A JOINT MODEL OF WEEKEND DISCRETIONARY ACTIVITY PARTICIPATION AND EPISODE DURATION

Kathryn Born
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin, TX 78712
Tel: 512-471-4535; Fax: 512-475-8744; Email: katieborn6@gmail.com

Shamsunnahar Yasmin
McGill University
Department of Civil Engineering & Applied Mechanics
Suite 483, 817 Sherbrooke St. W., Montréal, Québec, H3A 2K6, Canada
Tel: 514-398-6823; Fax: 514-398-7361; Email: shamsunnahar.yasmin@mail.mcgill.ca

Daehyun You
Arizona State University
School of Sustainable Engineering and the Built Environment
Room ECG252, Tempe, AZ 85287-5306.
Tel: (480) 965-3589; Fax: (480) 965-0557; Email: dae.you@asu.edu

Naveen Eluru
McGill University
Department of Civil Engineering and Applied Mechanics
Suite 483, 817 Sherbrooke St. W., Montréal, Québec, H3A 2K6, Canada
Tel: 514-398-6823; Fax: 514-398-7361; Email: naveen.eluru@mcgill.ca

Chandra R. Bhat (corresponding author)
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin, TX 78712
Tel: 512-471-4535; Fax: 512-475-8744; Email: bhat@mail.utexas.edu

Ram M. Pendyala
Arizona State University
School of Sustainable Engineering and the Built Environment
Room ECG252, Tempe, AZ 85287-5306.
Tel: (480) 727-9164; Fax: (480) 965-0557; Email: ram.pendyala@asu.edu

ABSTRACT
Research on travel demand modeling has primarily focused on weekday activity-travel patterns. However, weekend activities and travel constitute a major component of individuals' overall weekly activity-travel participation. This paper describes a modeling effort that focuses on discretionary-event weekend activity-travel demand. This study bridges the gap in the literature by modeling discretionary event type participation, duration of participation, and accompaniment type jointly in a simultaneous equations model system. A joint discrete-continuous modeling framework is formulated for analyzing these dimensions as a choice bundle. Specifically, the combination of event type and accompaniment type constitutes the discrete component while the duration of participation constitutes the continuous component. The model employs a copula-based sample selection approach that ties the discrete choice error component with the duration error component in a flexible manner. The data used in the paper is drawn from the 2008-2009 National Household Travel Survey sample of the Greater Phoenix metropolitan area in Arizona. The results from the estimation process clearly highlight the presence of sample selection in the joint modeling context. Furthermore, the results also highlight the flexibility of copula models in capturing such sample selection. The best copula model results were employed to generate hazard profiles for various alternative related duration intervals. The profiles generated highlight the inaccurate predictions obtained using approaches that ignore the presence of sample selection.

Keywords: Weekend activity participation, Sample selection, Copula, Duration modeling, Hazard profiles

1. INTRODUCTION

Research on travel demand modeling has primarily focused on weekday activity-travel patterns. However, weekend activities and travel constitute a major component of individuals' overall weekly activity-travel participation. Recent studies suggest that the level of weekend and weekday personal travel (both in terms of number of trips and person miles of travel) are about the same, and that the levels of mid-day weekend traffic congestion almost reach weekday commute traffic congestion levels in some locations (Sall and Bhat, 2007). This is not surprising, as discretionary special events (such as exhibitions, sporting events, and concerts) are more likely to be held over the weekend, leading to congestion in and around the discretionary special event sites. As cities move toward increasing weekend social, recreational, and cultural opportunities, it is important to understand and forecast the resulting weekend activity-travel patterns. The analysis of weekend activity-travel patterns is also important from the standpoint of policy analysis. For instance, there is evidence that individuals move activities to the weekend when their weekday schedules get busy or when traffic congestion during the commute periods leaves too little time for weekday discretionary activity participation (Bhat and Gossen, 2004; Lee, et al, 2009).

Overall, whether for demand forecasting or for analyzing the effects of transportation-related policy actions, it is important to examine special event-related weekend activity-travel patterns. In doing so, there are several key characteristics of weekend activity-travel patterns worthy of explicit recognition: (1) weekend activity-travel is highly leisure-oriented, (2) weekend activities may not have the same level of spatial and temporal rigidity as work activities (although some activities such as ballgames, church activities, and special events may have fixed locations and schedules), (3) inter-personal interactions in activity participation are stronger on the weekends due to greater joint activity participation and travel, and (4) the activity travel environment differs substantially between weekdays and weekends (for example, differences in facility operating hours, and lower transit level-of-service) (Lockwood, et al, 2005).

The current study augments the literature on weekend activity participation by modeling (1) discretionary special event type participation, (2) duration of participation and (3) accompaniment type (i.e., the travel party composition for the activity). The data used in the paper is drawn from the 2008-2009 National Household Travel Survey (NHTS) conducted in the United States. The survey sample residing in the Greater Phoenix metropolitan region in Arizona is extracted for use in this study. The study employs a copula-based joint discrete-continuous model of weekend activity participation to address the different dimensions involved in special event activity participation.

The rest of the paper is organized as follows. Section 2 provides a review of earlier literature on weekend activity participation while positioning the current study. In section 3, the methodology employed in the paper is outlined. Section 4 describes the data assembly process. Section 5 presents estimation results. Section 6 summarizes the paper and offers directions for future research in this arena.

2. EARLIER RESEARCH AND THE FOCUS OF THE CURRENT PAPER

Weekend activity participation and time-use has been receiving increased attention in the transportation field. These research efforts may be broadly classified into two categories: (1) studies that focus on comparison between weekend and weekday activity participation and (2) studies that exclusively examine weekend activity participation.

In the first category of studies, Bhat and Misra (1999), Yamomoto and Kitamura (1999) and Treuth, et al (2007), examine the allocation of time for discretionary activities on weekend days and weekdays while Sugie, et al (2003) examine the interdependency of shopping activity participation on weekday and weekend days. Some studies present comprehensive exploratory

analysis of the comparison between weekday and weekend activity-travel patterns (Lockwood, et al, 2005), while other studies highlight the differences in joint activity participation patterns between weekdays and weekend days (Srinivasan and Bhat, 2008). There has also been recent work applying time use models to study activity participation differences across weekdays and weekends; for example Lee, et al (2009) compared how weekday and weekend maintenance and discretionary activity time-use are affected by land-use and transportation network characteristics using multiple equation tobit models. For example, they found that in locations with suburban type density, individuals tend to lump discretionary activities for the weekends. Finally, some recent studies employ duration modeling approaches to study the differences between weekday and weekend activity patterns (see Zhong, et al, 2008 for separate duration models for various activity types). In summary, these studies highlight that weekday and weekend activity behavior differ in terms of activities pursued, activity duration (longer discretionary activity participation during weekends) and the activity travel environment experienced.

