NISS

Generation of Synthetic Daily Activity-Travel Patterns

Ryuichi Kitamura, Cynthia Chen, and Ram M. Pendyala

Technical Report Number 69
October, 1997

National Institute of Statistical Sciences
19 T. W. Alexander Drive
PO Box 14006
Research Triangle Park, NC 27709-4006
www.niss.org

GENERATION OF SYNTHETIC DAILY ACTIVITY-TRAVEL PATTERNS

Ryuichi Kitamura
Professor
Department of Civil Engineering Systems
Kyoto University, Sakyo-ku
Kyoto 606-01, Japan
Ph: 81-75-753-5134; Fax: 81-75-753-5916
Email: rkitamura@term.kuciv.kyoto-u.ac.jp

Cynthia Chen
Post-Graduate Research Engineer
Institute of Transportation Studies
University of California at Davis
Davis, CA 95616, USA
Ph: 916-752-2029; Fax: 916-752-6572
Email: qzchen@ucdavis.edu

Ram M. Pendyala
Assistant Professor
Department of Civil and Environmental Engineering
University of South Florida
Tampa, FL 33620-5350, USA
Ph: 813-974-1084; Fax: 813-974-2957
Email: pendyala@eng.usf.edu

ABSTRACT

Microsimulation approaches to travel demand forecasting are gaining increasing attention due to their ability to replicate the multitude of factors underlying individual travel behavior. The implementation of microsimulation approaches usually entails the generation of synthetic households and their associated activity-travel patterns to achieve forecasts with desired levels of accuracy. This paper develops a sequential approach to generate synthetic daily individual activity-travel patterns. The sequential approach decomposes the entire daily activity-travel pattern into various components, namely, activity type, activity duration, activity location, work location, and mode choice and transition. The sequential modeling approach offers practicality, provides a sound behavioral basis, and accurately represents individual's activity-travel patterns. In the proposed system, each component may be estimated as a multinomial logit model. Models are specified to reflect potential associations between individual activity-travel choices and such factors as time-of-day, socio-economic characteristics, and history dependence. As an example, the paper furnishes results for activity type choice models estimated and validated using the 1990 Southern California Association of Governments travel diary data set. Validation results show that the predicted pattern of activity choices conforms with observed choices by time of day. Thus the paper shows that realistic daily activity-travel patterns, which are requisites for microsimulation approaches, can be generated for synthetic households in a practical manner.

ACKNOWLEDGMENTS

The authors are grateful to the National Institute of Statistical Sciences for partially funding this research to RDC, Inc. through a grant from the National Science Foundation. The authors thank two anonymous referees for their helpful suggestions. The authors are solely responsible for all errors and omissions.

INTRODUCTION

Microsimulation of the behavior of a household or an individual is drawing attention as a new approach to travel demand forecasting (*1*). Microsimulation can replicate the behavior of complex systems or processes, and is therefore suited for the representation of travel behavior, which is a complex behavior. The factors that make travel behavior complex include: the multitude of contributing factors and decision rules involved; constraints that govern the behavior; inter-personal interactions; multiple planning horizons; and complexity of activity-travel decision making as a scheduling problem (*2*). Microsimulation is an effective approach to such a complex phenomenon which facilitates its practical, yet realistic, representation.

Achieving desired levels of accuracy in the outcome of travel demand forecasts produced by microsimulation of household behavior may require a large sample of households. This may happen when: high levels of spatial or temporal resolution are required of the outcome; sample households do not have a desirable geographical distribution; demand by small population segments is desired; or a high level of accuracy is desired. In such instances, the number of households available in the data set at hand may not be sufficiently large. As a result, the generation of synthetic households may be required. When the microsimulation expects daily travel patterns of household members as input data, the generation of synthetic daily travel patterns will be required.

An approach to the problem of synthetic travel pattern generation is presented in this paper. The proposed synthetic travel pattern generator has a sequential structure and can be decomposed into components to which certain aspects of observed activity-travel behavior correspond, thus establishing a link between mathematical models and observational data. The model components are each relatively simple and are estimated using commonly adopted estimation methods and existing data sets.

The problem of synthetic travel pattern generation is presented formally in the second section. Accumulated knowledge on the characteristics of daily travel patterns is reviewed in the third section. Following this, a discussion of several modeling issues is presented in the fourth section. The fifth section presents aspects of travel behavior incorporated into the model system. The formulation of the model components constitutes the sixth section, while estimation and validation results are furnished in the seventh section. Conclusions are drawn in the final section.

PROBLEM DESCRIPTION

Consider a household member, i, whose daily activity-travel pattern can be characterized as

$$(\mathbf{X}_i, \mathbf{T}_i, \mathbf{L}_i) = (X_{i0}, X_{i1}, \ldots, X_{in}; T_{i0}, T_{i1}, \ldots, T_{in}; L_{i0}, L_{i1}, \ldots, L_{in}) \qquad (1)$$

where

X_{ij} = the type of the j-th activity pursued by individual i,

T_{ij} = the duration of the j-th activity pursued by individual i,

L_{ij} = the location of the j-th activity pursued by individual i (if the activity is travel, then L_{ij} refers to the destination of trip j; in this case $L_{ij} = L_{i,j+1}$),

n = the number of activities involved in individual i's daily activity-travel pattern,

and (X_{i0}, T_{i0}, L_{i0}) is the initial condition. Note that travel is included here as one of the activity types. For simplicity, travel mode, which may be stored in another vector, say $\mathbf{M}_i$, is not included in the discussion here. The mode choice component is discussed in a subsequent section of this paper.

The development of a synthetic daily activity-travel pattern implies generating vectors $\mathbf{X}_i$, $\mathbf{T}_i$, and $\mathbf{L}_i$ given:

- attributes of individual i
- attributes of the household to which i belongs
- residence and work location of i
- demographic and socio-economic characteristics of the region
- land use characteristics of the region, and
- transportation network and travel time characteristics of the region.

Since it is most likely that synthetic activity-travel patterns will be generated for synthetic individuals and households, the first three items will comprise synthetic data. Generating synthetic individuals and households is, however, beyond the scope of this paper (for discussions on the generation of synthetic households, see *3*); it is assumed here that all personal and household attributes, as well as work location, are known for i. The latter three items will consist of projected values in cases where synthetic activity-travel patterns are generated for forecasting.

