

**A Multivariate Ordered Response Model System for Adults' Weekday Activity Episode  
Generation by Activity Purpose and Social Context**

**Nazneen Ferdous**

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

**Naveen Eluru**

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

**Chandra R. Bhat\***

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

**Italo Meloni**

The University of Cagliari,  
Department of Territorial Engineering  
Piazza d'Armi, 09123 Cagliari, Italy  
Phone: 39 70 675-5268, Fax: 39 70 675-5261  
E-mail: imeloni@unica.it

\*corresponding author

The research in this paper was completed when the corresponding author was a Visiting Professor at the Department of Territorial Engineering, University of Cagliari.

Original submission - August 2009; Revised submission - January 2010

## **ABSTRACT**

This paper proposes a multivariate ordered response system framework to model the interactions in non-work activity episode decisions across household and non-household members at the level of activity generation. Such interactions in activity decisions across household and non-household members are important to consider for accurate activity-travel pattern modeling and policy evaluation. The econometric challenge in estimating a multivariate ordered-response system with a large number of categories is that traditional classical and Bayesian simulation techniques become saddled with convergence problems and imprecision in estimates, and they are also extremely cumbersome if not impractical to implement. We address this estimation problem by resorting to the technique of composite marginal likelihood (CML), an emerging inference approach in the statistics field that is based on the classical frequentist approach, is very simple to estimate, is easy to implement regardless of the number of count outcomes to be modeled jointly, and requires no simulation machinery whatsoever.

The empirical analysis in the paper uses data drawn from the 2007 American Time Use Survey (ATUS) and provides important insights into the determinants of adults' weekday activity episode generation behavior. The results underscore the substantial linkages in the activity episode generation of adults based on activity purpose and accompaniment type. The extent of this linkage varies by individual demographics, household demographics, day of the week, and season of the year. The results also highlight the flexibility of the CML approach to specify and estimate behaviorally rich structures to analyze inter-individual interactions in activity episode generation.

*Keywords:* Composite Marginal Likelihood (CML) approach, social interactions, activity-based modeling, multivariate ordered probit model, American Time Use Survey (ATUS).

# 1. INTRODUCTION

## 1.1 Motivation

The emphasis of the activity-based approach to travel modeling is on understanding the activity participation characteristics of individuals within the context of their demographic attributes, activity-travel environment, and social interactions. In the activity-based approach, activity episodes rather than trip episodes take the center stage, with the focus being on activity episode generation and scheduling over a specified time period (Jones *et al.*, 1990, Bhat and Koppelman, 1999, Pendyala and Goulias, 2002, Arentze and Timmermans, 2004, and Pinjari and Bhat, 2010 provide extensive reviews of the activity-based approach). Several operational analytic frameworks for this activity analysis approach have also been formulated, and many metropolitan areas in the U.S. have implemented these frameworks (see Pinjari *et al.*, 2008 for a recent review). These frameworks have focused on a “typical” weekday frame of analysis, and follow a general structure where out-of-home work-related decisions (employed or not, duration of work, location of work, and timing of work) are modeled first followed by the generation and scheduling of out-of-home non-work episodes (in the rest of this paper, we will use the term “non-work episodes” to refer to out-of-home non-work episodes).

The generation and scheduling of non-work episodes entails the determination of the number of non-work episodes by purpose, along with various attributes of each episode and the sequencing of these non-work episodes relative to work and in-home episodes. In the context of episode attributes, one dimension that has been receiving substantial attention recently is the “with whom” dimension (or the social context). This is motivated by the recognition that individuals usually do not make their activity engagement decisions in isolation. For instance, within a household, an individual’s activity participation decisions are likely to be dependent on other members of the household because of the possible sharing of household maintenance responsibilities, joint activity participation in discretionary activities, and pick-up/drop-off of household members with restricted mobility (Gleibe and Koppelman, 2002, Kapur and Bhat, 2007). In a similar vein, outside the confines of the household, an individual’s activity participation might be influenced by non-household members because of car-pooling arrangements, social engagements, and joint recreational pursuits. In fact, Srinivasan and Bhat (2008), in their descriptive study of activity patterns, found that about 30% of individuals undertake one or more out-of-home (OH) activity episodes with household members on

weekdays, and about 50% pursue OH activity episodes with non-household companions on weekdays. These interactions in activity decisions across household and non-household members are important to consider to accurately predict activity-travel patterns. For instance, the spatial and temporal joint participation in dinner at a restaurant of a husband and a wife are necessarily linked. Thus, considering the husband's and wife's activity-travel patterns independently without maintaining the linkage in time and space in their patterns will necessarily result in less accurate activity travel pattern predictions for each one of them. Further, there is a certain level of rigidity in such joint activity participations (since such participations necessitate the synchronization of the schedules of multiple individuals in time and space), because of which the responsiveness to transportation control measures such as pricing schemes may be less than what would be predicted if each individual were considered in isolation (see Vovsha and Bradley, 2006 and Timmermans and Zhang, 2009 for extensive discussions of the importance of considering inter-individual interactions for accurately evaluating land-use and transportation policy actions).

To be sure, several recent studies have focused on explicitly accommodating inter-individual interactions in activity-travel modeling. The reader is referred to a special issue of *Transportation* edited by Bhat and Pendyala (2005), as well as a special issue of *Transportation Research Part B* edited by Timmermans and Zhang (2009), for recent papers on this topic. While these and other earlier studies have contributed in very important ways, they focus on intra-household interactions, and mostly on the interactions between the household heads (see, for example, Wen and Koppelman, 1999, Scott and Kanaroglou, 2002, Meka *et al.*, 2002, Srinivasan and Bhat, 2005, and Kato and Matsumoto, 2009). On the other hand, as discussed earlier in this paper, there is a significant amount of activity episode participations in the wider social network beyond the household (see also Goulias and Kim, 2005, Axhausen, 2005, Arentze and Timmermans, 2008, and Carrasco and Miller, 2009). Many earlier intra-household interaction studies in the literature also confine their attention to the single activity category of maintenance-oriented activities (see Srinivasan and Athuru, 2005 and Wang and Li, 2009). But, as indicated by PBQD (2000), over 75% of non-work episodes on a typical weekday are for discretionary purposes and, as pointed out by Srinivasan and Bhat (2008), a high percentage of these discretionary episodes involve one or more companions. This suggests the important need to consider inter-individual interactions in discretionary activity too (and not just in maintenance-oriented activity). Further, a significant fraction of existing studies on inter-individual

interactions focus on daily *time allocations* or *joint time-use* in activities over a certain time period (an extensive review of these time allocation/time-use studies is provided in Vovsha *et al.*, 2003 and Kato and Matsumoto, 2009). This is also true of the recent studies by Bhat and colleagues (Kapur and Bhat, 2007, Sener and Bhat, 2007) that use the multiple discrete-continuous extreme value (MDCEV) model to examine household and non-household companionship arrangement for each of several types of activities. While providing important insights, these studies of daily time-use do not directly translate to information regarding out-of-home episodes. On the other hand, it is the scheduling and sequencing of out-of-home episodes that get manifested in the form of travel patterns (Doherty and Axhausen, 1999, Scott and Kanaroglou, 2002, Vovsha *et al.*, 2003). Finally, even among those studies that consider inter-individual interactions at an episode level, almost all of them have adopted a framework that first generates activity episodes by activity purpose, and subsequently “assigns” each of these purpose-specific episodes to a certain accompaniment type (for example, alone versus joint), typically using a discrete choice model (see, for example, Wen and Koppelman, 1999, Gliebe and Koppelman, 2002, and Bradley and Vovsha, 2005). Unfortunately, such a sequential framework cannot accommodate general patterns of observed and unobserved variable effects that are specific to each activity purpose-accompaniment type combination (see also Scott and Kanaroglou, 2002).

## **1.2 The Current Paper**

The objective of the current paper, motivated by the discussion above, is to propose and estimate a joint modeling system for adult individuals’ (aged 15 years or over) non-work activity episodes (or simply “episodes” from hereon) by purpose that also explicitly incorporates companionship arrangement information. The six activity purpose categories considered in the paper are: (1) family care (including child care), (2) maintenance shopping (grocery shopping, purchasing gas/food, and banking), (3) non-maintenance shopping (window shopping, cloth shopping, electronics shopping, *etc.*), (4) meals, (5) physically active recreation (sports, exercise, walking, bicycling, *etc.*), and (6) physically inactive recreation (social, relaxing, movies, and attending

religious/cultural/sports events).<sup>1</sup> The companionship arrangement for episodes is considered in five categories: (1) alone, (2) only family (including children, spouse, and unmarried partner), (3) only relatives (parents, siblings, grandchild, *etc.*), (4) only friends (including friends, colleagues, neighbors, co-workers, peers, and other acquaintances), and (5) mixed company (a combination of family, extended family, and friends).<sup>2</sup> The total number of activity purpose-companionship type categories is 30, and the model system developed here jointly considers the number of episodes in each of these 30 categories. The data used in the empirical analysis is drawn from the American Time Use Survey (ATUS), which collects detailed individual-level activity information for one day from a randomly selected adult (15 years or older) in each of a subset of households responding to the Current Population Survey (CPS).

The paper uses a multivariate ordered-response model system for analyzing the number of episodes of each activity purpose-companionship type. In this system, we allow dependence between the number of episodes of different purpose-companionship types due to both observed exogenous variables as well as unobserved factors. The inclusion of dependence generated by unobserved factors allows *complementarity* and *substitution* effects in activity participation decisions (even after controlling for observed effects). For instance, individuals who are “go-getters” and “dynamic” in their lifestyle may have a higher participation propensity in sports-type activities (“physically active recreation”) and also in cultural/social activities (“physically inactive recreation”). This would constitute a complementary relationship between these two activity purpose categories. Similarly, individuals who are “sociable” may be more likely to participate in activity episodes with friends, but not alone. This represents a substitution relationship in the company types of ‘friends’ and “alone”. Besides, the presence of common unobserved factors among combination categories that share the same activity purpose or that share the same companionship type can also generate complementary effects. Thus, an individual who is “sociable” by personality may have a higher propensity to participate in dining out-with

---

<sup>1</sup> There is obviously some subjectivity in the classification adopted here, though the overall consideration was to accommodate differences between the disaggregate activity purposes along such contextual dimensions as location of participation, physical intensity level, duration of participation, amount of structure in activity planning, and company type of participation (see Srinivasan and Bhat, 2005).

<sup>2</sup> While we consider the companionship arrangement for episodes, the reader will note that we still consider the generation of episodes at the individual level. Future efforts should consider the generation of episodes at a higher level, such as a household level or a neighborhood level, so that there is consistency in activity episode generation across individuals. Thus, for example, if a husband has a joint out-of-home (OH) activity episode with his wife, it must also be true that the wife has a joint OH activity episode with her husband.

friends as well as a higher propensity to participate in physically-inactive recreation with friends. Overall, the extent of complementary and substitution relationships may be specific to the combinations of activity purpose category and company type, which is the general case modeled in the current paper.

The econometric challenge in estimating a joint multivariate ordered-response system with a large number of categories is that traditional classical and Bayesian simulation techniques become saddled with convergence problems, and are extremely cumbersome if not impractical to implement. An approach to deal with the estimation complication is the technique of composite marginal likelihood (CML), an emerging inference approach in the statistics field, though there has been little to no coverage of this method in econometrics and other fields (see Varin, 2008 and Bhat *et al.*, 2009). The CML is based on the classical approach, is very simple to estimate, is easy to implement regardless of the number of count outcomes to be modeled jointly, requires no simulation technique whatsoever, and usually provides accurate inferential conclusions. To the authors' knowledge, this is the first study to adopt a CML approach in the field of activity-travel modeling, though Bhat *et al.* (2009) use the CML approach in the context of a spatially dependent discrete choice model formulation. Very simply stated, the CML approach is based on developing the marginal log-likelihood of the joint distribution of a lower dimensional number of categories at one time (such as two categories at one time), while ignoring all other categories. Maximizing this marginal log-likelihood function provides a consistent estimator of the parameters identified by the lower dimensional marginal distribution. Then, by developing and maximizing a surrogate log-likelihood function that is the sum of the log-likelihood of each possible combination of the lower dimensional number of categories, one obtains a consistent estimator of all the relevant parameters characterizing the original high dimensional distribution.

The rest of the paper is organized as follows. Section 2 presents the model structure and highlights the important aspects of the CML approach, Section 3 undertakes a simulation exercise to demonstrate the ability of the CML technique to recover “true” parameters. Section 4 summarizes the data source and sample preparation procedure. Section 5 discusses the estimated results and demonstrates an application of the model, and the final section concludes the paper by summarizing the salient features and findings of the study and identifying potential future research directions.

