An Integrated Choice and Latent Variable Model for Multiple Discrete Continuous Choice Kernels: Application Exploring the Association between Day Level Moods and Discretionary Activity Engagement Choices

Annesha Enam
Department of Civil and Environmental Engineering, University of Connecticut,
261 Glenbrook Road, Unit 3037, Storrs, CT 06269-3037
Ph: 860-486-2992; Fax: 860-486-2298
Email: annesha.enam@uconn.edu

Karthik C. Konduri (*Corresponding Author*)
Department of Civil and Environmental Engineering, University of Connecticut
261 Glenbrook Road, Unit 3037, Storrs, CT 06269-3037
Ph: 860-486-2733; Fax: 860-486-2298
Email: karthik.konduri@ uconn.edu

Abdul R. Pinjari
Department of Civil Engineering
Center for infrastructure, Sustainable Transportation and Urban Planning (CiSTUP)
Indian Institute of Science (IISc)
Bangalore 560012, India
Ph: +91-80-2293 2043; Fax: +91-80-2360 0404
Email: abdul@civil.iisc.ernet.in

Naveen Eluru
Department of Civil, Environmental and Construction Engineering, University of Central Florida
12800 Pegasus Drive, Room 301D, Orlando, Florida 32816
Ph: 407-823-4815; Fax: 407-823-3315
Email: naveen.eluru@ucf.edu

ABSTRACT
In the recent years, multiple discrete continuous (MDC) models have emerged as a popular framework to simultaneously model the choice of multiple goods (that are imperfect substitutes to one another) and the associated consumption quantities. The paper presents a new integrated choice and latent variable (ICLV) model implementation called the Hybrid Multiple Discrete Continuous (HMDC) model that is capable of incorporating the influence of psychological factors (modeled as latent constructs) on MDC choice behaviors. Estimation of ICLV models (with single discrete choice kernels and MDC kernels) has been a challenge owing to the high dimensional integrals involved in the likelihood function. The typically used maximum simulated likelihood estimation (MSLE) approach becomes cumbersome when the dimensionality of integration increases. In this research, a composite marginal likelihood (CML) based estimation approach is proposed for parameter estimation of the HMDC framework. Unlike the ICLV model implementations with single discrete choice kernel, the dimension of the integral to be decomposed in the HMDC varies across observations. This necessitated the use of weights when decomposing the likelihood function using the CML approach. A simulation study was conducted using synthetic datasets to demonstrate the superiority of the weighted CML approach over its unweighted counterpart in the presence of MDC choice kernel. The applicability of the proposed model formulation and associated estimation routine was demonstrated using an empirical case study with data from the 2013 American Time Use Survey (ATUS). The empirical study identifies interesting association between day level moods and discretionary activity participation decisions.

Keywords: latent variable; multiple discrete continuous (MDC) choice; composite marginal likelihood (CML); integrated choice and latent variable (ICLV) model; hybrid choice model (HCM)

1. INTRODUCTION

There is a growing interest in the field of travel behavior research to incorporate psychological factors including attitudes, perceptions, beliefs, knowledge, emotions and learning for explaining the activity and travel behaviors exhibited by individuals (McFadden 1986, Gärling 1998, Hess 2012). This interest is in part motivated by theoretical and methodological advances in behavioral economics that support the notion that heterogeneity in behavior is not just attributable to the socio-economic and demographic differences but is also due to the differences in the underlying psychological factors.

Earliest efforts aimed at incorporating psychological factors for explaining individual behaviors in the transportation field can be traced back to the work by Golob et al. (1977). More recently with growing concerns of non-renewable energy consumption and greenhouse gas emissions, researchers have attempted to study the role of attitudes such as “pro environmental”, “addiction to car” on different dimensions of travel behaviors namely mode choice and vehicle type choice (Bolduc et al. 2005, Anable 2005, Daly et al. 2012, Glerum and Bierlaire 2012, Atasoy et al. 2013, Alvarez-Daziano and Bolduc 2013, Kamargianni and Polydoropoulou 2013, Hess and Spitz 2016).

In most studies, the random utility maximization framework proposed by McFadden (1986) is used. Psychological factors are constructed (“measured”) from associated indicators using either summary measures (e.g., mean of all indicators) (Koppelman and Hauser 1978, Harris and Keane 1998) or data reduction techniques (e.g., factor analysis) (Madanat et al. 1995). The constructed factors are then included as explanatory variables in the RUM based model to study the relationship between the factors and the choice variables. It can be noted that, the indicators do not capture all aspects of the underlying psychological factors and are often associated with measurement errors. Consequently, inconsistent and inefficient parameter estimates are obtained if the measurement errors in the indicator variables are not explicitly accounted for in the model formulation (Ashok et al. 2002). In an effort to address the measurement error issue (and other limitations of the RUM framework), the Hybrid Choice Modeling (HCM) framework was developed (Walker and Ben-Akiva 2002, and Ben-Akiva et al. 2002). In this paper, the specific variant of the HCM framework (also referred to as Integrated Choice and Latent Variable (ICLV) model in the literature) that combines the Multiple Indicator Multiple Cause (MIMIC) model for constructing psychological factors with RUM based model for representing the choice variables is of interest.

Over the years, a number of implementations of ICLV models have been developed and applied to study the role of psychological factors on different dimensions of activity and travel choices (see Kim et al. 2014 for a review of recent progress in HCM). In most ICLV implementations, the choice component has been limited to a “single discrete” choice dimension (wherein individual makes a choice of a single alternative from available alternative set). However, numerous activity-travel choice situations (and more generally in other consumer behavior research arenas) are characterized by “multiple discreteness”; i.e., individuals potentially choose more than one alternative from the available choice set of alternatives. Additionally, for the selected alternatives, they also make the choice of how much of the alternative to “consume” subject to resource constraint(s) (Bhat 2005). Such choice dimensions are characterized as multiple discrete-continuous (MDC). In the literature, activity-travel behaviors are increasingly being characterized and modeled as MDC variables to accurately account for the underlying decision-making process (e.g. choice of goods under presence of budget constraints, and satiation effect among others). Examples of MDC choices include study of vehicle fleet composition and usage

(Bhat et al. 2009, Jäggi et al. 2012, Pinjari et al. 2016), activity participation and time allocation choices (Sener et al. 2008, Bhat et al. 2010), vacation types and time spent (LaMondia et al. 2008, Lingling et al. 2011), vacation destination choices (Von Haefen et al. 2004, Van Nostrand et al. 2013), and land use choices (Pinjari et al. 2009, Kaza et al. 2011) among others. The study of individual behaviors as MDC choices is also widespread in other fields such as marketing and economics. For example, Shin et al. (2016) study commodity bundling in Korean telecommunications market as MDC choices. Richards et al. (2012) use MDC models for a study of shopping behaviors. Jeong et al. (2011) and Biying et al. (2012) study energy consumption behaviors as MDC choices. Despite the growing popularity of study of consumer choices as MDC variables, there is lack of ICLV implementation in the literature that is able to accommodate a MDC choice kernel. *In this research, a new Hybrid Multiple Discrete Continuous (HMDC) choice model formulation and associated estimation routine are presented that allows the study of the influence of psychological constructs on MDC choice dimensions*.

The Maximum simulated likelihood estimation (MSLE) technique has served as the workhorse for evaluating integrals involved in the ICLV model implementations (Kim et al. 2014). The computational intensity of the MSLE approach has limited empirical researchers from exploring the full breadth of ICLV model specification, such as the number of latent variables to explore, interactions between latent variables and sociodemographic variables, and correlations among latent variables and among choice alternatives. To overcome the limitations of MSLE, alternative estimation approaches such as composite marginal likelihood approach (Bhat and Dubey 2014) and Bayesian approach (Daziano 2015) have been proposed in the recent years. The current exploration proposes a CML based estimation approach with analytical approximation for normal cumulative density function (known as MACML in the literature, due to Bhat 2011) similar to Bhat et al. (2016). Unlike Bhat et al. (2016), however, the current research employs a *weighted* CML approximation for estimating ICLV models with MDC kernel. *The current paper is perhaps the first to highlight and demonstrate the importance of weights in the composite marginal likelihood (CML) estimation routine for ICLV models*. The presented research demonstrates substantial gain in the parameter consistency offered by the weighted CML routine over the unweighted CML routine. Further discussion on the choice of weight and the comparison of results between parameter estimates of the HMDC model using a weighted and unweighted CML estimation technique are provided in Section 3 of the paper.

It should be noted that, though ICLV framework has been a mainstay for analyzing the influence of psychological factors on different choice dimensions, it has received its fair share of scrutiny. Chorus and Kroesen (2014) note that cross-sectional data only offer evidence of inter-person variabilities, as opposed to changes in individual-level behavior. Consequently, policy interventions aiming at altering the level of latent variables for changing the choice outcomes are not supported by cross-sectional data driven ICLV implementations. Despite the criticism, researchers have continued to highlight the importance of the ICLV framework in terms of its ability to better reflect on consumer behavior (Bolduc and Daziano 2010). More recently, Vij and Walker (2016) conducted a systematic analysis based on multiple synthetic datasets to highlight the contribution of the ICLV framework over the traditional RUM based choice models that leave the source of heterogeneity to unobserved error components. The authors reemphasize the importance of the ICLV framework for lending structure to the underlying heterogeneity and for decomposing the influence of observed variables into constituent components, each of which might be attributable to different latent constructs. Given the statistical rigor and potential appeal in disentangling the structure in the unobserved heterogeneity, the current research attempts to add to the body of ICLV modeling and estimation approaches.

In addition to the aforementioned methodological contribution, this paper demonstrates the applicability of the proposed HMDC model using the 2013 American Time Use Survey dataset to explore the association between individuals' experienced moods (such as happiness, sadness, pain, stress and tiredness) and their discretionary activity engagement and time allocation in a day.

The rest of the paper is organized as follows. Section 2 presents the HMDC model formulation along with the proposed approach to parameter estimation and inference. Section 3 presents a simulation study to demonstrate the ability of the proposed estimation approach to recover consistent and efficient estimates of the parameters. Section 4 presents an empirical application of the HMDC model to explore association between individuals' moods and their time-use. The estimated HMDC model is also validated using a holdout sample in this section. Section 5 concludes the paper with a summary of contributions, findings, and avenues for future research.

2. METHODOLOGY

The HMDC formulation extends the existing ICLV model implementations by replacing the single discrete choice kernel with a MDC choice kernel. The choice kernel in the HMDC assumes the multiple discrete continuous probit (MDCP) structure proposed by Bhat et al. (2013). The estimation of the HMDC model proceeds by combining the pairwise composite marginal likelihood (CML) (Varin 2008) with the maximum approximated composite marginal likelihood (MACML) (Bhat 2011). *However, unlike Bhat et al. (2016), parameter estimation for HMDC employs a weighted version of the pairwise CML approximation*. Below, the model formulation is presented followed by a discussion of the approach for estimating model parameters.

2.1 Model Formulation

Similar to the ICLV framework, HMDC model formulation consists of three main components: 1) structural equation model of the latent variables, 2) measurement equation model of the latent variables and 3) MDC choice model. Figure 1 presents an overview of the proposed HMDC model formulation. In the traditional ICLV model, the utility in the choice model component is measured via one choice indicator whereas in the HMDC model, the utility of the choice model is measured via consumption quantities of multiple alternatives as indicators. In the remaining subsections (in 2.1.1, 2.1.2 and 2.1.3), the formulation of each of the three components of the HMDC model is presented in detail.

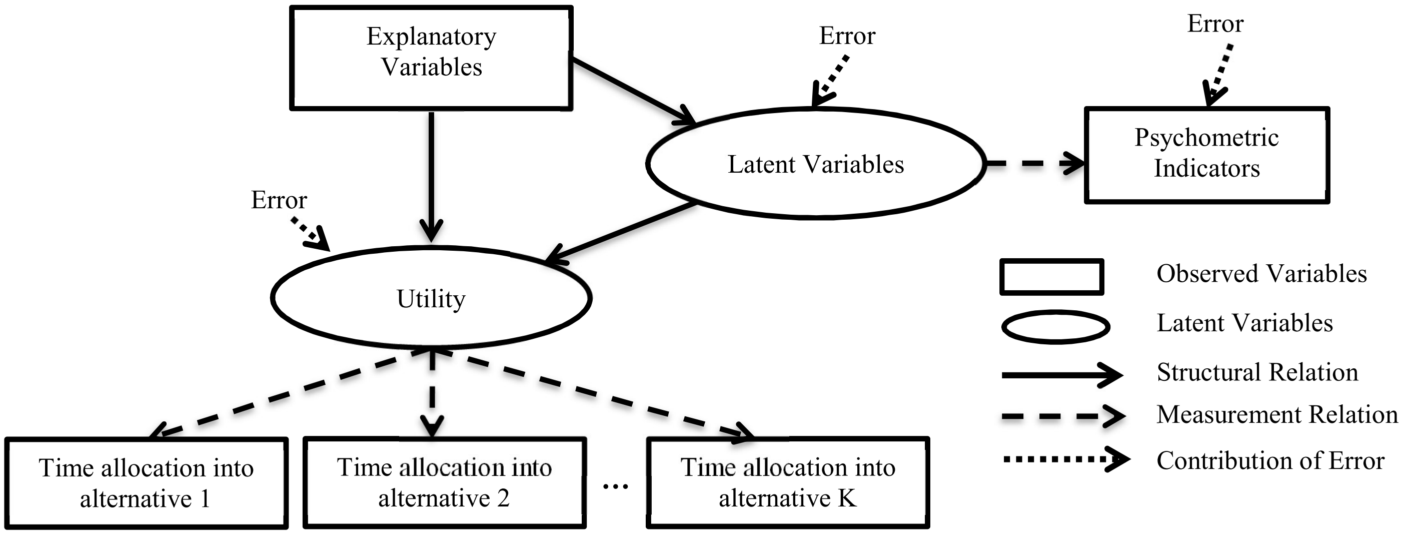

FIGURE 1: Hybrid multiple discrete continuous (HMDC) model framework.