ACCUMULATED KNOWLEDGE

The following discussions offer a brief summary of what is known about n, which is also a variable to be determined, and each of the three vectors, $\mathbf{X}_i$, $\mathbf{T}_i$, and $\mathbf{L}_i$. It is possible that additional information is available from the literature on time use. This literature is not well known in the transportation field, and needs to be explored further in the future.

Number of Activities Per Day: n

The total number of activity episodes captured in time use surveys tends to be 20 to 25 per person per day, including trips. In the transportation field, the average number of trips is between 3 to 5 per person per day. It is known that the number of trips captured varies greatly depending on the survey methodology. It is well established that total trip generation is associated with demographic and socio-economic attributes of the traveler.

Activity Type: X_i

There are certain regularities in the sequence with which individuals engage in different types of activities. For example, one may anticipate that the sequence of activities performed before leaving home for work or after coming back home from work, is fairly uniform across individuals. The literature in time use analysis needs to be explored to determine tendencies for activity sequences involving both in-home and out-of-home activities (*4*).

Kitamura (*5*) examined the sequence of trip purposes using standard trip diary data from Detroit. The trip purpose was used to identify the primary out-of-home activity type at each destination location. The analysis examined how out-of-home activities were sequenced in a home-based trip chain, i.e., the home-to-home series of trips which involves one or more stops. The results indicated that activities of more mandatory nature tend to be pursued first in a trip chain. The sequencing tendencies indicated the following hierarchy:

- work and school, work-related
- chauffeuring
- personal business (e.g., banking, dental and medical)

- shopping, and
- social and recreational

The presence of the same sequencing hierarchy was later found for activities throughout the day (*6*, *7*). Another important tendency is that activities pursued in the same trip chain tend to be similar (*5*).

Activity Duration: T_i

Several studies have investigated the duration of activity engagement. In a semi-markov process model of trip chaining, Lerman (*8*) used gamma distributions to represent the duration of sojourns at destination locations. Survival models have recently been applied to the time dimension in activity-travel patterns (*9*, *10*, *11*). These studies are typically based on the simplifying assumption that the durations of successive activities are independent.

Activity duration has been examined from the viewpoint of resource allocation. A theoretical model can be found in Kitamura, et al. (*12*) where the duration of an activity episode was analytically derived while assuming that the total daily activity pattern is optimized and that each activity episode has a logarithmic utility function. The model was estimated using a time use data set from the United States. Although the model is based on the assumption that daily time use is optimized as a whole, the resulting model applies to individual activity episodes. Golob and McNally (*13*) examined the allocation of time to different activity types using a structural equations model system. This approach facilitates the inference of causal relationships among activities of different types.

Critical in the analysis of activity duration is the correlation across the duration of respective activity episodes. As the total amount of time available is fixed at 24 hours a day, negative associations can be expected. In addition, the duration of each episode is also a function of n, the total number of episodes. The inter-relationships among duration of different types of activities and the number of activities, n, merit further exploration.

Activity Location: L_i

Non-home activity locations have traditionally been estimated using the gravity model of spatial interaction. The multinomial logit model of destination choice can be viewed as a special case of the gravity model family. In principle, these models depict that, *ceteris paribus*, more intense interaction exists between a pair of locations that are closer to each other, and the intensity of the interaction is positively related to the attraction level of the destination and the number of trips initiated at the origin.

One important issue is the characterization of location/destination choice for non-home-based trips, i.e., trips whose origin and destination are both non-home. For home-based destination choice underlying a simple trip chain involving only one stop (i.e., home-activity-home), the only spatial element to be considered is the separation between the destination and the home base. This does not hold true in the case of non-home-based choice. For example consider the choice of a shopping location on the way home from work; in this case, both the home location and the deviation from the regular commute route would be important considerations. Kitamura and Kermanshah (*7*) constructed a non-home-based destination choice model which included both the usual origin-to-destination travel time, t_{ij}, and the destination-to-home travel time, t_{ih}, in a multinomial logit choice model. Their estimation results clearly indicated that t_{ij} and t_{ih} are equally important for non-home-based destination choice. This finding is readily applicable to the generation of synthetic activity-travel patterns.

Travel Mode: M_i

There are numerous studies on travel mode choice. Most studies, however, are seriously limited because they are trip-based, i.e., they analyze each trip separately in isolation from other trips. Consider the choice of commuting by car because a car is needed for work. Then this mode choice behavior cannot be explained by solely examining the home-to-work commute trip and comparing the attributes of the travel modes available for that trip.

One of the critical requirements in synthetic pattern generation is to observe the constraints imposed on the transition between travel modes. For example, transition from public transit to driving alone is usually not possible unless the transition takes place at the home or work base where a private car is placed or at a special facility such as a park-and-ride lot. For a trip chain that originates and terminates at the home base, the sequence of travel modes tends to be governed by the boundary condition that the mode of the first trip from home is identical to that of the last trip to home. These regularities and tendencies serve as a set of constraints in the generation of activity-travel patterns.

MODELING CONSIDERATIONS

There are two broad classes of approaches to the generation of synthetic activity-travel patterns: sequential (incremental) approaches vs. simultaneous (holistic) approaches. The former adopt rules in order to generate, one by one, the activity that will immediately follow, given the history of activity generation so far. The latter approaches, on the other hand, deploy behavioral paradigms that are each concerned with the entire daily activity-travel pattern.

One paradigm for the simultaneous approaches is that an individual with given attributes has a probability vector that depicts the likelihoods with which he or she will exhibit respective activity-travel patterns. A study by Pas (*14*) is readily applicable to operationalize this paradigm. Another paradigm is utility maximization where an individual chooses that activity-travel pattern, from among a set of all feasible patterns, which offers the maximum utility. Studies based on this assumption include Adler and Ben-Akiva (*15*), Recker, et al. (*16*), and Recker (*17*). The two paradigms can be integrated to produce probabilities for alternative daily activity-travel patterns.

The simultaneous approaches have theoretical elegance. They can be expected to be more sensitive to parameters describing the travel environment than sequential approaches. In addition, simultaneous approaches can better reflect individuals' travel planning effort. Despite the advantages offered by simultaneous modeling approaches, a sequential approach is proposed in this study. There are three major reasons.