## 2. THE MODEL STRUCTURE

### 2.1 Background

The multivariate model system used in the paper assumes an underlying set of multivariate continuous latent variables whose horizontal partitioning maps into the observed set of count outcomes (number of episodes across purpose types and companionship types in the current context). Such an ordered-response system allows the use of a general covariance matrix for the underlying latent variables, which translates to a flexible correlation pattern among the observed count outcomes. On the other hand, the traditional approach in the econometric literature to address correlated counts is to start with a Poisson or negative binomial distribution for each univariate count and add a random component to the conditional mean specification. If these random components are allowed to be correlated across equations, the net result is a mixed count model that allows correlation across outcomes. Such a model can be estimated using classical or Bayesian simulation techniques (Egan and Herriges, 2006, Chib and Winkelmann, 2001). An important problem with this approach, however, is that the use of the Poisson or negative binomial distribution as the underlying kernel for mixing restricts “the amount of probability mass that can be accommodated at any one point” (see Herriges *et al.*, 2008). Thus, in cases with a high fraction of ‘0’ values, as in the current empirical context of the number of episodes in each activity purpose-companionship type combination, the count mixing models are not able to provide good predictions. The alternative of adding zero-inflated approaches to accommodate the high number of ‘0’ values, while easy to undertake in a univariate count model, becomes difficult in the multivariate count case.

Of course, the use of an ordered-response system for count outcomes is certainly not new in the transportation literature. In fact, it has a long history of use for modeling such travel count dimensions as household car ownership levels (Kitamura, 1987, 1988, Golob and van Wissen, 1989, Golob, 1990, Bhat and Guo, 2007) and trip generation/stop-making (see Meurs, 1989, Agyemang-Duah *et al.*, 1995, Agyemang-Duah and Hall, 1997, Bhat, 1999, Bricka and Bhat, 2006, and Carrasco and Miller, 2009 to list just a few). While the traditional ordered-response model was initially developed for the case of ordinal responses, and while count outcomes are cardinal, this distinction is really irrelevant for the use of the ordered-response system for count outcomes. This is particularly the case when the count outcome takes few discrete values, as in the current empirical case, but is also not much of an issue when the count outcome takes a large

number of possible values. A perceived problem in the latter case may be that the ordered-response model entails the estimation of  $K-1$  threshold values that horizontally partition the underlying continuous variable to map into the observed count values, where  $K$  is the largest possible count value. But, as has been demonstrated by Meyer (1990), there is little loss of efficiency due to the estimation of a large number of thresholds in the ordered-response model structure. As long as there are even a few observations in each of the  $K$  categories under consideration, it is straightforward to estimate the ordered-response structure.

The ordered-response applications in the transportation literature discussed above all focus on a univariate count outcome. Three earlier multivariate count studies using a multivariate ordered-response structure that are directly relevant to the current paper are Scott and Kanaroglou (2002), Bhat and Srinivasan (2005), and Herriges *et al.* (2008). These are discussed in turn below.

Scott and Kanaroglou use a trivariate normal distribution for the underlying latent continuous variables for three count outcomes, which correspond to the daily number of non-work episodes in couple households made by the male head, the female head, and jointly by both the heads. This leads to a trivariate integral for the probability expression for each household, which can be computed in a straightforward way using trivariate cumulative normal distribution functions. The restriction to three outcomes obviates the need for simulation, but also constrains the authors to consider all non-work episodes together without differentiating between activity types. Besides, the interaction in activity participation is confined to the household heads.

Bhat and Srinivasan appear to be the first to have proposed a modeling system and estimation approach that can conceptually accommodate any number of count outcomes. The authors use a logistic error term in each univariate ordered-response specification, and then also add a normally distributed mixing error term in the latent continuous equation. By allowing the mixing terms to be distributed multivariate normal, they effectively generate a flexible correlation structure across the outcome categories. They use a maximum simulated likelihood approach for evaluating the multi-dimensional integral in the resulting probability expression, using quasi-Monte Carlo simulation methods proposed by Bhat (2001; 2003). In addition, they develop a method to parameterize the likelihood function in terms of the elements of the Cholesky decomposed-matrix of the correlation matrix of the mixing normally distributed elements to ensure the positive definiteness of the matrix, and further parameterize the diagonal

elements of the Cholesky matrix to guarantee unit values along the diagonal. Bhat and Srinivasan apply their model system to analyze the number of episodes of participation of individuals in seven different activity purposes, but they do not focus on accompaniment type. While their simulation approach can be extended in principle to any number of count outcomes, numerical stability, convergence, and precision problems start surfacing as the number of dimensions increase.

Herriges *et al.* (2008) recently have proposed an alternate estimation approach for the multivariate ordered response system based on the posterior mode in an objective Bayesian approach as in Jeliazkov *et al.* (2008).<sup>3</sup> The approach of Herriges *et al.* (2008) is based on assuming prior distributions on the non-threshold parameters, reparameterizing the threshold parameters, imposing a standard conjugate prior on the reparameterized version of the error covariance matrix and a flat prior on the transformed threshold, obtaining an augmented posterior density using Baye’s Theorem for the reparameterized model, and fitting the model using a Markov Chain Monte Carlo (MCMC) method. Unfortunately, the method remains very cumbersome, requires extensive simulation, and is time-consuming. Further, convergence assessment becomes very difficult as the number of dimensions increase. In this regard, both the MSL and the Bayesian approach are “brute force” simulation techniques that are not straightforward to implement and can create convergence assessment problems. Herriges *et al.* apply their Bayesian estimation approach to examine the annual number of trips made by Iowa households to each of 29 lakes in the state.

In the current paper, we consider and use a third inference approach – the Composite Marginal Likelihood (CML) approach – that is very simple and requires no simulation machinery whatsoever. It entails the development of a surrogate likelihood function that involves easy-to-compute, low-dimensional, marginal likelihoods. Thus, in the multivariate ordered response system with  $K$  count outcomes (or categories), the individual likelihood contribution in the classic maximum likelihood approach involves the  $K$ -variate rectangular integral for the probability of the observed number of episodes of that individual across the  $K$  categories. However, the individual contribution in the composite marginal likelihood approach would

---

<sup>3</sup> It is interesting that Herriges *et al.* appear to be “discovering” the use of an ordered response structure for count outcomes, while such a structure has in fact been used extensively in the past for count outcomes in the transportation literature. Further, Herriges *et al.* do not seem to have been aware of the work of Bhat and Srinivasan (2005), which develops a frequentist inference approach for correlated counts.

involve the product of all lower dimensional combinations of probabilities of the observed number of episodes in each subset of categories (see de Leon, 2005, Varin and Czado, 2008, Bhat *et al.*, 2009). The CML approach can be applied using simple optimization software for likelihood estimation, is based on a classical frequentist approach, and is easy to apply and fast even for a large number of ordinal response variable categories and/or a large sample size. Under usual regularity conditions, the CML estimator is consistent, unbiased, and asymptotically normally distributed. The CML estimator (theoretically speaking) loses some efficiency relative to traditional maximum likelihood estimation, though this efficiency loss has been showed to be negligible in practice (see Varin and Vidoni, 2008).

## 2.2 Mathematical Formulation

Let  $q$  be an index for individuals ( $q = 1, 2, \dots, Q$ ), and let  $i$  be the index for episode category ( $i = 1, 2, \dots, I$ , where  $I$  denotes the total number of episode categories for each individual; in the current study,  $I = 30$ ). Let the number of episode count values for category  $i$  be  $K_i + 1$  (*i.e.*, the discrete levels, indexed by  $k$ , belong in  $\{0, 1, 2, \dots, K_i\}$  for category  $i$ ). In the usual ordered response framework notation, we write the latent propensity ( $y_{qi}^*$ ) for each episode category as a function of relevant covariates and relate this latent propensity to the observed count outcome ( $y_{qi}$ ) through threshold bounds (see McKelvey and Zavoina, 1975):

$$y_{qi}^* = \beta_i' x_{qi} + \varepsilon_{qi}, y_{qi} = k \text{ if } \theta_i^k < y_{qi}^* < \theta_i^{k+1}, \quad (1)$$

where  $x_{qi}$  is a  $(L \times 1)$  vector of exogenous variables (not including a constant),  $\beta_i$  is a corresponding  $(L \times 1)$  vector of coefficients to be estimated,  $\varepsilon_{qi}$  is a standard normal error term, and  $\theta_i^k$  is the lower bound threshold for count level  $k$  of episode category  $i$  ( $\theta_i^0 < \theta_i^1 < \theta_i^2 \dots < \theta_i^{K_i+1}$ ;  $\theta_i^0 = -\infty$ ,  $\theta_i^{K_i+1} = +\infty$  for each category  $i$ ). The  $\varepsilon_{qi}$  terms are assumed independent and identical across individuals (for each and all  $i$ ). For identification reasons, the variance of each  $\varepsilon_{qi}$  term is normalized to 1. However, we allow correlation in the  $\varepsilon_{qi}$  terms across episode categories  $i$  for each individual  $q$ . Specifically, define  $\varepsilon_q = (\varepsilon_{q1}, \varepsilon_{q2}, \varepsilon_{q3}, \dots, \varepsilon_{qI})'$ . Then,  $\varepsilon_q$  is multivariate normal distributed with a mean vector of zeros and a correlation matrix as follows:

$$\varepsilon_q \sim N \left[ \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \cdots & \rho_{1I} \\ \rho_{21} & 1 & \rho_{23} & \cdots & \rho_{2I} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{I1} & \rho_{I2} & \rho_{I3} & \cdots & 1 \end{pmatrix} \right], \text{ or} \quad (2)$$

$$\varepsilon_q \sim N[\mathbf{0}, \Sigma]$$

The off-diagonal terms of  $\Sigma$  capture the error covariance across the underlying latent continuous variables of the different episode categories; that is, they capture the effect of common unobserved factors influencing the propensity of choice of count level for each episode category. Thus, if  $\rho_{12}$  is positive, it implies that individuals with a higher than average propensity in their peer group to participate in the first episode category are also likely to have a higher than average propensity to participate in the second episode category. Of course, if all the correlation parameters (*i.e.*, off-diagonal elements of  $\Sigma$ ), which we will stack into a vertical vector  $\Omega$ , are identically zero, the model system in Equation (1) collapses to independent ordered response probit models for each episode category.

### 2.3 The Pairwise Marginal Likelihood Inference Approach

The parameter vector of the multivariate probit model is  $\delta = (\beta'_1, \beta'_2, \dots, \beta'_I; \theta'_1, \theta'_2, \dots, \theta'_I; \Omega)'$ , where  $\theta_i = (\theta_i^1, \theta_i^2, \dots, \theta_i^{K_i})'$  for  $i = 1, 2, \dots, I$ . Let the actual observed count level for individual  $q$  and episode category  $i$  be  $m_{qi}$ . Then, the likelihood function for individual  $q$  may be written as follows:

$$L_q(\delta) = \Pr(y_{q1} = m_{q1}, y_{q2} = m_{q2}, \dots, y_{qI} = m_{qI})$$

$$L_q(\delta) = \int_{v_1 = \theta_1^{m_{q1}} - \beta_1' x_{q1}}^{\theta_1^{m_{q1}+1} - \beta_1' x_{q1}} \int_{v_2 = \theta_2^{m_{q2}} - \beta_2' x_{q2}}^{\theta_2^{m_{q2}+1} - \beta_2' x_{q2}} \cdots \int_{v_I = \theta_I^{m_{qI}} - \beta_I' x_{qI}}^{\theta_I^{m_{qI}+1} - \beta_I' x_{qI}} \phi_I(v_1, v_2, \dots, v_I | \Omega) dv_1 dv_2 \dots dv_I \quad (3)$$

The likelihood function above requires the computation of an  $I$ -dimensional rectangular integral. While there are maximum simulated likelihood (MSL) approaches that can evaluate such multidimensional normal integrals using the Geweke-Hajivassiliou-Keane simulator (Hajivassiliou *et al.*, 1996), they can become problematic even for moderate  $I$  in terms of computational effort. Further, as noted in the previous section, such simulation methods do get

imprecise as the number of dimensions increase, leading to convergence problems during estimation.

In this paper, we employ a pairwise marginal likelihood estimation approach, which corresponds to a composite marginal approach based on bivariate margins (see Varin and Czado, 2008; Apanasovich *et al.*, 2008; Varin and Vidoni, 2008; and Bhat *et al.*, 2009 for the use of the pairwise likelihood approach in the past). The pairwise marginal likelihood function for individual  $q$  may be written as follows:

$$\begin{aligned}
L_{CML,q}(\delta) &= \prod_{i=1}^{I-1} \prod_{g=i+1}^I \Pr(y_{qi} = m_{qi}, y_{qg} = m_{qg}) \\
&= \prod_{i=1}^{I-1} \prod_{g=i+1}^I \left[ \begin{aligned} &\Phi_2(\theta_i^{m_{qi}+1} - \beta'_i x_{qi}, \theta_g^{m_{qg}+1} - \beta'_g x_{qg}, \rho_{ig}) - \Phi_2(\theta_i^{m_{qi}+1} - \beta'_i x_{qi}, \theta_g^{m_{qg}} - \beta'_g x_{qg}, \rho_{ig}) \\ &- \Phi_2(\theta_i^{m_{qi}} - \beta'_i x_{qi}, \theta_g^{m_{qg}+1} - \beta'_g x_{qg}, \rho_{ig}) + \Phi_2(\theta_i^{m_{qi}} - \beta'_i x_{qi}, \theta_g^{m_{qg}} - \beta'_g x_{qg}, \rho_{ig}) \end{aligned} \right], \quad (4)
\end{aligned}$$

$$\text{and } L_{CML}(\delta) = \prod_q L_{CML,q}(\delta)$$

The pairwise likelihood function above is easily maximized, and the effort involved is no more difficult than in a usual bivariate ordered probit model. Besides, note that a full likelihood estimator is not robust to misspecification of the assumed parametric high dimensional multivariate distribution, in addition to the computational and convergence problems noted earlier. On the other hand, the pairwise likelihood approach is typically more robust (as pointed out by Varin and Vidoni, 2008 and Varin, 2008), as well as is very simple computationally with literally no convergence-related issues. It can also be very easily coded in software packages that allow the computation of a bivariate normal cumulative distribution function and have an optimization procedure for maximizing a function with respect to embedded parameters.

