{\Omega}'_g$ representing the row vectorization of the upper diagonal elements of Ω_g . To formulate the estimation procedure, we will use the following notation: $f_S(\cdot; \boldsymbol{\eta}, \boldsymbol{\Sigma})$ for the multivariate normal density function of dimension S with mean vector $\boldsymbol{\eta}$ and covariance matrix $\boldsymbol{\Sigma}$; $\boldsymbol{\omega}_{\boldsymbol{\Sigma}}$ for the diagonal matrix of standard deviations of $\boldsymbol{\Sigma}$ (with its r^{th} element being $\omega_{\boldsymbol{\Sigma}, r}$); $\phi_S(\cdot; \boldsymbol{\Sigma}^*)$ for the multivariate standard normal density function of dimension S and correlation matrix $\boldsymbol{\Sigma}^*$ —such that $\boldsymbol{\Sigma}^* = \boldsymbol{\omega}_{\boldsymbol{\Sigma}}^{-1} \boldsymbol{\Sigma} \boldsymbol{\omega}_{\boldsymbol{\Sigma}}^{-1}$, $F_S(\cdot; \boldsymbol{\eta}, \boldsymbol{\Sigma})$ for the multivariate normal cumulative distribution function of dimension S with mean vector $\boldsymbol{\eta}$ and covariance matrix $\boldsymbol{\Sigma}$ —and $\Phi_S(\cdot; \boldsymbol{\Sigma}^*)$ for the multivariate standard normal cumulative distribution function of dimension S 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_{qgk} and V_{qgk} into $K \times 1$ vectors $\mathbf{y}_{qg} = (y_{qg1}, y_{qg2}, \dots, y_{qgK})'$ and $\mathbf{V}_{qg} = (V_{qg1}, V_{qg2}, \dots, V_{qgK})'$ respectively, and let $\mathbf{z}_q = (\mathbf{z}_{q1}, \mathbf{z}_{q2}, \dots, \mathbf{z}_{qK})'$ be a $K \times D$ matrix of variable attributes. Then, we may write, in matrix notation, $\mathbf{y}_{qg} = \mathbf{V}_{qg} + \mathbf{z}_q \tilde{\boldsymbol{\beta}}_{qg}$ and $\mathbf{y}_{qg}^* = \mathbf{M}_q \mathbf{y}_{qg} \sim MVN_{K-1}(\mathbf{H}_{qg}, \boldsymbol{\Psi}_{qg})$, where $\mathbf{H}_{qg} = \mathbf{M}_q \mathbf{V}_{qg}$ and

$\Psi_{qg} = \mathbf{M}_q \mathbf{z}_q \Omega_g \mathbf{z}'_q \mathbf{M}'_q$. Next, partition the vector \mathbf{y}_{qg}^* into a sub-vector $\tilde{\mathbf{y}}_{qg,NC}^*$ of length $L_{q,NC} \times 1$ ($0 \leq L_{q,NC} \leq K-1$) for the non-consumed goods, and another sub-vector $\tilde{\mathbf{y}}_{qg,C}^*$ of length $L_{q,C} \times 1$ ($0 \leq L_{q,C} \leq K-1$) for the consumed goods ($L_{q,NC} + L_{q,C} = K-1$). Let $\tilde{\mathbf{y}}_{qg}^* = \left(\left[\tilde{\mathbf{y}}_{qg,NC}^* \right], \left[\tilde{\mathbf{y}}_{qg,C}^* \right] \right)'$, which may be obtained from \mathbf{y}_{qg}^* as $\tilde{\mathbf{y}}_{qg}^* = \mathbf{R}_q \mathbf{y}_{qg}^*$, where \mathbf{R}_q is a re-arrangement matrix of dimension $(K-1) \times (K-1)$ 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 provided in Equation (8):

$$\mathbf{R}_q = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{q,NC} \\ \mathbf{R}_{q,C} \end{bmatrix}, \quad (8)$$

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 $\tilde{\mathbf{y}}_{qg,NC}^* = \mathbf{R}_{q,NC} \mathbf{y}_{qg}^*$ and $\tilde{\mathbf{y}}_{qg,C}^* = \mathbf{R}_{q,C} \mathbf{y}_{qg}^*$. $\mathbf{R}_{q,NC}$ 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 $\tilde{\mathbf{H}}_{qg} = \mathbf{R}_q \mathbf{H}_{qg}$, $\tilde{\mathbf{H}}_{qg,NC} = \mathbf{R}_{q,NC} \mathbf{H}_{qg}$, $\tilde{\mathbf{H}}_{qg,C} = \mathbf{R}_{q,C} \mathbf{H}_{qg}$, and $\tilde{\Psi}_{qg} = \mathbf{R}_q \Psi_{qg} \mathbf{R}'_q = \begin{bmatrix} \tilde{\Psi}_{qg,NC} & \tilde{\Psi}'_{qg,NC,C} \\ \tilde{\Psi}_{qg,NC,C} & \tilde{\Psi}_{qg,C} \end{bmatrix}$, where $\tilde{\Psi}_{qg,NC} = \mathbf{R}_{q,NC} \Psi_{qg} \mathbf{R}'_{q,NC}$, $\tilde{\Psi}_{qg,C} = \mathbf{R}_{q,C} \Psi_{qg} \mathbf{R}'_{q,C}$, and $\tilde{\Psi}_{qg,NC,C} = \mathbf{R}_{q,NC} \Psi_{qg} \mathbf{R}'_{q,C}$. Then, the

likelihood function corresponding to the consumption quantity vector \mathbf{x}_q^* for consumer q may be obtained from the KKT conditions in Equation (6), provided as Equation (9):

$$L_{qg} = \det(\mathbf{J}_{qg}) \int_{\mathbf{h}_{q,NC}=-\infty}^0 f_{K-1}(\mathbf{h}_{q,NC}, \mathbf{0}_{L_{q,C}} \mid \tilde{\mathbf{H}}_{qg}, \tilde{\Psi}_{qg}) d\mathbf{h}_{q,NC}, \quad (9)$$

where $\det(\mathbf{J}_{qg})$ is the determinant of the Jacobian of the transformation from \mathbf{y}_{qg}^* to the consumption quantities \mathbf{x}_q^* (see Bhat, 2008), as Equation (10) indicates:

$$\det(\mathbf{J}_{qg}) = \left\{ \prod_{k \in \mathcal{T}_q} \frac{1 - \alpha_{qgk}}{\mathbf{x}_{qk}^* + \gamma_{qgk}} \right\} \left\{ \sum_{k \in \mathcal{T}_q} \left(\frac{\mathbf{x}_{qk}^* + \gamma_{qgk}}{1 - \alpha_{qgk}} \right) \left(\frac{P_{qk}}{P_{qm_q}} \right) \right\}, \quad (10)$$

where \mathcal{T}_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 shown in Equation (11):

$$\begin{aligned} L_{qg} &= \det(\mathbf{J}_{qg}) \times f_{L_{q,C}}(\mathbf{0}_{L_{q,C}}; \tilde{\mathbf{H}}_{qg,C}, \tilde{\Psi}_{qg,C}) \times F_{L_{q,NC}}(\mathbf{0}_{L_{q,NC}}; \tilde{\mathbf{H}}_{qg,NC}, \tilde{\Psi}_{qg,NC}) \\ &= \det(\mathbf{J}_{qg}) \times \left(\prod_{s=1}^{L_{q,C}} \omega_{\tilde{\Psi}_{qg,C},s} \right)^{-1} \left(\phi_{L_{q,C}}(\omega_{\tilde{\Psi}_{qg,C}}^{-1}(-\tilde{\mathbf{H}}_{qg,C}), \tilde{\Psi}_{qg,C}^*) \right) \times \Phi_{L_{q,NC}}(\omega_{\tilde{\Psi}_{qg,NC}}^{-1}(-\tilde{\mathbf{H}}_{qg,NC}), \tilde{\Psi}_{qg,NC}^*), \end{aligned} \quad (11)$$

where $\tilde{\mathbf{H}}_{qg,NC} = \tilde{\mathbf{H}}_{qg,NC} + \tilde{\Psi}_{qg,NC,C}(\tilde{\Psi}_{qg,C})^{-1}(-\tilde{\mathbf{H}}_{qg,C})$, $\tilde{\Psi}_{qg,C}^* = \omega_{\tilde{\Psi}_{qg,C}}^{-1} \tilde{\Psi}_{qg,C} \omega_{\tilde{\Psi}_{qg,C}}^{-1}$,

$\tilde{\Psi}_{qg,NC} = \tilde{\Psi}_{qg,NC} - \tilde{\Psi}_{qg,NC,C}(\tilde{\Psi}_{qg,C})^{-1} \tilde{\Psi}_{qg,NC,C}'$, $\tilde{\Psi}_{qg,NC}^* = \omega_{\tilde{\Psi}_{qg,NC}}^{-1} \tilde{\Psi}_{qg,NC} \omega_{\tilde{\Psi}_{qg,NC}}^{-1}$, and ω_{Ξ} represents

the diagonal matrix of standard errors corresponding to matrix Ξ .

Then, the likelihood function for observation q is:

$$L_q(\boldsymbol{\theta}) = \sum_{g=1}^G \pi_{qg} L_{qg}(\boldsymbol{\theta}), \quad (12)$$

and the likelihood function is then given as:

$$L(\boldsymbol{\theta}) = \prod_q L_q(\boldsymbol{\theta}). \quad (13)$$

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 become inaccurate and time-consuming as K increases. An alternative is to use the MACML approach (Bhat, 2011),

in which the multiple integrals are evaluated using a fast analytic approximation method.⁵ 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 used to evaluate multidimensional integrals (see Bhat *et al.*, 2013 for an extended simulation analysis of the ability of the MACML method to recover parameters in the simple MDCP model).

One very important issue still needs to be dealt with: the positive definiteness of covariance matrices. The positive-definiteness of $\tilde{\Psi}_{qg}$ in the likelihood function can be ensured by applying a Cholesky decomposition to the matrices Ω_g ($g = 1, 2, \dots, G$), and estimating these Cholesky-decomposed parameters.⁶

3. SIMULATION EVALUATION

The simulation exercises undertaken in this section examine the ability of the MACML estimator to recover parameters from finite samples in an FDMN MDCP model by generating simulated data sets with known underlying model parameters. To examine the robustness of the MACML approach when applied to different numbers of mixtures, we consider both two- and three-

⁵ Note that in the current case, we use only the analytic approximation for the cumulative standard multivariate normal distribution embedded in the MACML; however, we will continue to refer to the approach as MACML for ease in presentation and also because the composite marginal likelihood (CML) inference approach subsumes the maximum likelihood (ML) inference approach used here as a special case.

⁶ Previous research in latent segmentation (see, for example, Bhat, 1997 and Sobhani *et al.*, 2013) has highlighted several estimation challenges in terms of stability and convergence. Most of these studies recommend the Expectation Maximization (EM) method to find good initial values to start the full information likelihood function iterations. However, the EM method also leads to long estimation times. In this paper, to obtain good start values as well as minimize estimation time, we implemented the following steps to estimate a model with S segments, with the parameters at the end of each step serving as the initial start values for the iterations associated with the subsequent step. In particular, we first estimated an MDCP model with only a single vector of constants in the baseline utilities of the alternatives (that is, we estimated a constants-only model as though there were only one latent segment). Second, we used the constants from the first step as initial values for the baseline utility constants in the first segment, made a random perturbation of these values by increasing or decreasing these values between 5% to 10% for the baseline utility constants in the remaining $S-1$ segments, used a constants-only specification for the segment membership probabilities with the constant for the first segment constrained to zero, and the other constants perturbed from zero in a way that adheres to the labeling restriction as discussed at the top of Section 4.3, and estimated a latent segmentation MDCP model (with only constants and no random parameters). Third, we used the results of the previous step as a starting point and added exogenous variables to the segment membership model to get a good segment membership specification. Fourth, we introduced exogenous variables in the baseline utilities of the MDCP-specific models for each segment to obtain a good latent segmentation model simultaneously with a refined segment membership model specification. Finally, we estimated the FDMN version of the model, allowing randomness in the parameters. The entire process was also aided and speeded up by the fact that we coded our own analytic gradient function.

mixture models. In addition, we examine the effects of (a) assuming that coefficients are fixed and not stochastic within each segment (that is, using the LC-MDCP model), and (b) assuming normality of the response coefficient when non-normality is present and thus using a single segment when multiple segments are present (that is, using the RC-MDCP model).

3.1 Experimental Design

In the design, we consider the case with three alternatives. In each of the two- and three-mixture cases, we consider two independent variables in the z_{qk} vector in the baseline utility for each alternative. That is, consider the following for the z_{qk} vectors:

$$z_{q1} = [0, 0, z_{q1,1}, z_{q1,2}], \quad z_{q2} = [1, 0, z_{q2,1}, z_{q2,2}], \quad \text{and} \quad z_{q3} = [0, 1, z_{q3,1}, z_{q3,2}], \quad (14)$$

where the last two variables in each z_{qk} ($k=1,2,3$) correspond to the two independent variables.

The first variable in z_{q2} is the constant specific to alternative 2, while the second variable in z_{q3} is the constant specific to alternative 3. The values of the two independent variables for each alternative (*i.e.*, $z_{q1,1}$ and $z_{q1,2}$ for the first alternative; $z_{q2,1}$ and $z_{q2,2}$ for the second alternative; and $z_{q3,1}$ and $z_{q3,2}$ for the third alternative) are drawn from standard univariate normal distributions. In particular, a synthetic sample of 5000 realizations of the exogenous variables is generated corresponding to $Q=5000$ consumers. Additionally, we generate budget amounts E_q ($q=1,2,\dots,Q$) from a univariate normal distribution with a mean of 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.