- *Practicality*: One important advantage of sequential approaches is the ease of implementation they offer. When viewed as an optimization problem, daily activity-travel behavior is very complex (*2*). Exact formulation of this behavior produces an overwhelmingly complex mathematical problem. The size of the problem at each step is much smaller in sequential approaches because a daily pattern is synthesized incrementally.
- *Behavioral Basis*: Sequential approaches do not lack a behavioral basis. For example, when proposing the paradigm of satisficing, Simon (*18*) noted that a person is not capable of enumerating all possible alternatives or discerning minute differences among them. Furthermore, a person often will not have complete information associated with all alternatives. As such, even though certain travel choices may be considered simultaneous, it

may be argued that people sequentially process "information elements" in order to reduce the size and dimensionality of the problem.

- *Contexts of Synthetic Activity-Travel Pattern Generation*: Synthetic activity-travel patterns are usually generated to represent baseline travel characteristics of the population under prevailing conditions. In this context, sequential model systems offer policy sensitivities that are consistent with the objectives of synthetic pattern generation.

The sequential approach adopted in this paper is based on the identity that, given n, the X-T-L triple can be expressed as:

$$\begin{aligned}
\Pr[X_i, T_i, L_i] &= \Pr[X_{i1}, X_{i2}, \ldots, X_{in}; T_{i1}, T_{i2}, \ldots, T_{in}; L_{i1}, L_{i2}, \ldots, L_{in}] \\
&= \Pr[X_{in}, T_{in}, L_{in} | X_{i1}, X_{i2}, \ldots, X_{i,n-1}; T_{i1}, T_{i2}, \ldots, T_{i,n-1}; L_{i1}, L_{i2}, \ldots, L_{i,n-1}] \\
&\times \Pr[X_{i,n-1}, T_{i,n-1}, L_{i,n-1} | X_{i1}, X_{i2}, \ldots, X_{i,n-2}; T_{i1}, T_{i2}, \ldots, T_{i,n-2}; L_{i1}, L_{i2}, \ldots, L_{i,n-2}] \\
&\times \ldots \\
&\times \Pr[X_{i1}, T_{i1}, L_{i1}] \qquad (2)
\end{aligned}$$

Each probability on the right-hand side can be formulated as a model for activity type, location and duration, given the past history of activity and travel. In adopting the sequential approach, the joint probability of an X-T-L triple needs to be decomposed into sequential elements. The following decompositions are possible:

$$\begin{aligned}
&\Pr[X_{ij}, T_{ij}, L_{ij} | \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}] \\
&= \Pr[L_{ij} | X_{ij}, T_{ij}, \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}] \Pr[T_{ij} | X_{ij}, \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}] \Pr[X_{ij} | \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}] \\
&= \Pr[T_{ij} | X_{ij}, L_{ij}, \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}] \Pr[X_{ij} | L_{ij}, \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}] \Pr[L_{ij} | \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}] \\
&= \ldots \qquad (3)
\end{aligned}$$

etc., where $\tilde{X}_{i,j-1} = (X_{i0}, X_{i1}, \ldots, X_{i,j-1})'$, etc.

Since all permutations of X_{ij}, T_{ij} and L_{ij} lead to the same joint probability, the model's replication capability should not depend on which permutation is adopted. Therefore that permutation which can be theoretically supported and/or which offers most modeling flexibility and sensitivity can be selected.

EXPLORATION OF POTENTIAL ASSOCIATIONS

As noted previously, knowledge has been accumulated on characteristics of activity-travel behavior. A few salient aspects of activity-travel behavior that merit inclusion in a synthetic generator are outlined in this section.

- *History Dependence*: History dependence has been found to be prevalent in studies of activity type choice (*5*, *19*) and location choice (*20*, *21*, *22*, *23*, *24*). While it is also likely that history dependence is prevalent for activity duration choice, the knowledge of the history dependence in activity duration appears to be extremely limited.
- *Time-of-Day Dependence*: Activity engagement is strongly dependent on the time of day. Tabulations of time use data (e.g., *25*) show surprising homogeneity in activity engagement across individuals. This is partly institutional (e.g., work and school) and partly physiological (e.g., meals, sleeping). The time-of-day dependence of activity engagement can be represented by formulating engagement probabilities as time-dependent functions (*6*).
- *Spatial and Temporal Constraints*: Different activities have different levels of constraints in terms of (i) engagement, (ii) duration, (iii) location, and (iv) timing. Higher levels of engagement and duration constraints are typically associated with work and school (mandatory) activities. It may be assumed that more flexible activities are organized around these constrained activities. Some types of activities may have tight constraints when they are

pursued with prior commitment, e.g., a medical appointment. In general, constraints associated with activity engagement vary significantly depending on institutional and situational factors (e.g., store hours), prior arrangement and commitment, as well as the type of activity. An issue in this effort is whether constraints associated with each activity should be explicitly considered and modeled, or treated as random elements. Considering data availability, only the latter approach is feasible. However, constraints on regular events such as work and school merit explicit consideration.

- *Planned vs. Unplanned Activities*: Some activities are routine, some are planned ahead, yet some are unplanned and are pursued in response to unanticipated events. It is desirable that the degree of planning be represented when synthesizing travel patterns as it allows the analysis of transportation policy impacts on an individual's travel plans. In the context of synthetic pattern generation, however, representing the level of planning in activity engagement is of lesser importance, given the constraints associated with activities are well understood. Also, data availability is an issue. Based on these considerations, the model system in this study does not explicitly incorporate the degree of planning.
- *Travel Time Budget*: History dependence in L_i as well as in T_i would arise if a traveler allocates a certain amount of time for traveling. This leads to the notion of travel time budgets (e.g., *26*). There have been disputes on whether individuals have a fixed time budget that is invariant across individuals. However, more recent results offer evidence that when the duration of a trip is reduced, then a portion of the time saving tends to be used to travel more (*13*, *27*).
- *Prism Constraints*: The spatial expanse that is accessible to an individual for activity engagement is determined by the speed of movement and the amount of time available.

Hagerstrand (*28*) defined this expanse in the time-space dimension as the time-space "prism." The prism contains all possible locations where activities can be engaged, and defines the amount of time available for activities at each location within it. Kondo and Kitamura (*29*) adopted the prism concept in the analysis of trip chaining behavior. Beckmann, et al. (*26*) used the concept to define accessibility measures. The prism concept is important because it defines the state space for the evolution of location choice.