The second category of studies focuses exclusively on weekend participation. Bhat and Gossen (2004) examined weekend activity and travel participation considering in-home and out-of-home activities, and worker and non-worker activity-travel patterns. Bhat, et al (2006) examine individual activity time-use participation in maintenance and leisure activities over the weekend. Kapur and Bhat (2007) use the American Time Use Survey to examine weekend discretionary activity participation patterns in terms of the "with whom" dimension (household and non-household members). Sall and Bhat (2007) examined the spatial and temporal characteristics of weekend work episodes. Zhong and Hunt (2010) proposed a duration model of weekend activity participation by different activity types. Habib (2011) formulated a random utility based approach to examine weekend activity participation and scheduling. These studies underscore the importance of socio-demographics, urban built environment, weekend day effects and seasonal effects on weekend activity participation.

Despite the recognition of the role of weekend activity-travel patterns in travel demand analysis and greenhouse gas emissions, there are virtually no research efforts focused on examining factors influencing discretionary special event activity participation, duration, and accompaniment type for the event (Jago, 1997). The importance of modeling activity duration and accompaniment type dimensions has been noted in the literature (e.g., Kato and Matsumoto, 2009; Bhat, 1996). Activity durations are directly related to vehicle soak times for automobile users - an important variable for emissions estimation. The accompaniment type strongly influences travel mode choice (and vehicle type choice for auto trips), choice dimensions with important energy and environmental implications.

This study bridges the gap in the literature by modeling (1) discretionary event type participation, (2) duration of participation, and (3) accompaniment type (i.e., with whom the individual participates in the special event). The study formulates a joint discrete-continuous framework for analyzing these dimensions. Specifically, the combination of event type and accompaniment type constitutes the discrete component while the duration of participation constitutes the continuous component. The discrete framework is quite appropriate in the current context because only a very small fraction of individuals pursue multiple discretionary special events within the same weekend day.

In terms of the continuous choice, the duration of the event for a particular discrete alternative is observed only if participation has occurred, necessitating the consideration of self-selection in the model structure. A copula-based sample selection approach that ties the discrete choice error component with the duration error component in a flexible manner is adopted in this study to account for this. In particular, the copula approach allows the testing of several flexible dependence structures rather than pre-imposing restrictive distributional assumptions (see

Bhat and Eluru, 2009 for an extensive discussion of the copula approach). The methodological details of the copula-based discrete-continuous model are presented in the next section.

3. METHODOLOGY
In this section, the structure of the copula-based joint generalized extreme value duration framework to jointly model weekend event type choice, accompaniment type choice, and activity duration is presented. The weekend event type and accompaniment type choice component takes the form of a generalized extreme value model while the event duration component takes the form of a non-parametric proportional hazard duration model. In addition to allowing a flexible parametric hazard distribution, the non-parametric duration model is ideally suited for hazard model estimation in the context of interval-level reporting of duration (where respondents round their activity and travel start and end times to the nearest five or 10 minutes when responding to travel or time use surveys).

3.1. Model Structure
The modeling of event type and company type choice is undertaken using a generalized extreme value framework. Specifically, let q be the index for individuals and let i be the index for the discrete choice combination of event type and company type. Also, let h_{qi}^* be the latent (indirect) utility accrued by individual q from participation in alternative i. Without any loss of generality, let $i=1$ represent the discrete category that corresponds to non-participation in any out-of-home event type. Then, write:

$$h_{qi}^* = \beta' x_{qi} + \varepsilon_{qi}, \tag{1}$$

where x_{qi} is a vector of independent variables, β is a corresponding vector of coefficients to be estimated, and ε_{qi} represents an idiosyncratic error term. Assume that the ε_{qi} terms are identically and independently Gumbel distributed across outcomes i and individuals q with a location parameter equal to 0 and a scale parameter equal to 1.[1] Individual q participates in alternative i if, and only if,

$$h_{qi}^* > \max_{\substack{j=1,2,\ldots I \\ j \neq 1}} h_{qj}^* \tag{2}$$

Let r_{qi} be a dichotomous variable; $r_{qi} = 1$ if the ith alternative is chosen by the qth individual, and $r_{qi} = 0$ otherwise. Defining

$$v_{qi} = \varepsilon_{qi} - \left\{ \max_{j=1,2,\ldots I,\ j \neq 1} h_{qj}^* \right\}, \tag{3}$$

and substituting the right side for h_{qi}^* from equation (1) in equation (2), it is possible to write:

$$r_{qi} = 1 \text{ if and only if } \beta' x_{qi} + v_{qi} > 0 \tag{4}$$

The implied marginal distribution of v_{qi} can be obtained from equation (3) and from the distributional assumptions on the ε_{qi}'s as follows:

---

[1] The methodology is also applicable to generalized GEV structures for the error terms. In fact, this is what we have done in our analysis and code. However, as we will discuss later, in the empirical context of the current paper, the discrete choice model consistently collapsed to the multinomial logit model for a whole range of variable specifications. Thus, to keep the presentation simple in this methodology section, we consider the error terms to be identically and independently distributed. Readers interested in getting the details of the more general methodology may obtain this by contacting the corresponding author.

$$F_i(v) = \text{Prob } (v_{qi} < v) = \frac{\sum\limits_{j \neq i} \exp(\beta' z_{qj})}{\exp(-v) + \sum\limits_{j \neq i} \exp(\beta' z_{qj})} \ , \quad i = 1,2,\ldots,I. \tag{5}$$

