## A Continuous Time Representation and Modeling Framework for the Analysis of Nonworker Activity-Travel Patterns: Tour and Episode Attributes

#### Rajul Misra, Chandra R. Bhat, and Sivaramakrishnan Srinivasan

Rajul Misra United Technologies Research Center 411 Silver Lane, MS 129-85 East Hartford, CT 06108 Phone: 860-610-1532, Fax: 860-610-7857, Email: MisraR@utrc.utc.com

Chandra R Bhat and Sivaramakrishnan Srinivasan The University of Texas at Austin, Department of Civil Engineering, 1 University Station C1761, Austin, Texas 78712-0278 Phone: 512-471-4535, Fax: 512-475-8744, Email: bhat@mail.utexas.edu, s.siva@mail.utexas.edu

TRB 2003: Paper # 03-3306

Final Submission for Publication:April 1, 2003Word Count:7,802 (including 2 figures and 5 tables)

Misra, Bhat, and Srinivasan

### ABSTRACT

This paper presents a set of four econometric models to examine the tour and episode-related attributes (specifically mode choice, activity duration, travel times, and location choice) of the activity-travel patterns of non-workers. The paper is a sequel to an earlier work by the authors [see Bhat and Misra (1)], which presented a comprehensive continuous-time framework for representation and analysis of the activity-travel choices of non-workers. That paper also presented detailed descriptions of the first two components of the modeling framework related to the number and sequence of activity episodes. The current paper estimates the proposed models using activity-travel data from the 1990 San Francisco Bay Area travel diary survey.

#### **1. INTRODUCTION**

The last decade has seen the emergence of the activity-based paradigm to travel demand modeling [see Bhat and Koppelman (2), Axhausen and Gärling (3), Kurani and Kitamura (4), and Arentze and Timmermans (5) for detailed discussions of this approach]. The activity-based paradigm, which views travel as a derived demand, overcomes a number of limitations of the previous trip-based paradigm for travel analysis. However, in contrast to the substantial literature on worker activity analysis [for example, see Bhat and Singh (6), Mahmassani *et. al.* (7), Hamed and Mannering (8), Pendyala *et al.* (9)], relatively little research exists that focuses on studying the activity-travel behavior of nonworkers. While some earlier studies [for example, Bowman and Ben-Akiva (10) and Kitamura and Fujii (11)] have developed analysis methods that may be applied to nonworker activity travel analysis, these studies do not model the temporal dimension of activities (except possibly for departure time of trips, categorized broadly into am peak, pm peak, midday and other times) and/or require the *a priori* designation of activities as primary and secondary or fixed and flexible.

Our earlier paper presented the conceptual foundations and structure of a comprehensive representation and analysis framework for nonworker activity-travel patterns that (a) considers all relevant activity travel attributes of the nonworker pattern, (b) includes both the generation and the scheduling of activity episodes, (c) considers time as an all-encompassing continuous entity in analysis, and (d) does not require the a priori designation of activity episodes as fixed or flexible or primary or secondary (1). The framework in our earlier paper represents a nonworker's activity-travel pattern as a series of out-of-home activity episodes (or stops) of different types interspersed with periods of in-home activity stays (the term "stops" is used to refer to out-of-home activity episodes in the rest of this paper; the chain of stops between two inhome activity episodes is referred to as a tour).

The characterization of a nonworker's daily activity travel pattern is accomplished by identifying a number of different attributes within the pattern. The attributes are classified on the basis of the level of representation with which they are associated; that is, whether they are associated with the entire daily pattern, a tour in the day, or an episode. Pattern-level attributes include whether or not the individual makes any stops during the day, the number of stops of each activity type if the individual leaves home during the day, and the sequencing of all episodes (both stops and in-home episodes). The only tour-level attribute is the travel mode for the tour. Episode-level attributes include the episode duration, travel time to the episode from the previous episode, and the location of out-of-home episodes (*i.e.*, stops). The representation system above then forms the basis for development of the analysis framework, which consists of a series of six econometric sub-models (see Figure 1).<sup>1</sup>

In Bhat and Misra (1), we provided detailed mathematical descriptions and empirical estimation results for the first two sub-models in the analysis framework related to the patternlevel attributes in Figure 1. In the current paper, we present the mathematical details and empirical estimation results of the remaining four sub-models in the proposed analysis framework (*i.e.*, the tour- and episode-level attributes in Figure 1).

The rest of this paper is organized into three sections. The next section (Section 2) provides the mathematical description of the four sub-models corresponding to the tour travel mode and the spatial-temporal attributes of activity episodes. Section 3 presents the empirical results obtained using the 1990 San Francisco Bay Area activity-travel survey data [see E.H.

<sup>&</sup>lt;sup>1</sup> Bhat and Misra (I) present conceptual and intuitive considerations supporting their proposed representation and analysis framework. They also discuss the limitations of the proposed framework. These are not replicated here due to space constraints.

White and Associates (12)]. The paper concludes with a summary of the important results and scope for future research.

#### 2. MATHEMATICAL FORMULATION

The analysis framework (see Figure 1) consists of one tour-level model (travel mode) and three episode-level models (morning home stay duration, activity duration/travel time and activity location). The tour-level mode choice model is modeled using a discrete choice formulation. The episode-level attributes include the activity duration of the episode, the travel time to the episode from the previous episode, and the location of each out-of-home episode (*i.e.*, stop). Since the duration of the first home-stay episode is likely to be different from that of other subsequent home-stay episodes because of life-style and sleeping habits, this first home-stay duration is modeled separately using a hazard model. The reader will also note that travel time to this first home-stay episode is undefined since the individual is at home at the beginning of the day. Next, the travel time to the episode from the previous episode and the activity duration of each out-of-home episode. Finally, the spatial location of each out-of-home episode (stop) is modeled using a disaggregate spatial destination choice model.