- *Trade-off between Activity Duration and Travel Time*: The trade-off between the duration of activity and the time spent to reach the activity location is also important. One may choose to visit a nearby opportunity and spend more time on the activity there, or visit a farther, but better opportunity and spend less time there. This consideration is adopted by Kitamura, et al. (*12*) in the formulation of time-utility functions. The model in this study accounts for this by making the probability of L_{ij} conditional on T_{ij}.
- *Modal Continuity, Permissible Transitions and Time-of-Day Dependence*: Despite the voluminous studies on travel mode choice, little is known on history dependence and time-of-day dependence of travel mode choice. Modal continuity and modal transition have rarely been addressed in the literature (a rare example can be found in *23*). In general, the travel modes used by an individual in a series of trips tend to be governed by the constraints surrounding modal transitions. In addition, as both transit and highway levels of service vary along the time of day, it is likely that mode choice is time-of-day dependent.
- *Relationships among Travel Choices*: It is now widely recognized that various dimensions of travel behavior are related to one another. For example, activity type choice influences destination choice as a traveler would choose a destination that fulfills the specific activity need. Similarly, inter-relationships exist between destination choice and mode choice, activity

type choice and departure time choice, and departure time choice and activity duration. The sequential model system developed in this study explicitly incorporates inter-dependencies among travel choice dimensions in synthesizing activity-travel patterns.

MODEL FORMULATION

For X_{ij} which is not travel, the following decomposition of the X-T-L triple may be adopted:

$$\Pr[X_{ij}, T_{ij}, L_{ij} | \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}]$$

$$= \Pr[L_{ij} | X_{ij}, T_{ij}, \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}] \Pr[T_{ij} | X_{ij}, \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}] \Pr[X_{ij} | \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}] \quad (4)$$

In this formulation, an activity type is selected first; given the type, its duration is determined; and finally, a location is chosen given the type and duration. Each of these decision elements is assumed to be dependent on the past history of behavior. This formulation is based on the view that activity engagement is the most fundamental decision that drives duration and location choice. While this may not hold true under all conditions, it may be regarded a typical activity engagement decision process.

When X_{ij} is travel, the following decomposition would be more appropriate:

$$\Pr[X_{ij}, T_{ij}, L_{ij} | \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}]$$

$$= \Pr[T_{ij} | X_{ij}, L_{ij}, \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}] \Pr[L_{ij} | X_{ij}, \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}] \Pr[X_{ij} | \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}] \quad (4')$$

Namely, the destination, L_{ij}, is determined before travel time, T_{ij}. This reflects the view that travel time can not be determined before destination and mode are determined.

Overview of the Synthetic Travel Pattern Generator

The components of the synthetic travel pattern generator are:

- activity-type choice models
  - home-based and non-home-based
  - workers and non-workers
- activity duration models
  - workers and non-workers
  - by activity type
- activity location choice models
  - home-based vs. non-home-based
  - workers and non-workers
  - by activity type
- mode choice and mode transition models
  - home-based and non-home-based
- initial departure timing models
  - workers and non-workers
- initial location models
  - workers and non-workers

where "worker" refers to an individual who is employed, either full-time or part-time, or a student. It is possible for a part-time workers' daily activity-travel pattern to not include a commute trip. At this stage, model components have been developed for weekdays only. The activity types used in the models are: work, work-related, school, return to work, social/recreation, shopping, personal business, eat out, home (transient), and home (absorbing). The remainder of this section describes each model component.

Activity-Type Choice Models

Activity-type choice models are concerned with $\Pr[X_{ij}|\tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}]$. These models probabilistically determine the next activity type to be engaged. Two types of models, home-based models and non-home-based models, are developed. The former is for an out-of-home activity that follows an in-home sojourn, while the latter is for an activity, whether in-home or out-of-home, that follows an out-of-home sojourn. While the latter includes "in-home activity" as an alternative in the choice set, the former excludes it. It is to be noted that the home-based vs. non-home-based distinction does not refer to the location where the choice is made. Both types of models are developed for workers and non-workers separately. The history dependence of activity type transition is represented by formulating the probability of an activity type as a function of the series of activities so far engaged, $\tilde{X}_{i,j-1}$, the time that has been allocated to them, $\tilde{T}_{i,j-1}$, and the current location, $L_{i,j-1}$.

Activity Duration Models

Consider an activity type, a. Given $X_{ij} = a$, T_{ij} will have a probability distribution function whose parameters are functions of t, $\tilde{X}_{i,j-1}$, $\tilde{T}_{i,j-1}$, $L_{i,j-1}$, and $\tilde{Z}_i$ as follows:

$$\Pr[T_{ij} \leq q|X_{ij} = a, \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}, \tilde{Z}_i] = G_a(q{:}\tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, L_{i,j-1}, \tilde{Z}_i), \quad q \geq 0, \ a = 1, 2, \ \ldots, \ k \qquad (5)$$

where

t = time of day when the (j-1)th activity ended

$\tilde{Z}_i$ = vector of person attributes and other explanatory variables.

G_a is a distribution function. Two sets of activity duration models are developed; one for workers and the other for non-workers. The same activity classification scheme as in the activity-type choice models is adopted and models are developed for all activity types except absorbing home (person returns home for the day).

Some distribution functions may be preferred over others for activity duration. For example, let an activity comprise n task elements, and let task completion times be identically and independently distributed (i.i.d.) with a negative exponential distribution for all task elements. Then the distribution of the duration of this activity is a type-n Erlang distribution. Other distributions, including negative exponential, Weibull, and log-normal distributions, have geneses that offer interpretations suitable for activity duration. The Weibull distribution is used in this modeling effort considering its goodness-of-fit and intuitively appealing interpretation in the context of activity duration modeling.

Activity Location Choice Models

The problem here is to determine the probability that the location of the j-th activity is g, given the type and duration of the activity, the completion time of the (j-1)th activity, t, $\tilde{X}_{i,j-1}$, $\tilde{T}_{i,j-1}$ and $\tilde{L}_{i,j-1}$. The models are formulated for all activity types, except in-home activity.