3.1.1 Two-Segment Case

For the coefficients on the z_{qk} variables, we assume hybrid coefficients as follows:

$$\lambda(\beta_q = \mathbf{a}) = \pi_1 \times f_4(\mathbf{a}; \mathbf{b}_1, \mathbf{\Omega}_1) + \pi_2 \times f_4(\mathbf{a}; \mathbf{b}_2, \mathbf{\Omega}_2), \quad (15)$$

where $\mathbf{b}_1 = (b_{11}, b_{12}, b_{13}, \tilde{b})' = (1.0, 2.0, 0.6, 0.5)'$ for segment 1, and $\mathbf{b}_2 = (b_{21}, b_{22}, b_{23}, \tilde{b})' = (2.0, 1.5, 0.2, 0.5)'$ for segment 2. Note that the dimension of b_1 and b_2 are the same as z_{q1} , z_{q2} , and z_{q3} (all of these are 4×1 vectors). That is, b_{11} is the mean constant

coefficient on the second alternative in segment 1, b_{12} is the mean constant coefficient on the third alternative in segment 1, b_{13} is the mean coefficient on the first independent variable in the first segment, and \tilde{b} is the mean coefficient on the second independent variable in the first segment. b_{21} through b_{23} are similar to b_{11} through b_{13} but for the second segment, and we maintain the same coefficient \tilde{b} in both segments for the second independent variable. For the covariance matrices $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$ of the coefficients we assume:

$$\mathbf{\Omega}_1 = \begin{bmatrix} 1 & 0.5 & 0.7 & 0 \\ 0.5 & 1 & 0.8 & 0 \\ 0.7 & 0.8 & 0.9 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathbf{L}_{\Omega_1} \mathbf{L}'_{\Omega_1} = \begin{bmatrix} 1 & 0.000 & 0.000 & 0 \\ 0.5 & 0.866 & 0.000 & 0 \\ 0.7 & 0.519 & 0.374 & 0 \\ 0.0 & 0.000 & 0.000 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0.500 & 0.700 & 0 \\ 0 & 0.866 & 0.519 & 0 \\ 0 & 0.000 & 0.374 & 0 \\ 0 & 0.000 & 0.000 & 0 \end{bmatrix}$$

$$\mathbf{\Omega}_2 = \begin{bmatrix} 0.81 & 0.54 & 0.72 & 0 \\ 0.54 & 1.0 & 0.8 & 0 \\ 0.72 & 0.8 & 0.89 & 0 \\ 0.00 & 0.00 & 0.00 & 0 \end{bmatrix} = \mathbf{L}_{\Omega_2} \mathbf{L}'_{\Omega_2} = \begin{bmatrix} 0.9 & 0.0 & 0.0 & 0 \\ 0.6 & 0.8 & 0.0 & 0 \\ 0.8 & 0.4 & 0.3 & 0 \\ 0.0 & 0.0 & 0.0 & 0 \end{bmatrix} \begin{bmatrix} 0.9 & 0.6 & 0.8 & 0 \\ 0.0 & 0.8 & 0.4 & 0 \\ 0.0 & 0.0 & 0.3 & 0 \\ 0.0 & 0.0 & 0.0 & 0 \end{bmatrix}$$

As indicated earlier, the positive definiteness of the $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$ matrices is ensured in the estimations by reparameterizing the likelihood function in terms of the lower Cholesky factor matrices \mathbf{L}_{Ω_1} and \mathbf{L}_{Ω_2} , and estimating the associated Cholesky matrix parameters. As should be obvious from the specification of $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$, we assume that the coefficient on the second independent variable (*i.e.*, \tilde{b}) is fixed in the simulations (note the zero entries in the last row and column of $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$).⁷ Then, in the two-mixture case, there are 11 Cholesky parameters to be estimated: $l_{\Omega_1,2} = 0.5$, $l_{\Omega_1,3} = 0.866$, $l_{\Omega_1,4} = 0.7$, $l_{\Omega_1,5} = 0.519$, $l_{\Omega_1,6} = 0.374$, $l_{\Omega_2,1} = 0.9$, $l_{\Omega_2,2} = 0.6$, $l_{\Omega_2,3} = 0.8$, $l_{\Omega_2,4} = 0.8$, $l_{\Omega_2,5} = 0.4$, and $l_{\Omega_2,6} = 0.3$.

The weight mixture values π_1 and π_2 are set by specifying the vector \mathbf{w}_q to include a constant and an independent variable w_{q1} drawn from a standard univariate normal distribution.

That is, $\mathbf{w}_q = [1, w_{q1}]'$. Also we specify $\boldsymbol{\mu}_1 = [0, 0]'$ for normalization and $\boldsymbol{\mu}_2 = [0.6, 0.1]'$ for the

⁷ We use the general presentation convention that the lower Cholesky matrix of a covariance matrix with a row/column with all zero values (that is, corresponding to a fixed parameter) is obtained by stripping out that row/column from the covariance matrix, obtaining the lower Cholesky matrix for the remaining sub-matrix, and then adding a row/column of zero values to the resulting lower Cholesky matrix.

second segment. Finally, we use a γ -profile in our estimations, and set the satiations parameters for all three alternatives to 1 in both segments. That is, $\gamma_{11} = \gamma_{12} = \gamma_{13} = 1$ for the first segment, and $\gamma_{21} = \gamma_{22} = \gamma_{23} = 1$ for the second segment.

Overall, the parameters to be estimated in the two-mixture case include the following: $b_{11}=1, b_{12}=2, b_{13}=0.6, b_{21}=2, b_{22}=1.5, b_{23}=0.2, \tilde{b}=0.5, l_{\Omega_1,2} = 0.5, l_{\Omega_1,3} = 0.866, l_{\Omega_1,4} = 0.7, l_{\Omega_1,5} = 0.519, l_{\Omega_1,6} = 0.374, l_{\Omega_2,1} = 0.9, l_{\Omega_2,2} = 0.6, l_{\Omega_2,3} = 0.8, l_{\Omega_2,4} = 0.8, l_{\Omega_2,5} = 0.4, \text{ and } l_{\Omega_2,6} = 0.3, \mu_{21} = 0.6, \mu_{22} = 0.1, \gamma_{11} = 1, \gamma_{12} = 1, \gamma_{13} = 1, \gamma_{21} = 1, \gamma_{22} = 1, \text{ and } \gamma_{23} = 1.$

3.1.2 Three-Segment Case

In this case, we assume the hybrid coefficients as follows:

$$\lambda(\boldsymbol{\beta}_q = \mathbf{a}) = \pi_1^* \times f_4(\mathbf{a}; \mathbf{b}_1^*, \boldsymbol{\Omega}_1^*) + \pi_2^* \times f_4(\mathbf{a}; \mathbf{b}_2^*, \boldsymbol{\Omega}_2^*) + \pi_3^* \times f_4(\mathbf{a}; \mathbf{b}_3^*, \boldsymbol{\Omega}_3^*), \quad (16)$$

where $\mathbf{b}_1^* = \mathbf{b}_1, \mathbf{b}_2^* = \mathbf{b}_2,$ and $\mathbf{b}_3^* = (3.0, 1.0, 2.0, 0.5)'$, $\boldsymbol{\Omega}_1^* = \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2^* = \boldsymbol{\Omega}_2,$ and

$$\boldsymbol{\Omega}_3^* = \begin{bmatrix} 4.0 & 1.0 & 1.2 & 0 \\ 1.0 & 1.25 & 1.1 & 0 \\ 1.2 & 1.1 & 1.81 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathbf{L}_{\Omega_3} \mathbf{L}'_{\Omega_3} = \begin{bmatrix} 2.0 & 0 & 0 & 0 \\ 0.5 & 1.0 & 0 & 0 \\ 0.6 & 0.8 & 0.9 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2.0 & 0.5 & 0.6 & 0 \\ 0 & 1.0 & 0.8 & 0 \\ 0 & 0 & 0.9 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The mixture weights $\pi_1^*, \pi_2^*,$ and π_3^* are set by specifying $\boldsymbol{\mu}_1^* = \boldsymbol{\mu}_1, \boldsymbol{\mu}_2^* = \boldsymbol{\mu}_2,$ and $\boldsymbol{\mu}_3^* = [0.4, 0]'$. Then, the parameters to be estimated in this three-mixture case include: $b_{11}^* = 1, b_{12}^* = 2, b_{13}^* = 0.6, b_{21}^* = 2, b_{22}^* = 1.5, b_{23}^* = 0.2, b_{31}^* = 2, b_{32}^* = 0.5, b_{33}^* = 0.1, \tilde{b}^* = 0.5, l_{\Omega_1,2}^* = 0.5, l_{\Omega_1,3}^* = 0.866, l_{\Omega_1,4}^* = 0.7, l_{\Omega_1,5}^* = 0.519, l_{\Omega_1,6}^* = 0.374, l_{\Omega_2,1}^* = 0.9, l_{\Omega_2,2}^* = 0.6, l_{\Omega_2,3}^* = 0.8, l_{\Omega_2,4}^* = 0.8, l_{\Omega_2,5}^* = 0.4, l_{\Omega_2,6}^* = 0.3, l_{\Omega_3,1}^* = 2, l_{\Omega_3,2}^* = 0.5, l_{\Omega_3,3}^* = 1.0, l_{\Omega_3,4}^* = 0.6, l_{\Omega_3,5}^* = 0.8, l_{\Omega_3,6}^* = 0.9, \mu_{21}^* = 0.6, \mu_{22}^* = 0.1, \mu_{31}^* = 0.4, \mu_{32}^* = 0.0, \gamma_{11}^* = 1, \gamma_{12}^* = 1, \gamma_{13}^* = 1, \gamma_{21}^* = 1, \gamma_{22}^* = 1, \gamma_{23}^* = 1, \gamma_{31}^* = 1, \gamma_{32}^* = 1, \gamma_{33}^* = 1.$

3.1.3 Data Generation

Using the design presented in the previous sections, we generate the consumption quantity vector \mathbf{x}_q^* for each individual using the forecasting algorithm proposed by Pinjari and Bhat (2011). The above data generation process is undertaken 100 times with different realizations of the β_q vector to generate 100 different data sets each for the two- and three-mixture cases.

We estimate two additional models on each of the 100 generated data sets for each of the two- and three-mixture cases. The first model ignores random coefficients on the independent variables in each mixture (latent segment), allowing random coefficients only on the constants. This corresponds to the Latent Class-MDCP (or LC-MDCP) model. Thus, the only Cholesky parameters estimated for the two-mixture case are $l_{\Omega_1,2} = 0.5$, $l_{\Omega_1,3} = 0.866$, $l_{\Omega_2,1} = 0.9$, $l_{\Omega_2,2} = 0.6$, and $l_{\Omega_2,3} = 0.8$. All other Cholesky parameters are effectively held to the value of zero. In Table 2b, for the LC-MDCP model, the only Cholesky parameters estimated are $l^*_{\Omega_1,2} = 0.5$, $l^*_{\Omega_1,3} = 0.866$, $l^*_{\Omega_2,1} = 0.9$, $l^*_{\Omega_2,2} = 0.6$, $l^*_{\Omega_2,3} = 0.8$, $l^*_{\Omega_3,1} = 2$, $l^*_{\Omega_3,2} = 0.5$, and $l^*_{\Omega_3,3} = 1.0$.

The second model assumes away non-normality by using a single segment for the entire sample (that is, assumes that μ_{21} in the two-mixture case, and μ_{21}^* and μ_{31}^* in the three-segments case, all go to the value of $-\infty$). This is the traditional normally-distributed random-coefficients MDCP (or RC-MDCP) model. Also, in this case $l_{\Omega_2,1} = 0.9$ in the two-segment case, and $l^*_{\Omega_2,1} = 0.9$ and $l^*_{\Omega_3,1} = 2$ in the three-segment case, are not estimable and fixed at 1.0. Additionally, in the two segment case, the following constraints are imposed: $b_{11} = b_{21}$, $b_{12} = b_{22}$, $b_{13} = b_{23}$, $l_{\Omega_1,2} = l_{\Omega_2,2}$, $l_{\Omega_1,3} = l_{\Omega_2,3}$, $l_{\Omega_1,4} = l_{\Omega_2,4}$, $l_{\Omega_1,5} = l_{\Omega_2,5}$, $l_{\Omega_1,6} = l_{\Omega_2,6}$, $\gamma_{11} = \gamma_{21}$, $\gamma_{12} = \gamma_{22}$, and $\gamma_{13} = \gamma_{23}$. In the three-segment case, the following constraints are imposed: $b_{11}^* = b_{21}^* = b_{31}^*$, $b_{12}^* = b_{22}^* = b_{32}^*$, $b_{13}^* = b_{23}^* = b_{33}^*$, $l^*_{\Omega_1,2} = l^*_{\Omega_2,2} = l^*_{\Omega_3,2}$, $l^*_{\Omega_1,3} = l^*_{\Omega_2,3} = l^*_{\Omega_3,3}$, $l^*_{\Omega_1,4} = l^*_{\Omega_2,4} = l^*_{\Omega_3,4}$, $l^*_{\Omega_1,5} = l^*_{\Omega_2,5} = l^*_{\Omega_3,5}$, $l^*_{\Omega_1,6} = l^*_{\Omega_2,6} = l^*_{\Omega_3,6}$, $\gamma_{11}^* = \gamma_{21}^* = \gamma_{31}^*$, $\gamma_{12}^* = \gamma_{22}^* = \gamma_{32}^*$, and $\gamma_{13}^* = \gamma_{23}^* = \gamma_{33}^*$.

We make the comparison between the proposed FDMN-MDCP model and the two restrictive formulations above (that is, the LC-MDCP and the RC-MDCP based on the ability to accurately recover model parameters as well as usual nested likelihood ratio tests).

The analytic approximation embedded in the MACML estimator is applied to two of the datasets 10 times with different permutations to obtain the approximation error. The approximation error is negligible, so only one set of permutations for computing the approximation will be considered in each of the 100 datasets. 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 parameters using the analytic approximation in the MACML for each data set s . Estimate the standard errors using the Godambe (sandwich) estimator.

(2) Compute the mean estimate for each model parameter across the data sets to obtain a **mean estimate**. Compute the **absolute percentage (finite sample) bias (APB)** of the estimator as:

$$APB(\%) = \left| \frac{\text{mean estimate} - \text{true value}}{\text{true value}} \right| \times 100.^8$$

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

(4) Compute the median standard error for each model parameter across the 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 increases).

(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:

$$APBASE(\%) = \left| \frac{ASE - FSSE}{FSSE} \right| \times 100$$

3.2 Simulation Results

3.2.1 Recoverability of Parameters in the MDCP with the Mixture Model

Tables 1a and 1b present the results for the simulation. Table 1a corresponds to the two-segment case, while Table 1b corresponds to the three-segment case. The second column presents the true values used in generating the data samples. The third column labeled “Parameter Estimates” provides the mean value (across the data sets) of each parameter as well as the corresponding

⁸ In case a true parameter value is zero, the APB is computed by taking the difference of the mean estimate from the true value (= 0), dividing this difference by the value of 1 in the denominator, and multiplying by 100.

APB measure, while the fourth broad column labeled “Standard Error Estimates” provides the ASE, FSSE, and the APBASE values for the parameter standard errors.

The APB values for the parameter estimates (third column) show that the MACML method does very well in recovering the parameters. The overall mean APB value across all parameters is 3.2% in the two-segment case (see the last row of the column labeled “APB” in Table 1a). The APB values are in general higher for the three-segment case (Table 1b), with an overall mean value of 6.4% across all parameters, probably due to the many additional parameters that have to be estimated relative to the two-segment model. In general, across the parameters, the APB values are relatively high for the γ satiation parameters in both the two- and three-segment cases. The satiation parameters are an important source of non-linearity in the overall utility function (see Equation 1), and make the likelihood surface more difficult to track computationally. The APB values of the μ parameters are also relatively high in both cases (two and three segments) relative to the APB values of the rest of the parameters. These μ parameters appear in the likelihood function through the mixture (π) probabilities, and it is well established in the literature (see, for example, Sobhani *et al.*, 2013) that these mixture probabilities are difficult to pin down because the likelihood surface can be relatively flat for a number of different combinations of the mixture probabilities near the likelihood optimal point.

The finite sample standard errors and the asymptotic standard errors (in the fourth broad column of Tables 1a and 1b) are close; the average absolute difference is 0.007 and 0.013 for the two- and three-segment cases, respectively. The mean APBASE value across all parameters is 7.8% for the two-segment case and 8.7% for the three-segment case. In both the two- and three-segment cases, the finite sample standard error estimates are generally higher (as a percentage of the mean estimates) for the γ and μ parameters relative to other sets of parameters, reinforcing the finding earlier that the γ and μ parameters are more difficult to recover than other parameters. Some elements of the Cholesky matrix also are difficult to pin down, again because the Cholesky elements enter the likelihood function in a very non-linear fashion as part of the evaluation of the cumulative multivariate normal density and distribution functions.