### 2.1 Travel Mode Choice (TMOD) Sub-Model

Travel mode is considered as a tour-level choice in our analysis framework because almost all tours maintained the same mode for all their trip legs [see Misra and Bhat (13)]. A variety of nested logit models [see Ben-Akiva and Lerman (14)] were tested for the choice of tour mode in our modeling framework, but we found a simple multinomial logit model to be adequate (*i.e.*, the log-sum parameters in the nested logit models were not statistically different from 1). The nested choice formulations tested included:

- Personalized modes (drive alone and non-motorized modes) grouped together in a nest and other modes (shared ride and transit) grouped in a separate nest.
- Private modes (drive alone, shared ride, and non-motorized) grouped together in a nest.
- Motorized modes (drive alone, shared ride, and transit) grouped together in a nest and non-motorized modes (walk and bike) grouped in another nest.

The alternatives in the mode choice model included driving alone, sharing a ride, transit, and non-motorized modes (bike/walk).

#### 2.2 Morning Home-stay Duration (MDUR) Sub-Model

Define a continuous variable  $T_i$  that represents the actual morning home-stay duration of nonworker *i* in the data set (this morning home-stay duration is measured in minutes from 3 a.m.). We consider  $T_i$  to be unobserved, because the observed home-stay durations are integral multiples of five minutes (*e.g.*, 5, 10, 15, 30, 60 minutes, *etc.*), leading to a substantial number of ties at these times. This is because respondents are reporting the timing of their home-stay durations by rounding-off to the nearest five-minute interval. Therefore, the observed home-stay duration data should be treated as interval-level data and a discrete model that retains an interpretation as an incompletely observed continuous-time hazard model should be used. Accordingly, let *u* represent some specified time on the continuous-time scale and let the discrete time interval be represented by an index k (k = 1,2,3,...,K) with k = 1 if  $u \in [0,u^1]$ , k = 2 if  $u \in [u^1, u^2],...,$  and k = K if  $u \in [u^{K-1}, \infty]$ . Let  $h_i$  represent the discrete period of failure for individual *i* (*i.e.*,  $h_i = k$  if the morning home-stay duration of individual *i* ends in discrete period *k*). The hazard function for individual *i* at some specified time *u* on the continuous time scale,  $\lambda_i(u)$ , can now be defined using the proportional hazard specification [see Kiefer (15)] as:

$$\lambda_i(u) = \lambda_0(u) \exp(-\beta' q_i + \zeta_i), \qquad (1)$$

where  $\lambda_0(u)$  is the baseline hazard at time u,  $q_i$  is a column vector of exogenous variables for individual i (not including a constant),  $\beta$  is a column vector of parameters, and  $\zeta_i$  is an unobserved heterogeneity term. The unobserved heterogeneity term takes into account unobserved differences among the morning home-stay durations of observationally equivalent individuals.

Bhat (16) shows that the above equation can be written in an equivalent integral form as follows:

$$\ln \Lambda_0(T_i) = \ln \int_0^{T_i} \lambda_0(u) du = \beta' q_i + \zeta_i + \varepsilon_i,$$
<sup>(2)</sup>

where  $\varepsilon_i$  has a standard Gumbel distribution given as:  $P(\varepsilon_i < s) = G(s) = 1 - \exp[-\exp(s)]$ . Rewriting Equation (2) in terms of the observed discrete time period of failure of individual *i* (*i.e.*,  $h_i = k$  if  $u \in [u^{k-1}, u^k]$ ), the probability that the discrete time period of failure of individual *i* is equal to *k* can be computed as:

$$P[h_{i} = k | \zeta_{i}] = G(\delta_{k} - \beta' q_{i} + \zeta_{i}) - G(\delta_{k-1} - \beta' q_{i} + \zeta_{i})$$
  
= exp[- { $I_{i,k-1} \exp(\zeta_{i})$ }] - exp[- { $I_{i,k} \exp(\zeta_{i})$ }], (3)

where  $\delta_k = \ln \Lambda_0(u^k)$  and  $I_{i,k} = \Lambda_0(u^k) \exp(-\beta' q_i)$ .

The probability above is conditional on the unobserved heterogeneity term  $\zeta_i$ . Let  $v_i$  [=exp( $\zeta_i$ )] be gamma distributed with a mean one (an innocuous identification assumption) and variance  $\sigma^2$ . Then, the unconditional probability of failure of individual *i* in the discrete time period *k* is [see Bhat for derivation (*16*)]:

$$P[t_i = k] = \int_{0}^{\infty} P[h_i = k \mid v_i] f(v_i) dv_i = \left\{ \left[ 1 + \sigma^2 I_{i,k-1} \right]^{-\sigma^{-2}} - \left[ 1 + \sigma^2 I_{i,k} \right]^{-\sigma^{-2}} \right\}$$
(4)

The parameters to be estimated in the morning home-stay duration model are the (*K*-1) integrated hazard elements  $\Lambda_0(u^k)$ 's, the column vector of parameters  $\beta$ , and the variance  $\sigma^2$  of the gamma mixing distribution [the shape of the hazard function can be obtained from the estimates of the integrated hazard elements; see Bhat (16)]. A maximum likelihood approach is used for estimation.

#### 2.3 Episode Duration/ Travel Time (EDUR-TT) Sub-Model

The episode duration and travel time equations for any individual *i* who undertakes an episode of activity j (j = 1, 2, ..., J) can be written as:

$$a_{ij} = \theta'_{j} y_{ij} + \omega_{ij} \text{ for } j = 1, 2, ..., J$$
  

$$t_{ij} = \gamma'_{j} x_{ij} + \eta_{ij} \text{ for } j = 1, 2, ..., J$$
(5)

where  $a_{ij}$  is the natural logarithm of the episode duration of participation in activity type *j* for individual *i*, and  $t_{ij}$  is the natural logarithm of the travel time duration to the episode from the previous episode.  $y_{ij}$  and  $x_{ij}$  are column vectors of exogenous variables, and  $\theta_j$  and  $\gamma_j$  are corresponding parameter vectors (including alternative specific constants).