Home-Based Models

The home-based location choice models take on a form that is similar to conventional destination choice models:

$$
\begin{aligned}
&\Pr[L_{ij} = g \mid h_i, X_{ij}, T_{ij}, \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}, \tilde{Z}_i, \tilde{A}, \tilde{S}] \\
&\qquad = \Pr[L_{ij} = g \mid t, h_i, X_{ij} = a, T_{ij} = q, \tilde{Z}_i, \tilde{A}, \tilde{S}] \\
&\qquad = H_a(g{:}\, t, h_i, q, \tilde{Z}_i, \tilde{A}, \tilde{S}) \qquad (6)
\end{aligned}
$$

where h_i denotes the residence zone, $\tilde{A}$ is a vector of attractiveness measures of alternative locations, and $\tilde{S}$ is a matrix of origin-destination travel times. Note the assumption that the location choice is conditionally independent of $\tilde{X}_{i,j-1}$, $\tilde{T}_{i,j-1}$ and $\tilde{L}_{i,j-2}$, given t, h_i (=$L_{i,j-1}$), X_{ij} (=a), and T_{ij} (=q).

Non-Home-Based Models

As will be discussed later, a travel mode is assigned in the procedure prior to the selection of destination location for a trip whose origin is not the home base. Let M_{ij} be the mode of the trip made to the j-th activity location. With the assumption that destination choice is conditionally independent of $\tilde{M}_{i,j-1}$ as well as $\tilde{X}_{i,j-1}$, $\tilde{T}_{i,j-1}$ and $\tilde{L}_{i,j-2}$, given t, h_i, $L_{i,j-1}$ (=f), X_{ij} (=a), T_{ij} (=q), and M_{ij} (=r),

$$
\begin{aligned}
&\Pr[L_{ij} = g | h_i, X_{ij}, T_{ij}, M_{ij}, \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}, \tilde{M}_{i,j-1} \tilde{Z}_i,\ \tilde{A},\ \tilde{S}] \\
&\qquad = \Pr[L_{ij} = g | t, h_i, X_{ij} = a, T_{ij} = q, L_{i,j-1} = f, M_{i,j-1} = r, \tilde{Z}_i, \tilde{A}, \tilde{S}] \\
&\qquad = Q_a(g{:}\, t, h_i, q, f, r, \tilde{Z}_i, \tilde{A}, \tilde{S}) \qquad (7)
\end{aligned}
$$

Mode Choice and Mode Transition Models

A travel mode is assigned to each trip using the following procedure:

- the travel mode for the first trip in each home-based trip chain is determined (home-based models), and
- a mode transition matrix is developed and applied to determine subsequent travel modes on a trip-by-trip basis (non-home-based models).

The model system incorporates a dummy variable that indicates whether a private car is parked at the work place, which makes the probability very high that a car will be used for a trip originating from the work place. Models are developed for workers and non-workers separately. Travel modes are grouped into: auto driver, auto passenger, public transit, and bicycle and walk.

Home-Based Models

The home-based models incorporate accessibility indices for the residence zone and, for workers, accessibility indices for the work zone. Accessibility indices by mode are defined as the "log-sum" variables of the utility functions of the destination choice models. Highway and transit travel times and distances to destination zones are also incorporated. Also included in the models are descriptors of the destination zone (e.g., percent retail) and the time of day when the trip starts. The models take on the form:

$$\Pr[M_{ij} = r | h_i, X_{ij}, T_{ij}, L_{ij}, \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}, \tilde{M}_{i,j-1} \tilde{Z}_i,\ \tilde{A},\ \tilde{S}]$$

$$= \Pr[M_{ij} = r | t, h_i, X_{ij} = a, T_{ij} = q, L_{ij} = f, \tilde{H}_{i,j-1}, \tilde{Z}_i, \tilde{A}, \tilde{S}] \qquad (8)$$

where $\tilde{H}_{i,j-1}$ is defined as $\tilde{H}_{i,j-1} = (H_{i1,j-1}, H_{i2,j-1}, \ldots, H_{im,j-1})'$, m being the number of modes and

$H_{ir,j-1} = 1$, if $\tilde{H}_{i,j-1}$ contains mode r, for $r = 1,2,\ldots,m$;

$H_{ir,j-1} = 0$, otherwise.

Non-Home-Based Models

The non-home-based models are transition models which determine the probability that a certain travel mode will be used for a trip given the mode of the previous trip. Additional explanatory variables include descriptors of the destination zone, car-parked-at-work dummy (for workers) and the time of day. The non-home-based mode choice models are trip-end models that are applied before destination location is determined. They can be summarized as follows:

$$\Pr[M_{ij} = r \mid h_i, X_{ij}, T_{ij}, \tilde{X}_{i,j-1}, \tilde{T}_{i,j-1}, \tilde{L}_{i,j-1}, \tilde{M}_{i,j-1} \tilde{Z}_i,\ \tilde{A},\ \tilde{S}]$$

$$= \Pr[M_{ij} = r \mid t, h_i, w_{i,j-1}, X_{ij} = a, T_{ij} = q, M_{i,j-1} = u, \tilde{H}_{i,j-1}, \tilde{Z}_i, \tilde{A}, \tilde{S}] \qquad (9)$$

where $w_{i,j-1}$ is the car-parked-at-work dummy. As noted earlier, non-home-based mode choice models are transition models, and replicate modal continuity conditions in the data set.

Initial Departure Timing Models

These models may be viewed as the duration models for the first activity of the day starting at say, 3:00 a.m., which is typically an in-home activity (most probably, sleeping). In this study, models are estimated for workers and non-workers separately, and applied in synthetic pattern generation to those sample individuals who are at home at 3:00 a.m.

Initial Activity Type/Location Models

These models determine the type and zonal location of the first activity. As noted earlier, this is usually an in-home activity and the location is the residence zone. Data sets thus do not typically offer rich information (in terms of variation across individuals) for these models. As a result, they tend to be simple frequency models without many explanatory variables.

Work Location Models

These models are equivalent of home-based work trip distribution models. The probability that a worker commutes to a certain zone is formulated as a function of network auto travel times, zonal attributes, and person and household attributes. The models are formulated as multinomial logit models.

7. SAMPLE ESTIMATION RESULTS

This section provides sample results of the estimation of activity type choice models and activity duration models for the following:

- Work activity for workers
- Social Recreation activity for workers and non-workers

For purposes of brevity, the presentation of results in this paper has been limited to two modules of the generator and two activity types. This section is intended to provide a representative indication of the performance of the model for mandatory (work) and discretionary (social recreation) activities and for workers and non-workers, thus covering a variety of behavioral conditions. The 1991 travel diary data set of the Southern California Association of Governments is used for model estimation and validation purposes. The data set provides a total of 136,640 trip records for 32,515 individuals. As such, it provides rich information with a sufficient sample size in each population segment considered in this effort. Sample estimation results for each of the two modules are summarized in the subsections that follow.