Overall, the MACML inference approach does well in accurately and precisely recovering parameters in both the two-segment and three-segment FDMN-MDCP model. The reported model estimation times are based on scaling to a desktop computer with an Intel(R)

Pentium(R) D CPU@3.20GHz processor and 4GB of RAM. The statistical software GAUSS was used for all the estimations reported in this paper.

3.2.2 Comparison between the Proposed Model and More Restrictive MDCP Models

Tables 2a and 2b present the results for the simulation exercise focusing on the comparison between the proposed FDMN MDCP model and two other, more restrictive versions of the model: the LC-MDCP and the RC-MDCP models. Table 2a corresponds to the two-segment case, while Table 2b corresponds to the three-segment case. The APB values of the parameters are in general higher in both cases (two and three segments) and in both alternative models relative to the APB values of the parameters in the original model (Tables 1a and 1b). In the two-segment model, the overall mean APB values across parameters are 28.5% and 26.0% for the LC-MDCP and RC-MDCP models, respectively—significantly higher in comparison with the mean APB value of 3.2% in the proposed model. The difference is even higher in the three-segment model with the overall mean APB values across parameters being 30.8% and 82.9% for the LC-MDCP and RC-MDCP models, respectively, relative to the overall mean APB value of 6.4% in the original model. The superior performance of the FDMN-MDCP model is also evidenced in the higher log-likelihood value, on average, for the FDMN-MDCP model across the 100 estimations (on the 100 data sets). In addition, for each of the 100 data sets, a likelihood ratio test comparing the FDMN-MDCP model with the two other models clearly rejects the other two model in favor of the FDMN-MDCP model (see last row of Tables 2a and 2b).

4. AN APPLICATION

In this paper, we demonstrate an application of the proposed model to analyze individual-level decisions regarding recreational destination locations and the number of trips to each destination, using data drawn from the 2012 New Zealand Domestic Travel Survey (DTS).

4.1 Background

Tourism has been an important contributor to New Zealand’s economy, thanks to the natural and beautiful landscape of the compact island country that also offers an extensive coastline for trekking, swimming, fishing, other water-based activities, and sports. In addition, New Zealand also boasts of some excellent wine factories, offers volcanic/geothermal excursion opportunities,

and its forests and pristine landscape have made it a much sought-after location for mainstream Hollywood movies (for example, the Fiordland and Southern Lakes in the southern part of New Zealand were the locations for the mythical Middle Earth in the "Lord of the Rings" trilogy). Overall, tourism contributes 9% of New Zealand's gross domestic product and is also an important source of employment; 10% of New Zealanders work in the tourism industry (see New Zealand Tourism Strategy 2015).

Although the international popularity of New Zealand has increased enormously in the past few years, domestic tourism continues to remain a significant source of income for the tourism industry. According to the New Zealand Tourism Industry Association (TIA, 2012), domestic travelers (New Zealand residents traveling within New Zealand) accounted for about 57% of New Zealand's total tourism industry spend of \$23 billion in 2012 (see Statistics New Zealand, 2013). The substantial amount of domestic tourism may be attributed to increased marketing efforts of leisure activity opportunities within the island nation and more control of the leisure vacation experience through on-line sites. However, it is also a result of a general trend across all countries around the globe of an increasingly compact geographic footprint of leisure travel, spurred by a shift from the traditional long period vacations undertaken during holidays or over the summer to short period leisure travel built around the work weeks (see, for example, White, 2011 and LaMondia and Bhat, 2012). This shift itself may be traced to easier schedule coordination opportunities for short duration leisure pursuits around work weeks, especially for the increasing number of families with multiple working individuals with school-going children.

The growing amount of short distance leisure trips, mostly undertaken using the personal auto mode, has led to increased attention on this leisure travel market among urban transportation planners because of the increased weekend day traffic on city streets and between cities in close proximity, and the concomitant effects on traffic congestion and air quality. Understanding these travel flow patterns can help planning and policy efforts to reduce the negative externalities of such travel. At the same time, unraveling the "push and pull" factors associated with individual and household leisure activity decisions helps cities and regions position themselves as unique and even exotic destinations, with an eye on generating jobs and revenue. This confluence of interest on leisure travel from the transportation and tourism domains has led to many studies in this space in the past decade, with a particular emphasis on destination choice for leisure pursuits. While the early literature in the area considered leisure

destination choices as repeated isolated (and independent) decision events for each leisure trip, the more recent literature has moved toward the more realistic representation of destination choices as inter-related decisions for multiple leisure trips over a longer-term period of a month or even a year. Examples of the latter string of multiple discrete-continuous (MDC) studies (with the discrete component being the choice of destination region, and the continuous component being the number of trips to each chosen destination region) include Kuriyama *et al.* (2010), (2011), Van Nostrand *et al.* (2013), von Haefen (2007), Whitehead *et al.* (2010), LaMondia *et al.* (2010), and Bhat *et al.* (2013). These studies explicitly accommodate variety-seeking and loyalty behavior by considering satiation effects based on Iso-Ahola's (1983) theory of vacation participation in which the individual/family balances needs for familiarity and novelty, within long period budget constraints, to provide an "optimally arousing experience" (see LaMondia *et al.*, 2008 for a detailed discussion). In this paper, we contribute to leisure destination choice modeling using the proposed FDMN MDCP model. To our knowledge, this is the first such application in the leisure travel literature.⁹

4.2 The Data

The data for this study is derived from three sources. The primary source, as mentioned earlier, is the 2012 New Zealand DTS, which asked survey respondents (New Zealand's residents) to provide information on all one-way trips 40 kilometers or longer from home, overnight trips from home, and flight or ferry trips from home made up to four weeks prior to the survey date (see Ministry of Business, Innovation and Employment, 2013). The survey was targeted at

⁹As with all the earlier MDC leisure studies, this study too focuses on the count of the number of times each leisure destination is visited. Thus, the "continuous" quantity used is actually a count variable, as opposed to a truly continuous measure as required by the theoretical model. But, as demonstrated by von Haefen and Phaneuf (2003), treating the integer count of trips as a continuous variable (within an MDC framework) does not lead to substantial bias in the results or the behavioral implications. This forms the basis for the use of the MDC framework in earlier studies, as well as in the current study, of leisure destination choices over a period of time. Similarly, as in earlier studies, the budget in the MDC formulation is the total number of leisure trips made over a given time period. This budget is "allocated" to the different possible discrete leisure destination locations. While a more reasonable approach would be to allocate a money budget, the operationalization of this alternative approach is extremely difficult because of the many assumptions that need to be made regarding monetary costs of participation per trip. Besides, a more practical problem is that expenditure information is rarely obtained in travel surveys. On the other hand, the number of trips to each destination is readily available from a sample of individuals in a survey, and the total trips (or "budget") is readily obtained by aggregating across the possible destination locations. A related issue in the use of total trips as the "budget" is that the MDC models of leisure destination choice focus on the count of trips to each destination, given the total number of leisure trips during a specified period. In forecasting mode, the latter "budget quantity" is itself predicted in an earlier "trip generation" step, including the choice of making no leisure trips at all.

individuals and not households in that only one randomly selected individual (over the age of 15 years) from each sampled household was interviewed. Telephone interviewing was used for the DTS and household telephone numbers were randomly selected from the white pages. Interviews were carried out according to pre-specified quotas for age, sex and region of origin. The process of data collection took place continuously throughout the year.

The survey obtained information on the resident city of the respondent, the city of destination for each trip, the primary reason of each trip, and the primary mode of transportation used to reach the destination. Additionally, the survey also obtained individual and household socio-demographic information. A second data source is a network level of service file that provided information on land travel distance and highway travel time between each city pair within New Zealand (see additional details in the next paragraph). The third data source is a disaggregate spatial land-cover characteristics data obtained from the 2012 Land Cover Database (LCDB) of the Land Resource Information System (LRIS) of New Zealand. The LCDB provides land-cover information at a 30 meters by 30 meter resolution. From this data, using a geographic information system based procedure, we developed total land area and acreage information for each 30x30 meter² grid and by six broadly defined land-cover categories: urban area (including central business districts, commercial and industrial areas, urban parklands, urban dumps, and housing and transportation-related land cover), water area (including rivers, land/ponds, freshwater, and estuarine open water), wetland area (context-dependent combinations of areas such as herbaceous freshwater vegetation, flaxland, and saline vegetation), agricultural area (including vineyards and orchards, perennial crops, short rotation cropland, and grasslands), bare-land area, and forest area (pine forests, mangroves, deciduous hardwoods and other exotic/indigenous forest areas).

The sample formation comprised several steps. First, we selected only leisure trips to primary destinations within New Zealand undertaken by a personal auto (personal auto trips comprise around 90% of all leisure domestic trips within New Zealand; see Ministry of Business, Innovation and Employment, 2008). Second, the leisure destination cities in New Zealand were mapped into one of 16 aggregate destination regions in the current analysis, as identified in Figure 1. Nine regions are in the North Island, while seven are in the South Island. This regional classification scheme is the same as that used by the Department of Tourism of New Zealand for its marketing campaigns, and is also the commonly used geo-political partitioning of the country.

Third, the total number of trips made by each individual to each region was obtained by appropriate aggregation across trips to cities within each region, and the individual-level trip budget is obtained as the total number of trips of the individual across all regions during the four week period. Fourth, we identified a centroidal city for each of the 16 destination regions, based on the city that attracted the most travelers within each region, and converted the city-to-city land-based travel distance and land-based travel time data to corresponding residence city-to-destination region skims. But travel from one region in one island to another region in another island by auto is possible only through the use of a ferry service (that transports vehicles too) across the Cook Strait between Wellington in the North (located in the Wellington region) and Picton in the South (located in the Marlborough region). On the other hand, the land-based travel time between two regions in different islands from earlier includes only the travel time from the origin point to one of the two ferry terminals plus the travel time from the other ferry terminal to the destination region. Thus, the total travel time between two regions in different islands should include the 3 hour 15 minute cruise (including ferry terminal times) between the north and south islands. At the end of this step, we obtain the land-based travel distance and the total travel time for each residence city-destination region pairing. Fifth, the travel cost skims were computed as a function of the respondent's reported household income, the estimated cost of vehicle fuel on land, the ferry cost if a ferry crossing is involved, and the land-based distance and total travel time skims (obtained in the previous step) between the respondent's residence city and the centroidal city of each destination region. To calculate the travel cost, we followed 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). Specifically, the travel cost was computed as:

$$\text{Cost (in NZ\$)} = 2 * (\text{one-way land travel distance in miles} * \text{fuel cost per mile} + \text{one-way total travel time in hours} * (0.5 * \text{hourly wage})) + \text{round-trip ferry cost (as applicable)}.$$

The fuel cost per mile is computed at NZ\$0.149 per mile based on a fuel cost of NZ\$1.75 per liter and a rather high vehicle efficiency factor of 5.3 liters for 100 km (5.3 liters for 62.1 miles or about 44 miles per gallon), given the long distance nature of trips under consideration. The round-trip ferry cost is NZ\$145. Sixth, the grid-based land-cover data were translated to a destination region-based land-cover data by suitable aggregation over cells within each destination region. Seventh, individual and household socio-demographic, as well as land cover data by region, were appended to the long distance travel records.

The final data sample used in the estimation included 3508 individuals. Table 3 provides the distribution of these individuals by the number of leisure trips made during the four week period before they were surveyed and by the number of distinct leisure destination regions visited. Although a sizeable fraction (72.3%) of the individuals in the sample make only one trip, a non-insignificant percentage of individuals (27.7%) make more than one trip. Most of the individuals who undertake more than one trip during the survey period prefer to travel to multiple destinations (see the second row and beyond in Table 3). For example, 53.3% of individuals making two trips during the survey period visit more than one distinct destination region, while 65% of individuals making three trips visit more than one distinct region. The corresponding numbers are 70.2% and 78.6% for individuals who make four and five or more trips, respectively, during the survey period. Clearly, this is a case of multiple discreteness for individuals who make more than one trip.

Table 4 provides descriptive statistics for each of the 16 destination regions. The third broad column presents the mean and standard deviations for the travel impedance skim measures of total travel time, travel distance, and travel cost for each destination region (computed from the residence city-destination region skims developed as discussed earlier in this section). Not surprisingly, the travel impedance measures are the highest for the Northland region in the North Island (the northernmost region) and the Southland region in the South Island (the southernmost region). As expected, the impedance measures decrease as one gets closer to the center of the country. Interestingly, the impedance measures are lower for the North Island regions compared to the South Island regions. This is because of two-interrelated factors. First, the North Island is more populated relative to the South Island (the North Island's population is about 3.2 million, while that of the South Island is about 1 million), which should result in more leisure trips generated from the North Island due to a sheer population size effect. Second, because of the compact nature of the North Island, there are more leisure trips generated per capita in the North than in the South, and most of these trips are destined to within the compact North Island. The net result is that, if one were to draw a horizontal "residential center of gravity" (RCG) line of tourists, it would go through the boundary of the Waikato and Manawaku-Wanganui (MW) regions in the North Island (see Figure 1). This is also evidenced in Table 4 in that the impedance measures are the smallest for the Waikato and MW regions, and increase as one goes farther away from the horizontal RCG line. Additionally, we should also note that, of the 3508

individuals in the sample, 2588 (73.7%) percent reside in the North Island, and 662 (18.9%) reside in the Waikato-MW regions. The fourth broad column in Table 4 provides the percentage of land in each region in each of the six land cover categories (the sum across all columns for each row add up to 100%). Of all the regions, Auckland has the highest percentage of urban land-cover, with Nelson and Wellington being the regions with the second and third highest urban land cover percentages. As we will see later, the high urban land cover is correlated with the intensity of tourist draw. In terms of wetland cover percentages, the highest are for Tasman, West Coast, Otago, and Southland. Nelson is the region with the highest forest land cover.

Table 5 provides additional descriptive statistics of the area of each region and destination region characteristics. The third column of the table presents the area of each region. As can be observed from this column and also from Figure 1, Canterbury in the south island is the largest region by size across all regions, while Waikato and MW are the largest regions in the North Island. The fourth column shows the number (and the corresponding percentage) of individuals who visited each region at least once. The Waikato region is clearly the one patronized by the most number of individuals, but Auckland, Bay of Plenty, and Canterbury also draw quite a few individuals. However, to get a better picture of attractiveness, the fifth column normalizes the number of people visiting by the area of each region (to accommodate for the fact that there are likely to be size effects here; that is, the larger a region, the more likely it is to be a destination). This column shows that on a per unit area basis, Auckland is by far the most popular destination, followed by Wellington and Nelson. Interestingly, as indicated earlier, these are the three destinations with the highest percentages of urban land cover, and Nelson is the region with the highest forest cover. The Auckland region includes the famous urban tourist attraction of the City of Auckland as well as such attractions as the Tiritiri Matangi Islands, a haven for nature hikers who want to experience the rich flora and fauna of the region up close (especially of a host of endangered species of birds, each with a unique bird call pattern). The Wellington region, with Wellington City that serves as the capital of the North Island, is well known for Mt. Victoria (that provides a nice walk trail and panoramic views of the city and the Wellington harbor), massage and waxing boutiques in the Lower Hutt area also overlooking the Wellington harbor, and an interactive national museum of New Zealand culture and heritage. Finally, the Nelson region in the north of the South Island, the smallest of all the regions but also the sunniest in all of New Zealand, includes the city of Nelson. Nelson is renowned for its Maori

(indigenous Polynesian tribe of New Zealand) arts and craftsmanship, water sports and activities (the Nelson region has the second largest amount of land percentage covered by water, and is liberally sprinkled with freshwater springs, especially near Takaka), and hiking/biking trails in the Abel Tasman National Park and other pristine forest land. Also interesting to note is that Tasman, West Coast, Otago, and Southland are some of the regions with the lowest number of visiting individuals per unit area, and these regions all have a relatively high wetland cover percentage as identified earlier, suggesting an inverse relationship between wetland cover percentage and tourist draw (perhaps because there is little to do within wetlands). The sixth broad column presents statistics on the number of visits to a destination region among those who visited the destination region at least once. The mean and maximum values from this column suggest that Auckland, Waikato, Bay of Plenty, Wellington, Canterbury, and Otago have the most loyal following.