2.1.1 Structural Equation Model of the Latent Variables

Equation 1 shows the structural equation of the latent variables in matrix form[1].

$$z^* = \omega\rho + \eta \tag{1}$$

where z^* is a $(L\times1)$ vector of latent psychological factors, ω is a $(L\times D)$ matrix of observed covariates for explaining the variability in the psychological factors, ρ is a $(D\times1)$ vector of coefficients associated with the observed covariates and η is a $(L\times1)$ vector of random error terms associated with the latent factors. η is assumed to be multivariate normally distributed: $\eta \sim N[0_L, \Gamma]$ with Γ representing the correlation matrix[2].

2.1.2 Measurement Equation Model of the Latent Variables

In the proposed HMDC formulation, latent factors can be constructed from both continuous and ordinal indicator variables. The measurement equation for the continuous indicators used to construct the latent variables (in matrix form) is shown in Equation 2.

$$y_c = \delta + dz^* + \xi \tag{2}$$

where y_c is a $(H\times1)$ vector of continuous indicators, δ is a $(H\times1)$ vector of constant terms, d is a $(H\times L)$ matrix of latent variable loadings onto the continuous indicators (commonly referred to as factor loadings), and ξ is a $(H\times1)$ vector of error terms. ξ is also assumed to be multivariate normally distributed: $\xi \sim N\left[0_H, \Sigma_{y_c}\right]$ with Σ_{y_c} representing the covariance matrix. For identification purposes Σ_{y_c} is assumed to be a diagonal matrix[2].

[1] In presenting the model formulation, the subscript for the individual is suppressed for the sake of brevity.

[2] The identification conditions are similar to those of a MIMIC model; please see Bollen (1983) for a detailed discussion about the identification conditions for the MIMIC model.

Equation 3 shows the measurement equation for the ordinal indicators used to construct the latent variable.

$$y_o^* = \tilde{\delta} + \tilde{d}z^* + \tilde{\xi} \text{ and } y_o = j \text{ if } \tau_{low} < y_o^* < \tau_{up} \tag{3}$$

where y_o is a $(G\times1)$ vector of ordinal indicators and y_o^* is the $(G\times1)$ vector of the continuous latent propensity variables underlying the ordinal indicators, $\tilde{\delta}$ is a $(G\times1)$ vector of constant terms, $\tilde{d}$ is a $(G\times L)$ matrix of latent variable loadings onto the ordinal indicators, τ_{low} and τ_{up} are both $(G\times1)$ vectors obtained by stacking the lower and the upper thresholds of the ordinal indicators respectively. j represents the ordinal indicator category and $j = \{1,2,\dots,J\}$. $\tilde{\xi}$ is a $(G\times1)$ vector of error terms associated with the underlying propensity of the ordinal indicators and is assumed to be multivariate normally distributed: $\tilde{\xi} \sim N[0_G, \Sigma_{y_o^*}]$ with $\Sigma_{y_o^*}$ representing the covariance matrix. For identification purposes $\Sigma_{y_o^*}$ is assumed to be an identity matrix[2].

By stacking the vector of the continuous indicators and the vector of the ordinal indicators and replacing the latent variable z^* with the structural equation shown in Equation 1, the reduced form expression for the measurement equation can be obtained as in equation 4:

$$\check{y} = \check{\delta} + \check{d}(\omega\rho) + \check{d}(\eta) + \check{\xi} \tag{4}$$

Where $\check{y} = (y_c', [y_o^*]')'$, $\check{\delta} = (\delta', \tilde{\delta}')'$, $\check{d} = (d', \tilde{d}')'$ and $\check{\xi} = (\xi', \tilde{\xi}')'$.

2.1.3 Multiple Discrete Continuous (MDC) Choice Model

Following Bhat (2008), the MDC choices can be formulated as an allocation problem wherein an individual consumes $x = \{x_1, x_2, \dots, x_K\}$ amounts of K goods to maximize his/her utility (U) subject to a budget constraint (E) as shown below:

$$\max U(x) = \sum_{k=1}^{K} \gamma_k \Psi_k \ln\left(\frac{x_k}{\gamma_k} + 1\right) \tag{5a}$$

$$\text{subject to } \sum_{k=1}^{K} x_k = E \tag{5b}$$

where x is a $(K\times1)$ vector of the quantity of goods consumed, $\gamma_k\ (> 0)$ is the translation (also satiation) parameter and $\Psi_k (> 0)$ is the baseline marginal utility. Ψ_k represents the marginal random utility at the point of zero consumption for good k. γ_k parameter serves to account for satiation effects associated with consuming goods. It should be noted that, to meet the budget constraint, every individual must consume at least one good (referred to with index m[3] from this point forward) from the available set of K goods. Both the baseline marginal utility and the translation parameter are parametrized in terms of exogenous explanatory variables. Further, the proposed HMDC framework parameterizes the baseline marginal utility (Ψ_k) in terms of latent psychological factors. Equation 6 shows the parameterized baseline marginal utility in the HMDC:

$$\Psi = \exp(\nu\beta + \lambda z^* + \varepsilon) \tag{6}$$

[3] For individuals who consume multiple goods, *m* can be assumed to be the good with the lowest index of *k* without any loss of generality.

where Ψ is a $(K\times1)$ vector of baseline marginal utilities associated with the different goods, ν is a $(K\times D)$ matrix of observed explanatory variables, β is a $(D\times1)$ vector of coefficients associated with the ν, λ is a $(K\times L)$ matrix of coefficients associated with the psychological factors and ε is a $(K\times1)$ vector of stochastic error terms which are assumed to be multivariate normally distributed: $\varepsilon \sim N[0_K, \Lambda]$ with Λ representing the covariance matrix. The optimization problem defined in Equation 5 can be solved by forming the Lagrangian and applying the Karush-Kuhn Tucker (KKT) conditions. From KKT first order conditions it follows that:

$$\mu_{km}^* = 0, if\ x_k{}^* > 0, k = 1,2, \dots, K, k\ \neq m \tag{7a}$$

$$\mu_{km}^* < 0\ if\ x_k{}^* = 0, k = 1,2, \dots, K, k\ \neq m \tag{7b}$$

where $\mu_{km}^* = \mu_k - \mu_m$ and $\mu_k = \beta' \nu_k - ln\left(\frac{x_k}{\gamma_k} + 1\right) + \lambda_k' z^* + \varepsilon_k$. By replacing the latent variable z^* with the corresponding structural equation (as shown in Equation 1), the $(K-1)$ sized vector μ can be expressed in the matrix notation as shown below:

$$\mu = \nu\beta + \lambda(\omega\rho) - \ln\left(\frac{x}{\gamma} + \begin{matrix} 1 \\ 1 \\ \vdots \\ 1 \end{matrix}\right) + \lambda(\eta) + \varepsilon \tag{8}$$

Finally it is assumed that the correlation between the measurement equation of the latent variables and the utility equations of the choice model arise only due to the common influence of the latent variables. As a result, $\check{\xi}$ (the error component in the measurement equation of the latent variables) and ε (the error component in the MDC choice model) are independent.

2.2 Model Estimation

The estimation of the HMDC model entails finding estimates for the following sets of parameter vectors: avec(ρ), avec(Γ), avec($\check{\delta}$), avec($\check{d}$), avec($\check{\Sigma}$), avec(τ_{low}), avec(τ_{up}), avec(β), avec(λ), avec(Λ) and avec(γ) where avec is used to represent the vector of the parameter inside the parentheses. The estimates can be obtained by applying the maximum likelihood estimation technique. The likelihood function of the HMDC model can be expressed as the joint probability of observing the vector of continuous indicator ($y_c = i$), the vector of ordinal indicators ($y_o = j$) and the vector of consumption quantities for the K goods (x). Furthermore, the probability of observing vector of ordinal indicator vectors (y_o) and the goods consumption (x) can be expressed in terms of the underlying propensity variables (y_o^*) and utility differences (μ_{km}^*) respectively. Denoting the vector of all parameters to be estimated as Θ, the likelihood function for the HMDC model formulation can be expressed as shown in Equation 9.

$$L(\Theta) = Pr(i, j, x|\Theta) = Pr\left(y_c = i, \tau_{low} < y_o^* < \tau_{up}, \mu_{cm}^* = 0, \mu_{nm}^* < 0|\Theta\right) \tag{9}$$

where μ_{cm}^* and μ_{nm}^* represents $(M-1)$ and $(K-M)$ sized partitions of the vector μ_{km}^* respectively with $c = \{1,2, \dots, M-1\}$ and $n = \{(M+1), (M+2), \dots, K\}$ and M being the total number of alternatives that are consumed.

The joint probability in Equation 9 can be broken down into a marginal probability density function (PDF) and a conditional cumulative density function (CDF) as shown in Equation 10.

$$L(\Theta) = Pr(y_c = i, \mu_{cm}^* = 0|\Theta) \times Pr(\tau_{low} < y_o^* < \tau_{up}, \mu_{nm}^* < 0|y_c = i, \mu_{cm}^* = 0, \Theta) \quad (10)$$

The dimension of the joint CDF in Equation (10) can vary from G (representing the case where all available goods are consumed i.e. $M = K$) to $(G + K - 1)$ (representing the case where only one good is consumed i.e. $M = 1$). This high dimensionality of integral in the likelihood function above is evaluated by adopting the composite likelihood estimation (Varin 2008) along with an analytical approximation for the multivariate CDF called maximum approximated composite marginal likelihood (MACML) proposed by Bhat (2011).

2.2.1 Composite Likelihood Estimation

The multivariate CDF component in Equation 10 is evaluated by applying the pairwise CML approach[4]. In the HMDC, decomposing the integral using pairwise CML entails treating the non-chosen alternatives as a single event (i.e. as a bundle). This process results in a total of ${}^{G}C_2$ (read as G choose 2 combinations and is evaluated as $\frac{G*(G-1)}{2}$) marginals of observing any two ordinal indicators and G marginals of observing an ordinal indicator along with the vector of non-chosen alternatives. Equation 11 presents the pairwise CML approximation of the likelihood function presented in Equation 10.

$$L_{PCML}(\Theta) = Pr(y_c = i, \mu_{cm}^* = 0|\Theta)$$
$$\times \left(\prod_{g=1}^{G-1}\prod_{g'=g+1}^{G} Pr\left(\tau_{low,g} < y_{o,g}^* < \tau_{up,g}; \tau_{low,g'} < y_{o,g'}^* < \tau_{up,g'}|y_c = i, \mu_{cm}^* = 0, \Theta\right)^W\right)$$
$$\times \left(\prod_{g=1}^{G} Pr(\tau_{low,g} < y_{o,g}^* < \tau_{up,g}; \mu_{nm}^* < 0|y_c = i, \mu_{cm}^* = 0, \Theta)^W\right) \quad (11)$$

The third probability component in the Equation 11 above is transformed so that the evaluation of orthant probability (i.e. bounded on both sides) is replaced with an evaluation of only cumulative probabilities (i.e. bounded on one side).

$$L_{PCML}(\Theta) = \Pr(y_c = i, \mu_{cm}^* = 0|\Theta)$$
$$\times \left(\prod_{g=1}^{G-1}\prod_{g'=g+1}^{G} Pr\left(\tau_{low,g} < y_{o,g}^* < \tau_{up,g}; \tau_{low,g'} < y_{o,g'}^* < \tau_{up,g'}|y_c = i, \mu_{cm}^* = 0, \Theta\right)^W\right)$$
$$\times \left(\prod_{g=1}^{G} \begin{pmatrix} Pr(y_{o,g}^* < \tau_{up,g}; \mu_{nm}^* < 0|y_c = i, \mu_{cm}^* = 0, \Theta) - \\ Pr(y_{o,g}^* < \tau_{low,g}; \mu_{nm}^* < 0|y_c = i, \mu_{cm}^* = 0, \Theta) \end{pmatrix}^W\right) \quad (12)$$

In Equations 11 and 12, W represents the weight. It can be seen that the second probability component of the pairwise likelihood expression in Equation 12 only involves the evaluation of bivariate normal CDF which is fairly easy to handle. However, the third probability expression still involves the evaluation of a multivariate normal CDF whose dimension can be as high as $(K - 1)$. This multivariate normal CDF in the expression above is evaluated using the MACML analytical approximation.

---

[4] Under regularity condition CML estimators are consistent and asymptotically normally distributed; for a formal proof see Xu and Reid (2011).

2.2.2 Choice of Weight

The dimension of the multivariate normal CDF to be approximated varies from one observation to another because of the presence of the MDC choice kernel wherein individuals choose a subset of the K goods. As a result, this requires that a weight other than unity be used to facilitate the recovery of the population parameters (Joe and Lee 2009). Generally speaking, in the pairwise treatment of CML, each random variable (event) appears in $(m_i - 1)$ (where m_i is the size of the random vector for the i^{th} observation) number of probability calculations i.e. the number of pairs for each observation varies across observations and as a result the contribution of each observation to the overall likelihood of the sample also varies if unit weights are assumed. Weighting (i.e. $W \neq 1$) allows one to ensure that the contribution of each observation to the overall likelihood is proportional to the size of the random vector of that observation.

There are a number of studies on the selection of optimal weights that will improve efficiency of the parameter estimates. A review of the literature suggests that one of the main considerations for the choice of weights is the dependency structure (Joe and Lee 2009) among the multivariate random vectors. The most widely recommended and implemented weight for a moderate dependency structure is $\frac{1}{(m_i-1)}$ (Kuk and Nott 2000, Zhao and Joe 2005). In estimating parameters of the HMDC, the following weights are proposed – this is analogous to $\frac{1}{(m_i-1)}$ weight for clustered data.

$$W = \begin{cases} \frac{1}{G} & where, M < K \\ \frac{1}{(G-1)} & where, M = K \end{cases} \tag{13}$$

2.2.3 Log-Likelihood Function

The log-likelihood function for the entire sample is shown in Equation 14.

$$LL(\Theta) = \sum_{n=1}^{N} \ln(L_{PCML,n}(\Theta)) \tag{14}$$

The above likelihood function and the associated gradients are implemented in matrix programming language GAUSS to obtain the parameter estimates $\hat{\Theta}_{PCML}$. Also, the variance-covariance matrix of the parameter estimates was obtained using the robust Gobambe sandwich estimator (Godambe 1960) shown in Equation 15 below.

$$V(\Theta) = (H[\Theta])^{-1}(J[\Theta])(H[\Theta])^{-1} \tag{15}$$

where, $H[\hat{\theta}] = -\left(\sum_{n=1}^{N} \frac{\delta^2 log L_{PCML,n}(\Theta)}{\delta\Theta\delta\Theta'}\right)_{\hat{\Theta}_{PCML}}$ and $J[\hat{\theta}] = \sum_{n=1}^{N}[(\frac{\ln(L_{PCML,n}(\Theta))}{\delta\Theta})(\frac{ln(L_{PCML,n}(\Theta))}{\delta\Theta'})]_{\hat{\Theta}_{PCML}}$. Section 3 presents a simulation study that demonstrates the ability of the proposed estimation technique for recovering consistent and efficient estimates of the model parameters.