Activity Type Models

The data set was randomly divided into two subsets; one subset was used for estimation purposes and the other for validation purposes. Multinomial logit models of activity type choice were estimated using standard maximum likelihood methods. Estimation and validation results are presented in Tables 1 and 2 respectively.

Table 1 presents estimation results for six activity type choice models. They are:

- Home-based and non-home-based models for workers
  - Return to work
  - Social-recreation
- Home-based and non-home-based models for non-workers
  - Social-recreation

Models estimated for workers set work as the reference alternative (utility is set to zero), while models for non-workers set shopping as the reference alternative. Explanatory variables include socio-economic characteristics of the person, dummy variables of time-of-day, and lagged dependent variables of history dependence. Those time periods which are not represented in the model are used as reference periods.

The models clearly show the time-of-day dependence of activity type choice. For example, in the home-based model for workers, the “return to work” activity peaks around 11:30 a.m. to 1:00 p.m.; this may be explained by workers having lunch and then returning to work. Social/recreation activities peak between 7:30 p.m. and 9:30 p.m. as evidenced by the larger coefficients associated with dummy variables representing evening hours. During the early morning, coefficients for social recreation are less than those for work indicating that the work activity peaks during that period.

Socio-economic variables also play important roles in determining activity type choice. The work and social recreation activity models presented in this paper do not include socio-economic variables as they were found to be statistically insignificant at the 0.05 level. However, other activity types including personal business, shopping, eat out, and school (not shown in this paper) were significantly influenced by socio-economic characteristics.

History dependence is represented by a lagged dummy variable, which takes on a value of 1 if the activity is performed earlier in the day and 0 otherwise. History dependence effects are found to be statistically insignificant for the return to work, but are found to be significant in explaining workers' social-recreation activity engagement. The coefficients have negative signs indicating that if a social-recreation activity was pursued previously in the day, then there is a reduced likelihood of repeating the activity. History dependence was also found to be significant for other activity types, notably shopping and personal business (not presented in this paper).

Table 2 presents validation results for the six models for which estimation results were presented in Table 1. The validation results are presented by time of day over a 24 hour period for return to work and social recreation activities. For each time period, the actual frequency of each activity type and the expected frequency (calculated as the product of mean probability and total frequency) are provided. χ^2 statistics are then calculated for each cell. In this table, if the χ^2 statistic is less than the critical value at n-1 degrees of freedom (where n is the number of time periods), the predicted frequency distribution is not significantly different from the actual frequency distribution. The χ^2 statistic associated with each cell also indicates the activity that contributes most to differences between the predicted and observed distributions.

Table 2 shows that when the model is applied to the validation set, the overall activity pattern by time-of-day is captured successfully. An examination of the χ^2 statistics indicates that, without exception, the actual and the expected frequency distributions are not significantly different for all of the six models presented. It is noteworthy that similar results were obtained for other activity types also (tables not shown).

Activity Duration Models

Table 3 presents estimation results for a few activity duration models. These models are estimated assuming that the duration of an activity episode is described by the Weibull distribution.

The Weibull distribution is often used to model the failure time distribution of manufactured components. Analogously, it may be used to model the distribution of the length (duration) of activity episodes. The Weibull density function is convenient in that it provides a wide variety of density curves to model real life failure time distributions. In addition, unlike the Gamma distribution, the Weibull distribution has a closed-form expression for its cumulative distribution function.

In all of the three models presented in the table, the model coefficients have the expected sign. History dependence is a significant factor influencing the length of an activity episode. For example, as the cumulative past time spent at work increases, the length of a "return to work" activity episode decreases. Similarly, as time spent at work or school increases, the duration of social-recreation activity episodes decreases.

Time-of-day is also found to be a significant factor explaining activity duration. Social-recreation activities are shorter in the morning and longer in the evening for workers. The reverse is true for non-workers who may often pursue social-recreation opportunities in the morning. Return to work activity durations are longest in the morning and late evening, but shorter during mid-afternoon. Finally, socio-economic variables such as gender, employment status, age, and household structure are found to influence activity durations.

Validation of the activity duration model (results not presented in the interest of brevity) may be done in a manner similar to that for activity type choice models. The observed and predicted frequency distributions of activity durations by activity type may be compared using χ^2 test-statistics to determine whether the activity duration models are statistically replicating observed activity sojourn patterns.

CONCLUSION

An analytical framework has been proposed in this paper for the development of a procedure for generation of synthetic activity-travel patterns. As more refined travel demand forecasting and policy analysis are demanded in the current transportation planning contexts, it is becoming inevitable that a new generation of travel demand models be adopted to satisfy planning needs. Microsimulation of travel behavior is emerging as a promising approach. Many issues, including the generation of synthetic activity-travel patterns, need to be resolved before its practical adaptation; yet only limited knowledge has been accumulated on these issues. In this study, attempts have been made to include a broad range of analytical issues and develop a rationale for the proposed approach. It is hoped that the paper has aided in paving the way for the development of a synthetic activity-travel pattern generator and toward the formulation of the next generation of travel demand models.

REFERENCES

1. Miller, E. Applications of Microsimulation to Activity Based Forecasting. Presented at the TMIP Conference on Activity Based Travel Forecasting, New Orleans, Louisiana, June, 1996.

2. Pas, E.I. Is Travel Demand Analysis and Modeling in the Doldrums? In P.M. Jones (ed.), *New Developments in Dynamic and Activity-Based Approaches to Travel Analysis*, Avebury, Aldershot, United Kingdom, 1990, pp. 3-27.

3. Beckman, R.J., K.A. Baggerly, and M.D. McKay. *Generating Synthetic Baseline Populations*. Report LA-UR 95-1985, Los Alamos National Laboratory, Los Alamos, New Mexico, 1995.

4. Bhat, C., and F.S. Koppelman. A Structural and Empirical Model of Subsistence Activity Behavior and Income. *Transportation*, Vol. 21, 1994, pp. 71-89.

5. Kitamura, R. A Sequential, History Dependent Approach to Trip-chaining Behavior. In *Transportation Research Record 944*, TRB, National Research Council, Washington, D.C., 1983, pp. 13-22.

6. Kitamura, R., and M. Kermanshah. Identifying Time and History Dependencies of Activity Choice. In *Transportation Research Record 944*, TRB, National Research Council, Washington, D.C., 1983, pp. 22-30.