The pairwise estimator  $\hat{\delta}_{CML}$  obtained by maximizing the logarithm of the function in Equation (4) with respect to the vector  $\delta$  is consistent and asymptotically normal distributed with asymptotic mean  $\delta$  and covariance matrix given by Godambe's (1960) sandwich information matrix (see Zhao and Joe, 2005):

$G(\delta) = [H(\delta)]^{-1} J(\delta) [H(\delta)]^{-1}$ , where

$$H(\delta) = E \left[ - \frac{\partial^2 \log L_{CML}(\delta)}{\partial \delta \partial \delta'} \right] \text{ and} \quad (5)$$

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

$H(\delta)$  and  $J(\delta)$  can be estimated in a straightforward manner at the CML estimate ( $\hat{\delta}_{CML}$ ):

$$\begin{aligned} \hat{H}(\hat{\delta}) &= - \left[ \sum_{q=1}^Q \frac{\partial^2 \log L_{CML,q}(\delta)}{\partial \delta \partial \delta'} \right] \\ &= - \left[ \sum_{q=1}^Q \sum_{i=1}^{I-1} \sum_{g=i+1}^I \frac{\partial \log \Pr(y_{qi} = m_{qi}, y_{qg} = m_{qg})}{\partial \delta} \frac{\partial \log \Pr(y_{qi} = m_{qi}, y_{qg} = m_{qg})}{\partial \delta'} \right]_{\hat{\delta}}, \text{ and} \quad (6) \end{aligned}$$

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

## 2.4 Positive-Definiteness of the Implied Multivariate Correlation Matrix

A point that we have not discussed thus far is how to ensure the positive-definiteness of the symmetric correlation matrix  $\Sigma$  (that is, all the eigenvalues of the matrix should be positive, or, equivalently, the determinant of the entire matrix and every principal submatrix of  $\Sigma$  should be positive). In particular, because the pairwise marginal approach does not estimate the entire correlation matrix as one single entity, it is not possible to impose positive-definiteness upfront with unconstrained optimization routines using, for example, Bhat and Srinivasan's (2005) technique.<sup>4</sup> But one of two alternative techniques may be used to ensure positive definiteness with the pairwise marginal approach. The first is to undertake the estimation with a constrained optimization routine by requiring that the implied multivariate correlation matrix for any set of pairwise correlation estimates be positive definite. However, such a constrained routine can be extremely cumbersome. The second is to use an unconstrained optimization routine, but check for positive-definiteness of the implied multivariate correlation matrix. The easiest method

<sup>4</sup> As indicated earlier, Bhat and Srinivasan (2005) parameterize the likelihood function in terms of the elements of the Cholesky decomposed-matrix of the entire correlation matrix and estimate these Cholesky-decomposed elements.

within this second technique is to allow the estimation to proceed without checking for positive-definiteness at intermediate iterations, but check that the implied multivariate correlation matrix at the final converged pairwise marginal likelihood estimates is positive-definite. This will typically work for the case of a multivariate ordered-response model if one specifies exclusion restrictions (*i.e.*, zero correlations between some error terms) or correlation patterns that involve a lower dimension of effective parameters. In fact, employing such exclusion restrictions or using a factor-analytic structure is sound econometric practice anyway (regardless of estimation approach used). Also, the number of correlation parameters in the full multivariate matrix explodes quickly as the dimensionality of the matrix increases, and estimating all these parameters becomes almost impossible (with any estimation technique) with the usual sample sizes available in practice. There is also one other reason for expecting that the pairwise estimation approach will automatically provide a positive-definite matrix for  $\Sigma$ . That is that  $\Sigma$  is pre-specified to be a correlation matrix (*i.e.*, its diagonals are normalized to the value of 1 for identification), and we expect the magnitude of many off-diagonal terms to be small relative to the unity diagonal (in part, due to exclusion restrictions, but also because a reasonably good specification of exogenous variables will ensure that the correlation in unobserved factors does not get too high in magnitude). In general, if the diagonal elements of a matrix are sufficiently high relative to the off-diagonal elements, the matrix will be positive-definite (as we allude to later, this is actually a general technique used to bring non-positive definite matrices to the positive-definite realm). However, if the above simple method of allowing the pairwise marginal estimation approach to proceed without checking for positive definiteness at intermediate iterations does not work, then one can check the implied multivariate correlation matrix for positive definiteness at each and every iteration. If the matrix is not positive-definite during a direction search at a given iteration, one can construct a “nearest” valid correlation matrix (for example, by replacing the negative eigenvalue components in the matrix with a small positive value, or by adding a sufficiently high positive value to the diagonals of a matrix and normalizing to obtain a correlation matrix; see Rebonato and Jaeckel, 1999, Higham, 2002, and Schoettle and Werner, 2004 for detailed discussions of these and other adjusting schemes; a review of these techniques is beyond the scope of this paper). The values of this “nearest” valid correlation matrix can be translated to the pairwise correlation estimates, and the analyst can

allow the iterations to proceed and hope that the final implied convergent correlation matrix is positive-definite.

In the empirical analysis of the current paper, we imposed a few exclusion restrictions, used an unconstrained optimization routine, and ensured that the implied multivariate correlation matrix at convergence is positive-definite.

### 3. SIMULATION STUDY

To evaluate the performance of the CML estimation technique, we undertake a simulation exercise for a trivariate case and a five-variate case. For each case, a sample of 1000 observations is generated using prespecified values for the  $\delta$  vector.

#### 3.1 Trivariate Case

For this case, we use three independent variables in the first episode category, four variables in the second category, and three variables again in the third category. The values for each of the independent variables are drawn from a standard univariate normal distribution. A fixed coefficient vector  $\beta_i$  ( $i=1,2,3$ ) is assumed on the variables, and the linear combination  $\beta_i'x_{qi}$  ( $q=1,2,\dots,Q; i=1,2,3$ ) is computed for each individual  $q$  and episode category  $i$ . Next, we generate  $Q$  trivariate realizations of the error term vector  $(\varepsilon_{q1}, \varepsilon_{q2}, \varepsilon_{q3})$  with a predefined positive-definite error correlation structure. To examine the potential impact of different correlation structures on the performance of the CML approach, we consider one correlation structure with low correlations and another with high correlations as follows:

$$\Sigma_1 = \begin{pmatrix} 1 & 0.30 & 0.20 \\ 0.30 & 1 & 0.25 \\ 0.20 & 0.25 & 1 \end{pmatrix}, \text{ and } \Sigma_2 = \begin{pmatrix} 1 & 0.90 & 0.80 \\ 0.90 & 1 & 0.75 \\ 0.80 & 0.75 & 1 \end{pmatrix} \quad (7)$$

The error term realization for each observation and each episode category,  $\varepsilon_{qi}$ , is then added to the systematic component ( $\beta_i'x_{qi}$ ) as in Equation (1) and then translated to “observed” values of  $y_{qi}$  (0, 1, 2,...) based on pre-specified threshold values. We assume four outcome levels for the first category, three for the second category, and five for the third episode category.

Correspondingly, we pre-specify a vector of three threshold values  $[\theta_1=(\theta_1^1,\theta_1^2,\theta_1^3)]$  for the first category, two for the second category  $[\theta_2=(\theta_2^1,\theta_2^2)]$ , and four for the third category  $[\theta_3=(\theta_3^1,\theta_3^2,\theta_3^3,\theta_3^4)]$ .

The above data generation process is undertaken 50 times with different realizations of the random error term to generate 50 different data sets. The CML estimation procedure is then applied to each dataset to estimate data-specific values.

### 3.2 The Five-variate Case

For the five-variate case, we adopt a procedure very similar to the above with two additional episode categories (for a total of five episode categories). We use 4 independent variables for the fourth category and 3 independent variables for the fifth category (maintaining the same number of independent variables as earlier for the first three episode categories). The positive-definite correlation structures considered are as follows:

$$\Sigma_1 = \begin{pmatrix} 1 & 0.30 & 0.20 & 0.22 & 0.15 \\ 0.30 & 1 & 0.25 & 0.30 & 0.12 \\ 0.20 & 0.25 & 1 & 0.27 & 0.20 \\ 0.22 & 0.30 & 0.27 & 1 & 0.25 \\ 0.15 & 0.12 & 0.20 & 0.25 & 1 \end{pmatrix}, \text{ and } \Sigma_2 = \begin{pmatrix} 1 & 0.90 & 0.80 & 0.82 & 0.75 \\ 0.90 & 1 & 0.85 & 0.90 & 0.72 \\ 0.80 & 0.85 & 1 & 0.87 & 0.80 \\ 0.82 & 0.90 & 0.87 & 1 & 0.85 \\ 0.75 & 0.72 & 0.80 & 0.85 & 1 \end{pmatrix} \quad (8)$$

We assume three outcome levels for the fourth category and four outcome levels for the fifth category.

### 3.3 Evaluation

The performance evaluation of the CML approach is undertaken based on the ability to recover the parameter vector  $\delta$  in the trivariate case and its corresponding counterpart in the five-variate case. The proximity of estimated and true values for each parameter is based on computing the following three metrics: (a) absolute bias (or the absolute difference between the mean of the relevant values across the 50 runs and the true values), (b) the absolute percentage bias (*i.e.*, the absolute bias as a percentage of the true value), and (c) the total error computed as the root

mean-squared error (RMSE) between the estimated and true values across all 50 runs

$$(RMSE = 100 \times \sqrt{\left( \sum_{r=1}^R (\hat{\delta}_r - \delta)^2 \right) / R}, \text{ where } R=50).$$

Table 1 presents the results for the trivariate case and Table 2 presents the results for the five-variate case. The results indicate that the CML approach recovers the parameters extremely well in terms of bias and relative bias for both low and high correlation cases and for the trivariate and five-variate cases. In fact, for both the low and high correlation case, the overall mean absolute percentage bias across all parameters (see last row) is less than 1% for the trivariate case and less than 1.5% for the five-variate case. These are incredibly good measures for the ability to recover parameter estimates. Further, the ability to recover parameters does not seem to be affected much by whether there is low correlation or high correlation. The RMSE values are also generally very small, though it is interesting to note that the RMSE values for the thresholds are consistently higher than for other parameters in both the low and high correlation case. The mean of the RMSE values across parameters shows little change between the trivariate case and the five-variate case. Overall, the simulation results clearly demonstrate the ability of the Composite Marginal Likelihood (CML) to recover the parameters in a multivariate ordered-response choice model context, independent of the correlation structure or the number of episode categories (*i.e.*, count outcome categories). Combined with its conceptual and implementation simplicity, the CML approach is an effective method in accounting for correlation in high-dimensional multivariate ordered response contexts.

## 4. DATA

### 4.1 Data Source

The data used for the empirical analysis in the paper is drawn from the 2007 American Time Use Survey (ATUS). The ATUS is a national level survey conducted and processed by the U.S. Census Bureau for the Bureau of Labor Statistics (ATUS, 2008). The household sample for the ATUS is drawn from the set of households that completed the Current Population Survey (CPS). Next, from each sampled CPS household, the ATUS randomly selects one individual of age 15 or over, and collects information on all episodes the individual participates in over the course of a single day. The episode-level information collected in the ATUS includes activity episode purpose, start and end time, location of participation (for example, grocery store, library, *etc.*),

and ‘with whom’ participated in. In addition, data on individual and household socio-demographics, individual labor force participation and employment-related characteristics, and regional location and characteristics of the survey day are also collected.