4.3 Variable Specification and Model Formulation

The number of destination region alternatives in the MDCP model is 16. Thus, rather than including 15 alternative-specific constants in the baseline preference and 16 region-specific satiation parameters (in addition to other explanatory variables) in each latent segment, we adopted an “unlabeled” MDCP specification in which the baseline preferences and satiations are captured through attributes of the individual regions. For identification in this unlabeled alternatives context, the constant for the first segment is constrained to zero, and the constants for other segments are constrained to be descending from the second segment forward.

4.3.1 Baseline Preference Specification

The first independent variable we used in the baseline preference (that is, as part of the z_{qk} vector in Equation (2)) is the logarithm of the area of each region, to proxy for the number of elemental destination opportunities within each aggregate region (see Bhat *et al.*, 1998). The expectation is that large regions are more likely to be chosen as a recreation destination based on a sheer “volume of opportunities” effect. The coefficient on this size variable may be viewed as an inclusive value characterizing the presence of common unobserved destination region attributes affecting the utility of elemental alternatives within each region. As in traditional discrete choice models, we expect this coefficient to be positive and less than one. If less than one, the

implication is that there are common unobserved region attributes that lead to higher sensitivity across elemental alternatives within a region than across different regions. The net effect is that there is an inelastic influence of increasing region size on the region's baseline utility. That is, compared to the case when the coefficient is one, the rise in the baseline utility of a region due to an increase in the region's size is much less when the coefficient is estimated to be less than one in magnitude (because of more redistribution of leisure trips across elemental destinations within the same region rather than across different regions).

The next set of variables we considered are land-cover effects, captured by interacting the land-cover percentage by category in each destination region with the travel time from each individual's residence city to the centroidal city of each destination region. We computed a land-cover accessibility measure of the Hansen-type (Fotheringham, 1983) for individual q and land-cover type i as presented by destination region k as $AC_{qki} = LC_{ki} / [f(TT_{qk})]$, where LC_{ki} is the percentage area in land-cover category i ($i = \text{urban, water, wetland, agricultural, bare-land, and forest}$) in destination region k , TT_{qk} is the travel time (in hours) from individual q 's residence city to the centroid of destination region k , and $f(\cdot)$ is a function.¹⁰ The accessibility measures proxy the intensity of opportunities for recreational participation specific to each land-use category in a destination region normalized by a measure of impedance (function of travel time) for individual q to reach those opportunities. In the empirical analysis, a host of functional forms can be tested for the travel time measure. In our specifications, we considered both a linear form, $f(TT_{qk}) = TT_{qk}$, as well as a logarithmic form, $f(TT_{qk}) = \ln(TT_{qk})$. The logarithmic form penalizes destination regions less for being far away from the residential location of the individual. In both cases, a positive coefficient on an accessibility measure implies that individuals are attracted toward proximal destination regions with a substantial percentage of area in the corresponding land use. Our expectation, based on the descriptive statistics, is a positive coefficient on the urban land cover accessibility variable, though things are less clear from the descriptive analysis regarding the nature of effects of other accessibility variables. Based on our specification tests, the linear form is the preferred functional form for $f(TT_{qk})$.

¹⁰ We do not introduce the land-cover percentages themselves directly in the baseline preference because these percentages do not vary across individuals in the sample. Thus, destination region land-cover percentages by themselves do not provide adequate variation to estimate parameters (because there are only 16 destination regions). But, by interacting these land cover percentages with individual-specific travel times to each region, we obtain rich variation across individuals in the resulting accessibility measures.

The land cover-based accessibility effects (which are specific to each land cover category) capture any preferences individuals have for specific types of activities that may be featured in each destination region (as manifested in the land-cover category percentages). However, these effects do not capture an overall diversity index for each destination region. That is, it is possible that some individuals may be drawn to destination regions that have a good diversity of activity participation opportunities as well as are relatively close by. We proxy this effect by constructing a diversity index of land-cover types for each destination region, based on generalizing a similar index proposed originally by Bhat and Gossen (2004). This land cover diversity index is computed as a fraction between 0 and 1 for each destination region. Regions with a value closer to one have a richer land-cover mix than regions with a value closer to zero. The actual form of the land-cover diversity index for destination region k is:

$$\text{Land-cover diversity index } D_k = 1 - \left[\frac{\sum_{i=1}^I \left| LC_{ki} - \frac{100}{I} \right|}{\left(\frac{200(I-1)}{I} \right)} \right], \quad (17)$$

where LC_{ki} is the percentage area in land-cover category i in destination region k (as earlier) and $I=6$ (that is, we have six land cover categories) in our empirical context. The functional form would assign the value of zero if a region's land-cover is only in one category, and would assign a value of 1 if a region's land-cover is equally split among the different land-cover categories. However, as in the case of the land-cover percentages, there is no variation in the diversity index for a region across individuals, and the only variation in the index is across the 16 destination regions. This is inadequate to estimate a parameter on the diversity index, and thus we introduce the diversity accessibility index by normalizing the diversity index by a function of travel time to obtain individual-specific diversity accessibility indices: $DA_{qk} = D_k / f(TT_{qk})$. As earlier, we test both a linear form and a logarithmic form for the effect of travel time in the denominator of this expression. The best data fit results were again obtained consistently with the linear form.

Another variable considered in the specifications was the travel cost to each destination region, with the expectation that a higher cost would deter visiting the corresponding region. Again, both a simple linear form as well as a logarithmic form were tested for this cost effect, with the linear form winning out as the preferred one in our empirical tests. In addition, we

included a dummy variable for the presence of a ferry ride. This accommodates any positive leisure/relaxation value of the ferry ride itself, after accounting for the total travel time effect.

A continuous random coefficient specification is considered on all of the above variables in the baseline preference for each discrete mixture (that is, each latent segment).

Finally, there is one other important issue with regard to the baseline preference specification. As discussed earlier, we use an unlabeled system for the alternatives, which essentially means that we constrain the mean coefficients on the alternative specific constants to be zero in the baseline utility for each destination region alternative. That is, the elements of \mathbf{b}_g (in the notation of Section 2.2) corresponding to the 15 alternative-specific dummy variables for each latent segment g in \mathbf{z}_{gk} are set to zero. However, we allow random covariance about this mean of zero. That is, the 15 elements of $\tilde{\boldsymbol{\beta}}_{gk}$ corresponding to the alternative-specific constants are included with a covariance matrix. Assume that the random coefficients on the alternative-specific constants (ASCs) are independent of the random coefficients on other independent variables. Let $\tilde{\boldsymbol{\beta}}_{g,ASC}$ be a vector that collects the random coefficients corresponding to the 15 ASCs for each segment g . Then, the simplest specification for the covariance matrix of the 15 ASCs (for each segment g) obtained as differences of the original 16 ASCs from the first ASC (corresponding to the Northland region) would be as below (which originates from a specification of independently and identically distributed (IID) random errors with a variance of 0.5 for each of the original 16 ASCs):

$$\tilde{\boldsymbol{\beta}}_{g,ASC} = \boldsymbol{\varepsilon}_{g,ASC} \sim f_{15}(\mathbf{0}, \boldsymbol{\Lambda}), \text{ where } \boldsymbol{\Lambda} = \begin{bmatrix} 1 & 0.5 & 0.5 & \dots & 0.5 \\ 0.5 & 1 & 0.5 & \dots & 0.5 \\ 0.5 & 0.5 & 1 & \dots & 0.5 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0.5 & 0.5 & 0.5 & \dots & 1 \end{bmatrix}_{15 \times 15 \text{ matrix}}. \quad (18)$$

However, there is likely to be spatial correlation across the utilities of the different regions because of similarity in unobserved attributes across proximally located regions. But, we have to assume that one region is not spatially correlated with all the other regions (because only differences in the baseline utilities matter). In our analysis, the first region (that is, the Northland region) will play this base role. We then accommodate spatial correlation across other regions

using a spatial autoregressive (SAR) error structure of order one for the random components of the ASCs of the other 15 regions as follows:

$$\tilde{\beta}_{g,ASC} = \rho \mathbf{W} \tilde{\beta}_{g,ASC} + \boldsymbol{\varepsilon}_{g,ASC}, \quad (0 < \rho < 1), \quad (19)$$

where ρ is the spatial autoregressive coefficient, \mathbf{W} is a distance-based spatial weight matrix with elements $w_{kk'}$ corresponding to regions k and k' (with $w_{kk} = 0$ and $\sum_{k'} w_{kk'} = 1$). With the

specification above, and defining $\mathbf{S} = [\mathbf{IDEN}_{15} - \rho \mathbf{W}]^{-1}$ [15×15 matrix], where \mathbf{IDEN}_K is the identity matrix of size K ($K=15$ in our case), we may then write:

$$\tilde{\beta}_{g,ASC} = \mathbf{S} \boldsymbol{\varepsilon}_{g,ASC}, \text{ and } \tilde{\beta}_{g,ASC} \sim f_{15}(\mathbf{0}, \mathbf{S} \boldsymbol{\Lambda} \mathbf{S}') \quad (20)$$

In the above expression, technically, we can allow the distribution of $\tilde{\beta}_{g,ASC}$ to vary across segments g by allowing a general specification for $\boldsymbol{\Lambda}$ that varies across segments (the only normalization requirement is that the first element for the first segment be 1) and/or by allowing the spatial autoregressive coefficient to vary across segments. However, the first specification leads to proliferation in the number of parameters (especially given the number of alternatives), while the second one is not intuitive because there is no reason for the intensity of spatial correlation in unobserved attributes to vary across segments. Thus, from a pragmatic standpoint, we use the same simple covariance matrix across all segments for the $\tilde{\beta}_{g,ASC}$ vector (as in Equation (18)). Doing so also allows a comparison of the magnitude of the mean of coefficients in the baseline preference across segments, as long as there are no substantial differences in the variance elements of the coefficients. A point to note in this discussion is that the expression in Equation (20) collapses to that of Equation (18) if there is no spatial correlation, as should be the case.

This leaves the specification of the weight matrix \mathbf{W} . Several weight matrix specifications were considered in our empirical analysis to characterize the nature of the dynamics of the spatial dependence across regions. These included (1) a contiguity specification that generates spatial dependence between the destination region alternatives based on whether or not two regions are contiguous (we considered the Marlborough and Wellington regions as being contiguous because they are the ferry landing points for travel between the two islands), (2) the inverse of a continuous travel time specification where the time between regions is

obtained from the skims discussed earlier, and (3) the inverse of the square of the continuous distance specification. In addition, for all the three specifications above, we also examined a specification that confines the spatial correlation to only the regions within each island (with zero spatial correlation between regions in different islands). Overall, the best data fit results were obtained consistently with the inverse of the continuous distance specification, which is the one used in the results discussed in the next section.

4.3.2 Satiation and Segmentation Specification

In our estimations, we considered both a γ -profile as well as an α -profile for introducing satiation. In all cases, the γ -profile provided superior results, so we will only discuss the specification for the γ -profile here. As discussed earlier, the parameter γ_{qgk} may be parameterized as $\exp(\theta'_g \mathbf{a}_{qk})$, where \mathbf{a}_{qk} is a vector of explanatory variables and θ_g is a corresponding vector of parameters specific to segment g in the mixture model. It is the specification of the \mathbf{a}_{qk} vector that we discuss here. In addition to a constant, we considered all the other variables discussed in the previous section. We particularly examined the effect of wetland land-cover accessibility on satiation behavior, based on the suggestion from the descriptive statistics that a higher wetland cover percentage leads to higher satiation effects (less trips).

All the variables associated with demographics characteristics were considered for characterizing different discrete segments (see Equation 7 earlier). These demographic variables included respondent age, respondent's household income, respondent's household size by number of adults (>18 years of age) and number of children (18 years or less), respondent's household structure (single person, couple, nuclear family, single parent, multi-family household, and non-family household), and respondent gender. Of these, the respondent's household structure provided a very good indication of the travel group, because almost all trips were made with family members in couple, nuclear family, single parent, and multi-family households. Also, in our specifications, we considered respondent gender only for single person, single parent, and non-family households, because the decision in other households is likely to be jointly made (and gender simply provides information on which respondent happened to be picked in the survey in these households, and should not provide any preference information).

All the segmentation variables were introduced as alternative-specific variables in the logit link function of Equation (7) with the first segment being the base.

4.4 Model Estimation Results

A number of different specifications were explored, with different sets of variables, different functional forms of variables, and different groupings. The final specification was based on having adequate observations in each category of categorical independent variables (such as for household structure), a systematic process of rejecting statistically insignificant effects, combining effects when they made sense and did not degrade fit substantially, and, of course, judgment and insights from earlier studies. To identify the optimal value for the number of latent segments (G), we estimated the model for increasing values of G ($G = 1, 2, 3, 4, \dots$) until we reached a point where an additional segment did not significantly improve model fit. The evaluation of model fit was based on the Bayesian Information Criterion (BIC):

$$\text{BIC} = -L(\boldsymbol{\theta}) + 0.5 \cdot R \cdot \ln(N).^{11} \quad (21)$$

The first term on the right side is the negative of the log-likelihood value at convergence; R is the number of parameters estimated and N is the number of observations (see Allenby, 1990, Bhat, 1997). As the number of segments, G , increases, the BIC value keeps declining till a point is reached where an increase in G results in an increase in the BIC value. Estimation is terminated at this point and the number of segments corresponding to the lowest value of BIC is considered the appropriate number for G . In our analysis, based on the Bayesian Information Criterion (BIC), the three-segment model was clearly the model with the best performance (the log-likelihood value at convergence for this model was -8,499.78 and, with 46 model parameters, the BIC was 8,687.52; the corresponding values for the model with one segment (that is, no

¹¹ Many measures have been suggested in the literature to evaluate model fit, especially in the context of the number of segments in latent segmentation models. These include the Akaike Information Criterion (AIC), the BIC, and many variants of both of these (see Fonseca, 2010 for a listing and description of these information criteria). In general, the criteria based on the AIC tend to favor complex models with many segments as the sample size increases, leading to potential overfit. On the other hand, the criteria based on the BIC tend to favor simpler models, with an adjustment for sample size (such as the $\ln(N)$ appearing in Equation (21)), to avoid overfit. More simply speaking, the BIC-based measures demand a higher strength of evidence to add complexity than do the AIC-based measures, and thus the BIC-based measures favor more parsimonious models with fewer segments than do the AIC-based measures (see Neath and Cavanaugh, 2012). In the context of latent segmentation models, where the number of parameters explodes as the number of segments increases, parsimony is a much desired property from an interpretation and simplicity perspective. Thus, most latent segment models adopt the BIC as the model selection criterion, as we also do in the current paper.

latent segmentation), two segments, and four segments were 8,872.21, 8,711.36, and 8,780.14, respectively.