Next, let s_{qi} represent the (continuous) duration time of participation in alternative i ($i=2,3,\ldots, I$) for individual q. The hazard function for individual q's duration in alternative i at some specified time t on the continuous-time scale T, $\lambda_{qi}(t)$, is defined using the proportional hazard specification as (Kiefer, 1988):

$$\lambda_{qi}(t) = \lim_{\delta \to 0^+} \frac{\text{prob}[t + \delta > s_{qi} \geq t \,|\, s_{qi} \geq a]}{\delta} = \lambda_{0i}(t) \exp(-\gamma' z_{qi}), \tag{6}$$

where $\lambda_{0i}(t)$ is the continuous-time baseline hazard at time t for special event type i, z_{qi} is a column vector of covariates for individual q and alternative i, and γ is a column vector of parameters to be estimated.[2] Equation (6) can be written in the equivalent form (Bhat, 1996),

$$s_{qi}^* = \ln \Lambda_{0i}(s_{qi}) = \gamma' z_{qi} + \eta_{qi} , \tag{7}$$

where $\Lambda_{0i}(.)$ is the integrated baseline hazard for alternative i ($i=2,3,\ldots, I$) and η_{qi} takes an extreme value form with distribution function given by:[3]

$$\Pr(\eta_{qi} < \eta) = G(\eta) = 1 - \exp[-\exp(\eta)] \tag{8}$$

The dependent variable in equation (7) is a continuous unobserved variable when duration data are available only in grouped form. However, the time interval of activity participation, t_{qi}, for individual q's chosen alternative i ($i=2,3,\ldots I$) is observed. Let the time intervals be represented by an index k ($k = 1,2,3,\ldots K$), with these time intervals being related to the continuous time dimension T in terms of bounds on the continuous time scale: $k = 1$ if $T \in [0,T_1]$, $k = 2$ if $T \in [T_1,T_2],\ldots, k = K$ if $T \in [T_{K-1},\infty]$ (the duration interval cut off points can be different for different alternatives; however, for ease in notation, they are assumed to be the same for all alternatives). Thus, $t_{qi} = k$ if the duration spell of individual q ends in time interval k for the chosen alternative i ($i=2,3,\ldots I$).

The overall equation system for the joint choice of discretionary event-accompaniment type discrete choice combination and duration can be written from equations (1) through (7) as:

$$r_{qi}^* = \beta' x_{qi} + v_{qi} , \ r_{qi} = 1 \text{ if } r_{qi}^* > 0 , \ r_{qi} = 0 \text{ otherwise}$$

$$s_{qi}^* = \gamma' z_{qi} + \eta_{qi}, \ t_{qi} = k \ \text{ if } \delta_{i,k-1} < s_{qi}^* \leq \delta_{i,k} ,$$

$$t_{qi} \text{ observed only if } r_{qi} = 1 \ (i = 2,3,\ldots I) , \delta_{i,k} = \ln \Lambda_{0i}(T^k) \tag{9}$$

In the above equation system, $\delta_{i,0} = -\infty$ and $\delta_{i,K} = +\infty$ for each alternative i ($i=2,3,\ldots I$). The probability that an individual q will choose alternative i ($i=2,3,\ldots I$) and participate for a grouped duration interval of k may be written as follows:

$$\Pr[r_{qi} = 1, t_{qi} = k] = \Pr[v_{qi} > -\beta' x_{qi}, \delta_{i,k-1} - \gamma' z_{qi} < \eta_{qi} < \delta_{i,k} - \gamma' z_{qi}]$$

$$= \Pr[\delta_{i,k-1} - \gamma' z_{qi} < \eta_{qi} < \delta_{i,k} - \gamma' z_{qi}] - \Pr[v_{qi} < -\beta' x_{qi}, \delta_{i,k-1} - \gamma' z_{qi} < \eta_{qi} < \delta_{i,k} - \gamma' z_{qi}]$$

$$= G(\delta_{i,k} - \gamma' z_{qi}) - G(\delta_{i,k-1} - \gamma' z_{qi}) -$$

$$\left(\Pr[v_{qi} < -\beta' x_{qi}, \eta_{qi} < \delta_{i,k} - \gamma' z_{qi}] - \Pr[v_{qi} < -\beta' x_{qi}, \eta_{qi} < \delta_{i,k-1} - \gamma' z_{qi}] \right) \tag{10}$$

---

[2] $\lambda_{0i}(t)$ may be viewed as the instantaneous probability that individual q's duration in event type i will be terminated in an infinitesimal time interval beyond time point t, given that the duration did not terminate before time point t.

[3] The distribution form is the same across event type regimes i ($i=2,3,\ldots I$), and hence the function G is not subscripted by the index i in the equation below.

The above probability depends upon the dependence structure between the random variables v_{qi} and η_{qi} for each special event type i. In the duration context of interest in this paper, Bhat (1998) proposed the use of Lee's sample selection method (Lee, 1983) for this coupling, which essentially corresponds to the use of a Gaussian copula (Bhat and Eluru, 2009). However, this approach imposes the restriction that the dependence surface is linear and symmetric. There is no reason to believe that such a Gaussian coupling is appropriate. In particular, it is important to test different parametric functional forms for this bivariate dependency surface and pick the one that empirically fits the data best. Recent developments related to copula techniques provide an effective way to do so.

3.2. General Bivariate Copula Structure

A copula is a device or function that generates a stochastic dependence relationship (i.e., a multivariate distribution) among random variables with pre-specified marginal distributions (see Bhat and Eluru, 2009; Trivedi and Zimmer, 2007). The precise definition of a copula is that it is a multivariate distribution function defined over the unit cube linking uniformly distributed marginals. In the bivariate case, let C be a *2*-dimensional copula of uniformly distributed random variables U_1 and U_2 with support contained in [0,1]. Then,

$$C_\theta(u_1, u_2) = Pr(U_1 < u_1, U_2 < u_2), \tag{11}$$

where θ is a parameter of the copula commonly referred to as the dependence parameter. A copula, once developed, allows the generation of joint bivariate distribution functions with given marginals. Thus, a bivariate distribution $J_i(v,\eta)$ can be generated for the two random variables v_{qi} (with margin F_i) and η_{qi} (with margin G) using the following expression (see Sklar, 1973):

$$J_i(v,\eta) = \Pr(v_{qi} < v, \eta_{qi} < \eta) = \Pr[U_1 < F_i(v), U_2 < G(\eta)] = C_\theta[u_1 = F_i(v), u_2 = G(\eta)] \tag{12}$$

A rich set of bivariate copulas $C_\theta(u_1, u_2)$ are available, including the Gaussian copula, the Farlie-Gumbel-Morgenstern (FGM) copula, and the Archimedean class of copulas (including the Clayton, Gumbel, Frank, and Joe copulas). For given functional forms of the margins, the precise bivariate dependence profile between the variables v_{qi} and η_{qi} is a function of the copula $C_\theta(u_1, u_2)$ used, and the dependence parameter θ. For detailed descriptions of the alternative copulas and their formulations, please see Bhat and Eluru (2009).