#### **4.2 Sample Formation and Description**

The 2007 ATUS micro data were processed in several steps to obtain the sample for the current analysis. First, only individuals who were surveyed on a weekday that was not a holiday were selected, because the focus of the current paper is to study individuals’ activity participation patterns on a typical weekday. Second, all work, work-related, education, education-related, travel, sleep, and in-home activity episodes (such as phone call, grooming, *etc.*) were removed from the list of activity episodes undertaken by the respondents on the survey day. Third, all out-of-home activity episodes, originally documented in over four hundred fine activity purpose types, were aggregated into six broad activity purpose type categories: (1) personal/family care (including personal care, caring for children in the household, pick-up/drop-off of children/adults, and caring for extended family members; for the sake of brevity, we will refer to personal/family care activities simply as “family care” activities from hereon), (2) maintenance shopping (such as grocery shopping, purchasing gas/food, and banking), (3) non-maintenance shopping, (4) meals, (5) physically active recreation (including sports, exercise, recreational and volunteer activities), and (6) physically inactive recreation (including social, relaxing, movies, and attending religious/sporting/recreational events). Subsequently, the companion types for each episode were classified into five mutually exclusive and collectively exhaustive categories: (1) alone, (2) only family (includes children, spouse or unmarried partner), (3) only relatives (parents, sibling, grandchild, *etc.*), (4) only friends (friends, co-workers, neighbors, *etc.*), and (5) mixed company (a combination of family, relatives, and friends). The activity type and companion type classification resulted in thirty episode categories. Fourth, the number of episodes undertaken during the survey day by an individual in each of the episode categories is obtained by aggregating all episodes of that category for the person. Fifth, data on household and individual socio-demographics, residential location, and zonal characteristics were appended to the person-level file. Finally, several screening and consistency checks were performed and records with missing or inconsistent data were eliminated.

The final sample for analysis includes out-of-home non-mandatory episode participation information for 4143 individuals (workers and non-workers, aged 15 years or older) on a typical weekday. Table 3 presents the percentage distribution of individuals' participation in episodes by activity type and companionship type. For example, the first entry in Table 3 indicates that 91.3% of individuals do not undertake family care activities alone. Across all the categories, we find that meals with friends is the most frequently undertaken episode category on weekdays, with over 27% of individuals in the sample participating in one or more episodes of this category. Other categories with relatively frequent participation (across individuals) include maintenance shopping alone, family care with family, meals alone, and physically inactive recreation with friends. The last of these is also the activity purpose that individuals are most likely (relative to other activity purposes) to undertake with relatives (8.9%) or with mixed company (7.2%).

## **5. EMPIRICAL ANALYSIS**

### **5.1 Variable Specification**

Several types of variables were considered in the model specification. These included (1) individual socio-demographics (gender, age, race, education level, employment status, student status, and indication of any disability), (2) household socio-demographics (household structure, presence of children, family income, and employment status of spouse/partner)<sup>5</sup>, and (3) day of the week and seasonal effect variables.

In addition to the three groups of variable discussed above, we also considered several interaction effects among the variables. The final specification was based on a systematic process of removing statistically insignificant variables and combining variables when their effects were not significantly different. The specification process was also guided by prior research and intuitiveness/parsimony considerations. We should also note here that, for the continuous variables in the data (such as age and income limits), we tested alternative functional forms that included a linear form, a spline (or piece-wise linear) form, and dummy variables for different ranges.

---

<sup>5</sup> The ATUS survey does not collect information on household vehicle ownership. As a result, this variable is not available for use in the empirical analysis.

## 5.2 Model Estimation Results

Table 4 presents the model estimation results. The columns in the table correspond to the explanatory variables, while the rows correspond to the episode categories. An empty cell indicates that the corresponding column exogenous variable does not have a statistically significant effect on the corresponding row episode category participation propensity. The t-statistic for each coefficient is provided beneath the coefficient in parentheses. The base category is listed in the heading of the column corresponding to that variable. The coefficients in the table indicate the effects of variables on the latent propensity of participation in each episode category (that is, they represent elements of the  $\beta_i$  vector in Equation (1)). Since all the variables in the model are dummy variables, the relative magnitudes of the coefficients also provide an estimate of the importance of the variables in influencing participation propensities and participation probabilities. The marginal impact of variables on the participation probabilities for each combination of number of episodes for the different episode categories varies across individuals because of the non-linear structure of the ordered probit formulation. Aggregate level marginal effects may be computed for each dummy variable by changing the value of the variable to one for the subsample of observations for which the variable takes a value of zero and to zero for the subsample of observations for which the variable takes the value of one. We can then sum the shifts in expected aggregate shares in the two subsamples after reversing the sign of the shifts in the second subsample and compute an effective marginal change in expected aggregate shares in the entire sample due to a change in the dummy variable from 0 to 1. We are not showing these marginal effects here because there are as many as 80 trillion aggregate marginal effects (one for each combination of episode levels across all the 30 episode categories) for each variable. But in Section 5.3, we demonstrate the application of the model due to changes in two variables. In the following sections, we discuss the effect of variables on the latent participation propensities by variable category.

### 5.2.1 Effect of Individual Socio-Demographic Variables

The results indicate the presence of distinct gender effects in activity type participation and accompaniment. Specifically, men are less likely than women to participate, across all companion types, in family care activities (except with “only friends”), maintenance activities (except “alone” and with “only friends”) and non-maintenance shopping activities (except with

“only friends”). These results reinforce the gender stereotype of women being more responsible for, and/or more vested and interested in, family care and shopping activities, a recurring finding in the literature (for example, see Yamamoto and Kitamura, 1999 and Frusti *et al.*, 2003). However, men have a higher propensity than women to (a) participate alone in discretionary activities (*i.e.*, meals out, physically active recreation, and physically inactive recreation), and (b) participate with “only friends” in meals out and physically inactive recreation. This is again consistent with the results found by Srinivasan and Bhat (2006) and Carrasco and Miller (2009), and suggests that men are more likely to undertake active and inactive leisure activities either alone or with friends on a weekday. Finally, men pursue physically inactive recreation with “mixed company” less than do women, potentially a reflection of the combination of family-centric responsibilities and social network level interactions of women relative to men (see Kapur and Bhat, 2007 for a similar result).

The effect of individual age on activity purpose and accompaniment type is accommodated in a non-linear fashion by introducing age in three categories: age less than 40 years, age 40 years or above but less than 60 years, and age 60 years or above (the base age category). The results suggest that, in general, individuals younger than 60 years are more disposed toward pursuing activities with “only family”, and are less likely to participate in physically active recreation with “only friends”. Further, individuals below the age of 40 years are the least likely (relative to other age groups) to participate in activity episodes alone and most likely to participate in episodes with mixed company. Overall, these patterns suggest a combination of the family orientation and larger social networks of younger individuals, perhaps due to household life cycle characteristics. For instance, compared to older individuals, younger individuals are likely to have more family responsibilities, have more social interactions with friends and co-workers, and also have a larger pool of individuals to interact as part of their extended family (parents, siblings, grandparents, *etc.*). Finally, individuals who are older than 60 years are most likely to participate in family care activities with “only relatives”, as evidenced by the negative coefficients corresponding to the age  $< 40$  and  $40 \leq \text{age} < 60$  columns for the “Family care-Only relatives” row of Table 4. This result may be attributable to such activities as care received by senior parents from their children, or child care provided by grandparents to grandchildren.

The race-related coefficients reveal that Caucasians are more likely than non-Caucasians to (1) participate in non-maintenance shopping and physically active recreation with “only family”, and (2) undertake meal episodes with only friends or with friends and family (we did not find statistically significant race differences in the group of non-Caucasians, and hence represent race differences by a simple binary representation between Caucasians and non-Caucasians). The above results are consistent with earlier studies that suggest that Caucasians have higher levels of participation in meals/recreational pursuits (see Bhat and Gossen, 2004 and Mallett and McGuckin, 2000), though our current study also introduces the “with whom” element that earlier studies do not. In this regard, our results also indicate that Caucasians tend to participate less than non-Caucasians in physically inactive recreation “alone”.

Education level also has an impact on the type of episodes pursued and accompaniment type. Specifically, individuals with an education level beyond high school have a higher propensity (than individuals with only a high school degree which is the base category) to participate alone in shopping activities (maintenance and non-maintenance) and physically active recreation. These results may be indicative of the tighter time constraints among individuals with high education, because of which it is easier to schedule shopping and physically active recreational activities (such as going to the gym) alone. Further, the results suggest that individuals with a bachelor’s degree or higher are more likely to pursue physically active recreation with relatives, and with friends. Overall, the results suggest an increased awareness among highly educated adults of the benefits of investing in health and fitness-enhancing pursuits, highlighting the importance of a good education for a healthy society.

Employment status, in the current study, is characterized as employed full-time, employed part-time and unemployed. The several negative coefficients in the “family care” and “maintenance shopping” panels of the table corresponding to the full-time employed variable reflect the lower propensity of full-time employees to pursue these activities (relative to other individuals). The same is true for non-maintenance shopping, though this is confined to the “alone” accompaniment type. Overall, full-time employed individuals have tight time constraints, which may explain their reduced participation in family care and shopping pursuits (see Goulias and Kim, 2001 for a similar result). However, full-time employed individuals have a high propensity to have meals out and physically inactive recreation episodes alone or with friends. The result regarding meals out alone or with friends is perhaps a manifestation of lunch

activity participation alone or with co-workers. Finally, full-time employees are less likely to participate in physically inactive recreation with “only family”, “only relatives”, and “mixed company”, potentially another reflection of tight time constraints (see also Yamamoto *et al.*, 2004). The results for part-time workers provide similar results as for full time workers, except for participation in maintenance shopping and physically inactive recreation.

The next variable in the table corresponds to student status. In this analysis, we defined an individual who is enrolled in high school, college, or university as a student. As expected, students have a high propensity to participate in discretionary activities (meals, physically active recreation, and physically inactive recreation) with friends, potentially a reflection of the combination of social opportunities to interact with friends as well as the social pressures to “fit in” within their peer group.

As one would expect, physical disability significantly affects activity episode participation. Individuals with a physical disability are likely to need assistance from their relatives or immediate family for activity participation, as indicated by the positive coefficients in the “only relatives” or “only family” rows of Table 4.

### ***5.2.2 Effects of Household Socio-Demographic Variables***

Household structure effects were considered by including several types of households, including nuclear family households (two adults of opposite/same sex with one or more children), couple families (two adults of opposite/same sex), single individual households, and “other” households (roommate households, returning young adult households, other related individual households, and all other types of households). The results show that adults in nuclear and couple family households are much more likely than adults in other households to pursue non-family care activities with their immediate family (as reflected in the positive coefficients for nuclear and couple families in the “only family” row for all non-family care activity purposes). Further, nuclear households are less likely than other households to participate in non-family care activities with friends or relatives. These results indicate the high levels of intra-household interactions within nuclear family households and, to a somewhat lesser degree, in couple family households. On the other hand, the results for “single individual” households shows that there is a relatively higher propensity of inter-household interactions with friends in the meal and physically inactive recreation activities of individuals who live alone (these individuals also

participate more in meals and recreation alone). Overall, the results reinforce the need to explicitly consider intra-household and inter-household interactions in activity-travel pattern modeling, as discussed in the first section of this paper. Clearly, the nature of the interactions varies by household structure, which also needs to be considered in the modeling. Besides, earlier studies, such as Bhat and Srinivasan (2005), indicate that nuclear and couple family households have a higher participation propensity in shopping and physically active and inactive recreation activities as a whole, but our current study reveals that this is the case only for episodes with the immediate family. In fact, as just indicated above, nuclear family households have a lower propensity than other households to participate in shopping and discretionary (meals/recreation) activities with friends and relatives. This underscores the need to consider accompaniment type at the level of generation of episodes (as done in this paper), and not further downstream in the modeling process where episodes are first generated purely by activity purpose and then assigned to one of many accompaniment types.

The effect of age of children is introduced in the model in three categories: presence of children 4 years old or younger (the base category), presence of children aged between 5 to 10 years, and presence of children aged between 11 to 15 years. As expected, adults in households with older children (aged 5 years or more) are more likely than adults in households with young children (less than 5 years of age) to have family care episodes with “only family”, a clear reflection of the chauffeuring of children to/from school and other non-school activities as children grow older (sometimes labeled in the popular press as the “soccer mom” and “tennis dad” responsibilities). Adults in households with children in the 5-10 age group partake more in maintenance shopping episodes with “only family”, which may be attributed to one or both parents pursuing maintenance shopping with the child “in tow”. This effect is not statistically significant for the oldest child group since these children have acquired a certain level of independence and do not need child care at all times. Besides, there is evidence from the social psychology literature that pre-teenagers and teenagers would rather not be seen with parents, since this is considered “uncool” (Thornton *et al.*, 1995, Williams, 2003).<sup>6</sup> Of course, the independence levels of children in the pre-teens and teens also enables the participation of parents in meals and physically inactive recreation activities with friends, as reflected by the

---

<sup>6</sup> This finding is also supported by message boards and parent blogs posted on a number of websites such as [life.familyeducation.com](http://life.familyeducation.com), [www.ParentsConnect.com](http://www.ParentsConnect.com), [www.theparentreport.com](http://www.theparentreport.com), [family.go.com](http://family.go.com), all dedicated to address and deal with pre-teen and teenage issues.

positive coefficients in the “meals-only friends” and “physically inactive recreation-only friends” rows of Table 4.