3. SIMULATION STUDY

A simulation study was performed to assess the ability of the estimation technique to recover the parameters. The simulation study was aimed at mimicking the subsequent empirical application (see Section 4). However, simplifications were made with regard to the model and parameter specification for the different components of the HMDC model to enable rapid testing and ease of interpretation. In the simulation study, following assumptions were made with regard to the different components of the HMDC model specification:

- With regard to the structural component of the latent variables, it was assumed that there are three latent variables (i.e. psychological factors) of interest. Further, each of the latent variables are assumed to be a function of two explanatory variables. All the 6 covariates of the structural equation of latent variable are generated from a $N[1,1]$ distribution.
- With regard to the measurement component of the latent variables, it was assumed that there are six indicators including two continuous and four ordinal indicators. Each latent variable is measured using a pair of indicator variables.
- With regard to the choice model component, it was assumed that there are five alternatives that the individual can consume. The baseline marginal utility equations for each of the five alternatives were assumed to be a function of a constant (normalized to zero for the first alternative), one observed explanatory variable and three latent variables. All the coefficients (including the coefficients of the latent variables) in the choice model are assumed to be alternative specific. The five covariates of the choice model utilities are generated from $N[0,1]$ distribution. Furthermore the budget for the MDC choice model was generated from $N[300,30]$ distribution.

The exogenous variables were generated only once and were kept fixed for the rest of the simulation study. The simulation study was conducted on simulated datasets of three different sizes namely: 1000, 2000 and 2500. For each sample size, 50 sets of observations were generated using different realizations of the error components η, $\breve{\xi}$, and ε. The variance covariance matrices assumed as well as the corresponding lower triangular Cholesky matrix for the error components η and ε are shown below. $\breve{\xi}$ is assumed to be an identity matrix and only the diagonal elements corresponding to the continuous indicators are estimated[5].

$$\Gamma = \begin{pmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.683 \\ 0.5 & 0.683 & 1 \end{pmatrix} \equiv C_\Gamma = \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 0.866 & 0 \\ 0.5 & 0.5 & 0.707 \end{pmatrix}$$

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1.21 & 0.66 & 0 & 0 \\ 0 & 0.66 & 1.17 & 0 & 0 \\ 0 & 0 & 0 & 0.64 & 0.80 \\ 0 & 0 & 0 & 080 & 1.81 \end{pmatrix} \equiv C_\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1.1 & 0 & 0 & 0 \\ 0 & 0.6 & 0.9 & 0 & 0 \\ 0 & 0 & 0 & 0.8 & 0 \\ 0 & 0 & 0 & 1 & 0.9 \end{pmatrix}$$

In the HMDC model estimation, for each of the three error components, the parameters of the variance-covariance matrices are not estimated directly instead the corresponding Cholesky factors are estimated. This was done to ensure that estimated parameter values result in positive definiteness of the variance-covariance matrices. Also, note that the first element in Λ was normalized to 1 for the purpose of identification (Train 2009). Lastly, as mentioned earlier in the formulation, all matrices associated with the error terms above represent variance-covariance

---

[5] No correlation has been allowed between the errors of the measurement equation, which is not a restrictive assumption rather normalization similar to that suggested by Bollen (1983) and others. The behavioral interpretation for this normalization could be that the indicators are correlated because of their dependency on the common latent variables, and once we account for those common latent variables, no other correlation exists between the indicators.

matrix except Γ which represents a correlation matrix. One must ensure that the parameters that are retrieved for Γ correspond to a correlation matrix. This can be achieved by only estimating the off diagonal elements in the lower triangular portion of the corresponding Cholesky matrix (C_{Γ}) and then using the Equation 16 below to retrieve the i^{th} diagonal element.

$$c_{ii} = \sqrt{1 - \sum_{j=1}^{i-1} c_{ij}^2}. \tag{16}$$

3.1 Simulation Results

In order to assess the ability of the proposed estimation procedure in recovering the parameters, two measures namely Absolute Percentage Bias (APB) and Relative Asymptotic Efficiency (RAE) are used. APB is used to assess the bias in the parameter estimates and RAE is used to evaluate the asymptotic efficiency of the estimates. Equations for the measures are shown in Equation 17.

$$\begin{gathered} APB = \left| \frac{\text{True value} - \text{Estimate}}{\text{True value}} \right| \times 100\% \\ \text{RAE} = \frac{ASE}{\text{FSSE}} \end{gathered} \tag{17}$$

where, ASE is the asymptotic standard error and is equal to the mean of the standard errors calculated using Godambe sandwich estimator given by Equation 16 across the fifty simulated datasets and $FSSE$ stands for the finite sample standard error and is equal to the standard deviation of the parameter estimates across the fifty sets of simulated data. Additionally, the confidence interval for the parameter estimates are also reported where the confidence interval is calculated using equation (18).

$$\text{Confidence Interval } (CI) = \text{ Estimated Value } \pm \text{ } 1.96 * \text{Asymptotic Standard Error } (ASE) \tag{18}$$

In general, the proposed estimation technique (and the estimators) appears to be promising with very good recovery of the parameter values both in terms of APB (indicating the degree of bias) and RAE (pointing to the asymptotic efficiency). Only the results obtained using the sample data set with size 2500 have been reported in Table 1.a and Table 1.b[6,7].The average values of the APB and RAE across all the 59 parameters were found to be 0.636% (value close to zero indicating no bias) and 1.099 (value close to one indicating good asymptotic efficiency) respectively. The average values of APB and RAE were also checked separately for different group of parameters of the latent variable model and the MDC choice model and it was found that for all groups of

---

[6] The average values (across all the parameters) of the APB and RAE for the 1000 sample data set are respectively 1.538% and 1.126; and the average values of the APB and RAE for the 2000 sample data set are respectively 1.112% and 1.131.

[7] The notation used to denote different set of parameters has been introduced in section 2, however we introduce suffix to differentiate between different parameters in Table 1.a and Table 1.b. For example $\check{d}_{ij}$ represents the factor loading of the i^{th}indicator on the j^{th}latent variable, similarly λ_{ij} represents the coefficient of the j^{th} latent variable on the i^{th} choice alternative. Similarly the Cholesky factors of all the error components are denoted using c_{ij}, where i and j represents row and column indices respectively.

parameters, the average value of APB was very close to 0 and the average RAE value varied in the acceptable range of 0.75 and 1.25 except for correlation parameters of the structural equation of latent variable. One plausible reason for this high RAE value of the correlation parameters might be owing to the fact that the Cholesky factors of these correlation parameters were further parameterized to ensure that we were estimating a correlation matrix instead of a variance covariance matrix. Also note that in the presented simulation study, the densest possible correlation matrix for the structural equation of the latent variable (i.e. allowed correlation between all the possible pairs of latent variables) was assumed. A less dense structure of the correlation matrix resulted in better RAE of the correlation parameters. We don't present results from this additional exploration for the sake of brevity. Additionally, the study of implications of error structures on the efficiency is an interesting avenue for future research. Nonetheless, the simulation study provides substantial evidence towards good recovery of the parameter values and the performance of the proposed estimation technique (and the estimators) is very promising.

In the simulation study, to examine the importance of weights in setting up the CML function, parameters were also estimated by assuming a unit weight. The results (presented in Appendix[8]) highlight the differences in recovery of the parameters using weighted CML and unweighted CML in the presence of MDC choice kernel. As can be observed from the Table A.1a and A.1b in the Appendix, large percentage bias values are observed (APB = 10.39%) when unit weight is assumed in CML approximation. Furthermore, the bias is much higher for most of the MDC choice kernel parameters (parameters with large bias percentages are highlighted in the Table). On the other hand, as reported in Table 1.a and Table 1.b, weighted CML approximation reduces the bias percentages significantly. Use of weights brings the APB values of MDC choice kernel parameters to a range that is comparable with parameters of other components of the HMDC. It can be noted that, for simulation results with weight, the true parameter value always falls within the 95% confidence interval and the 95% confidence interval is quite tight around the true parameter value. However, for simulation results without weight, the true value falls outside the confidence interval for a good number of (13 out of 26) MDC parameters. This further highlights the importance of using weights in setting up the estimator using CML technique in the current scenario. These observations (which are in line with the work by Varin et al. 2011) point to the importance of using weights in the CML approximation to ensure good recovery of the parameters when the size of the vector (to be dealt with CML) varies across observations.

[8] The results reported in Table A1.a and A1.b are comparable with those reported in Table 1a and 1b, because the same set of simulated data have been used for producing these two sets of results. Both the tables report the summary results obtained from 50 independent model runs.

Table 1.a: Simulation Results

Parameters	True Values	Estimated Values	Absolute Percentage Bias (APB)	Asymptotic Standard Error (ASE)	Finite Sample Standard Error (FSSE)	Relative Asymptotic Efficiency (RAE)	Confidence Interval	Contains True Value?
α_1	1.1	1.0981	0.1696	0.0300	0.0273	1.1007	(1.04) - (1.16)	Yes
α_2	1.7	1.7009	0.0516	0.0391	0.0379	1.0333	(1.62) - (1.78)	Yes
α_3	1.2	1.2114	0.9495	0.051	0.0654	0.7795	(1.11) - (1.31)	Yes
α_4	1.8	1.8023	0.1298	0.0693	0.0711	0.975	(1.67) - (1.94)	Yes
α_5	1.4	1.3985	0.1097	0.0899	0.0619	1.4539	(1.22) - (1.57)	Yes
α_6	1.6	1.5861	0.8698	0.1006	0.0683	1.4736	(1.39) - (1.78)	Yes
Γ_{21}	0.5	0.506	1.1954	0.0381	0.0281	1.3584	(0.43) - (0.58)	Yes
Γ_{31}	0.5	0.4999	0.0296	0.0452	0.0249	1.8157	(0.41) - (0.59)	Yes
Γ_{32}	0.5	0.5144	2.8809	0.0655	0.0425	1.5416	(0.39) - (0.64)	Yes
$\bar{\delta}_1$	-1.1	-1.1062	0.5652	0.0473	0.0427	1.1088	(-1.2) - (-1.01)	Yes
$\bar{\delta}_2$	-1.7	-1.7003	0.0156	0.0498	0.0498	1.0014	(-1.8) - (-1.6)	Yes
$\bar{\delta}_3$	-2.5	-2.4864	0.5441	0.1489	0.1588	0.9377	(-2.78) - (-2.19)	Yes
$\bar{\delta}_4$	-2	-1.9931	0.3432	0.1076	0.1184	0.9084	(-2.2) - (-1.78)	Yes
$\bar{\delta}_5$	-2	-1.9707	1.4635	0.1161	0.1005	1.1553	(-2.2) - (-1.74)	Yes
$\bar{\delta}_6$	-2.8	-2.7736	0.9433	0.1725	0.1636	1.0549	(-3.11) - (-2.44)	Yes
$\bar{d}_{11}$	1	1.0011	0.1087	0.0207	0.0202	1.0266	(0.96) - (1.04)	Yes
$\bar{d}_{21}$	1.1	1.1006	0.0589	0.0224	0.0212	1.0563	(1.06) - (1.14)	Yes
$\bar{d}_{32}$	1.2	1.2153	1.2737	0.0852	0.0876	0.9724	(1.05) - (1.38)	Yes
$\bar{d}_{42}$	1	1.0111	1.1134	0.061	0.0694	0.8788	(0.89) - (1.13)	Yes
$\bar{d}_{53}$	1.1	1.1201	1.8261	0.0928	0.0762	1.2179	(0.94) - (1.3)	Yes
$\bar{d}_{63}$	1.3	1.3206	1.5832	0.1224	0.0898	1.3635	(1.08) - (1.56)	Yes
Σ_{11}	1	0.9983	0.1675	0.0192	0.0169	1.1361	(0.96) - (1.04)	Yes
Σ_{22}	1	0.9954	0.4555	0.0209	0.0203	1.0301	(0.95) - (1.04)	Yes
β_1	-1	-1.0004	0.0355	0.1707	0.1499	1.1385	(-1.33) - (-0.67)	Yes
β_2	2	1.9907	0.4644	0.1693	0.1510	1.1206	(1.66) - (2.32)	Yes
β_3	-2	-1.9879	0.6059	0.1751	0.1578	1.1097	(-2.33) - (-1.64)	Yes
β_4	2.5	2.5106	0.4257	0.2129	0.1683	1.2653	(2.09) - (2.93)	Yes
β_5	-1	-0.9944	0.5576	0.0593	0.0678	0.8752	(-1.11) - (-0.88)	Yes
β_6	3	2.9991	0.0309	0.0604	0.0736	0.8215	(2.88) - (3.12)	Yes
β_7	-1	-0.9996	0.0353	0.0292	0.0329	0.8881	(-1.06) - (-0.94)	Yes