The estimation results for the three-segment mixture MDCP model are presented in Table 6. The first panel corresponds to the probabilistic assignment of individuals to each of the three segments (the first segment is the base segment). The second presents the parameter estimates on the independent variables in the baseline utility specifications of the MDCP model corresponding to each segment. The third provides the parameters in the satiation component. Each of these is discussed in turn in the next three sections.

4.4.1 Assignment of Individuals to Discrete (Latent) Segments

In the top panel of Table 6, the constants in the segmentation model contribute to the size of each segment and do not have any substantive interpretation. The other results in the top panel of Table 6 indicate that the second segment, relative to the other two segments, is more likely to consist of individuals with children (that is, the individuals are more likely to belong to nuclear or single parent households) and low-income individuals. This second segment also is less likely to comprise single person households relative to the first segment. The third segment comprises individuals who tend to be in couple households of middle age (48 years) or older, the least likely to be single person households, and less likely to be in the “lower than NZ\$50,000” annual income range relative to the second segment, but more likely to be in this income range relative to the first segment. A more intuitive way to characterize the different segments is to estimate the percentages of individuals in each category of the demographic variables in each segment (see Bhat, 1997 for the formula to do so). The results are presented in Table 7. For example, the first numerical value in the table indicates that 60.2% of individuals in the first segment are younger than 48 years, while the corresponding percentages are 61.8% and 35.7% in the second and third segments, respectively. In the overall sample, 46.4% of individuals are younger than 48 years. The figures in Table 7 support our previous observations regarding segment characteristics. Based on the relative characterizations of the segments, we will refer to the first segment as the “high-flyer low family commitments” (HFLFC) segment, the second as the “low income parents” (LIP) segment, and the third as the “couple baby-boomer” (CBB) segment (most individuals over 48 years of age in the sample were born between 1943 and 1964 and represent the post-war baby-boom generation of New Zealand). In terms of the relative sizes of the three

segments, this can also be estimated in a straightforward way by aggregating the individual segment-level probabilities (Equation 7) across all individuals. The sizes are estimated to be 11.4%, 57.9% and 30.7% for the HFLFC, LIP, and CBB segments, indicating a domination of the LIP segment in the population.

4.4.2 Baseline Utility Parameters

Referring back to Table 6, the effect of size (see the second panel) in the baseline utility function is positive and less than one. We specified different size coefficients across the segments, but the coefficients were not statistically different and were constrained to be equal. This was also our theoretical expectation, because we saw no reason that the size coefficient (representing the magnitude of region-specific unobserved factors affecting all elemental opportunities within the region) should vary across segments. The coefficient is statistically different from one, indicating the inelastic effect of size growth on the baseline utility.

The effect of the ferry dummy variable in the baseline utility is positive for the HFLFC and LIP segments, but not significant for the CBB segment. The absence of effect on the CBB segment may be a reflection of the relative lack of families with children in this segment, and the possibly intrinsic and positive “adventure” value of a ferry ride for families with children. The effect of travel cost on baseline utility is, as expected and on average, negative in all the segments. The LIP segment is the most cost-sensitive, followed by the CBB segment, in an inverse relationship of cost sensitivity to household income earnings of families across the segments. The results also show statistically significant heterogeneity (across individuals) in the responsiveness to cost within each latent segment, as manifested in the standard deviation estimates on the cost coefficient. The normal distribution assumption implies that some individuals do have a positive utility for cost, but the vast majority have a negative cost sensitivity. In particular, the mean and standard deviation estimates indicate that cost has a negative impact for 92% of individuals in the first and second segments, and for 96.5% of individuals in the final segment.

The land cover accessibility measures reinforce the findings from our descriptive analysis. Specifically, regions with high urban land cover “pull” leisure trips with about equal intensity from all three segments, though there is heterogeneity in the magnitude of the “pull” within each segment (as indicated by the statistically significant standard deviations on the urban

land cover variable in Table 6). Cities clearly offer a much higher density of tourism opportunities from regional events and festivals during the year to gastronomic indulgence opportunities, art galleries, museums, theaters and shopping centers. The effect of forest land-cover on baseline utility is also positive, suggesting a preference for destination regions with high forest land cover. This preference varies across the three discrete segments, with the HFLFC segment having the highest preference for forest-oriented leisure pursuits and the CBB having the lowest. The high preference of the first segment for regions with forest land cover is presumably a reflection of young, single individuals (with relatively little familial commitments) seeking adventurous hiking and bicycling trails through New Zealand's rough and rugged forest terrain. On the other hand, the relatively older CBB segment group may not prefer such physically-intensive leisure pursuits to the extent that their younger counterparts do. Also, there is a clear and generic tendency across all segments to stay away from regions with high wetland land cover. This is not surprising, given that wetlands offer little attraction for tourism and, in New Zealand, are typically associated with negative externalities such as pollution, drainage problems, and presence of invasive plant species (see Peters and Clarkson, 2010). The effect of the agricultural land-cover accessibility varies across segments; while the individuals in the third segment are attracted to agricultural areas, the individuals in the first and second segments tend to avoid agricultural areas. This is perhaps an indication of couple baby-boomers (CBB) being drawn to activities such as visiting vineyards for a relaxed wine-tasting escapade, activities that may not interest individuals with children (the LIP segment) or may be considered too "docile" by young individuals with little family commitments (the HFLFC segment).

The effects of the land-cover diversity accessibility index on the baseline function indicate that high-flying young individuals prefer regions with a good diversity of activities, while those in the LIP and CBB segments prefer regions with focused activities. Another interpretation is that those in the LIP and CBB segments are inclined to pursue very specific types of leisure activities (such as perhaps park entertainment for the LIP segment and wine tasting trips for CBBs), and then select regions that are heavily invested in opportunities of that specific leisure type.

Finally, the covariance estimate (not shown in Table 6) between the travel cost and urban accessibility random coefficients was 0.040 (t-statistic of 2.21), 0.035 (t-statistic of 2.28), and 0.042 (t-statistic of 2.03) for segments one, two and three respectively. This suggests that

individuals who are less sensitive (more sensitive) to travel costs also prefer (dislike) urban destination zones. That is, individuals who prefer recreation pursuits 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.

4.4.3 Satiation Effects

These effects are presented toward the bottom panel of Table 6. As indicated earlier, the satiation parameter is parameterized as $\gamma_{qgk} = \exp(\tilde{\theta}'_g \mathbf{a}_{qk})$, and the satiation coefficients in Table 6 are the $\tilde{\theta}_g$ parameters for each segment g . A positive parameter on a variable implies that an increase in the variable has the effect of increasing the γ_{qgk} parameter and decreasing satiation (that is, increasing repeat trips of the individual to a destination region), while a negative parameter has the effect of decreasing the γ_{qgk} parameter and increasing satiation (that is, decreasing repeat trips of the same individual to a destination region).

Everything else being equal, the constants indicate that satiation in the context of a destination region sets in fastest for the third CBB segment and slowest for the first HFLFC segment. That is, in general, individuals in the HFLFC segment are more willing to make repeat trips to a destination region than individuals in the LIP segment, and individuals in the LIP segment are more willing to make repeat trips to the same destination region than individuals in the CBB segment. The wetland land-cover accessibility measure has a negative effect in all segments, *i.e.*, destinations with higher wetland land cover lead to a higher satiation effect (less repeat visits to such regions by the same individual) than destinations regions with a lower wetland land cover. This is not surprising, given the negative characteristics associated with wetland areas in New Zealand. Finally, among the satiation parameters, the effect of the land-cover diversity accessibility index variable indicates that individuals in the first HFLFC segment get less satiated with (willing to make more repeat visits to) destination regions with a high diversity in activity type opportunities (as proxied by land cover percentages), while individuals in the third CBB segment get satiated very quickly with (are unlikely to make repeat visits to) destination regions with a high diversity.

4.4.4 Spatial Dependence

The spatial autoregressive coefficient, as expected, is positive, of the order of 0.10, and is different from zero at about the 7% level of significance for a one-tailed test.

4.4.5 Summary and Implications for Increasing Destination Competitiveness

A number of summary observations may be made from the model. First, the presence of a ferry leg appears to increase the attractiveness of a destination region for young single individuals and young parents (individuals in the HFLFC and LIP segments), but has relatively little attractive value for older baby-boomers. Of course, this is after controlling for the total cost of travel, which itself does have a very significant negative impact on destination region choice (especially for the LIP segment). Second, regions with high urban land cover are in general very attractive as a leisure trip destination. This is also true of regions with good forest cover; such regions have the highest attractive value for individuals in the first HFLFC segment and the least attractive value for individuals in the CBB segment. Third, regions with high wetland land cover lowers attractive value across the board, while regions with high agricultural land cover appeal substantially to middle-aged couples (individuals in the CBB segment) but “push away” young individuals in general and young parents in particular, presumably because agricultural lands in New Zealand correspond quite a bit to vineyards. Finally, the combined effects of the land-cover diversity index on the baseline and satiation function, as well as the constant coefficients in the satiation function, imply that individuals in the HFLFC segment place a premium on diversity of opportunities in terms of the types of activities offered by a destination region, and are much more willing to be loyal to a destination region that offers that diversity (if they make multiple leisure trips). On the other hand, the LIP and CBB segments are much less interested in diversity of activity type opportunities within a destination region, though they also look more for diversity in terms of destination regions visited in general. The individuals in the CBB group in particular are averse to repeat-visiting regions with high diversity of activity opportunities.

The kinds of insights above offered by our proposed model can be valuable in branding and marketing campaigns. As a simple illustration, consider two of the most popular destination regions: Auckland and Nelson. Auckland has a higher diversity in activity opportunities as proxied by land-cover percentages (a diversity index of 0.38) than does Nelson (a diversity index

of 0.29) which is heavily invested in forest land cover. Our results suggest that these two regions should use different strategies in their marketing and branding, as we discuss below.

Auckland should emphasize its “diversity uniqueness” when targeting the HFFLC group, perhaps by broadcasting customized media advertisements in high income neighborhoods all over New Zealand and having promotional flyers at bars and clubs where young singles spend quite a bit of time. This will serve Auckland well given that individuals in the HFFLC segment desire diversity and can be very loyal to regions that offer that diversity. While doing so, Auckland should also highlight its forest and urban land cover very specifically, because these will make the region more attractive in the perception map of individuals in the HFFLC group. At the same time, given the LIP and CBB segments are much larger in size, Auckland has to also target these segments appropriately. For the LIP and CBB groups, the strategy would be similar to the HFLFC group in its emphasis on urban and forest-related tourism opportunities. However, unlike promotions targeted at the HFLFC group, the Auckland promotion campaigns toward these two groups would do well not to speak about the diversity of types of activity opportunities, and retain a high intensity of coverage of the urban and forest-related tourism opportunities. For the CBB group, it would behoove Auckland campaigns to play up the vineyards and orchards for wine-tasting and consuming tours (Auckland, in addition to its diversity, has a large percentage of its land area invested in agricultural land-use).

Nelson is mainly invested in forest land-cover, with substantial opportunities for adventurous pursuits in rough and rugged forest terrain. This should be the main focus of promotional campaigns in all three segments as opposed to any diversity campaigns. In the CBB segment, Nelson can play up its vineyards and wine-tasting tourism outlets. Another important marketing strategy for Nelson is to highlight its geographic proximity to the ferry landing in Picton, which is only a two-hour drive on the Queen Charlotte Drive that also happens to be one of the most picturesque drives in all of New Zealand. When promoting the region to the first HFLFC segment and the second LIP segment, Nelson should play up the ferry crossing experience, given that the ferry experience has a positive influence on destination region choice for the first two segments. Playing up the scenic experience also can temper negative travel time effects in general.

Of course, in addition to targeting appropriate individuals for promoting current destination attributes, each region can also consider enhancing the accessibility to opportunities

located within the region. For instance, take the case of Waikato, and consider ways that Waikato can make itself more competitive. But before investing in changing the number and type of offerings, Waikato needs to undertake a cost-benefit analysis including an estimation of the additional tourism share that may be “pulled” to Waikato in response to such an investment. The proposed model can be used to provide information for such a cost-benefit analysis. Specifically, consider the case where Waikato realizes that it is not very much invested in urban activity opportunities, which, based on our model results, is a significant determinant of tourist “pull”. The model can then be used to evaluate the increase that may be expected in total tourist trip share to Waikato (including repeat trips) due to a 20% increase in its urban land cover (through additional urban activity opportunities). To do so, for each individual in the sample, we predict the number of trips attracted to Waikato in the base case and in the case of an increased urban land cover in the following steps: (1) for the base case, draw 500 realizations for all the stochastic terms in the utility function of Equation (1), (2) predict the number of trips to Waikato for each of the realizations using the prediction method of Pinjari and Bhat (2011), (3) average the predicted trips across the 500 realizations to obtain the individual prediction of the number of trips to Waikato, and (4) for the scenario case, increase the urban land cover percentage by 20%, drawing away an equivalent amount from agricultural land-use, (5) redo steps (1), (2), and (3) using the scenario sample, keeping the same 500 realizations for all the stochastic terms as in the base case. Then, from the individual-level predictions for the base and scenario cases, obtain the total Waikato trips in the two cases by aggregating across all individuals in the sample. Finally, we can obtain a pseudo-elasticity effect by taking the change in total trips to Waikato between the scenario and base cases as a percentage of the total trips to Waikato in the base case. This percentage turns out to be 16.1% (standard error of 1.7%) from the proposed model. As a point of reference, the corresponding percentage is estimated to be 13.3% (standard error of 1.2%) in the LC-MDCP model and 11.5% (standard error of 1.5%) in the RC-MDCP model.¹² Clearly, there are important differences among the models in the policy predictions, with the LC-MDCP and RC-MDCP models under-predicting the effectiveness of an increase in urban opportunities relative to the proposed FDMN-MDCP model. As we will see next, given that the proposed

¹² To be sure, in this part of the analysis, we did not just consider random coefficients on the variables in the baseline utility function, but also tested demographic variable interactions with the variables to obtain a RC-MDCP model that accommodates systematic heterogeneity (when found statistically significant) in the coefficients. This RC-MDCP specification is a much more appropriate “strawman” to compare with the proposed FDMN-MDCP model than a pure random coefficients specification that does not consider these interactions.

model fits the data much better than the other two models, the implication is that tourism policies to increase urban opportunities may be inappropriately discarded if the simpler LC-MDCP and RC-MDCP models were to be used.