The effect of income is captured using dummy variables for different income categories, which enables the accommodation of nonlinear impacts on the propensity to participate in episodes (the dummy variable representation was found to be superior to a continuous linear income effect in our specifications). The results in Table 4 show that household income influences participation in meals, physically active recreation, and physically inactive recreation. As expected, individuals in high-income households have a higher propensity to participate in these activity episodes because of their higher expenditure potential for discretionary pursuits. However, this is only true for episodes participated with “only family”. In fact, individuals in highest income group are less likely than individuals in other income groups to pursue physically inactive recreation alone (perhaps attributable to time constraints due to the level and intensity of work activity). Also, middle income individuals have a lower propensity to participate in non-maintenance shopping with “only friends”, a result that is not immediately intuitive and needs exploration in future studies. But, overall, such differential episode generation rates by accompaniment type can only be accommodated if accompaniment type is considered at the generation level, rather than later on in the modeling hierarchy.

Finally, in the category of household demographics, individuals in a household with a working spouse contribute more (less) than individuals without a working spouse to family care episodes alone or with immediate family (with friends).

### ***5.2.3 Day of Week and Season Variables***

The variables considered in this category include day of week variables and season variables (categorized as summer, fall, spring and winter). Clearly, there is a higher propensity of participation on Fridays in almost all non-physically active combinations of activity purpose and accompaniment. Further, it is unlikely that individuals pursue meals out activities alone on Fridays. For other activity purposes except maintenance shopping, there is no difference between Fridays and other days for solo-participation in episodes. Overall, individuals pursue more non-physical activity episodes on Fridays relative to other days of the week, and generally participate with family and friends.

The seasonal effects reflect a higher propensity to participate in physically active recreation with family and friends over the summer compared to other seasons. This may be attributable to better weather conditions for outdoor activities, more daylight time, and more schedule opportunities to pursue activities with family and friends.

#### ***5.2.4 Threshold Parameters***

The threshold parameters are not shown in the table, but are available on request from the authors. These parameters represent the cut-off points that map the latent propensity of individuals to participate in each activity purpose-accompaniment type category to the reported number of episodes for each category. As such, they do not have any substantive behavioral interpretations.

#### ***5.2.5 Correlation Estimates***

As indicated earlier in Section 2.4, it is not practical to estimate the parameters of the full correlation matrix (in the current case, the number of parameters in the full correlation matrix is 435). In our analysis, we specified several initial exclusion restrictions based on (1) intuitive considerations (for example, there is no reason why unobserved factors influencing participation in maintenance shopping with family should be correlated with unobserved factors influencing participation in physically active recreation with friends), and (2) the estimation of bivariate models for pairs of episode categories to determine if the corresponding correlations were statistically significant. These initial exclusion restrictions were used to estimate several alternative model specifications using the pairwise procedure proposed, and the final correlation matrix specification was obtained based on statistical fit and parsimony considerations.

The estimated covariances and their t-statistics (in parentheses) are shown in Table 5. Only the upper diagonal terms in the variance-covariance matrix are shown since the matrix is symmetric. As mentioned before, the variance of the error terms are set to one to normalize the scale (see Section 2.2). The covariance (correlation) matrix indicates several statistically significant correlations among the stop-making propensities of different activity type-accompaniment categories, highlighting the importance of accounting for common unobserved factors in modeling episode participation frequency. For the sake of conciseness, we focus only on the salient aspects of the covariance matrix structure in the discussion here. Specifically, the

following observations may be made from Table 5. First, the shaded matrices along the diagonal of the correlation matrix do not have many off-diagonal elements. This suggests the absence of common unobserved factors that affect participation across accompaniment types for any given activity purpose category. Thus, for example, a higher than average propensity to participate alone in non-maintenance activity (due to unobserved factors) does not increase or decrease the propensity to participate in non-maintenance activity with others. The main exception to this general observation is for meal activities, where there are significant substitution effects across accompaniment types. That is, an individual's propensity to pursue dining out with a particular companion type is negatively correlated with the individual's propensity to pursue dining out with other companion types. Second, the large number of parameters significant and consistently positive along the diagonals in each off-diagonal matrix of Table 5 highlights the preference for sticking to the same accompaniment (social) group for undertaking different types of activities. For example, individuals predisposed to participating in maintenance shopping activity with "only family" tend to participate in other activity purposes too with "only family". This preference (or stickiness) to pursue all types of activities with the same accompaniment group is particularly strong for the non-alone accompaniment categories. Third, some of the highest correlation values may be observed along the diagonal of the matrix corresponding to meals (row entry) and physically inactive recreation (column entry), suggesting that meals out and physically inactive recreation episodes are frequently combined (for instance, dinner out and a movie, or dinner out and a cultural event). This is reinforced by the fact that individuals who tend to have meals with "only friends" are not very likely to pursue physically inactive recreation alone or with "only relatives". In any case, there is a general complementary relationship between the propensities to participate in meals out and physically inactive recreation. Fourth, there are also quite high correlation values along the diagonals of the matrices corresponding to maintenance shopping and non-maintenance shopping, maintenance shopping and meals, and non-maintenance shopping and meals, highlighting the strong complementary tendencies among shopping/meal activities with the same accompaniment type. Fifth, the most number of off-diagonal correlation elements may be found in the matrix for non-maintenance shopping and physically inactive recreation, indicating substantial complementary effects in participation propensities for these two activity purpose categories across all types of accompaniment arrangements. Sixth, rather than the common perception that there is a substitution effect

between physically active and physically inactive recreation propensities, there is in fact a complementary effect. That is, individuals who participate more in physically inactive recreation are also more likely (after controlling for observed factors) to participate in physically active recreation. Finally, there is a general complementary relationship between participation with “mixed company” and participation with other company types for the non-maintenance shopping activity and other discretionary activity purposes (meals, physically inactive recreation, and physically active recreation).

### 5.2.6 Overall Measures of Fit

The log-composite likelihood value for the independent ordered response probit model (that is, independent ordered response probit models for each episode category) with only the threshold parameters is  $-1,136,772.91$ . The corresponding value at convergence for the fully specified independent ordered response probit model (IORP) is  $-1,083,191.5$  and that for the fully specified multivariate ordered response probit model (MORP) is  $-1,081,484.6$ . The composite likelihood ratio test (*CLRT*) statistic for comparing the MORP model with the IORP model is  $3413.83$ . However, the *CLRT* statistic does not have the standard chi-squared asymptotic distribution under the null hypothesis as in the case of the maximum likelihood inference procedure. In the current paper, we use bootstrapping to obtain the precise distribution of the *CLRT* statistic. The procedure is as follows (Varin and Czado, 2008):

1. Let the estimation sample be denoted as  $y_{obs}$ , and the observed *CLRT* value as  $CLRT(y_{obs})$ .
2. Generate  $B$  sample data (here,  $B = 50$ ) sets  $y_1, y_2, y_3, \dots, y_B$  using the CML convergent values under the null hypothesis.
3. Compute the *CLRT* statistic for each generated data set, and label it as  $CLRT(y_b)$ .
4. Calculate the p-value of the test using the following expression:

$$p = \frac{1 + \sum_{b=1}^B I\{CLRT(y_b) \geq CLRT(y_{obs})\}}{B + 1}, \text{ where } I\{A\} = 1 \text{ if } A \text{ is true.}$$

The estimated p-value based on 50 bootstrap samples is 0.0196 for the test between the MORP and IORP models. This low p-value rejects the null hypotheses of absence of correlations across the propensities of participation for the different episode categories, and highlights the

value of the MORP model estimated in the current paper. Of course, this should also be obvious from the many statistically significant parameters in the correlation matrix in Table 5.

Another more intuitive, but aggregate, approach to obtain a sense of measure of fit would be to compare the predicted versus the actual number of out-of-home episodes for each activity purpose-accompaniment combination level. In this paper, and to illustrate the data fit of the models while also conserving on space, we present the results only for the episode level combinations of two categories: meals with friends and physically inactive recreation with friends. These are two of the most common episode categories participated in during weekdays, as observed earlier in Section 4.2. Also, we select these two episode categories because they are helpful in demonstrating the application of the model in response to changes in socio-demographic variables (see next section). Table 6 presents the results, where the numbers in underlined font correspond to the actual number of individuals participating in each level of the two episode categories. The numbers in plain font are the predicted values from the MORP model, while the italicized numbers are the predicted values from the IORP model. A visual comparison of these numbers indicates the superiority in data fit of the MORP model. To quantify this, we develop a weighted mean absolute percentage error statistic that is computed as the absolute percentage error for each cell weighted by the fraction of individuals in each cell (based on the actual numbers in each cell). This statistic is 4.5% for the MORP model and 17.8% for the IORP model. One can also compute a more traditional root mean-squared error (RMSE) statistic between the predicted and actual values across all the cells for each of the MORP and IORP models. This statistic is 17.8 for the MORP and 76.4 for the IORP.

Overall, from the perspectives of both disaggregate and aggregate measures of fit, the MORP model clearly outperforms the IORP model.

### **5.3 Model Application**

The model estimated in this paper can be used to determine the change in the number of out-of-home episodes for each activity purpose-accompaniment type combination due to changes in independent variables over time. This is particularly important because of changing employment-related and demographic trends. For instance, the number of employed individuals is projected to continue to rise (albeit at a slower rate than in the past), despite the short-term slump due to the economy (see the latest national employment projections to 2016 by the Bureau of Labor

Statistics, 2007). Also, according to the US Census Bureau estimates from the Current Population Survey (CPS) (see US Census Bureau, 2009), the structure of the household is changing with a decrease in nuclear family households and an increase in single individual households. Such socio-demographic changes will have an effect on weekday episode participation, and the model in this paper can be used to assess these impacts and provide reliable information that can be used for activity-based travel demand forecasting and air quality analysis.

In this paper, we demonstrate the application of the model by studying the effect of two socio-demographic changes. The first is an increase in the number of full-time employed adults and the second is a decrease in nuclear family households along with a concomitant increase in single individual households. The increase in the number of full-time employed adults is reflected by randomly selecting current non-employed adults in the sample and designating them as full-time employees so that the number of full-time employees increases by 20% over the current full-time employment level. As indicated earlier, such a change mirrors the projected increase in employment levels in the U.S. population. The change in nuclear family households is similarly “implemented” by randomly selecting 20% of individuals who belong to nuclear families and placing them in single individual households. The impact of the two changes discussed above is evaluated by modifying exogenous variables to reflect the change, computing revised expected aggregate values for number of episodes in each combination category, and then obtaining a percentage change from the baseline estimates.

The effects of the changes in variables can be evaluated on each combination level of number of episodes across all the 30 episode categories. But there are about 80 trillion such combination levels. So, in this paper, we present the results only for the episode level combinations for two categories: meals with friends and physically inactive recreation with friends. These are two of the most common episode categories participated in during weekdays, as observed earlier in Section 4.2. Besides, the estimation results indicate that employment status and household structure, the two variables being examined here, have a direct influence on the “meals with friends” and “physically inactive recreation with friends” categories.

Table 7 presents the results from both the MORP (plain font) and IORP (italicized font) models. For each model, the predicted change in the number of individuals participating in each combination level of “meals with friends” and “physically inactive recreation with friends” is

computed as a percentage of the baseline (actual) numbers of individuals in each combination level. For ease in presentation, and also because the share of individuals participating in three or more episodes of physically active recreation with friends is very small, we have consolidated the 2 and 3 episode levels into a single 2+ episode level in Table 7. The results show a decrease in the (0,0) combination level due to an increase in full-time employed adults and decrease (increase) in nuclear family (single individual) households. This is, of course, because of the positive effect of full-time employed status on both the episode categories under consideration, and the negative (positive) effect of nuclear family household (single individual households) on both the episode categories (see Table 4). However, the percentage reduction in the number of individuals in the (0,0) cell is lower in the MORP case because of the positive correlation in the propensities of participation in the two episode categories. At the other extreme, both models show, as expected, an increase in the (2,2+) combination level. However, the MORP model indicates a substantially higher increase because of the complementary effect (positive correlation) in the unobserved propensities. The changes in the other cells, in general, also show a shift toward combinations of higher levels of episode participation in the two episode categories due to changes in the socio-demographic variables.

Overall, the exercise above demonstrates the application of the MORP model to predict the shifts in number of episodes of different activity purposes and accompaniment types due to changing socio-demographic characteristics of the population. In addition, the results also point to the biased results that can be obtained by ignoring the jointness in the propensity to participate in different episode categories.

## **6. SUMMARY AND CONCLUSIONS**

This paper proposes a multivariate ordered response system framework to model the interactions in activity episode decisions across household and non-household members at the fundamental level of activity generation. Such a system recognizes the dependence in the number of episodes generated for different purposes as well as with different accompaniment types, and explicitly allows *complementary* and *substitution* effects in activity episode participation decisions. The econometric challenge in estimating such a joint multivariate ordered-response system with a large number of episode categories is that traditional classical and Bayesian simulation techniques become saddled with convergence problems and imprecision; they are also extremely

cumbersome if not impractical to implement. We address this estimation problem by resorting to the technique of composite marginal likelihood (CML), an emerging inference approach in the statistics field that is based on the classical frequentist approach, is very simple to estimate, is easy to implement regardless of the number of count outcomes to be modeled jointly, and requires no simulation machinery whatsoever. It also represents a conceptually and pedagogically simpler procedure relative to simulation techniques, and has the advantage of reproducibility of the results. The simulation exercises undertaken for 3- and 5-dimensional multivariate ordered response models in the paper also indicate that the CML approach recovers the parameters extremely well. To our knowledge, this is the first study to adopt a CML approach in the field of activity-travel modeling, though Bhat *et al.* (2009) use the CML approach in the context of a spatially dependent discrete choice model formulation.