Table 1.b: Simulation Results

Parameters	True Values	Estimated Values	Absolute Percentage Bias (APB)	Asymptotic Standard Error (ASE)	Finite Sample Standard Error (FSSE)	Relative Asymptotic Efficiency (RAE)	Confidence Interval	Contains True Value?
β_8	3.5	3.4924	0.2164	0.0801	0.0801	0.9999	(3.34) - (3.65)	Yes
β_9	-3.5	-3.5231	0.6599	0.1043	0.1075	0.9701	(-3.73) - (-3.32)	Yes
λ_{21}	-1.5	-1.4902	0.6544	0.0555	0.0556	0.9982	(-1.6) - (-1.38)	Yes
λ_{22}	1.2	1.2063	0.5278	0.0575	0.0556	1.0333	(1.09) - (1.32)	Yes
λ_{23}	1.1	1.1095	0.8666	0.0773	0.0506	1.528	(0.96) - (1.26)	Yes
λ_{31}	-1.6	-1.5896	0.6489	0.056	0.0583	0.9596	(-1.7) - (-1.48)	Yes
λ_{32}	1.1	1.1029	0.2631	0.0537	0.0476	1.1291	(1) - (1.21)	Yes
λ_{33}	1	1.0083	0.8293	0.0713	0.0455	1.5668	(0.87) - (1.15)	Yes
λ_{41}	-1.4	-1.3945	0.3919	0.0543	0.0591	0.9183	(-1.5) - (-1.29)	Yes
λ_{42}	1.3	1.3028	0.2174	0.0608	0.0540	1.1244	(1.18) - (1.42)	Yes
λ_{43}	1.1	1.1145	1.3181	0.0778	0.0500	1.5545	(0.96) - (1.27)	Yes
λ_{51}	-3	-2.9975	0.0831	0.0901	0.0930	0.9685	(-3.17) - (-2.82)	Yes
λ_{52}	1	1.0009	0.0915	0.0551	0.0482	1.1432	(0.89) - (1.11)	Yes
λ_{53}	1	1.0147	1.4705	0.0748	0.0625	1.1968	(0.87) - (1.16)	Yes
γ_1	1.5	1.4898	0.6774	0.1061	0.1091	0.9723	(1.28) - (1.7)	Yes
γ_2	1.8	1.7972	0.1528	0.0811	0.0883	0.9187	(1.64) - (1.96)	Yes
γ_3	2.2	2.2339	1.5411	0.0815	0.1035	0.7879	(2.07) - (2.39)	Yes
γ_4	2.5	2.4809	0.7639	0.089	0.0821	1.0835	(2.31) - (2.66)	Yes
γ_5	2.8	2.7999	0.0051	0.1141	0.0997	1.1443	(2.58) - (3.02)	Yes
Λ_{22}	1.1	1.1002	0.0144	0.0551	0.0653	0.8432	(0.99) - (1.21)	Yes
Λ_{32}	0.6	0.5927	1.2234	0.0674	0.0662	1.0181	(0.46) - (0.72)	Yes
Λ_{33}	0.9	0.8875	1.3842	0.0272	0.0258	1.0554	(0.83) - (0.94)	Yes
Λ_{44}	0.8	0.7901	1.2316	0.0681	0.0613	1.1111	(0.66) - (0.92)	Yes
Λ_{54}	1	1.0016	0.1595	0.1046	0.0785	1.3327	(0.8) - (1.21)	Yes
Λ_{55}	0.9	0.9027	0.3016	0.0768	0.0710	1.0825	(0.75) - (1.05)	Yes
$\tau_{up,1}$	1.5	1.5114	0.7603	0.0815	0.0853	0.9548	(1.35) - (1.67)	Yes
$\tau_{up,2}$	1.5	1.4975	0.1636	0.0676	0.0852	0.7938	(1.37) - (1.63)	Yes
$\tau_{up,3}$	1.5	1.5161	1.0741	0.0735	0.0725	1.0129	(1.37) - (1.66)	Yes
$\tau_{up,4}$	1.5	1.5121	0.8038	0.0876	0.0845	1.0361	(1.34) - (1.68)	Yes
Mean			**0.6356**			**1.0989**		

4. EMPIRICAL STUDY

The primary purpose of the empirical study was to demonstrate the feasibility and applicability of the proposed HMDC model implementation for exploring the association between psychological factors and MDC choice dimensions. To this end, the association between moods experienced by an individual and their daily activity engagement choices were explored to understand the heterogeneity in individual activity participation and time allocation behaviors. The choice of the empirical study was motivated also in part due to gaps in the empirical literature. While there is a rich body of literature exploring the role of psychological factors on the different dimensions of travel choices (Anable 2005, Glerum and Bierlaire 2012, Atasoy et al. 2013, Alvarez-Daziano and Bolduc 2013, Kamargianni and Polydoropoulou 2013, Hess and Spitz 2016), literature exploring the relationship between psychological factors and the activity engagement choices of individuals is limited (Ettema et al. 2010, Abou-Zeid and Ben-Akiva 2012, Ravulaparthy et al. 2013). The study of the daily activity engagement choices is important because it helps better understand the factors influencing travel and subsequently allows the design of effective policies aimed at managing travel demand (Kitamura 1988, Pendyala and Bhat 2004, Chen and Mokhtarian 2006).

In the following subsections, the study motivation, data composition, model setup, estimation results, and validation analysis are presented.

4.1 Study Motivation

4.1.1 Moods and Behaviors

While traditional decision theories postulate decision making as a cognitive process, behavioral decision theories have increasingly emphasized the role of emotions/moods on decision making process as well as on the choice outcomes (Loewenstein and Lerner 2003). Loewenstein and Lerner (2003) identify two ways in which behavior can be influenced by the affect or emotions. According to authors, on one hand, individual behavior can be shaped by the expected emotion that would arise from the decision outcome. On the other hand, there is the immediate influence of the mood experienced at the time of making a choice which might not only impact the decision making process but also the decision outcome. Clark (2006) defines mood as a prevailing psychological state, feeling, or emotion which may be habitual or temporary. Decades of experimental work performed by behavioral psychologists show that positive and negative moods (emotions) have distinct effects on an individual's decision making process as well as on decision outcome (Fredrickson 2001, Isen 2001). For example, Fredrickson (2001) notes that a positive mood is associated with "broad, flexible cognitive organization and the ability to integrate diverse material" in the decision making process. On the other hand, a negative mood has been associated with narrowing individuals' attention while making decisions. Forgas (1989) studied the influence of both positive and negative moods in social decision making context. He notes that, sad people use comparatively direct search strategies at arriving decisions compared to happy people and also tend to prefer rewarding outcomes. In the current study, we explore the correlation between the moods that the individual experiences over the course of a day, and the activity participation and time allocation behaviors. This is in line with the exploration of influence of mood at the time of decision making postulated by Loewenstein and Lerner (2003).

4.1.2 Moods and Activity-travel Choices

There is research suggesting that cognitive and affective states of an individual contain both stable (Fredrick and Loewenstein 1999) and variable components (Oishi et al.1999). Also, researchers have shown that it is possible to identify the "stable" component of cognitive and affective states at the level of days or weeks (Gadermann and Zumbo 2007). Drawing on the work from the field of behavioral psychology and decision theory, current research aims to identify the association between "day level moods" and activity participation and time allocation decisions. Day level mood is defined as the "stable" state, feeling, or emotion that the individual experiences over the course of a day. It is the influence of this "stable" mood on activity engagement choices that is explored in this study. From this point forward, the "stable" component of the individual's mood will be referred to as merely moods.

In the context of activity-travel choices, there is recent research exploring how activity and travel choices impact the moods experienced. For example, Morris and Guerra (2015) explored the role of travel mode on the mood experienced. Mokhtarian et al. (2015) identify the influence of different trip attributes such as trip length, distance, purpose, mode on the fatigue experienced during travel. Similarly, Legrain et al. (2015) investigate most stressful mode of commute using a university wide travel survey. The current research attempts to explore the alternative association wherein the day level moods that sustain over the course of a day influence the daily discretionary activity engagement choices. This is done while controlling for the impact of other exogenous variables that contribute to heterogeneity in activity engagement choices[9]. Considering moods allows us to account for unobserved heterogeneity in the decision making process due to the differences in moods experienced (in addition to other observed explanatory variables) which would have been attributed to random error components otherwise (Hess 2012). Additionally, adopting the ICLV framework to include mood in exploring activity participation and time allocation behavior allows us to disentangle the influence of the observed explanatory variables into constituent components: 1) their direct influence on the activity participation and time allocation choices and 2) their indirect influence through their correlation with the latent mood variables (Vij and Walker 2016).

4.1.3 Activity Engagement Choices

In the empirical exploration, discretionary activity engagement choices are of interest. Discretionary activities offer the most flexibility in terms of their planning and scheduling when compared with other activities (e.g. work, education and maintenance activities to some extent). As a result, they are also the most amenable to being influenced by the factors of interest (including moods). The use of HMDC for the empirical exploration is appropriate because discretionary activity engagement requires handling multiple choice dimensions simultaneously. First, there is the discrete choice of participating in an activity and there is the continuous choice of amount of time spent in the activity, and second, there are multiple instances of these participation and time use variables because an individual could participate in multiple discretionary activity types over the course of a day. Thus discretionary activity engagement results in a multiple discrete continuous (MDC) choice situation.

[9] Also, the authors recognize that the relationship between the moods and the activity engagement choices is not one directional. In particular, the "variable" component of the moods also affect activity engagement choices, and activity participation and time allocation choices in turn affect "variable" component of moods. The evolution of the "variable" component of the moods (and other cognitive and affective states) are not the focus of this paper.

4.1.4 Study Objectives

The purpose of the case study is to investigate the influence of moods on choice outcome such as discretionary activity engagement behavior. More specifically, the study attempts to examine if higher levels of positive moods is associated with less passive leisure and higher participation in other types of discretionary activities. Alternatively, also of interest is whether high levels of negative moods would have the opposite association (i.e. more passive leisure and less participation in other discretionary activity types). This hypothesis is partly derived from the research that suggests positive association between negative moods and narrowing of attention while selecting between alternatives. It is postulated that, a direct search (under the influence of negative moods) would more often lead the individual into the most obvious choice of discretionary activity which is passive leisure, whereas a proactive search (under the influence of positive moods) would lead them to consider various options for discretionary activity participation and time allocation behavior. It is acknowledged that activity participation in turn can affect the moods experienced (in particular the "variable" component of the moods) (Ettema et al. 2010). However, it is posited that there is a "stable" component of the moods (both positive and/or negative) that may sustain over the course of the given day in an individual's life. It is the association between these sustained moods and the discretionary activity engagement choices that are of interest in this research.

4.2 Data Composition

The data used for the empirical study was drawn from the 2013 American Time Use Survey (ATUS). ATUS is cross-sectional survey collecting information about the activity engagement choices from a representative sample of individuals across the US since 2003. The survey follows an activity diary format asking a single individual (over the age of 15 years) from a household to report all the activities performed over a full 24 hour period. Individuals are also asked to provide a detailed account of the different activity characteristics including activity duration, location, and accompaniment type among other information. More recently, ATUS started administering supplemental modules to collect additional information regarding various psychological factors of interest. In the well-being module (that is of interest in this study), people are asked to report their general health and life satisfaction. Additionally, people are asked to rate their feeling with respect to 5 emotions: happiness, sadness, pain, stress and tiredness for three randomly chosen time intervals during the day on a scale of 0 to 6.

Respondents in the dataset for whom the total activity durations did not add up to 1440 minutes or those who had invalid responses for the questions regarding emotions were excluded from the analysis. This data preparation process resulted in 4002 observations. A quarter of the sample was set aside to perform a holdout sample validation. The remaining sample available for model estimation and subsequent empirical exploration comprised of 3025 observations. As noted earlier, the discretionary activity engagement choices were of interest in this study; fixed activities and maintenance activities were not considered in the analysis. The discretionary activities were categorized into six types namely: 1) active leisure, 2) passive leisure, 3) physical activity, 4) shopping for non-maintenance, 5) attending sports and arts events, and 6) social activity.

A brief description of the six discretionary activity types along with the percentage of respondents who participated in each of the particular discretionary activity types and the average amount of time spent in the discretionary activity type are reported in Table 2. It should be noted that the mean activity duration is the average across all respondents who have reported participating in the activity type on the survey day. As can be seen from Table 2, almost 90 percent of the respondents participated in some form of passive leisure during weekends; this activity type also was used as the reference activity type in the HMDC model specification. A little less than half of the respondents reported participating in some form of active leisure and social activity. Passive leisure had the highest mean duration, followed by attending sports and arts events. Similar to the participation rates, the average duration for active leisure and social activity appear to be similar. Finally, shopping for non-maintenance has the lowest mean duration across all discretionary activity types.

Table 2: Weekend Discretionary Activity Participation and Time Allocation

Activity Category	Activity Description	Participation (%)	Mean Duration (Minutes)[1]
Active leisure	Playing games, using computer for leisure, pursuing hobbies (arts and crafts, collecting), leisure reading, leisure writing	45	153
Passive leisure	Relaxing, thinking, using tobacco and drug, watching television, listening to the radio, listening to or playing music	88	256
Physical activity	Participating in sports, exercise, recreation	18	122
Shopping for non-maintenance	Shopping except for food, groceries and gas	25	81
Attending sports and arts events	Attending performing arts, attending museums, movies, films, gambling, other arts and entertainment, attending sporting and recreational events	6	190
Social activity	Socializing and communicating, attending and hosting social events	47	147

[1]Mean duration has been calculated on only across the individuals who have reported to participate into at least one episode of a particular type of activity

4.3 Exploratory Analysis

A descriptive analysis was first conducted to test the stability of the five types of mood variables across the day. As mentioned earlier, in the AUTS respondents were asked to report the five emotions at three random time points across the day on a scale of 0 to 6. Descriptive analysis revealed that the reported moods remained very stable across the day with minimal variation. For example for the negative emotions such as pain and tiredness, about 80 to 90 percent people showed a variation of 1 unit or less across the day. For the rest of the emotions, such as happiness, stress and tiredness the percentage of people showing a variation of less than or equal to 1 unit varied from about 70 to 80 percent. For all the five types of emotions less than 5 percent of people showed a variation of more than or equal to 3 units across the day. These results provide credence to our assumption that the mood variables represent stable, day-level moods that are not influenced by activity participation; instead, they can potentially influence daily activity participation behavior.