3.3. Estimation Procedure

The parameters to be estimated in the joint discrete-duration model include the β vector, the $(K - 1)$ $\delta_{i,k}$ parameters ($\delta_{i,0} = -\infty$ and $\delta_{i,K} = +\infty$) for each event type i (i=2,3,...I), and the vector γ. From equation (1), the probability of an individual choosing not to participate in any event type is simply the generalized extreme value probability expression; for example if the ε_{qi} error terms are assumed to be independent across discrete alternatives then the probability is given below:

$$\text{Prob}\,(r_{q1} = 1) = \frac{\exp(\beta' z_{q1})}{\sum_j \exp(\beta' z_{qj})}, \quad j = 1,2,...,I. \tag{13}$$

If some other GEV form is used for the ε_{qi} terms, then the implied distribution of v_{qi} will take the corresponding GEV probability form. In this paper, the multinomial logit and nested logit structures were examined. The probability of an individual choosing discrete choice combination i ($i=2,3,\ldots I$) and a duration interval k may be obtained from equation (10) and the appropriate copula expression as:

$$\text{Prob}[r_{qi} = 1, t_{qi} = k] = G(\delta_{i,k} - \gamma' z_{qi}) - G(\delta_{i,k-1} - \gamma' z_{qi}) - \left[C_\theta(u_{qi1}, u_{qi,k,2}) - C_\theta(u_{qi1}, u_{qi,k-1,2}) \right] \quad (14)$$

where $u_{qi1} = F_i(-\beta' x_{qi})$ and $u_{qi,k,2} = G(\delta_k - \gamma' z_{qi})$.

The corresponding probability for the original Lee's transformation will be slightly different and is given by:

$$\text{Prob}[r_{qi} = 1, t_{qi} = k] = \left[C_\theta(u_{qi1}, u_{qi,k,2}) - C_\theta(u_{qi1}, u_{qi,k-1,2}) \right] \quad (15)$$

where $u_{qi1} = F_i(\beta' x_{qi})$ and $u_{qi,k,2} = G(\delta_k - \gamma' z_{qi})$.

Next, let $1[.]$ be an indicator function taking the value of unity if the expression in parenthesis is true and 0 otherwise. Also, define the following dummy variables for $i=2,3,\ldots I$:

$$M_{qik} = 1[r_{qi} = 1] \times 1[t_{qi} = k]. \quad (16)$$

Then, the log likelihood function for the copula model takes the form

$$\log L = \sum_{q=1}^{Q} \left(r_{q1} \log[\text{Pr} ob(r_{q1} = 1)] + \sum_{i=2}^{I} \sum_{k=1}^{K} M_{qki} \log \left[\text{Pr} ob(r_{qi} = 1, t_{qi} = k) \right] \right) \quad (17)$$

All of the parameters in the model are consistently estimated by maximizing the log-likelihood function, which is accomplished using the GAUSS matrix programming language. Once the threshold parameters ($\delta_{i,k}$ parameters) are estimated for each alternative, the baseline hazard for discrete period k and alternative i ($i=1,2,\ldots I$), $\lambda_{0i}^*(k)$, can be computed using the expression below:

$$\lambda_{0i}^*(k) = \frac{G(\delta_{i,k}) - G(\delta_{i,k-1})}{1 - G(\delta_{i,k-1})} \quad (18)$$

4. DATA

The data used in this study is derived from the 2008-2009 National Household Travel Survey (NHTS) in the United States. The National Household Travel Survey (NHTS) is a survey of a random sample of households in the country that is conducted on a periodic basis by the US Department of Transportation to obtain detailed socio-economic, demographic, and personal travel data for the nation. In the 2008-2009 edition of the NHTS, individual jurisdictions including states and metropolitan areas were allowed to purchase additional samples for their respective areas to aid in model development and transportation planning activities at the local level. The Maricopa Association of Governments (MAG), the metropolitan planning organization for the Greater Phoenix metropolitan area, purchased such additional samples. The data set used for this study corresponds to survey records from the Greater Phoenix metropolitan area. The MAG add-on sample included more than 4,400 households who provided detailed socio-economic, demographic, and personal travel information (for a 24 hour travel diary period). The households were evenly distributed across the days of the week, such that about 2/7th of the 4,400 households responded to the survey and reported their trips for a weekend day (Saturday or Sunday).

The analysis in this paper is restricted to activity-travel participation on weekends for individuals 14 years of age or older. Four different discretionary event activity types are considered:

- Social/recreational activity
- Visit friends/family
- Go out/hang out
- Visit public place

As an individual may choose to participate in any combination of these activities over the course of a weekend day, discrete alternatives are constructed by considering all possible combinations of activity engagement. This leads to a total of 2^4-1=15 alternatives. In addition to these activity alternatives, three accompaniment options are defined:

- Alone
- With household members only
- With non-household members (may or may not include household members)

By crossing the 15 activity type participation alternatives with the three accompaniment type options, a total of 45 alternatives are obtained for the discrete component of the model system. Thus an individual may participate in multiple discretionary activity events, but cannot mix and match accompaniment types across the discretionary events. For example, a person who pursues social recreation alone and visits a public place alone would be included (because that person's choice appears within the 45 elemental choice set). However a person who pursues social recreation alone and visits a public place with household members would not be included in the analysis. This definition of the choice set is not too restrictive because the percent of individuals (responding to the survey on a weekend day) who engaged in multiple discretionary activity events with differing accompaniment types is extremely small at just over two percent of the sample. Thus, the choice set definition adopted in this study accounts for nearly 98 percent of the weekend survey respondents. The 46th and final alternative is the choice to not engage in any discretionary activity event.