Next, assume that the stochastic error terms  $\omega_{ij}$  and  $\eta_{ij}$  are distributed identically across all individuals for each activity type *j*. Furthermore, let  $\omega_{ij}$  and  $\eta_{ij}$  have a bivariate cumulative normal density function  $\phi_2(0,0,\sigma_{\omega_j}^2,\sigma_{\eta_j}^2,\rho_{\eta_j\omega_j})$  in each activity type *j*, where  $\sigma_{\omega_j}^2$  and  $\sigma_{\eta_j}^2$  are the variances of the error terms  $\omega_{ij}$  and  $\eta_{ij}$ , respectively, and  $\rho_{\eta_j\omega_j}$  is the correlation between the two error terms. We allow the error terms in the episode and travel time duration equations to be correlated to accommodate unobserved factors that impact these two decisions.

The equation system in (5) is in the form of a simultaneous regression equation system and can be estimated using a full information maximum likelihood method. The parameters to be estimated are the coefficient vectors  $\theta_j$  and  $\gamma_j$  corresponding to the exogenous variables  $y_{ij}$  and  $x_{ij}$ , respectively, variances  $\sigma_{\omega_j}^2$  and  $\sigma_{\eta_j}^2$ , and the covariances  $\rho_{\eta_j\omega_j}$  for each of the *J* activity types.

#### 2.4 Episode Location (ELOC) Choice Sub-Model

The ELOC model provides information on the spatial location at which each out-of-home episode is pursued by a nonworker. The model utilizes the distribution of travel times by the chosen mode to each activity episode (obtained in Section 2.3) to generate a probabilistic choice set of locations. Thus, the modeling system explicitly incorporates the spatial-temporal interactions in stop-making decisions, since episode duration and travel time to the out-of-home episode are jointly determined in the EDUR-TT model.

The (logarithm of) travel time to a stop is estimated as a continuous normally distributed variable in the equation system of (5). The first step in the probabilistic choice set generation model for location choice is to define discrete intervals on the logarithmic travel time scale. Let there be (M+1) discrete time intervals defined on the logarithmic time scale as follows:  $(-\infty, t_1), (t_1, t_2), (t_2, t_3)...(t_{m-1}, t_m)...(t_{M-1}, t_M), (t_M, +\infty)$ .

Consider an individual *i* at a particular zone and let  $t_i$  be the logarithm of travel time to her/his next stop (in the current presentation of the ELOC model, we suppress the index for activity type *j*; the reader will note that the ELOC model is specific to each activity type). Let  $C_{im}$  (m = 1, 2, ..., M) be the set of location zones *z* such that the travel time from the origin zone of the individual to these locations (by the chosen travel mode) falls within the interval  $[t_{m-1}, t_m]$ . By definition, each destination location *z* can belong to one and only one  $C_{im}$ . From the distribution of  $t_i$  determined in the EDUR-TT sub-model, we can write the probability of the choice set  $C_{im}$  as:

$$\pi_{im} = P(C_{im}) = \Phi\left(\frac{t_m - E(t_i)}{\sigma_{\eta_j}}\right) - \Phi\left(\frac{t_{m-1} - E(t_i)}{\sigma_{\eta_j}}\right).$$
(6)

where  $E(t_i)$  is the expected value of the logarithm of travel time duration for individual *i*, and  $\sigma_{\eta_i}$  is the estimated standard error of travel time duration. By construction,  $0 < \pi_{im} < 1 \quad \forall i, m, and \sum_{m} \pi_{im} = 1 \quad \forall i$ .

The conditional probability of choice of a particular location z from a given choice set  $C_{im}$  can be modeled using a multinomial logit formulation:

$$P_{i}(z|C_{im}) = \frac{e^{\alpha w_{iz}}}{\sum_{z'} e^{\alpha' w_{iz'}}} \quad if \ z \in C_{im}, \quad P_{i}(z|C_{im}) = 0 \quad if \ z \notin C_{im}$$
(7)

where  $\alpha$  is the column vector of the parameters to be estimated and  $w_{iz}$  is a column vector of exogenous variables corresponding to destination *z*.

The unconditional probability of choice of destination z is the product of the probabilities given by Equations (6) and (7). Let us now define a binary variable  $b_{iz}$  that takes the value 1 if nonworker *i* chooses location *z* and 0 otherwise. To estimate the model parameters, we maximize the following likelihood function:

$$= \prod_{i} \prod_{m} \left[ P_i(z \mid C_{im}) \pi_{im} \right]^{b_{iz}}$$

$$(8)$$

#### **3. EMPIRICAL RESULTS**

We now present the estimation results of the four sub-models presented in Section 2 for nonworker data obtained from the 1990 San Francisco Bay Area activity-travel survey. The sample comprises 2,048 nonworkers who are not students and who pursued at least one out-of-home activity episode. The total number of tours during the survey day across these 2,048 individuals is 3,095. The total number of episodes, excluding the initial morning in-home stays is 8,156, while the total number of stops is 4,852 [see Bhat and Misra for a detailed description of the survey and the sample formation process (1)].

#### 3.1 Tour Mode Choice (TMOD) Sub-Model

The results of the mode choice model are presented in Table 1. In the category of household composition, the coefficients on household size, the number of children between 12-16 years, and the number of individuals between 17 and 21 years have to be interpreted jointly. The estimates indicate that an increase in the number of small children (<12 years) or adults ( $\geq$  22 years) leads to higher use of ridesharing and non-motorized modes (this is reflected by the coefficient on household size). However, an increase in the number of children between 12-16 years of age increases transit use compared to the drive alone and non-motorized modes, and also reduces ridesharing propensity. The reduction in ridesharing associated with an increase in children between 12-16 years may be a reflection of teenagers wanting to be independent (and not wanting to be seen with their parents!). The results also indicate that an increase in adults between 17-21 years of age in the household leads to increased use of drive alone, perhaps due to the recently acquired ability to drive among individuals in this age group. Finally, in the category of household composition variables, an increase in the number of employed individuals reduces ridesharing of nonworkers in the household.