7. Kitamura, R., and M. Kermanshah. A Sequential Model of Interdependent Activity and Destination Choices. In *Transportation Research Record 987*, TRB, National Research Council, Washington, D.C., 1984, pp. 81-89.

8. Lerman, S.R. The Use of Disaggregate Choice Models in Semi-Markov Process Models of Trip Chaining Behavior. *Transportation Science*, Vol. 13, No. 4, 1979, pp. 273-291.

9. Mannering, F., E. Murakami, and S.G. Kim. Temporal Stability of Travelers' Activity Choice and Home-Stay Duration: Some Empirical Evidence, *Transportation*, Vol. 21, 1994, pp. 371-392.

10. Niemeier, D.A., and J.G. Morita. Duration of Trip-Making Activities by Men and Women: A Survival Analysis. *Transportation*, 1996, (forthcoming).

11. Bhat, C. A Hazard-Based Duration Model of Shopping Activity with Nonparametric Baseline Specification and Nonparametric Control for Unobserved Heterogeneity. *Transportation Research*, 1997, (forthcoming).

12. Kitamura, R., T. van der Hoorn, and F. van Wijk. A Comparative Analysis of Daily Time Use and the Development of an Activity-Based Traveler Benefit Measure. Presented at the EIRASS Conference on Activity Based Approaches, Eindhoven, The Netherlands, May, 1995.

13. Golob, T.F., and M.G. McNally. A Model of Household Interactions in Activity Participation and the Derived Demand for Travel. Presented at the EIRASS Conference on Activity Based Approaches, Eindhoven, The Netherlands, May, 1995.

14. Pas, E.I. A Flexible and Integrated Methodology for Analytical Classification of Daily Travel-Activity Behavior. *Transportation Science*, Vol. 17, No. 4, 1983, pp. 405-429.

15. Adler, T., and M. Ben-Akiva. A Theoretical and Empirical Model of Trip Chaining Behavior. *Transportation Research*, Vol. 13B, 1979, pp. 243-257.

16. Recker, W.W., M.G. McNally, and G.S. Root. A Model of Complex Travel Behavior: Part I--Theoretical Development, Part II--An Operational Model. *Transportation Research*, Vol. 20A, No. 4, 1986, pp. 307-318 and pp. 319-330.

17. Recker, W.W. The Household Activity Pattern Problem: General Formulation and Solution. *Transportation Research*, Vol. 29B, No. 1, 1995, pp. 61-77.

18. Simon, H. A Behavioral Model of Rational Choice. In H. Simon (ed.) *Models of Man*, John Wiley and Sons, Inc., New York, 1957.

19. Kitamura, R. Time-of-Day Characteristics of Travel: An Analysis of 1990 NPTS Data. In *Special Reports on Trip and Vehicle Attributes*, 1990 NPTS Report Series, Publication No. FHWA-PL-95-033, Federal Highway Administration, U.S. Department of Transportation, Washington, D.C., February, 1995, pp. 4-1-4-56.

20. Horton, F.E., and P.W. Shuldiner. The Analysis of Land Use Linkages. In *Highway Research Record 165*, TRB, National Research Council, Washington, D.C., 1967, 96-107.

21. Nystuen, J.D. A Theory and Simulation of Intra-Urban Travel. In W.L. Garrison and D.F. Marble (eds.) *Quantitative Geography - Part I: Economic and Cultural Topics*. Studies in Geography, Vol. 13, Northwestern University, Evanston, Illinois, 1967, pp. 54-83.

22. Sasaki, T. Estimation of Person Trip Patterns Through Markov Chains. In G.F. Newell (ed.) *Traffic Flow and Transportation*. Elsevier Science Publishers, New York, 1972, pp. 119-130.

23. Kondo, K. Estimation of Person Trip Patterns and Modal Split. In D.J. Buckley (ed.) *Transportation and Traffic Theory*, Elsevier Science Publishers, New York, 1974, pp. 715-742.

24. Hanson, S., and D.F. Marble. A Preliminary Typology of Urban Trip Linkages. *The East Lakes Geographer*, Vol. 7, 1971, pp. 49-59.

25. Robinson, J.P., R. Kitamura, and T.F. Golob. *Daily Travel in the Netherlands and California: A Time-diary Perspective*. Report prepared for Rijkswaterstaat, Dutch Ministry of Transport, The Hague, 1992.

26. Beckmann, M.J., T.F. Golob, and Y. Zahavi. Travel Probability Fields and Urban Spatial Structure: 1. Theory. *Environment and Planning A*, Vol. 15, 1983, pp. 593-606.

27. RDC, Inc. *Further Comparative Analysis of Daily Activity and Travel Patterns and Development of a Time-Activity-Based Traveler Benefit Measure*. Report prepared for Rijkswaterstaat, Dutch Ministry of Transport, the Hague, 1993.

28. Hagerstrand, T. What About People in Regional Science? *Ninth European Congress of the Regional Science Association Papers*, Vol. 24, 1970, pp. 7-21.

29. Kondo, K. and R. Kitamura. Time-space Constraints and the Formation of Trip Chains. *Regional Science and Urban Economics*, Vol. 17, 1987, pp. 49-65.

List of Tables

TABLE 1. Sample Estimation Results for Activity Type Choice Models

TABLE 2. Validation Results for Home-based Activity Type for Workers/Students

TABLE 3. Sample Estimation Results for Activity Duration Models (Weibull Distribution)