The empirical analysis in the paper uses data drawn from the 2007 American Time Use Survey (ATUS). Unlike conventional activity-travel surveys, the ATUS survey explicitly collects information on all accompanying family and non-family members for all activity episode participations. Thus, it is an ideal dataset for exploring the social context of adults' activity episode participations.<sup>7</sup> The empirical results provide important insights into the determinants of adults' weekday activity episode generation behavior. For instance, the results indicate the presence of distinct gender effects in activity type participation and accompaniment, with women being more responsible for, and/or more vested and interested in, family care and shopping activities, and men being more likely to undertake active and inactive leisure activities either alone or with friends. Further, there are also clear age-related effects. Individuals below the age of 40 years are the least likely (relative to other age groups) to participate in activity episodes alone and most likely to participate in episodes with mixed company, suggesting a combination of the family orientation and larger social networks of younger individuals. Race, education level, employment and student status, household structure and presence of children, household income, the day of week, and season of the year also have important effects on adults' weekday activity episodes by purpose and the social context of participation. In addition to estimating the coefficients of explanatory variables, the CML approach allows us to estimate the parameters

---

<sup>7</sup> A limitation of ATUS is that it does not collect locational information on household residences or activity episode participation locations. Hence, our analysis is unable to include built environment and locational effects on episode generation behavior. If available, this information can be incorporated as additional attributes in our multivariate ordered response system.

underlying the correlation due to unobserved factors in the propensity to participate in the 30 different purpose-accompaniment episode categories. Accommodating these unobserved correlation effects leads to a statistically superior data fit in the empirical context of this study and also provides useful insights into complementary and substitution effects among activity type and companionship type dimensions. Overall, the empirical estimation results underscore the ability of the CML approach to specify and estimate behaviorally rich structures to analyze inter-individual interactions in activity episode generation. The multivariate ordered-response system is applied in a demonstration exercise to evaluate the effect of changes in full-time employment and household structure “over time”.

In summary, the results underscore the substantial linkages in the activity episode generation of adults based on activity purpose and accompaniment type. The extent of this linkage varies by individual demographics, household demographics, day of the week, and season of the year. These inter- and intra-family linkages, and their variations across individuals, need to be accommodated within the framework of activity-based travel modeling for accurate travel forecasting and reliable transportation policy analysis.

## **ACKNOWLEDGEMENTS**

This research was partially funded by a Southwest Region University Transportation Center grant. The authors are grateful to Lisa Macias for her help in formatting this document. The corresponding author would like to acknowledge the support of an International Visiting Professorship from the University of Cagliari. The authors would also like to thank two anonymous referees for comments on an earlier version of the paper.

## REFERENCES

- Agyemang-Duah, K., Hall, F.L., 1997. Spatial transferability of an ordered response model of trip generation. *Transportation Research Part A* 31(5), 389-402.
- Agyemang-Duah, K., Hall, F.L., Anderson, W.P., 1995. Trip generation for shopping travel. *Transportation Research Record* 1493, 12-20.
- Apanasovich, T.V., Ruppert, D., Lupton, J.R., Popovic, N., Turner, N.D., Chapkin, R.S., Carroll, R.J., 2008. Aberrant crypt foci and semiparametric modelling of correlated binary data. *Biometrics* 64(2), 490-500.
- Arentze, T.A., Timmermans, H.J.P., 2004. A learning-based transportation oriented simulation system. *Transportation Research Part B* 38(7), 613-633.
- Arentze, T., Timmermans, H., 2008. Social networks, social interactions, and activity-travel behavior: A framework for microsimulation. *Environment and Planning B* 35(6), 1012-1027.
- ATUS, 2008. American time use survey user's guide understanding ATUS 2003 to 2007. <http://www.bls.gov/tus/atususersguide.pdf>.
- Axhausen, K.W., 2005. Social networks and travel: Some hypotheses. In *Social Aspects of Sustainable Transport: Transatlantic Perspectives*, Donaghy, K., Poppelreuter, S., Rudinger G., eds., Ashgate Publishing, Aldershot, England, Chapter 7, 90-108.
- Bhat, C.R., 1999. An analysis of evening commute stop-making behavior using repeated choice observations from a multi-day survey. *Transportation Research Part B* 33(7), 495-510.
- Bhat, C.R., 2001. Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research Part B* 35(7), 677-693.
- Bhat, C.R., 2003. Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. *Transportation Research Part B* 37(9), 837-855.
- Bhat, C.R., Gossen, R., 2004. A mixed multinomial logit model analysis of weekend recreational episode type choice. *Transportation Research Part B* 38(9), 767-787.
- Bhat, C.R., Guo, J.Y., 2007. A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels. *Transportation Research Part B* 41(5), 506-526.
- Bhat, C.R., Koppelman, F.S., 1999. Activity-based modeling of travel demand. In *The Handbook of Transportation Science*, Hall, R.W., ed., Kluwer Academic Publishers, Norwell, Massachusetts, 35-61.
- Bhat, C.R., Pendyala, R.M., 2005. Modeling intra-household interactions and group decision-making. *Transportation* 32(5), 443-448.
- Bhat, C.R., Srinivasan, S., 2005. A multidimensional mixed ordered-response model for analyzing weekend activity participation. *Transportation Research Part B* 39(3), 255-278.
- Bhat, C.R., Sener, I.N., Eluru, N., 2009. A flexible spatially dependent discrete choice model: Formulation and application to teenagers' weekday recreational activity participation. Technical Paper, Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, January 2009.

- Bradley, M., Vovsha, P., 2005. A model for joint choice of daily activity pattern types of household members. *Transportation* 32(5), 545-571.
- Bricka, S., Bhat, C.R., 2006. A comparative analysis of GPS-based and travel survey-based data. *Transportation Research Record* 1972, 9-20.
- Bureau of Labor Statistics, 2007. Employment projections 2006-2016, <http://www.bls.gov/news.release/ecopro.nr0.htm>, accessed May 3, 2009.
- Carrasco, J-A., Miller, E.J., 2009. The social dimension in action: A multilevel, personal networks model of social activity frequency between individuals. *Transportation Research Part A* 43(1), 90-104.
- Chib, S., Winkelmann, R., 2001. Markov chain Monte Carlo analysis of correlated count data. *Journal of Business and Economic Statistics* 19(4), 428-435.
- de Leon, A.R., 2005. Pairwise likelihood approach to grouped continuous model and its extension. *Statistics & Probability Letters* 75(1), 49-57.
- Doherty, S.T., Axhausen, K.W., 1999. The development of a unified modeling framework for the household activity-travel scheduling process. In *Traffic and Mobility: Simulation-Economics-Environment*, Brilon, W., Huber, F., Schreckengerg, M., Wallentowitz, H., eds., 35-56.
- Egan, K., Herriges, J., 2006. Multivariate count data regression models with individual panel data from an on-site sample. *Journal of Environmental Economics and Management* 52(2), 567-581.
- Frusti, T., Bhat, C.R., Axhausen, K.W., 2003. An exploratory analysis of fixed commitments in individual activity-travel patterns. *Transportation Research Record* 1807, 101-108.
- Gliebe, J.P., Koppelman, F.S., 2002. A model of joint activity participation between household members. *Transportation* 29(1), 49-72.
- Godambe, V., 1960. An optimum property of regular maximum likelihood equation. *Annals of Mathematical Statistics* 31, 1208-1211.
- Golob, T.F., 1990. The dynamics of household travel time expenditures and car ownership decisions. *Transportation Research Part A* 24(6), 443-463.
- Golob, T.F., van Wissen, L., 1989. A joint household travel distance generation and car ownership model. *Transportation Research Part B* 23(6), 471-491.
- Goulias, K.G., Kim, T.G., 2001. Multi-level analysis of activity and travel patterns: Accounting for person- and household- specific observed and unobserved effects simultaneously. *Transportation Research Record*, 1752, 23-31.
- Goulias, K.G., Kim, T.G., 2005. An analysis of activity type classification and issues related to the with whom and for whom questions of an activity diary. In *Progress in Activity-Based Analysis*, Timmermans, H.J.P., ed., Elsevier, Oxford, England.
- Hajivassiliou, V., McFadden, D., Ruud, P., 1996. Simulation of multivariate normal rectangle probabilities and their derivatives theoretical and computational results. *Journal of Econometrics* 72(1-2), 85-134.
- Herriges, J.A., Phaneuf, D.J., Tobias, J.L., 2008. Estimating demand systems when outcomes are correlated counts. *Journal of Econometrics* 147(2), 282-298.

- Higham, N.J., 2002. Computing the nearest correlation matrix – a problem from finance. *IMA Journal of Numerical Analysis* 22(3), 329-343.
- Jeliazkov, I., Graves, J., Kutzbach, M., 2008. Fitting and comparison of models for multivariate ordinal outcomes. In *Advances in Econometrics, Volume 23, Bayesian Econometrics*, Chib, S., Griffiths, W., Koop, G., Terrell, D., eds., Emerald Group Publishing Limited, Bingley, U.K., 115-156.
- Jones, P., Koppelman, F.S., Orfeuil, J., 1990. Activity analysis: State-of-the-art and future directions. In *Developments in Dynamic and Activity-based Approaches to Travel Analysis. A compendium of papers from the 1989 Oxford Conference*, Jones, P., ed., Avebury, U.K., 34-35.
- Kapur, A., Bhat, C.R., 2007. On modeling adults' daily time use by activity purpose and accompaniment arrangement. *Transportation Research Record* 2021, 18-27.
- Kato, H., Matsumoto, M., 2009. Intra-household interaction in a nuclear family: A utility-maximizing approach. *Transportation Research Part B* 43(2), 191-203.
- Kitamura, R., 1987. A panel analysis of household car ownership and mobility, infrastructure planning and management. *Proceedings of the Japan Society of Civil Engineers*, 383/IV-7, 13-27.
- Kitamura, R., 1988. A dynamic model system of household car ownership, trip generation, and modal split: Model development and simulation experiment. *Proceedings of the 14th Australian Road Research Board Conference, Part 3*, 96-111.
- Mallett, W.J., McGuckin, N., 2000. Driving to distractions: Recreational trips in private vehicles. *Transportation Research Record* 1719, 267-272.
- McKelvey, R.D., Zavoina, W., 1975. A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology* 4(summer), 103-120.
- Meka, S., Pendyala, R., Kumara, M., 2002. A structural equations analysis of within-household activity and time allocation between two adults. Presented at the 81<sup>st</sup> Annual Meeting of the Transportation Research Board, Washington, D.C.
- Meurs, H., 1989. Dynamic analysis of trip generation. Presented at the International Conference on dynamic behavior analysis, Kyoto, Japan, July 21-23, 1989.
- Meyer, B.D., 1990. Unemployment insurance and unemployment spells. *Econometrica* 58(4), 757-782.
- Parsons Brinckerhoff Quade and Douglas, Inc., 2000. Comparative analysis weekday and weekend travel with NPTS integration for the RT-HIS: Regional travel-household interview survey. Prepared for the New York Metropolitan Council and the North Jersey Transportation Planning Authority, February 2000.
- Pendyala, R.M., Goulias, K.G., 2002. Time use and activity perspectives in travel behavior research. *Transportation* 29(1), 1-4.
- Pinjari, A.R., Bhat, C.R., 2010. Activity-based travel demand analysis. *Handbook in Transport Economics*, de Palma, A., Lindsey, R., Quinet, E., Vickerman, R., eds., Edward Elgar Publishing, forthcoming.