The household race and structure variables indicate the following: (a) Caucasians are significantly less likely to use transit compared to other races, (b) nonworkers in nuclear family households are more likely to ride share probably because they tend to pursue activities with

their small children; however, the opposite is the case for single parents, perhaps reflecting a tendency to pursue non-work activities alone without children, (c) a nonworker in a single-parent family with at least one child greater than or equal to 22 years is predisposed toward transit use, presumably because of competition for the use of automobiles, and (d) nonworkers living alone are most likely to drive alone or bicycle/walk to activities, while nonworkers in roommate arrangements are more likely to use transit and non-motorized modes compared to drive alone and shared ride.

Among individual and household characteristics, the results show that women and older individuals (greater than 65 years) are more likely to share a ride. Income is a significant factor affecting mode choice, with individuals in low income-households more likely to use transit and non-motorized modes, while individuals in high income-households (>60k per year) are less likely to use a shared-ride mode relative to other modes. The impact of auto availability on mode choice was not found to be significant, after controlling for the effect of household income and number of workers in the household.

The pattern characteristics indicate that the presence of a high number of serve-passenger and personal-business activities during the day leads to a tendency to use a motorized non-transit mode for all tours of the day, while recreation activity (being often a group activity) leads to higher use of ride-sharing. An individual who performs only one tour during the whole day is very unlikely to use a non-motorized mode for this tour, since the expected heterogeneity of activity episodes in the tour may make the use of non-motorized modes inconvenient.

The impact of tour characteristics indicates a reluctance to use shared-ride and nonmotorized modes for a tour if personal-business or shopping stops are pursued in that tour, while tours with recreation stops are most likely to be pursued using shared-ride or non-motorized modes. This latter effect is likely to be a manifestation of joint recreational activity participation. Further, some of the recreational episodes may be exercise-related (walking, bicycling, jogging, *etc.*), which are pursued using non-motorized modes.

#### 3.2 Morning Home-stay Duration (MDUR) Sub-Model

Figure 2 presents the distribution of the baseline hazard on a continuous time scale for morning home-stay duration models for the cases without and with unobserved heterogeneity. We observe that while both baseline hazards are predominantly monotonic, the hazard for gamma heterogeneity increases at a considerably faster rate after a morning home-stay duration of around 480 minutes (corresponding to a clock time of 11 a.m.). Thus, the model without unobserved heterogeneity is not able to adequately capture the behavior of nonworkers for longer values of morning home-stay duration. This observation is also supported by the non-zero variance of the gamma distribution for unobserved heterogeneity in Table 2 (see toward the bottom of the table).

Table 2 also provides the parameter estimates for the effect of exogenous variables on the hazard rate. A positive coefficient indicates a lower hazard rate (*i.e.*, a longer morning home stay duration) and a negative coefficient implies a higher hazard rate (*i.e.*, a shorter home-stay duration). This is because of the negative coefficient on  $\beta$  in the relationship between the hazard and exogenous variables (see Equation 1).

The results of the effect of the household composition variables need to be considered jointly. Nonworkers in households with several active adults (>16 years of age and non-retired) tend to depart later in the day for their first out-of-home activity (see the positive coefficient on household size). However, the situation reverses for non-workers in households with children and retired individuals (the appropriate effects are obtained by adding the coefficient on household size and the coefficients on "Number of children of age 5 to 16 years" and "Number

of retired individuals in household"). The earlier departure from home for nonworkers in households with children and retired individuals may be attributed to time-critical serve passenger stops.

The effects of individual and household characteristics show that women and physically challenged individuals depart later in the day, possibly due to early morning engagements at home for women and longer preparation time to leave home for the physically challenged.

The effects of the pattern-level characteristics indicate that individuals with high activity levels (*i.e.*, number of tours in the day  $\geq$  3) leave home earlier to be able to achieve their activity participation objectives during the day. Also, nonworkers who begin their sojourns from home with a child care or serve passenger stop leave home early, reflecting the time-critical nature of such activities. This effect of child care/serve passenger stops is particularly pronounced if a nonworker participates in such stops in two separate tours in the day (a "constrained" tour is defined as a tour with at least one child care/serve passenger stop). Finally, nonworkers pursuing a high number of recreation stops in the first tour leave earlier in the day (perhaps reflecting the need to accommodate the long episode durations of recreation stops), while nonworkers pursuing shopping as the first stop leave later in the day.

A comparison of the log-likelihood values at convergence and the log-likelihood value at the Kaplan Meier Sample hazard using a likelihood ratio test yields a statistic of 293, which is far greater than the chi-squared table value with 11 degrees of freedom at any reasonable level of significance. Thus, there is significant duration dependence in morning home-stay duration as well as significant variations in home stay duration due to individual factors.

#### 3.3 Episode Duration - Travel Time (EDUR-TT) Sub-Model

The EDUR-TT sub-model provides estimates of the effect of variables on both episode duration as well as travel time to the episode from the previous episode. The log-likelihood value at convergence for the simultaneous linear equation system is -7467.366, while the log-likelihood value with no exogenous variables (except the constants) and independent episode and travel time estimations is -8894.69. The number of observations used in the estimation is 8156.

The parameters presented in the episode duration and travel time equations are estimated simultaneously and jointly for all the episodes of an individual. The sum of all the durations and travel times (across all episodes) was constrained to be no more than the time available from the departure of the first tour of the day until 3 a.m. the next day. The estimation results for episode duration and travel time duration are discussed in separate sections below for ease in presentation.