4.5 Data Fit Comparisons with the LC-MDCP and the RC-MDCP Models

The difference in policy sensitivity results between the FDMN-MDCP, LC-MDCP, and RC-MDCP models suggests the need to apply formal statistical tests to determine the structure that is most consistent with the data. In this section, we provide measures of fit for these models. For the RC-MDCP model, as we already indicated in a footnote earlier, we consider both observed and unobserved heterogeneity in the “strawman” specification

The LC-MDCP and the proposed model can be compared using the familiar likelihood ratio test, since the former is a restricted version of the latter with no continuous random heterogeneity in coefficients within each segment. For the test between the RC-MDCP and the proposed model, one can compute the adjusted likelihood ratio index with respect to the log-likelihood at equal shares:

$$\bar{\rho}^2 = 1 - \frac{\mathcal{L}(\hat{\boldsymbol{\theta}}) - M}{\mathcal{L}(c)}, \quad (22)$$

where $\mathcal{L}(\hat{\boldsymbol{\theta}})$ is the log-likelihood function at convergence, $\mathcal{L}(c)$ is the log-likelihood for the naïve unsegmented model with only the size measure in the baseline function, only the constant in the satiation function, no spatial dependence, and IID errors across regions as in Equation (18), and M is the number of parameters estimated in the model minus two (that is, minus the single size coefficient in the baseline utility and the single satiation constant estimated in the naïve unsegmented model). To test the performance of the two non-nested models (*i.e.* the proposed FDMN-MDCP and RC-MDCP models) statistically, the non-nested adjusted likelihood ratio test may be used. This test determines if the adjusted likelihood ratio indices of two non-nested models are significantly different. In particular, if the difference in the indices is $(\bar{\rho}_2^2 - \bar{\rho}_1^2) = \tau$, then the probability that this difference could have occurred by chance is no larger than $\Phi\{-[-2\tau\mathcal{L}(c) + (M_2 - M_1)]^{0.5}\}$ in the asymptotic limit. A small value of the probability of chance occurrence indicates that the difference is statistically significant and that the model with the higher value of adjusted likelihood ratio index is to be preferred.

The likelihood ratio test (for the comparison of the LC-MDCP and FDMN-MDCP models) and non-nested adjusted likelihood ratio test (for the comparison of the RC-MDCP and FDMN-MDCP models) constitute disaggregate measures of fit that consider performance at the multivariate and disaggregate level of all combinations of regions. While the best data fit measures, these are not very intuitive. So, we also evaluate the performance of the three models intuitively and informally at an aggregate level. However, since there are too many multivariate combinations possible of leisure trip-making to the destination regions and it is impossible to provide fit statistics for all these combinations, we compare the aggregate marginal bivariate predictions (with the true sample values) for combinations of two of the most visited regions – Waikato and Auckland. Specifically, we focus on the percentage of individuals who, during the four-week survey period, visit Waikato but not Auckland, Auckland but not Waikato, both Auckland and Waikato, and neither of the two. The prediction procedure is similar to the one used for undertaking the sensitivity analysis in the previous section, except that, for each individual, we compute the probability of visiting each of the four combinations of regions as the percentage of times in the 500 realizations that each of the combinations has a non-zero number of visits. The probabilities for each combination are added up across individuals to obtain the predicted number of individuals falling into each combination category and compared with the actual percentages using the mean absolute percentage error (MAPE) statistic.

The results of the data fit comparisons are presented in Table 8. The first row provides the log-likelihood for the naïve unsegmented model (that is, the $\mathcal{L}(c)$ value), which is, of course, the same across the three models. The second row indicates the superior performance of the proposed FDMN-MDCP model in terms of the convergent log-likelihood value, as does the adjusted likelihood ratio index in the fifth row (note that the small magnitude of this index is not surprising, given the multitude of different possible multivariate combinations). The sixth row formally shows the likelihood ratio test result of the comparison of the FDMN-MDCP model with the LC-MDCP model, indicating the clear dominance of the FDMN-MDCP data fit. The same result is obtained in the next row through a non-nested adjusted likelihood ratio test comparing the FDMN-MDCP model with the RC-MDCP model; the probability that the adjusted likelihood ratio index difference between these models could have occurred by chance is literally zero. Finally, the last panel of the table first shows the actual percentages of individuals falling in each combination of visiting/not visiting the Waikato and Auckland regions, followed by the

predicted percentages from the three different models. The MAPE values from the three models are provided in the last row of the table. The LC-MDCP models has a MAPE value that is about three times that of the FDMN-MDCP, while the RC-MDCP model has a MAPE that is about 3.5 times that of the FDMN-MDCP.

All the fit measures discussed thus far are based on model fit on the overall sample used in estimation. While taken together, these fit measures reveal the superiority of the proposed FDMN-MDCP model, there is still a small possibility that the better performance of our model is simply an artifact of overfitting and may not translate to predictive accuracy in other samples. To accommodate for this, we also evaluated the performance of the three models on various market segments of the estimation sample (such predictive fit tests are sometimes referred to as market segment prediction tests). The intent of using such predictive tests is to examine the performance of different models on sub-samples that do not correspond to the overall sample used in estimation. Effectively, the sub-samples serve a similar role as an out-of-sample for validation. The advantage of using the sub-sample approach rather than an out-of-sample approach to validation is that there is no reduction in the size of the sample for estimation. This is particularly an issue in models of the type estimated in this paper because of the need to use as much information as possible given the number of parameters to be estimated. If a model shows superior performance in the subsamples in addition to the overall estimation sample, it is indication that the model indeed provides a better data fit. To do so, we computed the mean absolute percentage error (MAPE) for the percentage of individuals predicted to visit the same four combinations of the two destinations as in Table 8 and for three segmentations of demographic variables: (1) income less than NZ\$50,000 and income greater than NZ\$50,000, (2) nuclear and non-nuclear households, and (3) age less than 48 years and age more than 48 years. The overall MAPE values for percentage of individuals predicted to visit the four destination combinations in the two income segments were 10.4% and 10.8% from the FDMN-MDCP model, 28.7% and 28.8% from the LC-MDCP model, and 36.1% and 36.3% from the RC-MDCP model. The corresponding values for the household structure segmentation were 10.7% and 10.6% from the FDMN-MDCP model, 29.0% and 28.7% from the LC-MDCP model, and 36.2% and 35.3% from the RC-MDCP model, and for the age segmentation were 9.8% and 9.5% from the FDMN-MDCP model, 28.3% and 27.9% from the LC-MDCP model, and 36.0% and 35.1%

from the RC-MDCP model. All in all, the FDMN-MDCP model clearly outperforms the other two models even in such a predictive exercise.

5. CONCLUSIONS

This paper has proposed a new econometric formulation and a complete blueprint of an associated estimation method for a finite discrete mixture of normals version of the multiple discrete-continuous probit (or FDMN-MDCP) model. The model allows consumers to choose multiple alternatives at the same time, along with the continuous dimension of the amount of consumption, and captures heterogeneity in the response coefficients of the baseline utility function. This is a very general way of including heterogeneity in the sensitivity to exogenous variables in the multiple discrete-continuous context, with the normally distributed random parameters approach and the latent class approach constituting special cases.

A simulation exercise is undertaken to evaluate the ability of the proposed approach to recover parameters from simulated datasets. The results from the experiments show that the proposed inference approach, which is computationally fast and straightforward to implement, does very well in recovering the true parameters used in the data generation. Also, the simulation results show that ignoring the continuous component of the mixing (as reflected in the LC-MDCP model) or ignoring the discrete component of the mixing (as in the RC-MDCP model) when the true data is generated using an FDMN MDCP structure leads to substantial parameter bias. The average absolute percentage bias (APB) for the LC-MDCP model is about 28.5%, and for the RC-MDCP model is 26%, relative to the APB for the correct FDMN-MDCP model which is of the order of 3%. Clearly, the repercussion of imposing incorrect restrictions is very severe on parameter bias.

The paper demonstrates the application of the proposed approach through a study of individuals' recreational (*i.e.*, long distance leisure trips of over 25 miles one-way) choice among alternative destination locations and the number of trips to each recreational destination location, using data drawn from the 2012 New Zealand Domestic Travel Survey (DTS). The Bayesian Information Criterion indicates that the preferred specification is a three-segment solution, with one segment loading on high flying low family commitment (HFLFC) individuals, the second on low income parents (LIP), and the third on couple baby-boomers (CBB). In a comparative

empirical assessment of the FDMN-MDCP with the simpler LC-MDCP and RC-MDCP models, the FDMN-MDCP came out clearly as the winner in terms of data fit.

The results of the preferred three-segment solution showed heterogeneity (in the form of a continuous normal distribution) in sensitivity to cost and urban land cover within each latent segment, and differences (across the three latent segments) in the response to the presence of a ferry ride, travel cost, land cover accessibility measures, and the land cover diversity accessibility index. These differences, in combination with the socio-demographic characteristics of individuals in each segment, provide important information for effective targeting and strategic positioning to increase destination competitiveness. More generally, the FDMN-MDCP formulation appears to be a valuable methodology for marketing and positioning in markets that are characterized by multiple discreteness. Future research should focus on applying the FDMN-MDCP formulation to other multiple discrete contexts. Also, while the application to recreational destination choice in this paper demonstrates the value of the formulation, future work should consider a much richer set of destination region attributes.

ACKNOWLEDGMENTS

This research was partially supported by the U.S. Department of Transportation through the Data-Supported Transportation Operations and Planning (D-STOP) Tier 1 University Transportation Center. The first author would like to acknowledge support from a Humboldt Research Award from the Alexander von Humboldt Foundation, Germany. The authors are grateful to Lisa Macias for her help in formatting this document, and to anonymous referees who provided useful comments on an earlier version of the paper.

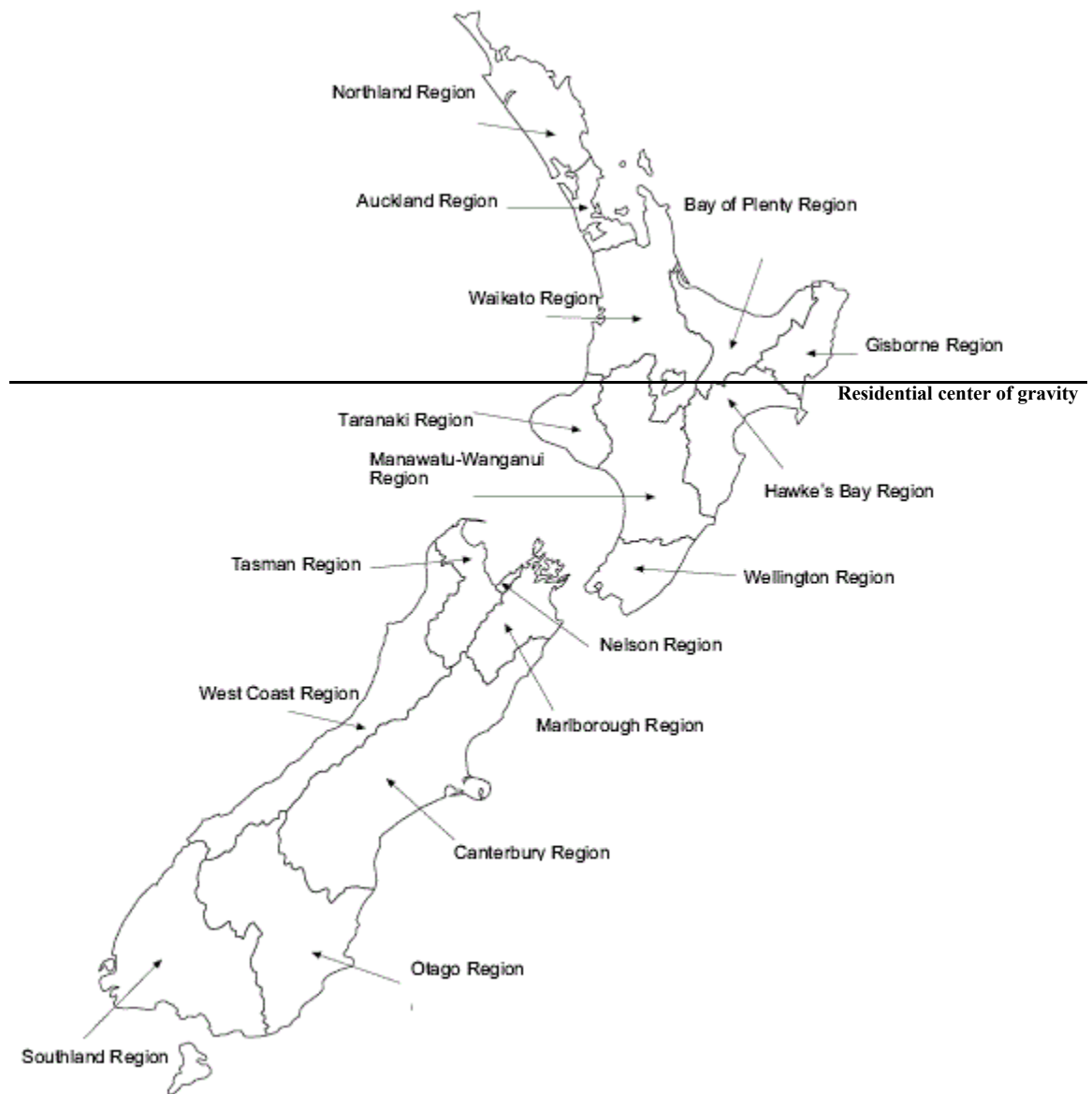
REFERENCES

- Amador, F.J., R. Gonzales and J. Ortuzar (2005). Preference heterogeneity and willingness to pay for travel time savings. *Transportation*, 32(6), 627-647.
- Allenby, G.M. (1990). Cross-validation, the Bayes theorem, and small-sample bias. *Journal of Business & Economic Statistics*, 8(2), 171-178.
- Balcombe, K., A. Chalak and I.M. Fraser (2009). Model selection for the mixed logit with Bayesian estimation. *Journal of Environmental Economics and Management*, 57(2), 226-237.
- Bastin, F., C. Cirillo and P.L. Toint (2010). Estimating non-parametric random utility models, with an application to the value of time in heterogeneous populations. *Transportation Science*, 44(4) 537-549.
- Bhat, C.R. (1997). An endogenous segmentation mode choice model with an application to intercity travel. *Transportation Science*, 31(1), 34-48.
- Bhat, C.R. (2000), A multi-level cross-classified model for discrete response variables, *Transportation Research Part B*, 34(7), 567-582.
- 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. (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, 923-939.
- Bhat, C.R., and R. Gossen (2004). A mixed multinomial logit model analysis of weekend recreational episode type choice. *Transportation Research Part B*, 38(9), 767-787.
- Bhat, C.R., and R. Sidharthan (2012). A new approach to specify and estimate non-normally mixed multinomial probit models. *Transportation Research Part B*, 46, 817-833.
- Bhat, C.R., M. Castro and M. Khan (2013). A new estimation approach for the multiple discrete-continuous probit (MDCP) choice model. *Transportation Research Part B*, 55, 1-22.
- Bhat, C.R., A. Govindarajan, and V. Pulugurta (1998). Disaggregate attraction-end choice modeling. *Transportation Research Record*, 1645, 60-68.
- Bujosa, A., A. Riera and R. L. Hicks (2010). Combining discrete and continuous representations of preference heterogeneity: a latent class approach. *Environmental and Resource Economics*, 47(4), 477-493.
- Campbell, D., E. Doherty, S. Hynes and T. Van Rensburg (2010). Combining discrete and continuous mixing approaches to accommodate heterogeneity in price sensitivities in environmental choice analysis. *84th Agricultural Economics Society Annual Conference*, March 29-31, Edinburgh, Scotland.