- Pinjari, A., Rajagopalan, B.S., Ferdous, N., Bhat, C.R., 2008. Activity-based model development: A review of current activity-based travel demand models. Report prepared for the North Central Texas Council of Governments (NCTCOG), March 2008.
- Rebonato, R., Jaeckel, P., 1999. The most general methodology for creating a valid correlation matrix for risk management and option pricing purposes. *The Journal of Risk* 2(2), 17-28.
- Schoettle, K., Werner, R., 2004. Improving “the most general methodology to create a valid correlation matrix”. In *Risk Analysis IV, Management Information Systems*, Brebbia, C.A., ed., WIT Press, Southampton, U.K., 701-712.
- Scott, D.M., Kanaroglou, P.S., 2002. An activity-episode generation model that captures interactions between household heads: development and empirical analysis. *Transportation Research Part B* 36(10), 875-896.
- Sener, I.N., Bhat, C.R., 2007. An analysis of the social context of children’s weekend discretionary activity participation. *Transportation* 34(6), 697-721.
- Srinivasan, K.K., Athuru, S.R., 2005. Analysis of within-household effects and between household differences in maintenance activity allocation. *Transportation* 32(5), 495-521.
- Srinivasan, S., Bhat, C.R., 2005. Modeling household interactions in daily in-home and out-of-home maintenance activity participation. *Transportation* 32(5), 523-544.
- Srinivasan, S., Bhat, C.R., 2006. Companionship for leisure activities: An empirical analysis using the American Time Use Survey. Presented at the Innovations in Travel Demand Modeling Conference, Transportation Research Board Conference Proceedings 42(2), 129-136.
- Srinivasan, S., Bhat, C.R., 2008. An exploratory analysis of joint-activity participation characteristics using the American Time Use Survey. *Transportation* 35(3), 301-328.
- Thornton, A., Orbuch, T.L., Axinn, W.G., 1995. Parent-child relationships during the transition to adulthood. *Journal of Family Issues* 16(5), 538-564.
- Timmermans, H.J.P., Zhang, J., 2009. Modeling household activity travel behavior: Examples of state of the art modeling approaches and research agenda. *Transportation Research Part B* 43(2), 187-190.
- US Census Bureau, 2009. FM-1. Families, by presence of own children under 18: 1950 to present, <http://www.census.gov/population/socdemo/hh-fam/fm1.xls>, accessed May 3, 2009.
- Varin, C., 2008. On composite marginal likelihoods. *Advances in Statistical Analysis* 92(1), 1-28.
- Varin, C., Czado, C., 2008. A mixed probit model for the analysis of pain severity diaries. Available at: [http://www-m4.ma.tum.de/Papers/Czado/varin\\_czado\\_pain\\_diaries.pdf](http://www-m4.ma.tum.de/Papers/Czado/varin_czado_pain_diaries.pdf)
- Varin, C., Vidoni, P., 2008) Pairwise likelihood inference for general state space models. *Econometric Reviews* 28(1-3), 170-185.
- Vovsha, P., Bradley, M., 2006. Advanced activity-based models in context of planning decisions. *Transportation Research Record* 1981, 34-41.

- Vovsha, P., Peterson, E., Donnelly, R., 2003. Explicit modeling of joint travel by household members: Statistical evidence and applied approach. *Transportation Research Record* 1831, 1-10.
- Wang, D., Li, J., 2009. A model of household time allocation taking into consideration of hiring domestic helpers. *Transportation Research Part B* 43(2), 204-216.
- Wen, C-H., Koppelman, F.S., 1999. An integrated system of stop generation and tour formation for the analysis of activity and travel patterns. *Transportation Research Record* 1676, 136-144.
- Williams, A., 2003. Adolescents' relationships with parents. *Journal of Language and Social Psychology* 22(1), 58-65.
- Yamamoto, T., Kitamura, R., 1999. An analysis of time allocation to in-home and out-of-home discretionary activities across working days and non-working days. *Transportation* 26(2), 211-230.
- Yamamoto, T., Kitamura, R., Pendyala, R.M., 2004. Comparative analysis of time-space prism vertices for out-of-home activity engagement on working and non-working days. *Environment and Planning B* 31(2), 235-250.
- Zhao, Y., Joe, H., 2005. Composite likelihood estimation in multivariate data analysis. *The Canadian Journal of Statistics* 33(3), 335-356.

## LIST OF TABLES

- Table 1 Evaluation of Ability to Recover “True” Parameters by CML Estimation Technique – Trivariate Case
- Table 2 Evaluation of Ability to Recover “True” Parameters by CML Estimation Technique – Five-variate Case
- Table 3 Percentage of Individuals in Each Number of Episodes Category by ‘With Whom’ and Activity Types (Weekday)
- Table 4 Model Estimation Results (t-statistics in parentheses)
- Table 5 Correlation in Unobserved Propensities Across the Choice Dimension (t-statistics in parentheses)
- Table 6 Number of Individuals Choosing “Meals with Friends” and “Physically Inactive Recreation with Friends” Episodes
- Table 7 Impact of Changes on the Percentage of Individuals Choosing Each Combination Level of “Meals with Friends” and “Physically Inactive Recreation with Friends” Episodes

**Table 1 Evaluation of Ability to Recover “True” Parameters by CML Estimation Technique – Trivariate Case**

Parameter	Low Correlation						High Correlation						
	True Value	Mean Estimate	Mean Standard Error	Absolute Bias	Absolute Percentage Bias	Root Mean Square Error (RMSE)	True Value	Mean Estimate	Mean Standard Error	Absolute Bias	Absolute Percentage Bias	Root Mean Square Error (RMSE)	
$\theta_1^1$	-1.0000	-1.0078	0.0553	0.0078	0.7821	0.0553	-1.0000	-1.0086	0.0532	0.0086	0.8644	0.0535	
$\theta_1^2$	1.0000	0.9998	0.0553	0.0002	0.0202	0.0556	1.0000	0.9977	0.0528	0.0023	0.2275	0.0524	
$\theta_1^3$	3.0000	3.0104	0.1301	0.0104	0.3473	0.1305	3.0000	3.0113	0.1208	0.0113	0.3756	0.1185	
$\theta_2^1$	0.0000	0.0074	0.0493	0.0074	-	0.0521	0.0000	-0.0081	0.0456	0.0081	-	0.0456	
$\theta_2^2$	2.0000	2.0100	0.0855	0.0100	0.5005	0.1023	2.0000	1.9976	0.0808	0.0024	0.1194	0.0638	
$\theta_3^1$	-2.0000	-1.9860	0.0765	0.0140	0.7018	0.0749	-2.0000	-2.0022	0.0758	0.0022	0.1118	0.0752	
$\theta_3^2$	-0.5000	-0.5002	0.0465	0.0002	0.0433	0.0443	-0.5000	-0.5026	0.0455	0.0026	0.5141	0.0381	
$\theta_3^3$	1.0000	1.0144	0.0528	0.0144	1.4445	0.0530	1.0000	1.0025	0.0517	0.0025	0.2542	0.0573	
$\theta_3^4$	2.5000	2.5218	0.1018	0.0218	0.8719	0.1100	2.5000	2.5144	0.0981	0.0144	0.5776	0.1143	
$\beta_{11}$	0.5000	0.4949	0.0400	0.0051	1.0283	0.0391	0.5000	0.4979	0.0309	0.0021	0.4216	0.0300	
$\beta_{21}$	1.0000	0.9976	0.0477	0.0024	0.2388	0.0508	1.0000	0.9975	0.0402	0.0025	0.2478	0.0384	
$\beta_{31}$	0.2500	0.2516	0.0375	0.0016	0.6541	0.0374	0.2500	0.2521	0.0278	0.0021	0.8247	0.0251	
$\beta_{12}$	0.7500	0.7515	0.0492	0.0015	0.1963	0.0519	0.7500	0.749	0.0395	0.0010	0.1299	0.0405	
$\beta_{22}$	1.0000	1.0013	0.0542	0.0013	0.1257	0.0701	1.0000	0.9913	0.0449	0.0087	0.8683	0.0523	
$\beta_{32}$	0.5000	0.5076	0.0447	0.0076	1.5247	0.0392	0.5000	0.5035	0.0349	0.0035	0.6967	0.0274	
$\beta_{42}$	0.2500	0.2605	0.0421	0.0105	4.1928	0.0440	0.2500	0.2523	0.0315	0.0023	0.9082	0.0289	
$\beta_{13}$	0.2500	0.2522	0.0352	0.0022	0.8743	0.0341	0.2500	0.2539	0.0278	0.0039	1.5671	0.0254	
$\beta_{23}$	0.5000	0.4974	0.0364	0.0026	0.5117	0.0357	0.5000	0.5011	0.0297	0.0011	0.2217	0.0293	
$\beta_{33}$	0.7500	0.7485	0.0395	0.0015	0.1937	0.0331	0.7500	0.7528	0.0330	0.0028	0.3787	0.0293	
$\rho_{12}$	0.3000	0.2995	0.0468	0.0005	0.1608	0.0400	0.9000	0.9031	0.0181	0.0031	0.3491	0.0194	
$\rho_{13}$	0.2000	0.1989	0.0399	0.0011	0.5720	0.0422	0.8000	0.8006	0.0201	0.0006	0.0704	0.0233	
$\rho_{23}$	0.2500	0.2644	0.0436	0.0144	5.7437	0.0477	0.7500	0.7565	0.0257	0.0065	0.8669	0.0255	
<b>Overall mean value across parameters</b>			0.0550	0.0063	0.9871%	0.0565	<b>Overall mean value across parameters</b>			0.0467	0.0043	0.5046%	0.0461

**Table 2 Evaluation of Ability to Recover “True” Parameters by CML Estimation Technique - Five-variate Case**

Parameter	Low Correlation						High Correlation						
	True Value	Mean Estimate	Mean Standard Error	Absolute Bias	Absolute Percentage Bias	Root Mean Square Error (RMSE)	True Value	Mean Estimate	Mean Standard Error	Absolute Bias	Absolute Percentage Bias	Root Mean Square Error (RMSE)	
$\theta_1^1$	-1.0000	-1.0176	0.0559	0.0176	1.76%	0.0647	-1.0000	-1.017	0.0540	0.0170	1.70%	0.0644	
$\theta_1^2$	1.0000	1.0085	0.0556	0.0085	0.85%	0.0530	1.0000	1.0077	0.0531	0.0077	0.77%	0.0495	
$\theta_1^3$	3.0000	2.9876	0.1261	0.0124	0.41%	0.1236	3.0000	2.9906	0.1194	0.0094	0.31%	0.1193	
$\theta_2^1$	0.0000	-0.0074	0.0492	0.0074	-	0.0454	0.0000	-0.0024	0.0454	0.0024	-	0.0399	
$\theta_2^2$	2.0000	2.0054	0.0849	0.0054	0.27%	0.0890	2.0000	2.021	0.0810	0.0210	1.05%	0.0943	
$\theta_3^1$	-2.0000	-2.0214	0.0788	0.0214	1.07%	0.0927	-2.0000	-2.036	0.0766	0.0360	1.80%	0.0833	
$\theta_3^2$	-0.5000	-0.5081	0.0468	0.0081	1.63%	0.0494	-0.5000	-0.5029	0.0452	0.0029	0.58%	0.0418	
$\theta_3^3$	1.0000	1.0053	0.0524	0.0053	0.53%	0.0593	1.0000	1.0085	0.0511	0.0085	0.85%	0.0534	
$\theta_3^4$	2.5000	2.5241	0.1008	0.0241	0.97%	0.0926	2.5000	2.5207	0.0956	0.0207	0.83%	0.1002	
$\theta_4^1$	1.0000	1.0124	0.0628	0.0124	1.24%	0.0560	1.0000	1.0122	0.0566	0.0122	1.22%	0.0601	
$\theta_4^2$	3.0000	3.0059	0.1362	0.0059	0.20%	0.0972	3.0000	3.0003	0.1269	0.0003	0.01%	0.1045	
$\theta_5^1$	-1.5000	-1.5114	0.0661	0.0114	0.76%	0.0766	-1.5000	-1.5204	0.0649	0.0204	1.36%	0.0664	
$\theta_5^2$	0.5000	0.5074	0.0499	0.0074	1.48%	0.0421	0.5000	0.5044	0.0476	0.0044	0.88%	0.0431	
$\theta_5^3$	2.0000	2.0215	0.0794	0.0215	1.08%	0.0690	2.0000	2.0082	0.0770	0.0082	0.41%	0.0775	
$\beta_{11}$	0.5000	0.5035	0.0401	0.0035	0.71%	0.0418	0.5000	0.5037	0.0318	0.0037	0.75%	0.0330	
$\beta_{21}$	1.0000	1.0068	0.0484	0.0068	0.68%	0.0454	1.0000	1.0051	0.0411	0.0051	0.51%	0.0424	
$\beta_{31}$	0.2500	0.2538	0.0379	0.0038	1.54%	0.0328	0.2500	0.2533	0.0289	0.0033	1.32%	0.0271	
$\beta_{12}$	0.7500	0.7568	0.0492	0.0068	0.90%	0.0547	0.7500	0.7624	0.0392	0.0124	1.65%	0.0426	
$\beta_{22}$	1.0000	0.9995	0.0543	0.0005	0.05%	0.0495	1.0000	1.0197	0.0455	0.0197	1.97%	0.0586	
$\beta_{32}$	0.5000	0.5006	0.0450	0.0006	0.13%	0.0371	0.5000	0.5083	0.0345	0.0083	1.65%	0.0362	
$\beta_{42}$	0.2500	0.2553	0.0426	0.0053	2.14%	0.0369	0.2500	0.2622	0.0316	0.0122	4.90%	0.0311	
$\beta_{13}$	0.2500	0.2602	0.0344	0.0102	4.06%	0.0337	0.2500	0.2574	0.0251	0.0074	2.96%	0.0287	
$\beta_{23}$	0.5000	0.5038	0.0366	0.0038	0.77%	0.0406	0.5000	0.5097	0.0279	0.0097	1.94%	0.0313	
$\beta_{33}$	0.7500	0.7538	0.0392	0.0038	0.51%	0.0437	0.7500	0.7575	0.0318	0.0075	1.00%	0.0293	
$\beta_{14}$	0.7500	0.7633	0.0577	0.0133	1.78%	0.0572	0.7500	0.7636	0.0451	0.0136	1.82%	0.0428	
$\beta_{24}$	0.2500	0.2482	0.0491	0.0018	0.72%	0.0426	0.2500	0.2456	0.0358	0.0044	1.74%	0.0307	
$\beta_{34}$	1.0000	1.0083	0.0637	0.0083	0.83%	0.0610	1.0000	1.0067	0.0516	0.0067	0.67%	0.0453	
$\beta_{44}$	0.3000	0.3053	0.0493	0.0053	1.77%	0.0483	0.3000	0.3026	0.0363	0.0026	0.85%	0.0365	
$\beta_{15}$	0.4000	0.4057	0.0383	0.0057	1.44%	0.0370	0.4000	0.4056	0.0304	0.0056	1.41%	0.0265	
$\beta_{25}$	1.0000	1.0035	0.0465	0.0035	0.35%	0.0475	1.0000	1.0067	0.0404	0.0067	0.67%	0.0417	
$\beta_{35}$	0.6000	0.6053	0.0405	0.0053	0.89%	0.0469	0.6000	0.6125	0.0332	0.0125	2.09%	0.0389	
$\rho_{12}$	0.3000	0.2988	0.0468	0.0012	0.40%	0.0522	0.9000	0.905	0.0180	0.0050	0.56%	0.0200	
$\rho_{13}$	0.2000	0.1951	0.0401	0.0049	2.46%	0.0401	0.8000	0.7973	0.0205	0.0027	0.34%	0.0221	
$\rho_{14}$	0.2200	0.227	0.0559	0.0070	3.19%	0.0565	0.8200	0.8231	0.0290	0.0031	0.38%	0.0323	
$\rho_{15}$	0.1500	0.1411	0.0432	0.0089	5.96%	0.0382	0.7500	0.7501	0.0251	0.0001	0.01%	0.0249	
$\rho_{23}$	0.2500	0.2477	0.0441	0.0023	0.92%	0.0403	0.8500	0.8457	0.0200	0.0043	0.50%	0.0195	
$\rho_{24}$	0.3000	0.2978	0.0606	0.0022	0.75%	0.0581	0.9000	0.9021	0.0240	0.0021	0.23%	0.0252	
$\rho_{25}$	0.1200	0.1227	0.0482	0.0027	2.27%	0.0541	0.7200	0.7205	0.0300	0.0005	0.06%	0.0308	
$\rho_{34}$	0.2700	0.263	0.0508	0.0070	2.58%	0.0578	0.8700	0.8662	0.0220	0.0038	0.43%	0.0195	
$\rho_{35}$	0.2000	0.1884	0.0393	0.0116	5.79%	0.0404	0.8000	0.7987	0.0198	0.0013	0.16%	0.0184	
$\rho_{45}$	0.2500	0.2575	0.0540	0.0075	3.00%	0.0469	0.8500	0.8517	0.0256	0.0017	0.20%	0.0249	
<b>Overall mean value across parameters</b>			0.0561	0.0076	1.47%	0.0549	<b>Overall mean value across parameters</b>			0.0448	0.0082	1.06%	0.0453