#### 3.3.1 Episode Duration

Separate episode duration models for each activity type are estimated and presented in Table 3. The results for home-stay duration indicate that the non-first home stay durations are determined exclusively by the overall activity and travel characteristics of individuals (i.e., the pattern and episode characteristics) rather than sociodemographics. This is different from the results for the first home-stay duration in the day (see Section 3.2), and supports our decision to model the first home-stay duration separately from subsequent home-stay durations in the day. The impact of the pattern characteristics indicates shorter home-stay episode durations for individuals with a high level of travel activity in the day. On the other hand, home-stay duration is longer if the number of serve passenger stops in the day increases, presumably reflecting long home-stay duration model, the results indicate a shorter duration of non-first in-home episodes if the first home-stay in the day is long.

The duration of serve passenger episodes is not related to exogenous variables, except for the number of serve passenger stops in the day (see the third main column in Table 3). The results for personal-business episodes suggest longer durations for high-income individuals, perhaps reflecting larger and more time-consuming business transactions (see the fourth main column in Table 3). The total number of stops in the day has the expected negative effect for personal-business episode durations. Personal business episodes pursued using non-motorized modes are likely to be of short duration, perhaps because such episodes are pursued very close to home and can, therefore, be participated-in more frequently (the results for travel duration in the next section support this hypothesis). Finally, and as expected, the duration of personal-business episodes is shorter if the individual begins out-of-home activities late in the day.

The shopping episode results (fifth main column in Table 3) suggest a lower duration for older individuals, a higher duration for women (probably because they are primarily responsible for grocery shopping responsibilities in the household), and the usual negative effects on duration of high overall activity levels (see the coefficients on total daily stops and stops in the current tour) and a late start in the day. The modal effects suggest longer durations when shopping is undertaken using a shared-ride mode (reflecting, perhaps, longer durations of joint shopping activities) or transit modes (presumably due to transit-schedule constraints). The shorter shopping duration associated with non-motorized modes indicates secondary shopping episodes to pick up critical non-bulky grocery and other shopping items.

The results for the recreation activity purpose indicate that individuals who are mobilitychallenged spend lesser time on recreation activities. The other results are consistent with the notion of shorter durations for individuals with high overall activity levels and who begin their out-of-home activities later in the day.

#### 3.2.2 Travel Time to Episode

The purpose-specific estimation results for travel time to an episode are presented in Table 4. The results for travel time to a home-stay episode (*i.e.*, travel time of the return-home trip in a tour) indicate short durations for non-workers with small children and long duration for Hispanic/Asian Americans. The former finding is perhaps a reflection of short trips due to efficient activity-travel patterns with young children. This interpretation is reinforced by the short travel time duration for non-workers in nuclear families (see the coefficient on nuclear family under household structure). The results also suggest the following: (a) retirees travel shorter durations on their return home, (b) high income individuals travel longer durations, (c) individuals with high activity levels travel shorter durations (this can be attributed to more efficient patterns, to reduce "dead-time" traveling back home), (d) individuals using shared ride travel for long durations on their return home, possibly reflecting a conscious decision to travel jointly for trips to activities far away from home, (e) individuals traveling by transit have a long duration of travel, which is likely to be an indication of the slower transit mode, and (f) individuals starting their out-of-home activities later in the day have a shorter travel duration.

As in the case of episode duration, travel time to serve-passenger episodes are not substantially affected by exogenous variables. For personal-business activity, African-Americans travel longer, possibly due to lower accessibility levels. The effects of all other variables are similar to those for the return home trip.

The results for shopping indicate longer travel time durations for Hispanic and Asian Americans, perhaps due to visits to specialized shops that sell ethnic goods. Other results are similar to those already discussed earlier for return home trips. The same is true for the results of travel time duration to recreation activity episodes.

The covariance between the episode and travel time durations due to unobserved effects is provided in the final row of Table 4. These estimates indicate significant correlations for home-stay, serve passenger, and personal-business activity types. The covariance for home-stay episodes is negative, indicating that individuals responsible for substantial in-home chores (leading to long home-stay durations) organize their tours to minimize the "dead" travel time on the return home. The covariances for serve passenger and personal-business activities are positive, indicating that people may be willing to travel longer for activities of longer duration.

#### 3.3 Activity Location (ALOC) Sub-Model

The travel time component of the EDUR-TT sub-model provides the distribution of the desired travel time for participation in episodes. As discussed in Section 2.4, we use this distribution to circumscribe the locations that an individual considers for participation. But the actual location is based on a choice model from among the locations in the individual's consideration set. The ALOC sub-model provides results characterizing this location choice. For this study, we considered all the zones within the San Francisco Bay Area to be potential candidate locations.

Table 5 presents the estimation results of the ALOC sub-model. The population density of the attraction zone is used as an indicator of the zonal attractiveness of the candidate location. A zone size indicator (specific to each activity type) is used to represent the total number of available opportunities at each candidate location. The zone size is represented by the total area of the attraction zone (in acres) for serve-passenger activity, the total zonal employment (retail plus non-retail) for personal-business activity, the zonal retail employment for shopping activity, and the total zonal household population of the attraction zone for recreation activity. Travel impedance is defined based on the mode chosen and the time of occurrence of the activity. The impedance measure assumes a ratio of 1.75 between the values of out-of-vehicle and in-vehicle travel time based on an estimated trip-based mode choice model. In addition, a value of \$4 per hour is used to convert all the costs into in-vehicle time units for computation of the impedance measure.

The estimation results for all activity purposes indicate that non-workers tend to choose locations that are not densely populated for their episode participations. This is perhaps because zones with lower population density are invested more in non-residential land uses and present more variety in opportunities for participation in non-work activities. The effect of the size variable indicates larger probabilities of choice of zones with larger size. Travel impedance has the expected negative effect on location choice, with the effect being most pronounced for personal-business and shopping activities and least for recreational activities. The interaction effects of impedance for the serve passenger purpose indicate lower travel impedance sensitivity for women and individuals with disabilities impacting transit use. For personal business, African-Americans participate in locations which are farther away, perhaps a reflection of stronger locational preferences and/or less accessible locations. For shopping activity, we find that non-Caucasians and high income individuals (>60k) are willing to select locations with higher impedances. The results for recreation indicate choice of locations that are closer for Hispanic and Asian American individuals, individuals with disabilities, and high income individuals. These may reflect accessibility and mobility constraints.