After extensive cleaning and data preparation, the final sample used for model estimation included 1773 individuals, 481 of whom made at least one discretionary event on the weekend day. The other individuals engaged in absolutely no special discretionary activity event (they may have done other activities, but they did not engage in the four specific activities of interest in this paper). Table 1 offers a detailed description of the discretionary activity participation and duration for the survey sample. It is found that 9.4 percent pursue social/recreational activities alone, 28.9 percent pursue only "go out/hang out", 8.1 percent pursue only visit public place, 49.1 percent pursue only visit friends/family, and 4.6 percent pursue two discretionary activity events (but with the same accompaniment type choice). No individual pursues more than two discretionary special events on the same weekend day in this data set. It is found that 34.1 percent of the activities are undertaken alone, 20.8 percent involve non-household members, and 45.1 percent involve exclusively household members. This suggests that a large proportion of weekend discretionary activities are pursued only with household members. The go out/hang out activity exhibits the longest average duration at about three hours, followed by social/recreational activity at about 2.5 hours. The shortest duration is exhibited by visit public place at just about 1.5 hours. The table also furnishes the number of observations that fall into different duration bands and the number of individuals pursuing the different discretionary activity event alternatives.

Overall, the data set is quite suitable for modeling weekend discretionary activity participation, time allocation, and accompaniment type. The socio-economic and demographic profile of the respondent sample does not present any unusual anomalies. In the interest of brevity, a detailed tabulation and description of the survey sample is not provided here. Suffice it to say that the sample was thoroughly checked to ensure that it was suitable for use in this study.

Table 1. Descriptive Analysis of Weekend Discretionary Activity Participation

Special Event Activity Participation Rate and Average Duration (minutes)				
Special Event Type	**Accompaniment**			
	Alone	**With family**	**With non-family**	**Across Accompaniment Type**
	Avg. Duration (%)	Avg. Duration (%)	Avg. Duration (%)	Avg. Duration (%)
Social / recreational	160.85 (4.2%)	150.22 (3.7%)	130.71 (1.5%)	151.91 (9.4%)
Go out / hang out	189.61 (7.9%)	181.89 (13.1%)	171.24 (7.9%)	181.09 (28.9%)
Visit public place	102.00 (1.5%)	76.75 (5.0%)	115.00 (1.7%)	89.13 (8.1%)
Visit friends / families	147.89 (18.5%)	123.93 (22.2%)	159.23 (8.3%)	138.95 (49.1%)
Aggregation of multiple special events	212.50 (2.1%)	264.60 (1.0%)	297.00 (1.5%)	251.23 (4.6%)
Across Event Type	161.12 (34.1%)	140.96 (45.1%)	167.90 (20.8%)	153.43 (100.0%)
Discrete Choice Component				
Attributes		**Accompaniment**		
		Alone	**With household members**	**With non-household members**
Sample Size		1773		
No Special Event		1292		
Special Events (Frequency)	Social/recreational	26	20	9
	Go out/hang out	44	66	40
	Visit Public Place	8	26	12
	Visit Friend/family	96	110	46
Duration Component				
Sample Size		481		
Duration Categories (Frequency)	≤30 minutes	70		
	≤60 minutes	45		
	≤90 minutes	54		
	≤120 minutes	43		
	≤150 minutes	61		
	≤180 minutes	56		
	≤240 minutes	68		
	≤360 minutes	55		
	>360 minutes	29		

5. MODEL ESTIMATION RESULTS

This section presents the model estimation results. Several categories of exogenous variables were considered in the model specification including (1) individual characteristics, (2) household characteristics, and (3) location attributes. Individual characteristics considered include gender, age, and employment status. The household characteristics considered were household size, presence and number of children, number of adults, household composition, residential tenure, and vehicle ownership. In terms of location attributes, residential location (categorized as urban and rural) was considered.

As discussed earlier, the choice context being examined consists of 46 alternatives. Estimating all potential exogenous variable effects for all of the alternatives would result in a cumbersome and likely inefficient model specification. Hence, in this paper, variable effects are considered across the event type and accompaniment type dimensions. This allows capturing a majority of the exogenous variable impacts while retaining a parsimonious specification. The overall variable specification process was also guided by behavioral intuitiveness and prior research.

5.1. Model Specification

The empirical analysis involved estimating six different copula-based structures (Gaussian, FGM, Frank, Gumbel, Clayton, and Joe) as well as the independent structure. For asymmetric copulas the traditional Lee's transform based dependency and the non-traditional dependency were also estimated. The Bayesian Information Criterion (BIC) test was employed to determine the preferred model structure as the different model structures are not nested within one another. The BIC for a given empirical model is equal to $-2\ln(L) + K \ln(Q)$, where $\ln(L)$ is the log-likelihood value at convergence, K is the number of parameters, and Q is the number of observations. The model with the _lower_ BIC value is the preferred model. For the different copula models with the same number of parameters, the BIC effectively translates into a comparison of the log-likelihood. The Joe copula with the traditional dependency offered the superior fit among the various models considered. The log-likelihood value at convergence for the Joe copula-based model was found to be -3159.65. The likelihood value at convergence for the independent model structure was -3161.86. The BIC values for the Clayton copula-based model and the independent model were 6558.67 and 6563.09 respectively. Clearly, the BIC values confirm that the copula-based model offers statistically superior fit compared to the independent model. The log-likelihood and BIC values for the Gaussian copula were -3161.21 and 6561.79 respectively, clearly highlighting the increased flexibility offered by the copula approach. In addition to the Joe and Gaussian copulas, the Clayton and Gumbel copula model with traditional dependency also outperformed the independent model. In the interest of brevity, only a discussion of results for the Joe copula model is presented here.

5.2. Discrete Component

For the discrete component of the model (joint consideration of the type of discretionary activity pursued and the accompaniment type chosen), several GEV structures and the multinomial logit specifications were estimated. The GEV models were estimated with correlation across the accompaniment type dimension, i.e., alternatives of the same accompaniment type were postulated as exhibiting stronger substitution patterns within distinct activity types. Alternative model specifications were tested for nested dependencies across the discrete choice combination alternatives. After extensive testing, it was found that the GEV structures we tested were not supported in this empirical model estimation exercise; the model repeatedly collapsed to the

standard multinomial logit specification. The remainder of the discussion in this paper is based on the multinomial logit model specification for the discrete component of the model.