**Table 3 Percentage of Individuals in Each Number of Episodes Category by ‘With Whom’ and Activity Types (Weekday)**

'With Whom' Dimension	Number of Episodes	Activity Type Dimension					
		Family care	Maintenance shopping	Non-maintenance shopping	Meals	Physically active recreation	Physically inactive recreation
Alone	0	91.3	73.7	86.7	79.1	90.2	87.4
	1	7.3	18.6	10.9	18.6	7.9	9.7
	2	1.4	5.6	2.4	2.3	1.9	2.8
	≥ 3		2.2				
Only family (children/spouse/partner)	0	78.8	90.5	92.5	91.6	96.4	95.7
	1	11.2	7.6	5.6	7.4	3.0	3.6
	2	6.5	1.9	1.9	1.0	0.6	0.7
	3	2.1					
≥ 4	1.4						
Only relatives (includes parents, brother, sister, and other related persons)	0	91.5	95.8	96.5	93.5	97.5	91.1
	1	5.6	3.4	2.9	5.6	2.1	6.6
	≥ 2	3.0	0.8	0.6	0.9	0.4	2.3
Only friends (includes friends, co-workers, neighbors, etc.)	0	96.7	96.5	98.5	72.9	94.9	81.6
	1	2.4	2.9	1.5	22.3	4.2	12.9
	2	0.9	0.5		4.8	0.9	4.1
	≥ 3						1.4
Mixed company (i.e., with family and/or relatives and/or friends)	0	94.2	98.3	99.3	95.6	97.9	92.8
	1	4.0	1.7	0.7	4.4	2.1	5.8
	≥ 2	1.8					1.4

**Table 4 Model Estimation Results (t-statistics in parentheses)**

		Individual socio-demographics variables									
		Male (base: female)	Age (base: age ≥ 60)		Caucasian (base: non-Caucasian)	Education level (base: high school graduate)		Employment status (base: not employed)		Student (base: not student)	Have a disability (base: no disability)
			Age <40	40 ≤ Age <60		Education < bachelors	Education ≥ bachelors	Full time employed	Part time employed		
Family care	Alone	-0.179 (-3.05)	-0.122 (-2.12)					-0.318 (-5.05)	-0.174 (-2.03)		
	Only family	-0.509 (-10.31)	0.339 (3.76)	0.438 (4.76)							
	Only relatives	-0.195 (-3.24)	-0.347 (-4.68)	-0.220 (-2.88)					-0.327 (-5.00)		0.406 (3.02)
	Only friends										
	Mixed company	-0.398 (-5.81)	0.314 (5.00)						-0.113 (-1.62)	-0.182 (-1.81)	
Maintenance shopping	Alone		-0.363 (-8.55)			0.277 (5.47)	0.298 (6.23)				
	Only family	-0.218 (-3.82)	0.305 (3.62)	0.376 (4.37)				-0.278 (-4.69)			
	Only relatives	-0.226 (-3.00)						-0.323 (-4.34)			0.398 (2.67)
	Only friends							-0.126 (-1.72)			
	Mixed company	-0.210 (-2.07)	0.303 (3.13)					-0.190 (-1.95)			
Non-maintenance shopping	Alone	-0.101 (-1.97)	-0.430 (-6.45)	-0.188 (-2.76)		0.241 (3.92)	0.162 (2.65)	-0.198 (-3.19)	-0.134 (-1.62)		
	Only family	-0.233 (-3.84)			0.322 (3.45)						
	Only relatives	-0.310 (-3.89)									0.419 (2.33)
	Only friends										
	Mixed company	-0.289 (-2.01)	0.394 (2.91)								
Meals	Alone	0.290 (6.34)						1.075 (16.02)	0.600 (6.57)		
	Only family		0.161 (1.85)	0.131 (1.52)							
	Only relatives										
	Only friends	0.088 (2.10)			0.107 (1.89)			0.867 (16.53)	0.447 (6.10)	0.489 (7.36)	
	Mixed company		0.287 (4.19)		0.310 (2.95)						
Physically active recreation	Alone	0.190 (3.54)	-0.261 (-3.57)	-0.242 (-3.38)		0.136 (1.91)	0.451 (7.21)				
	Only family		0.168 (2.16)		0.254 (2.12)						0.589 (3.21)
	Only relatives						0.189 (2.20)				
	Only friends		-0.281 (-3.00)	-0.403 (-4.56)			0.156 (2.25)			0.461 (4.59)	
	Mixed company		0.205 (2.10)							0.331 (2.62)	
Physically inactive recreation	Alone	0.080 (1.58)			-0.188 (-2.98)			0.386 (7.18)			
	Only family		0.212 (2.87)					-0.465 (-6.34)			
	Only relatives							-0.250 (-4.54)			
	Only friends	0.123 (2.73)						0.225 (4.78)		0.377 (5.50)	
	Mixed company	-0.205 (-3.35)	0.284 (4.85)					-0.192 (-3.25)			

**Table 4 (Continued) Model Estimation Results**

		Household socio-demographics variables						Day-of-the-week and seasonal effect variables			
		Household (HH) structure (base: "other" HH)			Presence of children (base: age ≤ 4)		HH income ( base: < 30k)		Spouse/partner employed (base: unemployed)	Friday (base: other days of the week)	Summer (base: fall, spring, and winter)
		Nuclear family HH	Couple HH	Single individual HH	4< Age ≤ 10	10< Age ≤ 15	30k ≤ Income < 75k	Income ≥ 75k			
Family care	Alone							0.161 (2.82)			
	Only family	0.381 (6.39)	-0.418 (-4.60)		0.579 (10.98)	0.308 (5.87)		0.537 (9.78)			
	Only relatives								0.183 (2.80)		
	Only friends							-0.477 (-5.20)			
	Mixed company										
Maintenance shopping	Alone								0.098 (1.92)		
	Only family	0.487 (6.81)	0.364 (4.92)		0.161 (2.27)				0.210 (3.26)		
	Only relatives	-0.423 (-4.59)							0.184 (2.23)		
	Only friends										
	Mixed company										
Non-maintenance shopping	Alone										
	Only family	0.614 (9.04)	0.393 (4.96)						0.202 (2.93)		
	Only relatives	-0.321 (-3.38)							0.237 (2.81)		
	Only friends						-0.216 (-2.02)				
	Mixed company										
Meals	Alone			0.278 (5.31)						-0.143 (-2.42)	
	Only family	0.576 (8.14)	0.532 (7.15)				0.345 (4.36)	0.294 (3.49)		0.124 (1.83)	
	Only relatives	-0.451 (-5.82)	-0.215 (-2.77)								
	Only friends	-0.166 (-3.37)		0.202 (3.78)		0.189 (3.60)				0.141 (2.80)	
	Mixed company									0.275 (3.47)	
Physically active recreation	Alone			0.169 (2.67)							
	Only family	0.357 (3.99)	0.187 (1.80)				0.206 (1.99)	0.348 (3.20)		0.222 (2.77)	
	Only relatives										
	Only friends	-0.191 (-2.29)								0.157 (2.09)	
	Mixed company						0.384 (2.70)	0.546 (3.85)			
Physically inactive recreation	Alone			0.146 (2.56)				-0.273 (-4.70)			
	Only family	0.405 (4.73)	0.210 (2.12)				0.327 (3.19)	0.411 (3.86)		0.161 (2.01)	
	Only relatives	-0.462 (-6.63)								0.150 (2.26)	
	Only friends	-0.225 (-4.09)		0.247 (4.37)		0.342 (6.16)				0.088 (1.63)	
	Mixed company									0.275 (4.05)	



**Table 6 Number of Individuals Choosing “Meals with Friends” and “Physically Inactive Recreation with Friends” Episodes**

Number of “meals with friends” episodes	Number of “physically inactive recreation with friends” episodes			
	0	1	2	3
0	<u>2667.00</u> <sup>a</sup>	<u>269.00</u>	<u>67.00</u>	<u>19.00</u>
	2650.14 <sup>b</sup>	285.22	70.17	16.87
	<i>2501.61</i> <sup>c</sup>	<i>375.46</i>	<i>117.54</i>	<i>38.47</i>
1	<u>597.00</u>	<u>207.00</u>	<u>92.00</u>	<u>28.00</u>
	638.12	188.62	69.20	24.84
	<i>729.13</i>	<i>127.76</i>	<i>42.60</i>	<i>14.92</i>
2	<u>117.00</u>	<u>58.00</u>	<u>12.00</u>	<u>10.00</u>
	100.39	55.14	28.72	15.57
	<i>152.35</i>	<i>29.27</i>	<i>10.16</i>	<i>3.72</i>

<sup>a</sup> The actual number of individuals participating in each combination level of episode category.

<sup>b</sup> The predicted number of individuals from the MORP model participating in each combination level of episode category.

<sup>c</sup> The predicted number of individuals from the IORP model participating in each combination level of episode category.

**Table 7 Impact of Changes on the Percentage of Individuals Choosing Each Combination Level of “Meals with Friends” and “Physically Inactive Recreation with Friends” Episodes**

Change	Number of “meals with friends” episodes	Number of “physically inactive recreation with friends” episodes		
		0	1	2+
Increase in full-time employed adults by 20% (and corresponding decrease in the number of non-employed adults)	0	-3.99 <sup>a</sup>	-3.79	-4.14
		-4.38 <sup>b</sup>	-1.70	1.24
	1	10.52	8.24	6.54
		10.92	7.33	6.62
	2	13.67	15.56	34.07
		20.12	9.48	13.46
Decrease in nuclear family households by 20% (and corresponding increase in the number of single individual households)	0	-1.57	2.07	4.08
		-1.80	2.66	9.03
	1	0.82	4.12	5.84
		0.94	4.18	5.46
	2	2.80	6.86	22.66
		5.47	5.83	10.87

<sup>a</sup> Percentage change in the number of individuals from the MORP model participating in each combination level of episode category.

<sup>b</sup> Percentage change in the number of individuals from the IORP model participating in each combination level of episode category.