Descriptive analysis was followed by exploratory factor analysis to explore the structure of the latent constructs of overall positive and negative moods that sustained throughout the day. There could be multiple constructs of positive and negative emotions. The latent constructs were developed based on indicators regarding the levels of five emotions: happiness, sadness, pain, stress and tiredness reported at three random time periods during the day. An exploratory factor analysis was performed using the fifteen indicator variables without specifying any prior structure for the factors. The process resulted in five latent constructs of moods with the indicators of the same emotion at the different time periods loading onto the same latent construct[10]. Therefore, the five latent constructs can be described as capturing the five emotions of happiness, sadness, pain, stress and tiredness and also they seem to sustain throughout the day with little variability. This result was not surprising. The stability of moods throughout the day may partly be attributed to the data collection approach. In the survey, users were asked to provide information about the moods not during the act of participating in the activity but after the fact. It is reasonable to assume that in such a context, it is only the feelings that they experienced/sustained throughout the day that will be remembered and thus reported.

Following the exploratory factor analysis, the HMDC model was estimated with five latent variables identified using three indicators each. Further, the choice model consisted of a multiple discrete continuous kernel that models both the participation and time use decisions for the six discretionary activity types. Section 4.4 presents the HMDC model estimation results.

4.4 HMDC Model Estimation Results

Table 3 summarizes the parameter estimates for the structural equations of the latent variables. Results from the measurement equations of the latent variables are presented in Table 4. Finally, parameter estimates for the multiple discrete continuous choice model are reported in Table 5. The t-statistics for the coefficient estimates are reported in the parentheses. The total number of parameters estimated in the model is 210 and the mean value of the log-likelihood function at convergence is -34.6150. The model estimation results obtained were behaviorally plausible and consistent with expectations. A detailed discussion of the results is presented in the following subsections.

4.4.1 Structural Equation Model of Latent Variables

The estimates of parameters in the structural equation (SE) provide valuable information regarding the variation of the latent construct with changes in observed explanatory variables. The choice of the explanatory variables used was based on a review of previous research from the field of happiness (or the lack of it) (Clark 2006 and Gerdtham and Johannesson 2001). The different variables used in the SE model include socio-economic characteristics such as gender, age, household income, education level, presence of spouse or partner as well as unemployment indicator. Additionally, it was hypothesized that the overall health and life satisfaction (which can be thought of as a proxy for the overall well-being of individual) will also have a strong influence on the daily moods experienced/exhibited by individuals.

[10] The results from the exploratory factor analysis are not reported in the paper in the interest of space.

Table 3: Estimation Results for the Structural Equation Model of Latent Variables

	Happiness	Pain	Sadness	Stress	Tiredness
Coefficients to the exogenous variables					
Female indicator	0.1749 (3.971)	0.0832 (2.193)	0.0773 (1.873)	0.2595 (6.038)	0.3261 (7.876)
Middle Income ($25 - $50 Thousand)[a]		-0.2311 (-3.951)	-0.2094 (-3.521)	-0.1228 (-2.081)	-0.0732 (-1.578)
High Income ($50 - $100 Thousand) indicator[a]		-0.1801 (-3.241)	-0.1749 (-2.918)	-0.1409 (-2.522)	
Very High Income (>$100 Thousand) indicator[a]		-0.2854 (-4.856)	-0.2047 (-3.142)	-0.1609 (-2.593)	
Age 35 to 54 indicator[b]		0.2744 (6.377)	0.1877 (4.238)		-0.0935 (-2.018)
Age 55 to 64 indicator[b]	0.1313 (2.219)	0.4186 (6.947)	0.2207 (3.409)	-0.286 (-5.079)	-0.2462 (-3.905)
Age 65 & above indicator[b]	0.2225 (3.702)	0.3749 (6.59)	0.1723 (2.882)	-0.3957 (-7.326)	-0.4445 (-7.305)
High school graduate indicator[c]				0.1268 (2.684)	
College graduate indicator[c]	-0.2021 (-4.245)			0.2083 (3.938)	
Post graduate indicator[c]	-0.4316 (-6.659)	-0.1129 (-2.457)		0.336 (5.125)	
Presence of spouse or partner indicator	0.2122 (4.873)	-0.0614 (-1.657)	-0.1392 (-3.741)		
Unemployment indicator					-0.1996 (-2.865)
Health condition very good indicator[d]	0.681 (5.021)	-1.5286 (-9.793)	-1.0292 (-5.881)	-0.9066 (-5.553)	-1.0505 (-8.806)
Health condition good indicator[d]	0.4808 (3.619)	-1.0949 (-7.035)	-0.7944 (-4.525)	-0.7053 (-4.385)	-0.8001 (-6.926)
Life condition poor indicator[e]	-1.2488 (-9.043)	0.6225 (4.717)	1.2525 (6.984)	1.2329 (7.875)	0.9511 (8.637)
Life condition good indicator[e]	-0.6474 (-11.788)	0.2598 (5.436)	0.4882 (9.177)	0.5321 (9.476)	0.3186 (6.491)
Lower triangular Cholesky factors of the correlation matrix					
Happiness	1				
Pain	-0.1422 (-5.676)	0.9898			
Sadness	-0.3948 (-11.663)	0.3001 (9.666)	0.8684		
Stress	-0.3866 (-11.627)	0.3542 (11.34)	0.5786 (13.889)	0.6248	
Tiredness	-0.2211 (-8.556)	0.3743 (14.488)	0.2006 (7.593)	0.3623 (10.454)	0.7997

[a]Base:Income below $25,00 Thousand, [b]Base: Age below 35 years old, [c]Base: Less than high school education, [d]Base: Poor health condition, [e]Base: Very good life condition

Table 4: Estimation Results for the Measurement Equation Model of Latent Variables

	Constants	Standard Deviation	Loading on Happiness	Loading on Pain	Loading on Sadness	Loading on Stress	Loading on Tiredness
Indicator1	4.1239 (31.944)	1.1216 (37.093)	0.8981 (30.723)				
Indicator2	3.9851 (26.996)	1.0867 (32.521)	1.0445 (31.174)				
Indicator3	3.9767 (27.061)	1.0959 (37.702)	1.03 (31.87)				
Indicator4	2.2478 (11.826)	0.8649 (27.194)		1.2136 (40.646)			
Indicator5	2.2877 (11.344)	0.5876 (16.561)		1.3048 (43.426)			
Indicator6	2.2463 (11.741)	0.8391 (25.621)		1.2244 (40.613)			
Indicator7	1.2093 (7.629)	0.8942 (27.294)			0.8687 (25.302)		
Indicator8	1.1819 (7.511)	0.8017 (22.911)			0.881 (25.347)		
Indicator9	1.1502 (7.59)	0.8268 (25.181)			0.851 (25.525)		
Indicator10	1.7383 (9.529)	1.1459 (39.53)				1.0932 (21.778)	
Indicator11	1.6611 (8.692)	1.0158 (33.18)				1.153 (21.847)	
Indicator12	1.5193 (8.837)	1.0894 (38.23)				1.0357 (23.19)	
Indicator13	2.8038 (18.647)	1.3196 (47.503)					1.2022 (35.132)
Indicator14	3.0506 (17.109)	0.9496 (26.92)					1.4854 (34.745)
Indicator15	3.4224 (21.773)	1.4169 (53.65)					1.2972 (32.984)

Most of the coefficients are statistically significant and provide plausible behavioral interpretations. Females appear to have a higher level of both positive and negative emotions. Individuals with higher income are generally found to be associated with lower negative emotions. However, it was interesting to find that income didn't have a significant impact on happiness itself. This observation is in line with earlier research from the field of happiness where it was also found that higher income does not necessarily make people happier despite general belief that it would (Kahneman et al. 2006). It was found that positive and negative emotions seem to vary in differing ways across various age groups. For example, people above 55 years old seem to be happier as well as less stressed and less tired compared to others. On the other hand with regard to the negative emotions of pain and sadness it appears like they are increasing with aging in general. Education attainment was found to significantly impact happiness and stress. Individuals who have high levels of educational attainment are found to be less happy – it may be likely that individuals who are highly educated may generally be more critical about their feeling of happiness. Also, there is a significant trend of increased stress with higher levels of education attainment.

Presence of spouse or partner in the household appears to have a positive impact on the happiness and negative impact on the feelings of pain and sadness. The effect of unemployment was found to be significant only for tiredness. It is plausible that people who are unemployed for long durations tend to get used to their circumstances and do not let their employment status influence their general moods. There is also evidence to this end in the area of happiness (e.g. Clark 2006). Finally, both the conditions of health and life were found to have a very substantial influence (both in terms of statistical significance and magnitude of the coefficient estimates) on moods. As one would expect, good health was found to be negatively associated with all 4 negative emotions and positively associated with feeling of happiness. Similar association was also observed for evaluation of overall life satisfaction on the different emotions wherein poor life satisfaction was associated with higher levels of negative emotions and also with lower level of happiness. It should be noted that, the significant contribution of health and life condition of individual on the latent constructs (i.e. mood) further lend evidence on the stability of these affective states of individual and supports the validity of the day level construction of mood in this particular empirical context.

One of the many desirable features of the HMDC formulation is its ability to accommodate correlations between error terms due to unobserved explanatory variables. A full correlation matrix across the five latent constructs was explored and the estimates of the lower triangular Cholesky matrix corresponding to the correlation matrix are reported in Table 3. It can be seen that all estimates of the lower triangular Cholesky values are very significant. The correlation matrix corresponding to the Cholesky values is reported in Equation 19 below:

$$\Gamma = \begin{pmatrix} & \textbf{Happiness} & \textbf{Pain} & \textbf{Sadness} & \textbf{Stress} & \textbf{Tiredness} \\ \textbf{Happiness} & 1 & -0.1422 & -0.3948 & -0.3866 & -0.2211 \\ \textbf{Pain} & -0.1422 & 1 & 0.3532 & 0.4056 & 0.4019 \\ \textbf{Sadness} & -0.3948 & 0.3532 & 1 & 0.7614 & 0.3738 \\ \textbf{Stress} & -0.3866 & 0.4056 & 0.7614 & 1 & 0.5605 \\ \textbf{Tiredness} & -0.2211 & 0.4019 & 0.3738 & 0.5605 & 1 \end{pmatrix} \tag{19}$$

As expected, the feeling of happiness is negatively correlated with all the four negative emotions while the four negative emotions are positively correlated to each other. Also, among the five moods, stress seems to have the strongest correlation with the rest of the emotions. The magnitude of correlation between stress and sadness is the highest.

4.4.2 Measurement Equation Model of Latent Variables

The purpose of the measurement equation is to help define the underlying latent constructs. In the empirical study the indicators are treated as continuous indicators. The measurement equation parameter estimates themselves do not provide any interesting behavioral insights. As noted earlier as part of the exploratory factor analysis, all the indicators load positively and significantly on each of the 5 latent moods further validating the construction/definition of the latent variables as moods that sustain over the course of a day.

4.4.3 Multiple Discrete Continuous (MDC) Choice Model

The parameter estimates for the MDC choice model are presented in this subsection. The influence of observed exogenous variables are presented first followed by a discussion of the association between moods and the weekend discretionary activity engagement behaviors.

The choice of the exogenous variables in the MDC model was motivated by previous research on the topic of activity engagement (Garikapati et al. 2014, Pinjari and Bhat 2010, and Srinivasan and Bhat 2006 among others). The findings are in line with the earlier literature on the topic. Also it should be noted that, some of the exogenous variables explored in the MDC model were also included in the structural equation model of the latent variables. In other words there is a direct influence of the observed explanatory variables and there is also an indirect effect of these variables mediated through the latent variable. In this section, only the direct influence of the observed exogenous variables on the discretionary activity participation and time allocation decisions are discussed[11]. A number of household- and person-level exogenous variables were explored. Additionally, built environment variables and day of week for which activities are reported are used to further explain the heterogeneity in the activity engagement behaviors. Lastly, the latent constructs are used to understand the role of moods.

4.4.3.1 Baseline Marginal Utility

Results for the baseline marginal utility (see Table 5) which provide insights into the participation choices (i.e. what activities to participate in) of the individuals are discussed in this subsection. Passive leisure was chosen as the reference alternative. The constants of the MDC choice model capture the influence of the average unexplained effect after accounting for different exogenous and endogenous variables. Both the signs and the magnitudes of the constants are consistent with expectations. All else being equal, passive leisure was found to be the most popular discretionary activity to participate in followed by social activity and active leisure; individuals appear to have the least propensity to participate in attending sports and arts events.

[11] An analysis of the total effect of different exogenous variables (calculated from the direct and indirect effect) did not reveal any change in the sign of the coefficient corresponding to different exogenous variables.