Estimation results are furnished in Table 2 for the MNL discrete model component. Among individual attributes, it is found that males are less likely to pursue social/recreational activities. Teenagers in the age group of 14-19 years are more likely to pursue "go out/hang out", a finding that is consistent with expectations. Among household demographics, several behaviorally intuitive findings can be discerned. Car ownership is positively associated with "go out/hang out"; on the other hand, the presence of children and the presence of multiple adults have a negative impact on the pursuit of "go out/hang out" activities. This finding is consistent with expectations in that individuals in families with children and multiple adults are less likely to "hang out". On the other hand, individuals in households with multiple adults are more likely to visit a public place. The presence of children is positively associated with joint activity engagement, whether with household members alone or including non-household members as well, an indication that children spur joint activities in the family and beyond. Individuals who live in households that own their residence are likely to be in family households; as a consequence residential tenure (own) is positively associated with family member accompaniment choice. There is a weak, albeit positive relationship between number of workers in the household and pursuit of social/recreational activities. Single adults (with no children) are more likely to pursue discretionary activity events with non-household members (relative to the base alternative "alone"), a sign of the need for social contact. Couples and single parents are more likely to pursue discretionary activity events on weekends with family members. Couples are also less likely to engage in go out/hang out as evidenced by the negative coefficient associated with the two adult-no children family composition in this activity type choice.

Household income has a positive, albeit statistically insignificant impact on go out/hang out. Income has a negative impact on visit friend/family activity type choice and on pursuing discretionary activity events with non-family members. It is likely that high income households are also multi-adult families with or without children. Individuals in such households are likely to pursue activities jointly with other household members as opposed to non-family members. It is found that individuals living in urban areatype are more likely to pursue social/recreational, go out/hang out, and visit friend/family activities. This is consistent with the notion that access to such activity destinations is likely to be superior in urban environments. Individuals residing in urban environments appear less inclined to pursue activities with family members, presumably because urban dwellers are young and single.

5.3. Continuous Duration Component

Results of the continuous duration model component are presented in Table 3. The threshold parameters (for the group duration interval categories) do not have a behavioral interpretation, but simply serve as cutoff points to map the grouped intervals to the continuous underlying latent duration propensity measure which is unobserved. It should be noted that a single set of thresholds for all 15 discretionary event participation alternatives was estimated in this effort (instead of estimating 15x8=120 parameters). This restriction provided an efficient and parsimonious model specification. However, to account for possible differences among the different combinations, deviations or shifts in thresholds were estimated for selected activity types and accompaniment type dimensions. These shift coefficients account for differences in thresholds across different alternative combinations.

Table 2. Joe Copula Model Estimates: MNL Component of Activity Participation Choice

Variables	Special Event Type Dimension								Company Type Dimension			
	Social/ recreational		Go out/hang out		Visit Public Place		Visit Friend /family		With Family Member		With non-family member	
	Coef	t-stat	Coef	t-stat	Coeff.	t-stat	Coef	t-stat	Coef	t-stat	Coef	t-stat
Individual Characteristics												
Gender												
Male	-0.832	-2.590	-	-	-	-	-	-	-	-	-	-
Age												
14-19	-	-	0.999	2.875	-	-	-	-	-	-	-	-
Household Demographics												
Number of Cars	-	-	0.214	2.065	-	-	-	-	-	-	-	-
Number of Children	-	-	-0.163	-1.663	-	-	-	-	0.178	2.043	0.370	4.870
Number of Adults	-	-	-0.274	-1.951	0.774	2.699	-	-	-	-	-	-
Residential Tenure												
Own residence	-	-	-	-	-	-	-	-	0.696	1.885	-	-
Number of worker	0.177	1.318	-	-	-	-	-	-	-0.379	-3.357	-	-
Family composition												
One adult and no children	-	-	-	-	-	-	-	-	-	-	0.738	2.019
Two adult and no children	-	-	-0.439	-1.635	-	-	-	-	0.487	2.094	-	-
Single parent	-	-	-	-	-	-	-	-	0.901	1.788	-	-
Number of driver	-	-	-	-	-0.635	-1.693	-	-	0.274	1.859	-	-
Household income	-	-	0.017	1.100	-	-	-0.016	-1.551	-	-	-0.024	-1.530
Location Attributes												
Urban	1.179	2.215	0.852	2.518	-	-	0.538	2.035	-0.625	-2.173	-	-

Table 3. Joe Copula Model Estimates: Activity Duration Component

Variables	Coefficient	t-stat
Threshold parameters		
30 minutes	-1.760	-7.188
60 minutes	-1.005	-4.318
90 minutes	-0.546	-2.381
120 minutes	-0.166	-0.727
150 minutes	0.201	0.865
180 minutes	0.576	2.420
240 minutes	1.114	4.658
360 minutes	1.707	7.731
Threshold Shifts		
Event type dimension		
Go out/hang out	0.856	4.811
Visit Friend/family	0.444	3.189
Individual Characteristics		
Age		
26 – 40	-0.351	-2.003
Employment status		
Employed	0.187	1.676
Household Demographics		
Number of cars	0.159	2.449
Number of adults	-0.129	-1.572
Household income	-0.035	-3.936
Copula Parameter (same for all alternatives)	1.059	40.653

The model estimation effort showed that the choice alternatives of "go out/hang out" and "visit friends/family" have thresholds that are different from the common set of thresholds presented in the top section of Table 3. These deviations are found to be positive and significant implying that alternative combinations that involve these activity type categories are likely to have longer durations than alternative combinations that do not involve these activity type alternatives.

An examination of the impact of exogenous variables suggests that those aged 26-40 years pursue discretionary activity events for shorter durations relative to other age groups. It is likely that this young-middle age group consists of busy individuals, possibly with young children in the household. As a result, their activity participation duration is shorter. Employed individuals (who presumably work primarily on weekdays) spend longer duration at discretionary activity events. Car ownership is positively associated with activity duration; with higher levels of car ownership, vehicle use constraints are eliminated. If there is no competition for the use of household vehicles, then individuals can afford to spend longer duration at activities. Household income has a negative impact on activity duration, possibly suggesting that higher income individuals pursue other types of activities (such as shopping and eat meal), thus leaving less time for discretionary activity duration.