#### 4. SUMMARY AND CONCLUSIONS

The analysis of the activity-travel behavior of non-workers and a study of the factors influencing their travel-related decisions is an important element of a comprehensive household-level activity-based travel demand modeling system. This paper presents the formulations for four sub-models that form a part of an activity-travel pattern modeling framework for non-workers. The paper also presents empirical estimation results for data from the San Francisco Bay Area.

The empirical analysis results indicate the following:

- Women, non-workers in households with young children, non-workers undertaking a higher number of tours/stops, and those departing late for the first tour exhibit greater temporal efficiency in their activity-travel patterns. The temporal efficiency is often achieved by undertaking lesser number of tours, departing later for the first stop of the day, choosing destinations closer to home, and reducing the return-home time at the end of tours.
- Presence of a serve-passenger activity tends to "constrain" the activity-travel pattern leading to an earlier departure from home in the morning, while nonworkers performing more shopping stops show greater flexibility in their activity-travel pattern.
- Caucasians are significantly less likely to use transit and do not have strong locational preferences for shopping activities. Recreational activity participation exhibits the highest usage of shared-ride and non-motorized modes. Individuals with disabilities tend to spend less time on recreation activity.
- Nonworkers choose destinations that are not densely populated. Travel impedance has a significant effect on location choice for all activity episodes.

It is important to note that the primary focus of this paper is to develop a better understanding of how the different activity-travel decisions of non-workers are influenced by their household characteristics, socio-demographics, and personal taste preferences. Subsequent research can extend this study by incorporating supply-side variables to study the impact of policy changes on the individuals' activity-travel behavior. Furthermore, the sequential organization of the three models in this paper is one of several alternative analysis sequences. The detailed examination of these various alternative sequences is an important area for future work on activity scheduling mechanisms.

In addition to the above extensions, several other venues for future research are also available. First, an econometric micro-simulator for analyzing and forecasting the daily activitytravel patterns of non-workers may be developed. The micro-simulator can be linked to a geographic information system (for data acquisition, storage, and retrieval) and can be coupled with a graphical user interface to serve as a comprehensive transportation planning tool. The authors have completed an initial version of such an econometric micro-simulator and have applied it to the Dallas-Fort Worth area [see Misra and Bhat (17)]. Second, alternate model structures and/or error distributions may be tested in the mathematical formulations to achieve a better behavioral representation. The parametric structure of most of the models used in the current paper may be relaxed to develop semi-parametric or non-parametric models. Some of the medium-term decisions (e.g., residential location choice and auto ownership) that are assumed to be exogenous could also be incorporated within the modeling framework. Third, specific studies may be designed to further explore the approaches for achieving temporal efficiency, substitution effects between in-home and out-of-home activities, and day-to-day dynamics. Fourth, an integrated household-level representation and analysis framework that fully incorporates all the different inter-personal linkages between workers, students, and non-workers in a household would be a valuable avenue for future research. Future studies need to consider a more extensive set of land-use and urban form factors in influencing individual activity and travel patterns.

## ACKNOWLEDGEMENTS

The authors are grateful to Lisa Weyant for her help in typesetting and formatting this document. Thanks are also due to four reviewers for their constructive suggestions on an earlier version of the paper.

### REFERENCES

- 1. Bhat, C. R. and R. Misra. A Representation and Analysis Framework for Analysis of the Activity-Travel Patterns of Nonworkers: Organization of Activities. *Transportation Research Record 1777*, TRB, National Research Council, Washington, D.C., 2001, pp16-24.
- 2. Bhat, C. R. and F. S. Koppelman. A Retrospective and Prospective Survey of Time-Use Research. *Transportation*, Vol. 26, No. 2, 1999, pp. 119-139.
- 3. Axhausen, K. and T. Gärling. Activity-based Approaches to Travel Analysis: Conceptual Frameworks, Models and Research Problems. *Transport Reviews*, Vol. 12, 1992, pp. 324-341.
- 4. Kurani, K. S. and R. Kitamura. Recent Developments and the Prospects for Modeling Household Activity Schedules. *Report prepared for the Los Alamos National Laboratory*, Institute of Transportation Studies, University of California, Davis, CA, 1996.
- 5. Arentze, T. A., and H. J. P. Timmermans. Albatross A Learning-Based Transportation Oriented Simulation System. *Transportation Research*, 2002 (forthcoming).
- 6. Bhat, C. R. and S. K. Singh. A Comprehensive Daily Activity-Travel Generation Model System for Workers. *Proceedings of the 1998 Annual Transportation Research Board Meeting*, Washington, D.C., 1998.
- Mahmassani, H. S., Hatcher, S. G. and C. G. Caplice. Daily variation of trip chaining, scheduling, and path selection behavior of work commuters. *Understanding Travel Behavior in an Era of Change* (P. Stopher and M. Lee-Gosselin, eds.), Pergamon, Elsevier Science Ltd., Oxford, UK, 1997, pp. 351-379.
- 8. Hamed, M. M., and F. L. Mannering. Modeling Travelers' Postwork Activity Involvement: Toward a New Methodology. *Transportation Science*, Vol. 27, No. 4, 1993, pp. 381-394.
- 9. Pendyala, R. M., Yamamoto, T. and R. Kitamura. On the Formulation of Time-Space Prisms to Model Constraints on Personal Activity-travel Engagement. *Transportation*, Vol. 29, No. 1, 2002, pp. 73-94.
- Bowman, J. L., and M. E. Ben-Akiva. Activity-Based Disaggregate Travel Demand Model System with Activity Schedules. *Transportation Research Part A*, Vol. 35, Elsevier Science, New York, 2000, pp. 1-28.
- Kitamura, R., and S. Fujii. Two Computational Process Models of Activity Travel Behavior. *Theoretical Foundations of Travel Choice Modeling* (T. Garling, T. Laitila, and K. Westin, eds.), Elsevier Science, Oxford, England, 1998.
- 12. E.H. White and Associates. 1990 Bay Area Travel Survey. *Final Report to the San Francisco Bay Area Metropolitan Transportation Corporation,* January 1991.
- 13. Misra, R. and C. R. Bhat. Activity-Travel Patterns of Non-Workers in the San Francisco Bay Area: An Exploratory Analysis. *Transportation Research Record 1718*, TRB, National Research Council, Washington, D.C., 2000, pp 43-51.
- 14. Ben-Akiva, M. E. and S. R. Lerman. *Discrete Choice Analysis: Theory and Application to Travel Demand*. The MIT Press, Cambridge, 1985.
- 15. Kiefer, N. M. Economic duration data and hazard functions. *Journal of Economic Literature*, Vol. 26, No.2, 1988, pp. 646-679.
- 16. Bhat, C. R. A Hazard-based Duration Model of Shopping Activity With Nonparametric Baseline Specification and Nonparametric Control for Unobserved Heterogeneity. *Transportation Research*, Vol. 30B, No. 3, 1996, pp. 189-207.
- 17. Misra, R. and C. R. Bhat. A Comprehensive and Operational Econometric Modeling Framework for the Analysis of the Activity-Travel Patterns of Nonworkers. Accepted for publication in *The 9<sup>th</sup> WCTR selected proceedings*, 2001.