Table 5: Estimation Results for the Multiple Discrete Continuous Choice Model (Baseline Marginal Utility)

	Active leisure	Physical activity	Non maintenance shopping	Attending sports and arts events	Social activity
Coefficients to the exogenous variables					
Constants	-1.6013 (-19.15)	-2.2099 (-17.311)	-2.1471 (-18.364)	-3.1597 (-11.309)	-1.4192 (-17.035)
Individual level characteristics					
Female indicator	0.3787 (6.851)		0.2624 (4.314)	0.2044 (2.672)	0.4217 (7.613)
Young age indicator	0.1472 (2.292)	0.1266 (1.589)		0.2003 (2.311)	
Old Age indicator	0.2725 (4.474)	-0.1935 (-2.18)			
Student indicator		0.3107 (3.087)			
Employment indicator		0.1122 (1.679)	0.2744 (4.6)	0.2716 (3.432)	0.1515 (3.198)
Disability indicator	-0.1908 (-1.987)	-0.485 (-3.493)	-0.4889 (-4.055)	-0.3065 (-1.874)	-0.3554 (-3.941)
Household level characteristics					
HH income indicator ($25 to $50 Thousand)				0.1188 (1.392)	
HH income indicator ($50 to $100 Thousand)	0.1795 (2.871)	0.1827 (2.486)			
HH income indicator (More than $100 Thousand)	0.429 (4.417)	0.5086 (4.81)	0.2304 (2.359)	0.3739 (3.21)	0.113 (1.321)
Spouse/partner indicator				0.1568 (2.097)	
Presence of kid indicator (Age 0-5)	0.2424 (4.019)		0.0931 (1.319)		
Presence of kid indicator (Age6- 12)				-0.2787 (-2.529)	
Presence of kid indicator (Age 13-17)			0.1575 (1.783)		0.1078 (1.511)
Built environment characteristic					
Metropolitan indicator			-0.1732 (-2.28)		
Day of week indicator					
Saturday indicator	0.0791 (1.622)	0.1265 (2.093)	0.2985 (5.224)	0.358 (4.397)	0.0659 (1.421)
Coefficients to the endogenous latent variable					
Happiness	-0.0533 (-1.656)	0.1375 (3.41)	0.0041 (0.109)	0.1391 (2.687)	0.1127 (3.572)
Pain	-0.0715 (-2.242)	-0.0923 (-2.339)	-0.0358 (-1.019)	-0.1382 (-2.451)	-0.0543 (-1.852)
Sadness	-0.0747 (-1.329)	0.0193 (0.248)	-0.2854 (-4.144)	-0.2322 (-2.509)	-0.1111 (-2.039)
Stress	0.1106 (1.679)	0.0463 (0.516)	0.3788 (4.955)	0.3546 (3.4)	0.2202 (3.53)
Tiredness	-0.0557 (-1.398)	0.0421 (0.822)	-0.1057 (-2.331)	-0.0084 (-0.135)	-0.0301 (-0.802)

Table 6: Estimation Results for the Multiple Discrete Continuous Choice Model (Satiation Parameter)

	Active leisure	Passive leisure	Physical activity	Non maintenance shopping	Attending sports and arts events	Social activity
Coefficients to the exogenous variables						
Constants	5.2008 (53.88)	5.1603 (59.266)	5.4227 (39.278)	4.4034 (36.711)	7.5567 (13.302)	5.3042 (46.659)
Individual level characteristics						
Female indicator	-0.169 (-1.882)		-0.4885 (-4.002)	0.2183 (2.24)		-0.2001 (-2.101)
Young age indicator	0.2537 (2.429)			0.177 (1.717)		0.2833 (2.727)
Old Age indicator						-0.3797 (-3.635)
Household level characteristics						
HH income indicator ($25 to $50 Thousand)		-0.2565 (-3.166)	0.2859 (1.858)			
HH income indicator ($50 to $100 Thousand)	-0.1967 (-1.919)	-0.1913 (-2.243)				
HH income indicator (More than $100 Thousand)	-0.3889 (-3.549)	-0.2818 (-2.186)		0.1606 (1.466)		

Female respondents appear to participate more in social activity, active leisure, non-maintenance shopping and sports and arts events compared to male respondents. Younger individuals (ages 15 to 34 years) appear to have a higher propensity to engage in active leisure compared to individuals belonging to middle age group (35 to 64 years). This propensity seems to be even higher for the elderly (above 65 years). As expected, those who are in the youngest age group appear more inclined to participate in physical activity compared to those in the middle age group. The opposite seems to hold true for the elderly. Younger individuals also appear to have a higher propensity for sports and arts events compared to other age groups. Students were found to exhibit a higher tendency to participate in more physical activity compared to those who are not enrolled. Individuals who are employed seem to have an inclination to participate more into almost all types of discretionary activities other than active leisure compared to passive leisure. This is not surprising because those who are employed may have additional disposable income thus allowing them to seek discretionary activities other than passive leisure.

The disability indicator was found to have significant influence on the weekend discretionary activity participation. The coefficient was found to be negative for all discretionary activity types and highly significant. This is reasonable because it is likely that these individuals may be suffering from mobility restrictions and as a result participating less in different types of discretionary activities compared to those who do not have any disabilities. Among the different household level characteristics, those with high income appear to have higher propensity to participate in different types of discretionary activities compared to passive leisure. Presence of children was found to have a differing effect based on the age of the children. This is reasonable because older children may not be dependent on their parents' as much as younger children possibly leading to different types of discretionary activity engagement. It was interesting to note that whether the respondents reported their time use on Saturday or Sunday had a significant influence. Saturday indicator had a positive effect on participation in all five discretionary activity types compared to passive leisure. This is plausible because most individuals use Sunday as a day to relax and prepare for a new work week.

The association between moods and weekend discretionary activity participation is discussed below. It can be seen that the coefficients of all five moods: happiness, sadness, pain, stress and tiredness on all the discretionary activity types are shown in the table even though some of the coefficients are insignificant. This was done because examining the association between moods and discretionary activity engagement choices was the primary focus of the empirical study so even the insignificant coefficients are reported for the sake of completeness. It must be noted that no inferences are drawn for the moods with insignificant coefficient values; all insignificant coefficients of moods are highlighted in the table. The coefficient estimates provide support to the a priori hypothesis that people with high levels of positive emotions engage more in discretionary activities other than passive leisure. On the other hand, those individuals who suffer from negative emotions were found to do the opposite by participating more in passive leisure; one exception to this was the relationship between those who experience higher levels of stress (a negative emotion) on their discretionary activity participation choices.

In general it appears like people who are happy want to participate more in physical activity, sports and arts events and social activities compared to passive leisure. On the other hand those suffering from high levels of pain and sadness tend to participate less in discretionary activities other than passive leisure. Similar observations were also made for tiredness but it was found to significantly associate with participation in two activity categories namely active leisure and non-maintenance shopping. It is interesting to note that unlike other negative emotions, higher levels of stress were not associated with lower levels of participation in discretionary activities when compared to passive leisure. One plausible explanation to this observation may be how people cope when faced with stress – individuals may seek out opportunities (including look for moral and social support) to deal with stress (Scheier et al. 1986). It is also interesting to note that even though stress and other negative emotions were highly correlated, the association between these latent constructs and the activity participation choices are very different and quite the opposite.

It can be noted that some of the findings obtained from the current exploration can potentially be explained using an alternative direction of causality. For example, the positive association between positive moods and higher participation in physical activity, sports and arts events can also plausibly be because these activities can make people happy. Similarly, the positive association between stress and shopping activity may be because shopping is considered as a stressful activity by some individuals.

However, it is worth noting that the current analysis focuses on the association between individuals' moods that are "stable" over the day and their activity engagement choices on that day. Since these moods do not vary across the day, we believe that the plausibility of the causality we are testing (that stable moods on a day influence activity engagement on that day) is greater than that for the reverse causality (that activity engagement on a day influences stable moods on that day). Of course, it is likely that activity engagement habits over a long period of time influence stable moods people experience on a given day. Exploration of such long-term relationships between moods and activity engagement is not a focus of this study; albeit certainly worthy of future research and so is the exploration of relationship between moods that vary across a day and activity engagement.

4.4.3.2 Correlation Structure

Finally, different error structures were tested for the error components associated with the baseline marginal utilities of the different discretionary alternatives. More specifically, the presence of heteroscedasticity as well as correlation across different alternatives was explored. It must be noted that, theoretically it is possible to estimate all the $(\frac{n*(n-1)}{2} - 1)$ Cholesky factors corresponding to the error component of the choice model; where n is the number of choice alternatives. However, estimating the full covariance matrix (after normalization) does not allow inferring the underlying correlations among different alternatives. For this reason, in the current study, different correlation structures were assumed a priori (to test out different hypothesis) and the corresponding Cholesky factors were estimated. In particular, the presence of following correlation structures were explored:

- Correlation among non-maintenance shopping, attending sports and arts events and social activity
- Correlation between active and passive leisure
- Correlation among active leisure and the rest of the discretionary activities other than passive leisure

The estimation results indicate the presence of significant correlation between active leisure and three other discretionary activities namely physical activity, attending sports and arts events and social activity (with the corresponding Cholesky factors estimated as 0.1718, 0.1193 and 0.1050 respectively). As can be seen, the correlation structure of the choice model seems to be relatively sparse. This is likely due to the fact that the inclusion of latent constructs may have accounted for the error correlations due to the common unobserved factors resulting in relatively sparse correlation structure for the choice model (Hess 2012). This is another advantage of using the HMDC (and more generally the ICLV model) i.e. to be able to isolate and parse out the factors contributing to the correlation across different choice alternatives rather than relegating the correlations to the unobserved random factors.

4.4.3.3 Satiation Parameter

It must be noted that in addition to the baseline marginal utility, the satiation coefficients were also parameterized as a function of different exogenous variables to gain insights into the second dimension of activity engagement namely the time use dimension (i.e. amount of time spent in the discretionary activity types). The corresponding results are presented in Table 6. The coefficient of the exogenous variables in the satiation parameter indicates presence of statistically significant variation in satiation based on gender, age and income. Specifically, females exhibit higher satiation (meaning lower amount of consumption) for active leisure, physical activity and attending sports and arts events compared to males, while the opposite is true for non-maintenance shopping activities. Those who are in the young age group exhibit lower level of satiation (higher amount of consumption) for active leisure, non-maintenance shopping and social activities, while those in the old age group exhibit high level of satiation for social activities. In terms of income, the effect was found to be statistically significant for the two types of leisure activities. Those with higher income show higher satiation meaning lower level of consumption for both types of leisure activities.

4.5 Validation Study

This section briefly introduces the forecasting steps for the proposed model formulation and also highlights the results of a validation study using holdout sample. The validation sample consisted of 977 observations from ATUS dataset. For forecasting the activity participation and time use choices, one needs to use the structural equation of the latent variable and the MDC choice model only. The measurement equations of the latent variable are not needed for the forecasting. More specifically, in forecasting the activity engagement choices, the below steps were carried out:

- Predict the latent variables using the structural equation of the latent variables (i.e. Equation 1).
- Using the predicted latent variables and other exogenous variables of the MDC choice model, predict the activity participation and time use choices (i.e. consumption quantities for vector x) using the forecasting procedure proposed by Pinjari and Bhat (2011).

However, due to the presence of random error component in both the structural equation of the latent variable and the MDC choice model, the activity engagement choices are predicted with multiple draws of error (100, 200, 500, 1000, 2000, 5000, and 10000). For each set of draws, the average participation rate and average amount of time allocated to various activities are calculated for each individual. The average value of the participation rate and consumption are found to be very stable across different error draws. Also, the calculated standard deviations across draws are found to be very small even for 100 draws of error. Finally, the forecasted values of the participation rate (in percentage) and consumption average (in minutes) are compared against the corresponding observed values from the hold out sample. The forecasting errors are calculated using equation (20) and (21).

$$Mean\ Absolute\ Error\ (MAE) = \frac{\sum_{k=1}^{K}\left|\frac{(y_k - \widehat{y_k})}{y_k}\times 100\right|}{K} \tag{20}$$

$$Root\ Mean\ Square\ Error\ (RMSE) = \sqrt{\frac{\sum_{k=1}^{K}(y_k - \widehat{y_k})^2}{K}} \quad (21)$$

where, y_k is the observed average participation rate (or consumption amount) for alternative k and $\widehat{y_k}$ is the predicted average participation rate (or consumption amount) for alternative k. The calculated MAE and RMSE along with the predicted and observed participation rate and consumption amount are presented in Table 7[12]. It can be seen that, the HMDC model provides reasonable forecasts with approximately 10 percent MAE for participation rate and approximately 9 percent MAE for the average amount of time allocated. The low values of RMSE (approximately 2 and 3.5 for participation and time allocated respectively) also points to the good predicting ability of the estimated HMDC model.

Table 7: Validation Results of the HMDC Model using Hold Out Sample

		Activity Categories					
		Active leisure	**Passive leisure**	**Physical activity**	**Non maintenance shopping**	**Attending sports and arts events**	**Social activity**
Observed	Participation Rate (%)	41.965	86.285	19.959	26.510	7.984	48.516
	Consumption (in minutes)	61.552	232.127	23.226	19.205	14.444	71.311
Predicted	Participation Rate (%)	45.038	87.764	17.668	24.716	5.754	46.62
	Consumption (in minutes)	68.618	228.426	21.608	21.657	11.774	69.782
MAE (%)	Participation Rate	9.853					
	Consumption	8.907					
RMSE	Participation Rate	2.186					
	Consumption	3.691					

5. SUMMARY AND FUTURE RESEARCH

In the travel behavior arena, researchers often explain the heterogeneity in activity-travel choices across individuals using a variety of observed explanatory variables such as socio-economic, demographic, and built environment factors in models of the choices. With theoretical and methodological advances in the behavioral economics, there is a growing recognition that heterogeneity in the individual behaviors arises also due to differences in individual psychological factors (e.g. attitudes, preferences, and moods among others). The paper develops a new hybrid

[12] Reported results were produced using 1000 draws of errors. Any further increase in the number of draws did not change the predicted values of activity participation and time allocation.

multiple discrete continuous (HMDC) model formulation. HMDC is an Integrated Choice and Latent Variable (ICLV) implementation which allows simultaneous estimation of latent variable model and choice model in the presence of MDC kernel.

Apart from the HMDC model formulation, a major challenge in the research was to come up with a simulation free, analytical estimator for estimating the parameters of the model. CML approximation technique was employed to decompose the high dimensional integrals into lower dimensional marginal densities that can be evaluated using analytical approximation techniques. Another challenge was the variable size of the integral across observation because of the presence of the MDC kernel which resulted in varying number of marginal densities to be evaluated for each individual. To normalize the contribution from each observation to the likelihood function, a non-unit weight was used that was proportional to the size of the integral to be decomposed. The use of the *weight* significantly improved the consistent recovery of the true parameters.

In general, the proposed estimation routine provides a very good recovery of parameters both in terms of bias and asymptotic efficiency. An average absolute percentage bias (APB) value of 0.64% and an average relative asymptotic efficiency (RAE) of 1.099 were obtained across all parameter values. Further, a comparison of the simulation results between weighted and unweighted CML approach reveals striking differences in recovering unbiased estimates of the parameters. An unweighted CML resulted in an average APB of 10.39% as opposed to less than 1% bias obtained using *weighted* CML. This demonstrates the importance of *weights* in setting up the CML function when the dimension of the integral to be evaluated varies across observations.

The empirical study conducted to demonstrate the feasibility of the proposed framework investigates the association between moods and the discretionary activity engagement choices of individuals on weekends. In particular, the study attempts to examine if higher levels of positive emotion (such as happiness) would be associated with participating more in activities other than passive leisure. Alternatively, the study also wanted to explore if higher levels of negative emotions (such as pain, sadness, tiredness and stress) are associated with participating more in passive leisure.