5.4. Copula Parameter

The copula parameter for the Joe-copula model with the traditional dependency structure highlights the presence of significant unobserved component that affects the discretionary activity event participation choice as well as the activity duration. The significant traditional dependency parameter indicates that the correlation between the random error term in the discrete choice combination and the duration choice exhibits a negative relationship. In summary, the result clearly underscores the importance of accommodating unobserved heterogeneity in modeling the special event participation.

5.5. Elasticity Estimates

Based on the model parameter estimates, it is possible to compute elasticities of activity engagement and time allocation. The elasticity values are furnished in Tables 4 and 5 for the entire survey sample. The elasticities are computed across the activity type dimensions rather than across all 46 alternatives in the choice set. Thus, elasticity estimates are generated for the five broad activity type choice categories.

The elasticity estimates in the MNL component (Table 4) reflect the percent change in the likelihood of pursuing an activity type for a unit (percent) change in the exogenous variable. It is found that social/recreation activity participation is highly influenced by urban residential location choice. Those living in urban areas are 61.8 percent more likely to pursue social/recreational activity, and about 40 percent more likely to pursue "go out/hang out" activity, than those not living in urban environments. Equivalently, those living in non-urban areas are much more unlikely than urban dwellers to participate in discretionary activity, which has important implications for social exclusion considerations (see next section). Similarly, those in the age group of 14-19 years old are much more likely to engage in go out/hang out in comparison to other age groups. On the other hand, males are much less likely to engage in social/recreational activities. A unit increase in the number of adults in the household greatly increases the propensity to visit a public place. Those residing in an urban environment are more likely to go out/hang out, and less likely to visit a public place. All of these findings are consistent with expectations.

Elasticity estimates in Table 5 pertain to those for the continuous duration component. A positive elasticity implies an increased likelihood of falling into a certain duration interval and a negative elasticity implies just the opposite. Consistent with model estimation results, it is found that those aged 26-40 years, in households with more adults, and in high income households are more likely to choose shorter durations. Those who are employed and reside in households with more vehicles are likely to allocate larger durations to their weekend discretionary activity episodes.

6. CONCLUSIONS

Despite the growing recognition of the importance of weekend activity-travel demand, there have been a limited number of studies devoted to the study of such behavior. The current study augments existing literature on weekend activity-travel pattern analysis by examining individual participation and time allocation to discretionary activity events during the weekend days. Specifically, the current study develops a comprehensive framework to model (1) event type participation, (2) duration of participation and (3) accompaniment type (i.e., with whom the individual participates in the special event).

The study formulates a joint discrete-continuous framework for analyzing these dimensions. Specifically, the combination of event type and accompaniment type constitutes the discrete component while the duration of participation constitutes the continuous component. As a non-zero duration is observed for a particular discrete alternative only if a participation has

occurred, it is critical to consider self-selection in the model structure. This is accomplished in this study using a copula-based sample selection approach that ties the discrete choice error component with the duration error component in a flexible manner. The copula approach allows the testing of several flexible dependence structures rather than pre-imposing restrictive distributional assumptions.

An important observation from our results is that non-urban dwellers participate very little in social/recreational activity and in "go out/hang out" activity. That is, non-urban dwellers may be facing social exclusion issues (broadly defined as the "inability to participate fully in society", one aspect of which is not being able to participate in the "normal activities of daily life"; see Farber, et al, 2011). This is a rather worrisome issue, because it is well established in the sociological literature that out-of-home discretionary and rejuvenating relaxation activities are important for human mental and physical well-being (see, for example, Voorpostel, et al, 2010). The implication is that there is a need to examine the reasons for this relative lack of participation in discretionary activities among non-urban dwellers, and address this situation. This is particularly important as the rural population in many areas of the United States is increasing, as the so-called baby-boomers look for retirement places and appear to be transplanting themselves in rural areas in the south, especially at the fringes of metropolitan areas (see LaMondia and Bhat, 2010). The result is that not only is the population of rural areas in some parts of the US increasing, but it is also getting increasingly gentrified with retired people. Such individuals, in general, are time-rich but also physically less mobile than their younger peers, which underscores the need to design socially inclusive communities in non-urban areas.

The primary source of the data used in the paper is the 2008-2009 National Household Travel Survey for the Greater Phoenix metropolitan region in Arizona, comprising a sample of 1773 individuals who reported their activity-travel patterns for one weekend day. The empirical analysis involved estimating six different copula-based structures (Gaussian, FGM, Frank, Gumbel, Clayton, and Joe) as well as the independent structure with the traditional dependency structure (Lee's form) and a non-traditional dependency structure. The Bayesian Information Criterion test is employed to determine the preferred model as the different model structures are not nested within one another.

The Joe copula model offered the superior fit among the various copula models considered. The model results clearly highlight the presence of significant self-selection effects of activity participation decisions on activity duration. A host of exogenous variables including individual characteristics, household characteristics, and location area type were considered in the model specification. The results highlight the significance of these variables in discretionary event participation, accompaniment type, and activity duration. The most significant variables influencing the choice of participation and duration include age and gender at the person level, and number of adults, vehicle ownership, presence of children, and family composition at the household level. The copula parameter is significant suggesting that the prediction of activity durations will be adversely affected if the potential endogeneity across the discrete choice participation alternative and the duration are neglected. Further, the results also highlight the flexibility of the copula approach in capturing various forms of dependencies in the decision process. Future work in this domain should explore the use of more complex generalized extreme value structures for representing the dependencies across the discrete and continuous choice dimensions.