- Castro, M., C.R. Bhat, R.M. Pendyala and S.R. Jara-Diaz (2012). Accommodating multiple constraints in the multiple discrete-continuous extreme value (MDCEV) choice model. *Transportation Research Part B*, 46(6), 729-743.
- Cherchi, E., C. Cirillo and J. Polak (2009). User benefit assessment in presence of random taste heterogeneity: comparison between parametric and nonparametric models. *Transportation Research Record*, 2132, 78-86.
- Ferdous, N., R.M. Pendyala, C.R. Bhat, and K.C. Konduri (2011) Modeling the Influence of Family, Social Context, and Spatial Proximity on Use of Nonmotorized Transport Mode. *Transportation Research Record*, 2230, 111-120.
- Fonseca, J. R. (2010). On the performance of information criteria in latent segment models. *World Academy of Science, Engineering and Technology*, 63.
- Fosgerau, M. (2005). Unit income elasticity of the value of travel time savings. Presented at 8th NECTAR Conference, Las Palmas G.C., June 2-4.
- 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.
- Frühwirth-Schnatter, S. (2011). Label switching under model uncertainty. *Mixtures: Estimation and Application*, 213-239.
- Geweke, J. and M. Keane (1999) Mixture of normals probit models. In C. Hsiao, M.H. Pesaran, K.L. Lahiri, and L.F. Lee (eds.) *Analysis of Panel and Limited Dependent Variables*, 49-78, Cambridge University Press, Cambridge.
- Greene, W.H. and D.A. Hensher (2003). A latent class model for discrete choice analysis: contrasts with mixed logit. *Transportation Research Part B*, 37(8), 681-698.
- Greene, W.H. and D.A. Hensher (2013). Revealing additional dimensions of preference heterogeneity in a latent class mixed multinomial logit model. *Applied Economics*, 45(14), 1897-1902.
- Greene W.H., D.A. Hensher and J.M. Rose (2006). Accounting for heterogeneity in the variance of the unobserved effects in mixed logit models (NW transport study data). *Transportation Research Part B*, 40(1), 75-92.
- 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., L. Pendleton, C. Mohn, J. Hilger, K. Kuriyama, D. Layton, C. Busch, and F. Vasquez (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.
- Hensher, D.A., J.M. Rose and W.H. Greene (2005). *Applied Choice Analysis: A Primer*. Cambridge University Press, Cambridge, U.K.
- Iso-Ahola, S.E. (1983). Towards a social psychology of recreational travel. *Leisure Studies*, 2(1), 45-56.

- Kim, J., G.M. Allenby and P.E. Rossi (2002). Modeling consumer demand for variety. *Marketing Science*, 21(3), 229-250.
- Kuriyama, K., W.M. Hanemann and J.R. Hilger (2010). A latent segmentation approach to a Kuhn-Tucker model: an application to recreation demand. *Environmental Economics and Management*, 60(3), 209-220.
- Kuriyama, K., Y. Shoji, and T. Tsuge (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., and C.R. Bhat (2012). A conceptual and methodological framework of leisure activity loyalty accommodating the travel context. *Transportation*, 39(2), 321-349.
- LaMondia, J.J., C.R. Bhat, and D.A. Hensher (2008). An annual time use model for domestic vacation travel. *Journal of Choice Modelling*, 1(1), 70-97.
- LaMondia, J.J., T. Snell, and C.R. Bhat (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.
- Ministry of Business, Innovation and Employment (2008). Tourism sector profile. Available at: <http://www.med.govt.nz/sectors-industries/tourism/tourism-research-data/other-research-and-reports/sector-profiles>
- Ministry of Business, Innovation and Employment (2013). Statement of Intent. May 2013. Available at: <http://www.mbie.govt.nz/about-us/publications/soi/Ministry-of-Business-Innovation-and-Employment-SOI-2013.pdf>
- Mittelhammer, R. C., and G. Judge (2011). A family of empirical likelihood functions and estimators for the binary response model. *Journal of Econometrics*, 164(2), 207-217.
- Neath, A. A. and Cavanaugh, J. E. (2012). The Bayesian information criterion: Background, derivation, and applications. *WIREs Computational Statistics* 4, 199-203.
- New Zealand Tourism Strategy (2015). Available at: http://www.tianz.org.nz/content/library/FINAL_NZTS2015_HR.pdf
- Peters, M., and B.R. Clarkson (2010). Wetland restoration: a handbook for New Zealand freshwater system. Lincoln, N.Z: Manaaki Whenua Press. ISBN 978-0-478-34706-7.
- 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., and C.R. Bhat (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.
- Sobhani, A., N. Eluru and A. Faghih-Imani (2013). A latent segmentation based multiple discrete continuous extreme value model. *Transportation Research Part B*, 58, 154-169.
- Statistics New Zealand (2013). Tourism Satellite Account. Tauranga Aotearoa, Wellington, New Zealand.

- Torres, C., N. Hanley and A. Riera (2011). How wrong can you be? Implications of incorrect utility function specification for welfare measurement in choice experiments. *Journal of Environmental Economics and Management*, 62(1), 111-121.
- Tourism Industry Association (2012). New Zealand Tourism Industry Association Annual Report. Available at:
https://www.tianz.org.nz/content/library/TIA_Annual_Report_1112v11_ONLINE.pdf
- Train, K.E. (1998). Recreation demand models with taste variation. *Land Economics*, 74, 230-239.
- Train, K.E. (2008). EM algorithms for nonparametric estimation of mixing distributions. *Journal of Choice Modelling*, 1(1), 40-69.
- Train, K.E. and G. Sonnier (2005). Mixed logit with bounded distributions of correlated partworths. In *Applications of Simulation Methods in Environmental and Resource Economics*, Scarpa, R., Alberini, A., (eds.), Ch. 7, pp. 117-134, Springer, Dordrecht, The Netherlands.
- Van Nostrand, C., V. Sivaraman, and A. R. Pinjari (2013). Analysis of long-distance vacation travel demand in the United States: a multiple discrete-continuous choice framework. *Transportation*, 40(1), 151-171.
- 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., and D.J. Phaneuf (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.
- Wafa, Z., C.R. Bhat, R.M. Pendyala, and V.M. Garikapati (2015). Latent-segmentation-based approach to investigating spatial transferability of activity-travel models. *Transportation Research Record*, 2493, 136-144.
- 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., D.J. Phaneuf, C.F. Dumas, J. Herstein, J. Hill, and B. Buerger (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.
- Xiong, Y. and F.L. Mannering (2013). The heterogeneous effects of guardian supervision on adolescent driver-injury severities: A finite-mixture random-parameters approach. *Transportation Research Part B*, 49, 39-54.



Source: www.stats.govt.nz

Figure 1. Boundaries of New Zealand Regions

Table 1a. Evaluation of the ability to recover true parameters for the two-segment case

Parameter	True Value	MACML Method				
		Parameter Estimates		Standard Error Estimates		
		Mean Estimate	Absolute Percentage Bias (APB)	Asymptotic Standard Error (ASE)	Finite Sample Standard Error (FSSE)	Absolute Percentage Bias of Asymptotic Standard Error (APBASE)
b_{11}	1.000	1.063	6.3%	0.157	0.136	14.8%
b_{12}	2.000	1.997	0.2%	0.387	0.438	11.7%
b_{13}	0.600	0.586	2.4%	0.063	0.061	3.2%
b_{21}	2.000	1.901	4.9%	0.419	0.407	2.9%
b_{22}	1.500	1.503	0.2%	0.141	0.136	3.5%
b_{23}	0.200	0.196	2.2%	0.032	0.035	8.2%
\tilde{b}	0.500	0.500	0.0%	0.013	0.012	12.4%
$l_{\Omega_1,2}$	0.500	0.476	4.7%	0.055	0.057	2.5%
$l_{\Omega_1,3}$	0.866	0.865	0.2%	0.040	0.044	9.6%
$l_{\Omega_1,4}$	0.700	0.666	4.9%	0.049	0.036	35.6%
$l_{\Omega_1,5}$	0.519	0.529	2.0%	0.051	0.049	3.0%
$l_{\Omega_1,6}$	0.374	0.378	1.0%	0.026	0.028	5.9%
$l_{\Omega_2,1}$	0.900	0.898	0.2%	0.023	0.021	10.7%
$l_{\Omega_2,2}$	0.600	0.598	0.4%	0.031	0.032	3.5%
$l_{\Omega_2,3}$	0.800	0.796	0.5%	0.021	0.020	4.0%
$l_{\Omega_2,4}$	0.800	0.795	0.6%	0.025	0.025	0.8%
$l_{\Omega_2,5}$	0.400	0.392	2.1%	0.021	0.018	17.2%
$l_{\Omega_2,6}$	0.300	0.299	0.3%	0.015	0.014	4.5%
μ_{21}	0.600	0.505	15.8%	0.206	0.183	12.9%
μ_{22}	0.100	0.112	11.9%	0.046	0.050	8.2%
γ_1	1.000	1.038	3.8%	0.125	0.115	8.4%
γ_2	1.000	1.008	0.8%	0.146	0.136	7.0%
γ_3	1.000	1.103	10.3%	0.396	0.395	0.2%
Overall Mean Value Across Parameters		3.2% (for APB)		0.115	0.113	7.8%
Mean Time (mins)	18.3					
Std. dev of Time	7.5					
% of Runs Converged	100%					

Table 1b. Evaluation of the ability to recover true parameters for the three-segment case

Parameter	True Value	MACML Method				
		Parameter Estimates		Standard Error Estimates		
		Mean Estimate	Absolute Percentage Bias (APB)	Asymptotic Standard Error (ASE)	Finite Sample Standard Error (FSSE)	Absolute Percentage Bias of Asymptotic Standard Error (APBASE)
b_{11}^*	1.000	1.003	0.3%	0.296	0.285	4.0%
b_{12}^*	2.000	1.885	5.8%	0.433	0.421	2.7%
b_{13}^*	0.600	0.547	8.8%	0.215	0.192	11.7%
b_{21}^*	2.000	1.844	7.8%	0.599	0.566	5.9%
b_{22}^*	1.500	1.432	4.5%	0.295	0.282	4.5%
b_{23}^*	0.200	0.206	3.2%	0.093	0.092	1.3%
b_{31}^*	3.000	3.378	12.6%	0.082	0.080	3.0%
b_{32}^*	1.300	1.235	5.0%	0.032	0.038	14.1%
b_{33}^*	0.300	0.346	15.3%	0.175	0.155	12.7%
\tilde{b}^*	0.500	0.499	0.3%	0.023	0.024	3.5%
$l_{\Omega_1,2}^*$	0.500	0.506	1.3%	0.137	0.124	10.1%
$l_{\Omega_1,3}^*$	0.866	0.863	0.4%	0.109	0.104	4.5%
$l_{\Omega_1,4}^*$	0.700	0.675	3.5%	0.082	0.072	14.7%
$l_{\Omega_1,5}^*$	0.519	0.500	3.7%	0.081	0.090	9.6%
$l_{\Omega_1,6}^*$	0.374	0.383	2.4%	0.056	0.052	7.6%
$l_{\Omega_2,1}^*$	0.900	0.920	2.3%	0.098	0.093	5.4%
$l_{\Omega_2,2}^*$	0.600	0.582	3.1%	0.075	0.074	0.4%
$l_{\Omega_2,3}^*$	0.800	0.790	1.3%	0.047	0.054	13.1%
$l_{\Omega_2,4}^*$	0.800	0.784	2.0%	0.105	0.084	24.5%
$l_{\Omega_2,5}^*$	0.400	0.401	0.3%	0.087	0.086	1.4%
$l_{\Omega_2,6}^*$	0.300	0.304	1.3%	0.058	0.057	0.9%
$l_{\Omega_3,1}^*$	2.000	2.030	1.5%	0.057	0.061	6.1%
$l_{\Omega_3,2}^*$	0.500	0.575	14.9%	0.067	0.053	25.8%
$l_{\Omega_3,3}^*$	1.000	0.986	1.4%	0.070	0.072	3.7%
$l_{\Omega_3,4}^*$	0.600	0.580	3.3%	0.050	0.067	25.5%
$l_{\Omega_3,5}^*$	0.800	0.886	10.7%	0.341	0.400	14.7%
$l_{\Omega_3,6}^*$	0.900	1.060	17.8%	0.060	0.055	8.5%
μ_{21}^*	0.600	0.687	14.6%	0.541	0.503	7.6%
μ_{22}^*	0.100	0.112	12.3%	0.148	0.149	1.0%
μ_{31}^*	0.400	0.342	14.4%	0.225	0.201	11.9%
μ_{32}^*	0.000	0.010	10.0%	0.008	0.007	14.2%
γ_1	1.000	1.041	4.1%	0.267	0.255	4.7%
γ_2	1.000	1.145	14.5%	0.329	0.308	6.8%
γ_3	1.000	1.156	15.6%	0.520	0.473	9.8%
Overall Mean Value Across Parameters		6.4% (for APB)		0.172	0.165	8.7%
Mean Time (mins)	72.4					
Std. dev of Time	19.6					
% of Runs Converged	100%					

Table 2a. Effects of ignoring continuous heterogeneity and non-normality in the two-segment model

Parameter	True Value	Latent Class MDCP (LC-MDCP) Model		Random Coeffs. MDCP (RC-MDCP) Model	
		Mean Estimate	Absolute Percentage Bias (APB)	Mean Estimate	Absolute Percentage Bias (APB)
b_{11}	1.000	1.203	20.3%	1.349	34.9%
b_{12}	2.000	1.543	22.9%	1.690	15.5%
b_{13}	0.600	0.890	48.3%	0.337	43.8%
b_{21}	2.000	1.293	35.4%	1.349	32.6%
b_{22}	1.500	1.402	6.5%	1.690	12.7%
b_{23}	0.200	0.289	44.5%	0.337	68.7%
\tilde{b}	0.500	0.654	30.8%	0.427	0.4%
$l_{\Omega_1,2}$	0.500	0.592	18.4%	0.376	24.8%
$l_{\Omega_1,3}$	0.866	0.965	11.4%	0.572	34.0%
$l_{\Omega_1,4}$	0.700	-- ^a	--	0.942	34.6%
$l_{\Omega_1,5}$	0.519	-- ^a	--	0.626	20.6%
$l_{\Omega_1,6}$	0.374	-- ^a	--	0.407	8.8%
$l_{\Omega_2,1}$	0.900	0.782	13.1%	-- ^b	--
$l_{\Omega_2,2}$	0.600	0.329	45.2%	0.376	37.3%
$l_{\Omega_2,3}$	0.800	0.764	4.5%	0.942	17.8%
$l_{\Omega_2,4}$	0.800	-- ^a	--	0.572	28.5%
$l_{\Omega_2,5}$	0.400	-- ^a	--	0.626	56.4%
$l_{\Omega_2,6}$	0.300	-- ^a	--	0.407	35.6%
μ_{21}	0.600	0.431	28.2%	-- ^c	--
μ_{22}	0.100	0.140	40.0%	-- ^c	--
γ_1	1.000	1.209	20.9%	1.030	3.0%
γ_2	1.000	1.823	82.3%	1.405	40.5%
γ_3	1.000	1.117	11.7%	1.017	1.7%
Overall Mean Value Across Parameters		28.5% (for APB)		26.0% (for APB)	
Mean (across 100 data sets) log-likelihood value at convergence		-39,517.923		-39,561.115	
Number of times the likelihood ratio test statistic favors the FDMN-MDCP model^d		All one hundred times when compared with $\chi^2_{6,0.95} = 12.59$		All one hundred times when compared with $\chi^2_{13,0.95} = 22.36$	

^a These parameters are not estimated and are fixed at 0.0 (see Section 3.1.3).