## LIST OF FIGURES

FIGURE 1 Analysis framework for nonworker activity-travel patterns.

FIGURE 2 Morning Home-Stay Duration Model: Nonparametric Baseline Hazard.

## LIST OF TABLES

TABLE 1 Tour Mode Choice Model Estimation Results [Base = Drive Alone]

TABLE 2 MDUR Model Estimation Results: Parameters on Explanatory Variables

 TABLE 3 Episode Duration Model Estimation Results

TABLE 4 Travel Time to Episode Model Estimation Results

 TABLE 5 ALOC Model Estimation Results





FIGURE 1 Analysis framework for nonworker activity-travel patterns.



FIGURE 2 Morning Home-Stay Duration Model: Nonparametric Baseline Hazard.

TABLE 1 Tour Mode Choice Model Estimation Results [Base = Drive Alone]
------------------------------------------------------------------------

Variable Name	Shared	Ride	Tra	nsit	Non-motorized Modes			
	Estimates	Est./s.e.	Estimates	Est./s.e.	Estimates	Est./s.e.		
Constant	-0.823	-4.679	-2.114	-7.907	-1.590	-7.942		
Household Composition								
Household size	0.289	5.042	-	-	0.263	4.771		
Number of children between 12 and 16 years	-0.410	-3.404	0.868	3.660	-	-		
Number of individuals between 17 and 21 years	-0.457	-3.527	-0.753	-2.072	-0.974	-4.101		
Number of employed Individuals in HH	-0.424	-5.294	-	-	-	-		
Household Race								
Caucasian	-	-	-1.373	-5.201	-	-		
Household Structure								
Nuclear family with all children < 22 years	0.415	3.411	-	-	-	-		
Single parent family with all child < 22 years	-0.842	-3.615	-	-	-	-		
Single parent family with 1 child $\geq$ 22 years	-	-	1.475	2.976	-	-		
Single member family	-1.652	-10.766	-1.107	-2.878	-	-		
Roommate family type	-	-	1.508	2.533	1.252	3.423		
Individual and Household Characteristics								
Sex (female = 1)	0.414	4.518	-	-	-	-		
Age > 65 years	0.239	2.494	-	-	-	-		
HH income less than 20k	-	-	1.159	4.232	0.450	3.195		
HH Income greater than 60k	-0.369	-3.634	-	-	-	-		
Pattern Characteristics								
Total number of serve-passenger stops in the day	0.122	2.755	-0.567	-2.553	-0.510	-5.036		
Total number of personal-business stops in the day	-	-	-0.647	-3.557	-	-		
Total number of recreation stops in the day	0.221	2.200	-	-	-	-		
Individual performs only one tour in the day	-	-	-	-	-0.543	-4.130		
Tour Characteristics								
Number of stops for personal-business in current tour	-0.264	-4.384	-	-	-0.541	-5.067		
Number of stops for shopping in current tour	-0.186	-3.053	-	-	-0.466	-4.257		
Number of stops for recreation in current tour	0.221	2.200	-	-	0.296	2.664		
Log-likelihood value at convergence	-2671.725							
Log-likelihood value at market shares	-3011.717							
Number of cases	3095							

## TABLE 2 MDUR Model Estimation Results: Parameters on Explanatory Variables

Explanatory Variables	Estimates	Est./s.e.	
Household Composition			
Household size	0.078	1.583	
Number of children (5 to 16 years)	-0.240	-2.390	
Number of retired individuals in the household	-0.163	-2.414	
Individual and Household Characteristics			
Female	0.250	2.091	
Has disability impacting transit use	0.725	3.044	
Pattern Characteristics			
Number of tours in the day $\geq 3$	-1.278	-7.919	
Child care/serve passenger is first stop of day	-0.753	-3.577	
Two constrained tours in the day <sup>1</sup>	-0.674	-3.341	
Number of recreation stops in first tour of the day	-0.302	-3.542	
Shopping is first stop of day	0.722	4.144	
MU (variance of gamma distribution)	1.941	4.394	
Log-likelihood value at convergence	-6068.572		
Log-likelihood value with no exogenous variables and no unobserved heterogeneity ( <i>i.e.</i> , the Kaplan-Meier log-likelihood value)	-6215.025		
Number of cases	2048		