Table 4. Elasticity Estimates: MNL Activity Type Choice Component

Variables	No special event	Social/recreational	Go out/hang out	Visit Public Place	Visit Friends/family
Individual Characteristics					
Gender					
Male	2.309	-74.463	1.384	1.346	1.412
Age					
14-19	-11.966	-7.058	128.656	-6.631	-7.076
Household Demographics					
Number of Cars	-1.947	-1.162	20.950	-1.074	-1.158
Number of Children	-3.212	13.223	-3.238	13.326	13.344
Number of Adults	-0.972	-0.515	-23.345	111.099	-0.626
Residential Tenure					
Own residence	-6.732	17.287	17.725	18.526	18.097
Number of worker	3.445	6.472	-10.794	-11.160	-10.940
Family composition					
One adult and no children	-5.332	14.676	13.820	14.362	14.361
Two adult and no children	-3.053	19.330	-21.568	20.340	20.203
Single parent	-14.733	37.918	38.930	40.539	39.508
Number of driver	6.160	-14.366	-14.778	-35.405	-14.596
Household income	0.207	-0.318	1.291	-0.298	-1.749
Location Attributes					
Urban	-8.245	61.807	40.840	-36.951	16.193

Table 5. Elasticity Estimates: Activity Duration Component

Variables	Duration								
	30 minutes	60 minutes	90 minutes	120 minutes	150 minutes	180 minutes	240 minutes	360 minutes	>360 minutes
Individual Characteristics									
Age									
26 – 40	32.364	20.757	12.568	6.068	0.142	-5.529	-11.754	-18.059	-25.890
Employment status									
Employed	-15.779	-10.889	-7.210	-4.113	-1.125	1.910	5.485	9.408	14.890
Household Demographics									
Number of cars	-12.696	-9.206	-6.418	-3.966	-1.509	1.078	4.252	7.892	13.308
Number of adults	11.368	7.552	4.789	2.535	0.421	-1.665	-4.042	-6.556	-9.882
Household income	2.959	2.027	1.329	0.745	0.185	-0.380	-1.039	-1.753	-2.735

7. REFERENCES

Bhat, C.R. (1996) A hazard-based duration model of shopping activity with nonparametric baseline specification and nonparametric control for unobserved heterogeneity, *Transportation Research*, 30B, pp. 189-207.

Bhat, C.R. (1998) A model of post-home arrival activity participation behavior. *Transportation Research*, 32B(6), pp. 387-400.

Bhat, C.R. and N. Eluru (2009) A copula-based approach to accommodate residential self-selection effects in travel behavior modeling. *Transportation Research*, 43B(7), pp. 749-765.

Bhat, C.R. and R. Gossen (2004) Mixed multinomial logit model analysis of weekend recreational episode type choice. *Transportation Research*, 38B(9), pp. 767-787.

Bhat, C.R. and R. Misra (1999) Discretionary activity time allocation of individuals between in-home and out-of-home and between weekdays and weekends, *Transportation*, 26(2), pp. 193-209,

Bhat, C.R., S. Srinivasan, and S. Sen (2006) A joint model for the perfect and imperfect substitute goods case: application to activity time-use decisions. *Transportation Research*, 40B(10), pp. 827-850.

Farber, S., A. Paez, R.G. Mercado, M. Roorda, and C. Morency (2011) A Time-use Investigation of Shopping Participation in Three Canadian Cities: Is there Evidence of Social Exclusion? *Transportation* 38(1), 17-44.

Habib, K.M.N. (2011) A random utility maximization (RUM) based dynamic activity scheduling model: Application in weekend activity scheduling. *Transportation*, 38(1), pp. 123-151.

Jago L.K. (1997) Special events and tourism behaviour: a conceptualisation and an empirical analysis from a values perspective. Ph.D. Thesis, Department of Hospitality, Tourism and Marketing, Faculty of Business, Victoria University.

Kapur, A. and C.R. Bhat (2007) On modeling weekend adults' daily time use by activity purpose and accompaniment arrangement. *Transportation Research Record*, 2021, pp. 18-27.

Kato, H. and M. Matsumoto (2009) Intra-household interaction in a nuclear family: a utility maximizing approach. *Transportation Research*, 43B(2), pp. 191-203.

Kiefer, N.M. (1988) Economic duration data and hazard functions. *Journal of Economic Literature*, 27, June, pp. 646-679.

LaMondia, J., and C.R. Bhat (2010) Development of a microsimulation analysis tool for paratransit patron accessibility in small and medium communities. *Transportation Research Record,* 2174, 29-38.

Lee Y., S. Washington, and L.D. Frank (2009) Examination of relationships between urban form, household activities, and time allocation in the Atlanta Metropolitan Region. *Transportation Research*, 43A, pp. 360–373.

Lee, L.-F. (1983) Generalized econometric models with selectivity. *Econometrica* 51(2), pp. 507-512.

Lockwood, A., S. Srinivasan, and C.R. Bhat (2005) An exploratory analysis of weekend activity patterns in the San Francisco Bay Area. *Transportation Research Record*, 1926, pp. 70-78.

Sall, E.A. and C.R. Bhat (2007) An analysis of weekend work activity patterns in the San Francisco Bay Area. *Transportation*, 34, pp. 161–175.

Sklar, A. (1973) Random variables, joint distribution functions, and copulas. *Kybernetika*, 9(6), pp. 449-460.

Srinivasan, S. and C.R. Bhat (2008) An exploratory analysis of joint-activity participation characteristics using the American Time Use Survey. *Transportation* 35(3), pp. 301-328.

Sugie, Y., J. Zhang, and A. Fujiwara (2003) A weekend shopping activity participation model dependent on weekday shopping behaviour. *Journal of Retailing and Consumer Services*, 10, pp. 335–343.

Trivedi, P.K. and D.M. Zimmer (2007) Copula modeling: An introduction for practitioners. *Foundations and Trends in Econometrics* 1(1), Now Publishers.

Treuth, M.S., D.J. Catellier, K.H. Schmitz, R.R. Pate, J.P. Elder, R.G. McMurray, R.M. Blew, S. Yang, and L. Webber (2007) Weekend and weekday patterns of physical activity in overweight and normal-weight adolescent girls. *Obesity*, 15(7), 1782-1788.

Voorpostel, M., T. van der Lippe, and J. Gershuny (2010) Spending Time Together—Changes Over Four Decades in Leisure Time Spent with a Spouse. *Journal of Leisure Research* 42(2), 243-265.

Yamamoto, T. and R. Kitamura (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), pp. 211-230.

Zhong, M., J.D. Hunt, and X. Lu (2008) Studying differences of household weekday and weekend activities: a duration perspective. *Transportation Research Record*, 2054, pp.28-36.

Zhong, M. and J.D. Hunt (2010) Exploring best-fit hazard functions and lifetime regression models for urban weekend activities: Case study. *ASCE Journal of Transportation Engineering*, 136(3), 255-266.