To this end, data from the 2013 American Time Use Survey was used. The HMDC formulation was employed to explore the role of five moods: happiness, sadness, pain, stress and tiredness on participation and time allocation to six discretionary activities: active leisure, passive leisure, physical activity, shopping for non-maintenance, attending sports and arts events, and social activity with passive leisure serving as the reference activity type. The empirical exploration provided statistically significant evidence in support of the association between positive and negative moods and the weekend discretionary activity engagement choices after controlling for the effect of various observed explanatory variables. A validation exercise was performed using a holdout sample technique to demonstrate the validity and applicability of the HMDC model for forecasting. The results (low forecasting error) point to the ability of the HMDC model to provide valid predictions.

It can be noted that, the current empirical study explores the association between moods and discretionary activity participation propensity through the parameterization of the baseline marginal utility. Future research is needed to investigate the association between moods and satiation patterns of different discretionary activities. This would help investigate if positive (negative) moods are associated with seeking more (less) diversity in the choice of discretionary activities as posited by previous research on the influence of positive moods in variety seeking behavior (Kahn and Isen 1993).

The research presented in this paper has both methodological and empirical contributions. First, on the methodological front, the HMDC comprises one of the first attempts to implement a MDC choice kernel into the ICLV framework. To the authors' knowledge, generalized heterogeneous data model (GHDM) proposed by Bhat et al (2016) is the only other attempt to estimate simultaneous equation system using composite marginal likelihood technique (CML) that involves latent variable model and MDC choice outcomes. However, the current paper is perhaps the first to highlight and demonstrate the importance of using weights in setting up the CML function to estimate the parameters of an ICLV framework with MDC choice kernel. HMDC is general enough and allows for exploration of complex error structures to accommodate correlations across latent variables, and correlations across alternatives. The formulation of HMDC is also flexible and allows for treating indicator variables used in constructing the latent variables as both ordinal and continuous.

It should be emphasized that, the empirical exploration conducted as part of the study does not intend to recommend policy interventions based on the findings – rather identifying and characterizing the additional heterogeneity (through the addition of latent constructs of emotions) in the activity time allocation behavior after accounting for traditional exogenous variables was the main objective of the empirical exploration. The empirical study sheds light into the interrelationships among different types of emotions throughout the day (namely happiness, pain, sadness, stress and tiredness) as well as highlight the association between daily moods and daily activity time allocation after accounting for other traditional exogenous variables. Additionally, the endogenous treatment of latent mood variables allowed the study of variation in individual moods as a function of different exogenous explanatory variables which is a topic of interest in the field of happiness and hedonic psychology.

There exist a number of avenues for future research both on the methodological and empirical fronts based on the research presented in the paper. Research is warranted on the appropriate choice of weight in the proposed estimator. Exploring the suitable choice of weight (in terms of relative efficiency) based on different dependency structure would be a valuable addition leading to more efficient estimator. On the empirical side, the proposed HMDC formulation and estimation technique can be readily employed to explore the association between other types of psychological factors (such as life style choice and personality type) and the activity engagement choices. Also, it is believed that, with the increasing interest in studying the role of individual attitude (and other psychological factors) on various activity-travel choice decisions of interest namely household energy consumption (Abrahamse and Steg 2009, Hartmann and Apaolaza-lbanez 2012, Azadeh et al. 2014), vehicle holding and vehicle usage behavior (Siriwardena 2010, Wang et al. 2016), physical and leisure activity participation (Deforche et al. 2006) the proposed HMDC formulation and associated estimation routine can be used due to its statistical rigor and richness in behavioral representation.

6. ACKNOWLEDGEMENTS

The authors would like to acknowledge Dr. Chandra Bhat at the University of Texas at Austin for making various GAUSS codes publicly available including those for MDCP (Bhat et al. 2013), ICLV model formulation (Bhat and Dubey 2014) and MACML analytical approximation (Bhat et al. 2011). These codes were greatly useful for developing the HMDC model formulation and estimation routine presented in this study. Authors are also thankful to two anonymous reviewers for their invaluable feedback on an earlier version of the paper.

7. FUNDING

The third author gratefully acknowledges partial support provided by the Center for Teaching Old Models New Tricks (TOMNET), a University Transportation Center sponsored by the US Department of Transportation through Grant No. 69A3551747116.

REFERENCES

1. Abou-Zeid, M. & M. Ben-Akiva.(2012). Well-being and activity-based models. Transportation 39(6),1189-1207.
2. Abrahamse, W., & Steg, L. (2009). How do socio-demographic and psychological factors relate to households' direct and indirect energy use and savings?. Journal of economic psychology, 30(5), 711-720.
3. Alvarez-Daziano, R., and D. Bolduc.(2013). Incorporating Pro-environmental Preferences towards Green Automobile Technologies through a Bayesian Hybrid Choice Model. Transportmetrica A: Transport Science 9(1),74-106.
4. Anable, J. (2005). 'Complacent car addicts' or 'aspiring environmentalists'? Identifying travel behaviour segments using attitude theory. Transport Policy, 12(1), 65-78.
5. Ashok, K., Dillon, W. R., & Yuan, S. (2002). Extending discrete choice models to incorporate attitudinal and other latent variables. Journal of Marketing Research, 39(1), 31-46.
6. Atasoy, B., Glerum, A., & Bierlaire, M. (2013). Attitudes towards mode choice in Switzerland. disP-The Planning Review, 49(2), 101-117.
7. Azadeh, A., Narimani, A., & Nazari, T. (2014). Estimating household electricity consumption by environmental consciousness. International Journal of Productivity and Quality Management, 15(1), 1-19.
8. Ben-Akiva, M., McFadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., Bolduc, D., Borsch-Supan, A., Brownstone, D., Bunch, D. S., Daly, A., De Palma, A., Gopinath, D., Karlstrom, A., & Munizaga, M. A.(2002). Hybrid choice models: Progress and challenges. Marketing Letters, 13(3),163-175.
9. Bhat, C. R. (2005). A multiple discrete–continuous extreme value model: formulation and application to discretionary time-use decisions. Transportation Research Part B: Methodological, 39(8), 679-707.
10. 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: Methodological, 42(3),274-303.
11. Bhat, C. R.(2011). The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models. Transportation Research Part B: Methodological, 45(7),923-939.
12. Bhat, C. R., & Dubey, S. K.(2014). A new estimation approach to integrate latent psychological constructs in choice modeling. Transportation Research Part B: Methodological, 67,68-85.
13. Bhat, C. R., Astroza, S., Bhat, A. C., & Nagel, K. (2016). Incorporating a multiple discrete-continuous outcome in the generalized heterogeneous data model: Application to residential self-selection effects analysis in an activity time-use behavior model. Transportation Research Part B: Methodological, 91, 52-76.
14. Bhat, C. R., Castro, M., & Khan, M.(2013). A new estimation approach for the multiple discrete–continuous probit (MDCP) choice model. Transportation Research Part B: Methodological, 55,1-22.
15. Bhat, C. R., Sen, S., & Eluru, N. (2009). The impact of demographics, built environment attributes, vehicle characteristics, and gasoline prices on household vehicle holdings and use. Transportation Research Part B: Methodological, 43(1), 1-18.
16. Bhat, C. R., Sener, I. N., & Eluru, N. (2010). A flexible spatially dependent discrete choice model: formulation and application to teenagers' weekday recreational activity participation. Transportation research part B: methodological, 44(8), 903-921.

17. Biying, Y., Zhang, J., & Fujiwara, A. (2012). Analysis of the residential location choice and household energy consumption behavior by incorporating multiple self-selection effects. Energy Policy, 46, 319-334.
18. Bolduc, D., & Alvarez-Daziano, R. (2010). On estimation of hybrid choice models. In Choice Modelling: The State-of-the-Art and the State-of-Practice: Proceedings from the Inaugural International Choice Modelling Conference. Emerald Group Publishing (p. 259).
19. Bolduc, D., M. Ben-Akiva, J. Walker, A. Michaud.(2005). Hybrid Choice Models with Logit Kernel: Applicability to Large Scale Models. In: Lee-Gosselin, M., Doherty, S. (eds.) Integrated Land-Use and Transportation Models: Behavioural Foundations, Elsevier, Oxford, pp.275-302.
20. Bollen, K.(1983). Structural equations with latent variables. New York: John Wiley, 19891, 612-21.
21. Chen, C., and Mokhtarian., P. L. (2006). Tradeoffs Between Time Allocations to Maintenance Activities/Travel and Discretionary Activities/Travel. Transportation, 33(3), 223–240.
22. Chorus, C. G., & Kroesen, M. (2014). On the (im-) possibility of deriving transport policy implications from hybrid choice models. Transport Policy, 36, 217-222.
23. Clark, A.(2006). A note on unhappiness and unemployment duration. Discussion Paper, Institute for the Study of Labor (Website: http://ftp.iza.org/dp2406.pdf, Accessed: July 29, 2015).
24. Daly, A., S. Hess, B. Patruni, D. Potoglou, and C. Rohr(2012). Using Ordered Attitudinal Indicators in a Latent Variable Choice Model: a study of the impact of security on rail travel behavior." Transportation, 39(2),pp.267-297.
25. Daziano, R. A. (2015). Inference on mode preferences, vehicle purchases, and the energy paradox using a Bayesian structural choice model. Transportation Research Part B: Methodological, 76, 1-26.
26. Deforche, B. I., De Bourdeaudhuij, I. M., & Tanghe, A. P. (2006). Attitude toward physical activity in normal-weight, overweight and obese adolescents. Journal of adolescent health, 38(5), 560-568.
27. Ettema, D., Gärling, T., Olsson, L. E., & Friman, M.(2010). Out-of-home activities, daily travel, and subjective well-being. Transportation Research Part A: Policy and Practice, 44(9),723-732.
28. Forgas, J. P. (1989). Mood effects on decision making strategies. Australian Journal of Psychology, 41(2), 197-214.
29. Frederick, S., Loewenstein, G., 1999. Hedonic adaptation. In: Kahneman, D., Diener, E., Schwarz, N. (Eds.), Foundations of Hedonic Psychology: Scientific Perspectives on Enjoyment and Suffering. Russell Sage Foundation, New York, pp. 302–329.
30. Fredrickson, B. L.(2001). The role of positive emotions in positive psychology. American psychologist, 56(3),218-226.
31. Gadermann, A.M., Zumbo, B.D., 2007. Investigating the inter-individual variability and trajectories of subjective well-being. Social Indicators Research 81, 1–33.
32. Garikapati, V., You, D., Pendyala, R., Vovsha, P., Livshits, V., & Jeon, K.(2014). Multiple Discrete-Continuous Model of Activity Participation and Time Allocation for Home-Based Work Tours. Transportation Research Record: Journal of the Transportation Research Board, (2429),90-98.
33. Gärling, T.(1998). Behavioural assumptions overlooked in travel choice modelling. Travel Behaviour Research: Updating the state of play, 3-18.

34. Gerdtham, U. G., & Johannesson, M.(2001). The relationship between happiness, health, and socio-economic factors: results based on Swedish microdata. The Journal of Socio-Economics, 30(6),553-557.
35. Glerum, A., & Bierlaire, M.(2012). Accounting for response behavior heterogeneity in the measurement of attitudes: an application to demand for electric vehicles. In Proceedings Swiss Transport Conference.
36. Godambe, V.P.,(1960). An optimum property of regular maximum likelihood estimation. The Annals of Mathematical Statistics, 31(4),1208-1211.
37. Golob, T. F., Horowitz, A. D., & Wachs, M.(1977). Attitude-behavior relationships in travel demand modelling (No. GMR-2398).
38. Harris, K. M., & Keane, M. P. (1998). A model of health plan choice: Inferring preferences and perceptions from a combination of revealed preference and attitudinal data. Journal of Econometrics, 89(1), 131-157.
39. Hartmann, P., & Apaolaza-Ibáñez, V. (2012). Consumer attitude and purchase intention toward green energy brands: The roles of psychological benefits and environmental concern. Journal of Business Research, 65(9), 1254-1263.
40. Hess, S.(2012). Rethinking heterogeneity: the role of attitudes, decision rules and information processing strategies. Transportation Letters, 4(2),105-113.
41. Hess, S., & Spitz, G. (2016). How much do attitudes and values matter in mode choice?. In Transportation Research Board 95th Annual Meeting (No. 16-4839).
42. Isen, A. M.(2001). An influence of positive affect on decision making in complex situations: Theoretical issues with practical implications. Journal of consumer psychology, 11(2),75-85.
43. Jäggi, B., Erath, A., Dobler, C., & Axhausen, K. (2012). Modeling household fleet choice as function of fuel price by using a multiple discrete-continuous choice model. Transportation Research Record: Journal of the Transportation Research Board, (2302), 174-183.
44. Jeong, J., Kim, C. S., & Lee, J. (2011). Household electricity and gas consumption for heating homes. Energy Policy, 39(5), 2679-2687.
45. Joe, H. and Lee, Y.(2009), On weighting of bivariate margins in pairwise likelihood, Journal of Multivariate Analysis 100(4),670-685.
46. Kahn, B. E., & Isen, A. M.(1993). The influence of positive affect on variety seeking among safe, enjoyable products. Journal of Consumer Research,257-270.
47. Kahneman, D., Krueger, A. B., Schkade, D., Schwarz, N., & Stone, A. A.(2006). Would you be happier if you were richer? A focusing illusion. science, 312(5782),1908-1910.
48. Kamargianni, M., & Polydoropoulou, A. (2013). Hybrid choice model to investigate effects of teenagers' attitudes toward walking and cycling on mode choice behavior. Transportation Research Record: Journal of the Transportation Research Board, (2382), 151-161.
49. Kaza, N., Towe, C., & Ye, X. (2011). A hybrid land conversion model incorporating multiple end uses. Agricultural and Resource Economics Review, 40(3), 341.
50. Kim, J., Rasouli, S., & Timmermans, H. (2014). Hybrid choice models: principles and recent progress incorporating social influence and nonlinear utility functions. Procedia Environmental Sciences, 22, 20-34.
51. Kitamura, R.(1988). An evaluation of activity-based travel analysis. Transportation, 15(1-2),9-34.
52. Koppelman, F. S., & Hauser, J. R. (1978). Destination choice behavior for non-grocery-shopping trips. Transportation Research Record, (673).
53. Kuk, A. Y., & Nott, D. J. (2000). A pairwise likelihood approach to analyzing correlated binary data. Statistics & Probability Letters, 47(4), 329-335.