<sup>1</sup> A "constrained" tour is defined as a tour with at least one serve passenger/child-care stop

## TABLE 3 Episode Duration Model Estimation Results

Variable Name	Home Stay		Serve-passenger		Personal-business		Shopping		Recreation	
·	Estimates	Est./s.e.	Estimates	Est./s.e.	Estimates	Est./s.e.	Estimates	Est./s.e.	Estimates	Est./s.e.
Constant	7.690	112.534	1.726	13.388	4.844	38.529	4.063	33.770	5.317	29.313
Household Composition										
Number of individuals over 65 years	-	-	-	-	-	-	-0.149	-4.089	-	-
Individual & Household Characteristics										
Sex (female = 1)	-	-	-	-	-	-	0.227	3.565	-	-
Has disability impacting transit use	-	-	-	-	-	-	-	-	-0.492	-2.437
HH income greater than 60k	-	-	-	-	0.173	1.789	-	-	-	-
Pattern Characteristics										
Total number of stops in the day	-0.092	-6.263	-	-	-0.311	-14.208	-0.183	-7.517	-	-
Total number of tours in the day	-0.713	-27.590	-	-	-	-	-	-	-0.278	-5.259
Total number of serve-passenger stops in the day	0.107	5.428	-0.099	-2.172	-	-	-	-	-	-
Tour Characteristics										
Number of stops in current tour	-	-	-	-	-	-	-0.063	-2.263	-0.223	-7.132
Shared ride tour mode	-	-	-	-	-	-	0.266	4.374	-	-
Transit tour mode	-	-	-	-	-	-	0.651	3.440	-	-
Non-motorized tour mode	-	-	-	-	-0.684	-5.338	-0.586	-5.990	-1.523	-13.155
Episode Characteristics										
Morning home-stay duration	-0.034	-9.278	-	-	-0.068	-8.825	-0.027	-4.503	-0.021	-2.663
Standard deviation of error term	1.021	78.676	1.645	37.148	1.468	55.964	1.131	57.131	1.448	48.436

# TABLE 4 Travel Time to Episode Model Estimation Results

	Home Stay		Serve-passenger		Personal-business		Shopping		Recreation	
	Estimates	Est./s.e	Estimates	Est./s.e	Estimates	Est./s.e	Estimates	Est./s.e	Estimates	Est./s.e
Constant	2.886	55.278	2.546	67.578	3.176	39.081	2.338	55.132	3.412	35.033
Household Composition										
Number of children between 5 and 11 years	-0.099	-3.428	-	-	-	-	-	-	-	-
Household Race (Caucasian is base)										
African American	-	-	-	-	0.183	2.047	-	-	-	-
Hispanic/Asian American	0.133	3.22	-	-	-	-	0.165	2.69	-	-
Household Structure										
Nuclear family with all children < 22 yrs	-0.15	-3.928	-0.311	-5.769	-0.133	-2.368	-0.109	-2.13	-	-
Retired couple	-0.07	-2.109	-	-	-0.16	-3.473	-	-	-	-
Individual and Household Characteristics										
Sex (female = 1)	-	-	-	-	-	-	-	-	-0.117	-2.334
HH income less than 20k	0.083	2.556	-	-	-	-	0.145	3.268	-	-
Number of vehicles in HH	-	-	-	-	-	-	-	-	-0.061	-2.917
Pattern Characteristics										
Total number of stops	-	-	-	-	-0.073	-5.743	-	-	-	-
Total number of tours	-0.12	-7.906	-	-	-0.081	-3.14	-	-	-0.205	-7.105
Tour Characteristics										
Number of stops in current tour	-	-	-	-	-	-	-0.028	-2.11	-	-
Shared ride tour mode	0.211	7.511	-	-	0.292	6.773	0.236	6.096	0.287	5.172
Transit tour mode	0.858	11.559	-	-	0.842	6.341	0.865	6.784	0.725	4.84
Non-motorized tour mode	-	-	-	-	-0.17	-2.447	-	-	0.268	3.601
Episode Characteristics										
Morning home-stay duration	-0.017	-6.157	-	-	-0.026	-6.04	-	-	-0.031	-7.056
Standard deviation of error term	0.735	78.68	0.708	37.148	0.771	55.959	0.749	57.131	0.818	48.436
Covariance between episode and travel duration error terms	-0.062	-3.472	0.080	2.112	0.152	6.139	0.024	0.970	-0.004	-0.140

## **TABLE 5 ALOC Model Estimation Results**

Variable	Serve-Passenger		Personal-Business		Shopping		Recreation	
Variable	Estimates	Est./s.e.	Estimates	Est./s.e.	Estimates	Est./s.e.	Estimates	Est./s.e.
Density and Size Effects								
Population density of attraction zone	-0.001	-2.512	-0.001	-2.512	-0.001	-2.512	-0.001	-2.512
Logarithm of zone size indicator	-	-	0.280	10.290	0.386	16.249	0.119	3.857
Impedance Effects								
Primary effect	-0.041	-3.505	-0.058	-14.271	-0.088	-12.004	-0.005	-2.661
Interaction Effects								
African American	-	-	0.037	3.062	0.048	3.501	-	-
Hispanic/ Asian American	-	-	-	-	0.026	2.460	-0.032	-2.798
Sex (female = 1)	-0.057	-3.719	-	-	0.003	0.303	-	-
Age over 65 years	-	-	-	-	-	-	0.003	1.336
Has disability impacting transit use	-0.104	-1.481	-	-	-	-	-0.030	-1.992
Household income < 20k	-	-	-0.004	-0.465	-	-	-	-
Household income > 60k	-	-	-	-	0.031	3.602	-0.037	-5.343
Log-likelihood value at convergence	-13837.661							
Log-likelihood value with no exogenous variables	-14656.873							
Number of cases	4852							