^b This parameter is fixed to 1.0 for identification.

^c These parameters are implicitly fixed to the value of minus infinity.

^d The mean (across data sets) log-likelihood value at convergence for the FDMN-MDCP model with a two-segment mixture is -38,927.438.

Table 2b. Effects of ignoring continuous heterogeneity and non-normality in the three-segment model

Parameter	True Value	Latent Class MDCP Model		Random Coeff. MDCP (RC-MDCP) Model	
		Mean Estimate	Absolute Percentage Bias (APB)	Mean Estimate	Absolute Percentage Bias (APB)
b_{11}^*	1.000	1.543	54.3%	1.738	73.8%
b_{12}^*	2.000	1.276	36.2%	1.277	36.1%
b_{13}^*	0.600	0.320	46.7%	0.246	59.0%
b_{21}^*	2.000	1.652	17.4%	1.738	13.1%
b_{22}^*	1.500	1.724	14.9%	1.277	14.9%
b_{23}^*	0.200	0.102	49.0%	0.246	23.1%
b_{31}^*	3.000	1.592	46.9%	1.738	42.1%
b_{32}^*	1.300	1.035	20.4%	1.277	1.8%
b_{33}^*	0.300	0.472	57.3%	0.246	17.9%
\tilde{b}^*	0.500	0.366	26.8%	0.418	16.4%
$l_{\Omega_{1,2}}^*$	0.500	0.411	17.8%	0.063	87.4%
$l_{\Omega_{1,3}}^*$	0.866	0.599	30.8%	0.926	6.9%
$l_{\Omega_{1,4}}^*$	0.700	-- ^a	--	0.322	54.0%
$l_{\Omega_{1,5}}^*$	0.519	-- ^a	--	0.618	19.1%
$l_{\Omega_{1,6}}^*$	0.374	-- ^a	--	0.546	46.1%
$l_{\Omega_{2,1}}^*$	0.900	0.724	19.6%	-- ^b	--
$l_{\Omega_{2,2}}^*$	0.600	0.326	45.7%	0.063	89.5%
$l_{\Omega_{2,3}}^*$	0.800	0.598	25.3%	0.322	59.7%
$l_{\Omega_{2,4}}^*$	0.800	-- ^a	--	0.926	15.7%
$l_{\Omega_{2,5}}^*$	0.400	-- ^a	--	0.618	54.5%
$l_{\Omega_{2,6}}^*$	0.300	-- ^a	--	0.546	82.1%
$l_{\Omega_{3,1}}^*$	2.000	1.396	30.2%	-- ^b	--
$l_{\Omega_{3,2}}^*$	0.500	0.398	20.4%	0.063	87.4%
$l_{\Omega_{3,3}}^*$	1.000	0.733	26.7%	0.322	67.8%
$l_{\Omega_{3,4}}^*$	0.600	-- ^a	--	0.926	54.3%
$l_{\Omega_{3,5}}^*$	0.800	-- ^a	--	0.618	22.7%
$l_{\Omega_{3,6}}^*$	0.900	-- ^a	--	0.546	39.3%
μ_{21}^*	0.600	0.467	22.2%	-- ^c	--
μ_{22}^*	0.100	0.156	56.0%	-- ^c	--
μ_{31}^*	0.400	0.298	25.5%	-- ^c	--
μ_{32}^*	0.000	0.017	17.0%	-- ^c	--
γ_1	1.000	1.327	32.7%	2.723	172.3%
γ_2	1.000	1.201	20.1%	2.959	195.9%
γ_3	1.000	1.102	10.2%	2.871	187.1%
Overall Mean Value Across Parameters		30.8% (for APB)		82.9% (for APB)	
Mean (across 100 data sets) log-likelihood value at convergence		-39,599.201		-39,797.634	
Number of times the likelihood ratio test favors the FDMN-MDCP model^d		All one hundred times when compared with $\chi_{9,0.95}^2 = 16.92$		All one hundred times when compared with $\chi_{27,0.95}^2 = 40.11$	

^a These parameters are not estimated and are fixed at 0.0 (see Section 3.1.3).

^b This parameter is fixed to 1.0 for identification.

^c These parameters are implicitly fixed to the value of minus infinity.

^d The mean (across data sets) log-likelihood value at convergence for the FDMN-MDCP model with a three-segment mixture is -39,001.232.

Table 3. Recreational Travel Number of Trips

Number of trips	Number of individuals	Number (%) of individuals visiting ^a				
		1 region	2 regions	3 regions	4 regions	5 regions
1	2,535 (72.3%)	2,535 (100%)	0	0	0	0
2	732 (20.9%)	342 (46.7%)	390 (53.3%)	0	0	0
3	180 (5.0%)	63 (35%)	87 (48.3%)	30 (16.7%)	0	0
4	47 (1.3%)	14 (29.8%)	23 (48.9%)	7 (14.9%)	3 (6.4%)	0
5	7 (0.2%)	1 (14.3%)	2 (28.6%)	3 (42.8%)	0	1 (14.3%)
6	3 (0.1%)	2 (66.7%)	1 (33.3%)	0	0	0
7	2 (0.1%)	0	1 (50%)	1 (50%)	0	0
10	2 (0.1%)	0	0	2 (100%)	0	0

^aPercentages add up to 100% in each row.

Table 4. Destination Region Characteristics

Island	Region	Travel Impedance Measures (Std. Dev.)			Land Cover Percentage					
		Travel Time (hours)	Travel Distance (miles)	Cost (NZ\$)	Urban	Water	Wetland	Agricultural	Bare-land	Forest
NORTH ISLAND	Northland	8.31 (6.53)	397.6 (303.9)	314.1 (334.2)	0.75	2.44	0.92	47.92	1.18	46.79
	Auckland	6.53 (6.64)	306.4 (295.6)	265.8 (316.4)	10.68	2.85	0.62	49.04	0.92	35.90
	Waikato	5.98 (5.97)	273.9 (257.8)	241.0 (281.3)	1.14	3.57	0.88	53.10	0.70	40.61
	Bay of Plenty	6.74 (5.60)	313.3 (237.2)	276.9 (272.4)	1.32	2.39	0.27	23.17	0.28	72.57
	Gisborne	7.82 (4.95)	366.2 (205.5)	322.1 (260.1)	0.35	0.36	0.41	46.44	1.55	50.89
	Taranaki	6.41 (4.32)	294.5 (169.9)	261.9 (219.1)	0.98	0.39	0.08	53.81	0.43	44.31
	Manawatu-Wanganui	6.07 (3.72)	279.7 (152.5)	249.2 (198.3)	0.67	0.48	0.32	60.22	0.82	37.50
	Hawke's Bay	6.30 (4.40)	290.5 (175.1)	258.7 (222.5)	0.59	0.92	0.22	53.93	0.64	43.70
	Wellington	6.46 (3.19)	301.5 (151.4)	266.7 (195.5)	2.53	1.34	0.23	47.82	0.72	47.36
SOUTH ISLAND	Tasman	9.70 (3.74)	392.2 (162.5)	386.1 (266.7)	0.34	1.34	1.29	19.58	3.37	74.08
	Nelson	9.58 (3.61)	388.0 (161.0)	381.6 (261.8)	6.48	3.12	0.25	13.72	0.93	75.49
	Marlborough	8.34 (3.43)	337.4 (156.3)	332.1 (237.3)	0.28	0.56	0.20	43.34	9.97	45.65
	West Coast	10.86 (4.83)	474.7 (210.6)	447.2 (322.4)	0.14	1.43	1.35	15.79	9.41	71.88
	Canterbury	10.48 (5.52)	443.6 (241.6)	425.3 (338.6)	0.71	2.09	0.36	65.67	12.12	19.05
	Otago	14.07 (6.56)	629.5 (291.1)	580.0 (425.9)	0.45	2.76	1.50	73.49	4.73	17.07
	Southland	16.40 (6.86)	749.7 (307.3)	680.9 (471.0)	0.24	2.99	0.98	43.97	4.50	47.32

Table 5. Recreational Travel Destination Choice and Number of Trips

Island	Destination Region	Area (miles ²)	Total number (%) of individuals visiting each region*	Number of visiting individuals per unit area (per miles ²)	Number of trips among those who visit each destination			
					Mean	Min.	Max.	Std. Dev.
NORTH ISLAND	Northland	5,383	290 (8.3%)	0.0539	1.16	1	4	0.44
	Auckland	2,162	575 (16.4%)	0.2660	1.17	1	6	0.49
	Waikato	9,883	788 (22.5%)	0.0798	1.19	1	7	0.53
	Bay of Plenty	4,806	454 (12.9%)	0.0945	1.20	1	8	0.61
	Gisborne	3,224	42 (1.2%)	0.0129	1.17	1	4	0.53
	Taranaki	2,808	104 (3.0%)	0.0370	1.12	1	3	0.35
	Manawatu-Wanganui	8,577	288 (8.2%)	0.0337	1.13	1	4	0.38
	Hawke's Bay	5,469	185 (5.3%)	0.0339	1.09	1	3	0.31
	Wellington	3,137	328 (9.4%)	0.1046	1.18	1	4	0.47
SOUTH ISLAND	Tasman	3,778	70 (2.0%)	0.0186	1.16	1	3	0.50
	Nelson	172	31 (0.9%)	0.1805	1.06	1	2	0.25
	Marlborough	4,820	74 (2.1%)	0.0153	1.07	1	2	0.25
	West Coast	9,010	77 (2.2%)	0.0085	1.13	1	4	0.47
	Canterbury	17,508	465 (13.3%)	0.0267	1.21	1	6	0.52
	Otago	12,351	260 (7.4%)	0.0210	1.22	1	6	0.56
	Southland	13,261	80 (2.3%)	0.0060	1.06	1	2	0.24

* Total percentage across all rows in this column add up to more than 100% because some travelers visit more than one destination region.

Table 6. Three Segments FDMN-MDCP Model Estimation Results

Variable	First Segment		Second Segment		Third Segment	
	Estimate	t-stat ⁺	Estimate	t-stat	Estimate	t-stat
<i>Segment Probabilities</i>						
Alternative specific constant	-	-	1.020	3.50	0.314	2.33
Age: 48 years or older	-	-	-	-	0.880	3.40
Single person household	-	-	-0.501	-3.69	-0.646	-3.50
Couple household	-	-	-	-	0.467	2.70
Nuclear family household	-	-	0.542	4.77	-	-
Single parent household	-	-	0.229	2.61	-	-
Income less than NZ \$50,000	-	-	1.250	2.30	0.604	3.12
<i>Baseline utilities</i>						
Logarithm of the area (miles ²) – mean	0.797*	4.72*	0.797*	4.72*	0.797*	4.72*
Ferry (dummy) – mean	0.102	2.40	0.121	3.20	-	-
Travel cost (\$/100) –mean	-0.700	-15.22	-0.821	-35.04	-0.780	-4.95
Travel cost (\$/100) – standard deviation	0.501	3.00	0.573	2.89	0.442	3.14
Land cover accessibility measure specific to						
Urban (/10 ⁴) –mean	0.431	2.43	0.429	2.09	0.457	2.64
Urban (/10 ⁴) – standard deviation	0.119	2.28	0.100	2.23	0.091	2.17
Forest (/10 ⁴) –mean	0.450	5.09	0.360	6.66	0.210	4.44
Wetland (/10 ⁴) –mean	-4.210	-3.23	-4.195	-5.10	-4.030	-2.69
Agricultural (/10 ⁴) –mean	-0.112	-4.91	-0.498	-9.15	0.212	3.59
Land-cover diversity accessibility index	0.270	2.69	-0.443	-2.32	-0.213	-2.16
<i>Satiation parameters ($\tilde{\theta}_g$ parameters)</i>						
Constant	1.802	27.20	1.789	25.42	1.672	23.11
Land cover accessibility measure specific to						
Wetland (/10 ⁴) –mean	-2.535	-3.56	-2.367	-4.10	-2.055	2.17
Land-cover diversity accessibility index	0.770	2.09	-	-	-0.231	-2.04
Spatial autoregressive coefficient (t-stat)	0.096 (1.56)					
Log-Likelihood at Convergence	-8,499.78					

* The size coefficient (coefficient corresponding to the logarithm of the area in miles²) is constrained to be equal across all segments. The t-statistic for this coefficient is with respect to the hypothesis that the coefficient is equal to one.

+ All coefficients are different from zero (or different from one in the case of the size variable) at the 95% confidence level or higher (or a p-value of 0.05 or lower). The 95% confidence level corresponds to an absolute t-statistic value of 1.96.

Table 7. Quantitative Characterization of the Three Segments

Segmentation Variable		First Segment	Second Segment	Third Segment	Overall Market
Age	Younger than 48	60.2%	61.8%	35.7%	46.4%
	48 years or older	39.8%	38.2%	64.3%	53.6%
Household structure	Single person	15.4%	13.8%	13.7%	14.0%
	Couple	22.9%	18.3%	41.5%	26.0%
	Nuclear family	45.0%	52.1%	31.6%	45.0%
	Single parent	5.1%	7.4%	4.7%	6.3%
	Multi family or non-family	11.6%	8.4%	8.5%	8.7%
Income	Less than NZ\$50,000	22.9%	42.6%	34.5%	37.8%
	NZ\$50,000 or more	77.1%	57.4%	65.5%	62.2%

Table 8. Measures of Fit $\Phi[-16.22] \ll 0.0001$

Summary Statistic	Estimation Sample											
	FDMN-MDCP		LC-MDCP		RC-MDCP							
Log-likelihood of the naïve unsegmented model	-15,783.21											
Log-likelihood at convergence	-8,499.78		-8,550.03		-8,648.46							
Number of parameters	49		40		15							
Number of observations	3,508											
Adjusted likelihood ratio index	0.458		0.455		0.451							
Predictive likelihood ratio test between FDMN-MDCP and LC-MDCP models	Test statistic $[-2*(LL_{LC-MDCP}-LL_{FDMN-MDCP})]=102 >$ Chi-Squared statistics with 9 degrees of freedom at any reasonable level of significance											
Non-nested adjusted likelihood ratio test between the FDMN-MDCP and RC-MDCP models	$\Phi[-16.22] \ll 0.0001$											
Percentage of individuals (trips) predicted to visit....	Actual percentage		Predicted percentage									
	Individuals	Trips	Individuals	Trips	Individuals	Trips	Individuals	Trips				
Waikato but not Auckland	16.9	17.9	17.6	18.4	20.0	20.5	21.4	22.3				
Auckland but not Waikato	10.8	8.1	12.3	9.5	14.0	11.8	14.5	12.4				
Both Auckland and Waikato	5.6	6.4	6.7	7.7	8.6	9.4	9.3	9.7				
Neither Auckland nor Waikato	66.7	67.6	63.4	64.4	57.4	58.3	54.8	55.6				
Mean Absolute Percentage Error	10.7%		11.3%		28.9%		30.2%		36.2%		36.7%	