54. LaMondia, J., Bhat, C. R., & Hensher, D. A. (2008). An annual time use model for domestic vacation travel. Journal of Choice Modelling, 1(1), 70-97.
55. Legrain, A., Eluru, N., & El-Geneidy, A. M. (2015). Am stressed, must travel: the relationship between mode choice and commuting stress. Transportation research part F: traffic psychology and behaviour, 34, 141-151.
56. Lingling, W. U., Zhang, J., & Fujiwara, A. (2011). Representing Tourist's Time Use Behavior Based on a Multiple Discrete-Continuous Extreme Value Model. Journal of the Eastern Asia Society for Transportation Studies, 9, 796-809.
57. Loewenstein, G., & Lerner, J. S. (2003). The role of affect in decision making. Handbook of affective science, 619(642), 3.
58. Madanat, S. M., Yang, C. T. D., & Ying-Ming, Y. E. N. (1995). Analysis of stated route diversion intentions under advanced traveler information systems using latent variable modeling. Transportation Research Record, (1485), 10-17.
59. McFadden, D.(1986). The choice theory approach to market research. Marketing Science, 5(4),275–297.
60. Mokhtarian, P. L., Papon, F., Goulard, M., & Diana, M. (2015). What makes travel pleasant and/or tiring? An investigation based on the French National Travel Survey. Transportation, 42(6), 1103-1128.
61. Morris, E. A., & Guerra, E. (2015). Mood and mode: does how we travel affect how we feel?. Transportation, 42(1), 25-43.
62. Oishi, S., Diener, E. F., Lucas, R. E., & Suh, E. M. (1999). Cross-cultural variations in predictors of life satisfaction: Perspectives from needs and values. Personality and social psychology bulletin, 25(8), 980-990.
63. Pendyala, R. M., and Bhat., C. R. (2004). An Exploration of the Relationship Between Timing and Duration of Maintenance Activities. Transportation, 31(4), 429–456.
64. Pinjari, A. R., & Bhat, C. (2011). An efficient forecasting procedure for Kuhn-Tucker consumer demand model systems. Technical paper. Department of Civil & Environmental Engineering, University of South Florida.
65. Pinjari, A. R., & Bhat, C.(2010). A Multiple Discrete–Continuous Nested Extreme Value (MDCNEV) model: formulation and application to non-worker activity time-use and timing behavior on weekdays. Transportation Research Part B: Methodological, 44(4),562-583.
66. Pinjari, A. R., Augustin, B., Sivaraman, V., Imani, A. F., Eluru, N., & Pendyala, R. M. (2016). Stochastic frontier estimation of budgets for Kuhn–Tucker demand systems: Application to activity time-use analysis. Transportation Research Part A: Policy and Practice, 88, 117-133.
67. Pinjari, A. R., Bhat, C. R., & Hensher, D. A. (2009). Residential self-selection effects in an activity time-use behavior model. Transportation Research Part B: Methodological, 43(7), 729-748.
68. Ravulaparthy, S., Yoon, S.Y., and Goulias, K.G. (2013). Linking elderly transport mobility and subjective well-being. Transportation Research Record: Journal of the Transportation Research Board, 2382, 28-36.
69. Richards, T. J., Gómez, M. I., & Pofahl, G. (2012). A multiple-discrete/continuous model of price promotion. Journal of Retailing, 88(2), 206-225.
70. Scheier, M. F., Weintraub, J. K., & Carver, C. S. (1986). Coping with stress: divergent strategies of optimists and pessimists. Journal of personality and social psychology, 51(6), 1257.

71. Sener, I. N., Copperman, R. B., Pendyala, R. M., & Bhat, C. R. (2008). An analysis of children's leisure activity engagement: examining the day of week, location, physical activity level, and fixity dimensions. Transportation, 35(5), 673-696.
72. Shin, J., Choi, J. Y., & Lee, D. (2016). Model for studying commodity bundling with a focus on consumer preference: Evidence from the Korean telecommunications market. Simulation, 92(4), 311-321.
73. Siriwardena, S. D. (2010). Effectiveness of eco-marketing on green vehicle purchase behavior in Maine: A nested-logit model approach (Doctoral dissertation, The University of Maine).
74. Srinivasan, S., & Bhat, C. R. (2006). A multiple discrete-continuous model for independent- and joint-discretionary-activity participation decisions. Transportation, 33(5), 497-515.
75. Train, K. (2003). Discrete choice methods with simulation. Cambridge university press.
76. 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.
77. Varin, C.(2008). On composite marginal likelihoods. AStA Advances in Statistical Analysis, 92(1),1-28.
78. Varin, C., Reid, N., & Firth, D. (2011). An overview of composite likelihood methods. Statistica Sinica, 21(1), 5-42.
79. Vij, A., & Walker, J. L. (2016). How, when and why integrated choice and latent variable models are latently useful. Transportation Research Part B: Methodological, 90, 192-217.
80. von Haefen, R.H., D.J. Phaneuf, and G.R. Parsons (2004). Estimation and welfare analysis with large demand systems. Journal of Business and Economic Statistics, 22(2), 194-205.
81. Walker, J., & Ben-Akiva, M.(2002). Generalized random utility model. Mathematical Social Sciences, 43(3),303-343.
82. Wang, S., Fan, J., Zhao, D., Yang, S., & Fu, Y. (2016). Predicting consumers' intention to adopt hybrid electric vehicles: using an extended version of the theory of planned behavior model. Transportation, 43(1), 123-143.
83. Xu, X., Reid, N.(2011). On the robustness of maximum composite likelihood estimate. Journal of Statistical Planning and Inference, 141(9),3047-3054.
84. Zhao, Y., & Joe, H.(2005). Composite likelihood estimation in multivariate data analysis. Canadian Journal of Statistics, 33(3), 335-356.

APPENDIX A
Table A.1a: Simulation Results (Without weight)

Parameters	True Values	Estimated Values	Absolute Percentage Bias (APB)	Asymptotic Standard Error (ASE)	Finite Sample Standard Error	Relative Asymptotic Efficiency (RAE)	Confidence Interval	Contains True Value?
α_1	1.1	1.0897	0.9341	0.0375	0.0314	1.1911	(1.02) - (1.16)	Yes
α_2	1.7	1.6805	1.1489	0.047	0.0472	0.9944	(1.59) - (1.77)	Yes
α_3	1.2	1.2403	3.3614	0.0561	0.0568	0.9877	(1.13) - (1.35)	Yes
α_4	1.8	1.8498	2.7656	0.0772	0.0668	1.1554	(1.7) - (2)	Yes
α_5	1.4	1.4343	2.4496	0.0989	0.0658	1.5037	(1.24) - (1.63)	Yes
α_6	1.6	1.6188	1.1753	0.1106	0.0763	1.4493	(1.4) - (1.84)	Yes
Γ_{21}	0.5	0.5091	1.8114	0.0423	0.0264	1.6039	(0.43) - (0.59)	Yes
Γ_{31}	0.5	0.5042	0.8348	0.0493	0.0288	1.7078	(0.41) - (0.6)	Yes
Γ_{32}	0.5	0.5098	1.9626	0.0712	0.0504	1.4137	(0.37) - (0.65)	Yes
$\bar{\delta}_1$	-1.1	-1.097	0.2717	0.0565	0.0532	1.0612	(-1.21) - (-0.99)	Yes
$\bar{\delta}_2$	-1.7	-1.6881	0.6975	0.0603	0.0545	1.1072	(-1.81) - (-1.57)	Yes
$\bar{\delta}_3$	-2.5	-2.4073	3.7073	0.1418	0.1595	0.8888	(-2.69) - (-2.13)	Yes
$\bar{\delta}_4$	-2	-1.9215	3.9241	0.1047	0.1206	0.8679	(-2.13) - (-1.72)	Yes
$\bar{\delta}_5$	-2	-1.9113	4.4373	0.113	0.0932	1.2127	(-2.13) - (-1.69)	Yes
$\bar{\delta}_6$	-2.8	-2.6779	4.36	0.1636	0.1723	0.9498	(-3) - (-2.36)	Yes
$\bar{d}_{11}$	1	1.0084	0.8357	0.0235	0.0216	1.0861	(0.96) - (1.05)	Yes
$\bar{d}_{21}$	1.1	1.1083	0.7557	0.0254	0.0212	1.2028	(1.06) - (1.16)	Yes
$\bar{d}_{32}$	1.2	1.1618	3.1862	0.0832	0.0852	0.9762	(1) - (1.32)	Yes
$\bar{d}_{42}$	1	0.9697	3.0289	0.061	0.0658	0.9272	(0.85) - (1.09)	Yes
$\bar{d}_{53}$	1.1	1.0809	1.7408	0.0943	0.0719	1.3114	(0.9) - (1.27)	Yes
$\bar{d}_{63}$	1.3	1.2624	2.8941	0.1211	0.0913	1.3271	(1.03) - (1.5)	Yes
Σ_{11}	1	0.9909	0.9104	0.0238	0.0242	0.985	(0.94) - (1.04)	Yes
Σ_{22}	1	0.9885	1.1523	0.0261	0.0279	0.9365	(0.94) - (1.04)	Yes
β_1	-1	-0.4639	53.6118	0.2269	0.2239	1.0132	(-0.91) - (-0.02)	No
β_2	2	2.5258	26.291	0.2219	0.2147	1.0336	(2.09) - (2.96)	No
β_3	-2	-1.7737	11.3173	0.2423	0.2328	1.0409	(-2.25) - (-1.3)	Yes
β_4	2.5	2.9306	17.2239	0.2895	0.2758	1.0496	(2.36) - (3.5)	Yes
β_5	-1	-1.1127	11.2674	0.0835	0.091	0.9181	(-1.28) - (-0.95)	Yes
β_6	3	3.0254	0.8469	0.0938	0.1144	0.8196	(2.84) - (3.21)	Yes
β_7	-1	-1.0204	2.0402	0.0413	0.0467	0.8854	(-1.1) - (-0.94)	Yes

Table A.1b: Simulation Results (Without weight)

Parameters	True Values	Estimated Values	Absolute Percentage Bias (APB)	Asymptotic Standard Error (ASE)	Finite Sample Standard Error (FSSE)	Relative Asymptotic Efficiency (RAE)	Confidence Interval	Contains True Value?
β_8	3.5	3.8435	9.813	0.1358	0.1349	1.0069	(3.58) - (4.11)	No
β_9	-3.5	-4.1934	19.8112	0.1981	0.2005	0.9881	(-4.58) - (-3.81)	No
λ_{21}	-1.5	-1.7127	14.1815	0.085	0.0831	1.0231	(-1.88) - (-1.55)	No
λ_{22}	1.2	1.3356	11.302	0.0791	0.0674	1.1743	(1.18) - (1.49)	Yes
λ_{23}	1.1	1.2342	12.1976	0.0988	0.0806	1.2256	(1.04) - (1.43)	Yes
λ_{31}	-1.6	-1.8031	12.6915	0.0859	0.0866	0.9927	(-1.97) - (-1.63)	No
λ_{32}	1.1	1.2249	11.3508	0.0746	0.0633	1.1776	(1.08) - (1.37)	Yes
λ_{33}	1	1.1263	12.6285	0.0919	0.0723	1.2714	(0.95) - (1.31)	Yes
λ_{41}	-1.4	-1.6142	15.3007	0.083	0.0828	1.0027	(-1.78) - (-1.45)	No
λ_{42}	1.3	1.4567	12.0521	0.0846	0.075	1.1274	(1.29) - (1.62)	Yes
λ_{43}	1.1	1.2482	13.4704	0.1003	0.0774	1.2955	(1.05) - (1.44)	Yes
λ_{51}	-3	-3.5056	16.8532	0.1483	0.1466	1.0113	(-3.8) - (-3.21)	No
λ_{52}	1	1.1296	12.9568	0.0777	0.0583	1.334	(0.98) - (1.28)	Yes
λ_{53}	1	1.1545	15.449	0.0994	0.0832	1.1949	(0.96) - (1.35)	Yes
γ_1	1.5	2.3809	58.7244	0.129	0.1024	1.2593	(2.13) - (2.63)	No
γ_2	1.8	2.5996	44.4234	0.095	0.1146	0.8291	(2.41) - (2.79)	No
γ_3	2.2	3.2027	45.5769	0.0982	0.1207	0.8133	(3.01) - (3.4)	No
γ_4	2.5	3.1685	26.74	0.1037	0.0999	1.0379	(2.97) - (3.37)	No
γ_5	2.8	3.6457	30.2044	0.1491	0.1578	0.9445	(3.35) - (3.94)	No
Λ_{22}	1.1	1.0966	0.3059	0.0748	0.0814	0.9188	(0.95) - (1.24)	Yes
Λ_{32}	0.6	0.5895	1.7479	0.0846	0.078	1.0841	(0.42) - (0.76)	Yes
Λ_{33}	0.9	0.7879	12.4534	0.0375	0.0394	0.9518	(0.71) - (0.86)	No
Λ_{44}	0.8	0.8266	3.3244	0.0879	0.0924	0.9517	(0.65) - (1)	Yes
Λ_{54}	1	1.0993	9.9325	0.1541	0.1403	1.0983	(0.8) - (1.4)	Yes
Λ_{55}	0.9	1.0123	12.4738	0.1183	0.117	1.0109	(0.78) - (1.24)	Yes
$\tau_{up,1}$	1.5	1.4671	2.1965	0.0781	0.078	1.0009	(1.31) - (1.62)	Yes
$\tau_{up,2}$	1.5	1.4791	1.3957	0.0663	0.0803	0.8262	(1.35) - (1.61)	Yes
$\tau_{up,3}$	1.5	1.4886	0.7632	0.0718	0.067	1.0702	(1.35) - (1.63)	Yes
$\tau_{up,4}$	1.5	1.4699	2.0044	0.0834	0.0862	0.9669	(1.31) - (1.63)	Yes
Mean			**10.3944**			**1.0989**		