TABLE 1
Sample Estimation Results for Activity Type Choice Models

HB for Workers/Students		HB for Non-workers/Non-students		NHB for Workers/Students		NHB for Non-workers/non-students	
Return to Work		*Return to Work*		*Return to Work*		*Return to Work*	
Variables	Estimates	Not Applicable		Variables	Estimates	Not Applicable	
Constant	-0.32 (-1.2)			Constant	-1.56 (-2.4)		
D(3-6:30am)	-0.82 (-0.6)			D(7:30-8:30am)	1.05 (1.5)		
D(6:30-7:30am)	-2.01 (-1.7)			D(8:30-9:30am)	1.67 (2.4)		
D(7:30-8:30am)	-1.35 (-1.6)			D(9:30-10:30am)	2.94 (4.4)		
D(8:30-9:30am)	-0.88 (-1.2)			D(10:30am-12:30pm)	3.83 (5.9)		
D(9:30-11:30am)	0.87 (2.2)			D(12:30-2:30pm)	4.75 (7.3)		
D(11:30am-1:30pm)	2.52 (8.2)			D(2:30-4:30pm)	3.91 (6.0)		
D(1:30-3:30pm)	1.49 (4.9)			D(4:30-6:30pm)	1.91 (2.9)		
D(3:30-5:30pm)	-0.10 (-0.3)			D(6:30-8:30pm)	2.13 (3.0)		
D(5:30-7:30pm)	-0.78 (-2.5)			D(8:30-10:30pm)	3.19 (3.6)		
				D(10:30pm-3am)	1.37 (1.5)		
Social/Recreation		*Social/Recreation*		*Social/Recreation*		*Social/Recreation*	
Variables	Estimates	Variables	Estimates	Variables	Estimates	Variables	Estimates
Constant	1.22 (4.9)	Constant	-0.19 (-0.7)	Constant	-3.08 (-14.5)	Constants	0.75 (2.4)
D(3-6:30am)	-3.53 (-13.0)	D(3-6:30am)	1.46 (3.3)	D(7:30-8:30am)	-0.37 (-1.3)	D(3:00-8:30am)	-0.69 (-1.6)
D(6:30-7:30am)	-4.42 (-16.1)	D(6:30-8:30am)	0.92 (3.0)	D(8:30-9:30am)	0.85 (3.4)	D(8:30-10:30am)	-1.55 (-4.6)
D(7:30-8:30am)	-4.70 (-16.9)	D(8:30-10:30am)	0.03 (0.1)	D(9:30-10:30am)	1.87 (7.5)	D(10:30-12:30am)	-2.03 (-6.2)
D(8:30-9:30am)	-3.55 (-13.1)	D(10:30am-2:30pm)	-0.36 (-1.3)	D(10:30am-12:30pm)	2.81 (12.2)	D(12:30-2:30pm)	-1.70 (-5.2)
D(9:30-11:30am)	-2.30 (-8.8)	D(2:30-6:30pm)	-0.07 (-0.2)	D(12:30-2:30pm)	3.03 (13.1)	D(2:30-4:30pm)	-1.78 (-5.4)
D(11:30am-1:30pm)	-1.46 (-5.5)	D(6:30-7:30pm)	0.88 (2.9)	D(2:30-4:30pm)	4.30 (18.5)	D(4:30-6:30pm)	-1.30 (-3.8)
D(1:30-3:30pm)	-0.99 (-3.7)	D(7:30-8:30pm)	0.86 (2.7)	D(4:30-6:30pm)	4.19 (16.9)	D(6:30-8:30pm)	-0.55 (-1.6)
D(3:30-5:30pm)	0.20 (0.8)	D(mid to old couples)	0.39 (5.0)	D(6:30-8:30pm)	5.50 (16.6)		
D(5:30-7:30pm)	0.07 (0.3)			D(8:30-10:30pm)	6.33 (10.2)		
D(7:30-9:30pm)	1.29 (5.7)			D(10:30pm-3am)	4.70 (8.4)		
D(history)	-0.80 (-10.2)			D(history)	-0.11 (-2.8)		
Summary Statistics	*N=20,928*	*Summary Statistics*	*N=6,108*	*Summary Statistics*	*N=52,478*	*Summary Statistics*	*N=11,929*
Final Likelihood	-28232.33	Final Likelihood	-8507.53	Final Likelihood	-90093.69	Final Likelihood	-19052.58
Initial Likelihood	-43518.55	Initial Likelihood	-9830.45	Initial Likelihood	-124685.22	Initial Likelihood	-23212.76
Likelihood w. Const	-38135.64	Likelihood w. Const	-9302.57	Likelihood w. Const	-102731.17	Likelihood w. Const	-20806.40
1-L(F)/L(0)	0.35	1-L(F)/L(0)	0.13	1-L(F)/L(0)	0.28	1-L(F)/L(0)	0.18
1-L(F)/L(C)	0.26	1-L(F)/L(C)	0.09	1-L(F)/L(C)	0.12	1-L(F)/L(C)	0.08

Note: Values in the parentheses are t-ratios. D refers to dummy variable, coded as 1 or 0; brief descriptions in the parentheses identify the condition(s) for which the dummy variable is equal to 1.

TABLE 2
Validation Results for Home-based Activity Type for Workers/Students

	Workers: HB Return to Work			*Workers: NHB Return to Work*			*Workers: HB Social Recn*			*Workers: NHB Social Recn*		
Time Period	Actual	Expected	χ^2	Actual	Expected	χ^2	Actual	Expected	χ^2	Actual	Expected	χ^2
3:00 am - 5:59 am	0	0	0.00	0	0	0.38	24	31	1.56	0	1	0.81
6:00 am - 7:59 am	1	1	0.00	1	1	0.03	95	101	0.31	3	6	1.76
8:00 am - 9:59 am	7	6	0.26	18	14	1.72	114	122	0.58	29	22	2.28
10:00 am - 11:59 am	14	21	2.13	58	57	0.02	100	100	0.00	42	48	0.63
12 noon - 1:59 pm	149	141	0.41	298	274	2.14	110	108	0.05	73	69	0.17
2:00 pm - 3:59 pm	68	68	0.00	120	109	1.23	163	179	1.39	118	124	0.33
4:00 pm - 5:59 pm	28	25	0.39	31	39	1.59	352	383	2.57	150	154	0.11
6:00 pm - 7:59 pm	18	36	8.91	6	12	3.19	531	555	1.01	149	126	4.10
8:00 pm - 9:59 pm	7	10	1.01	3	6	1.19	152	149	0.07	53	55	0.09
10:00 pm - 2:59 am	10	5	5.56	2	2	0.00	23	33	2.86	27	24	0.30

	Non-Workers: HB Social Recn			*Non-Workers: NHB Social Recn*[a]		
Time Period	Actual	Expected	χ^2	Actual	Expected	χ^2
3:00 am - 6:59 am	19	20	0.07			
7:00 am - 8:59 am	72	72	0.00	14	16	0.21
9:00 am - 10:59 am	101	105	0.13	51	45	0.89
11:00 am - 12:59 pm	65	67	0.06	52	58	0.58
1:00 pm - 2:59 pm	36	63	11.52	58	58	0.01
3:00 pm - 4:59 pm	56	59	0.17	48	57	1.40
5:00 pm - 6:59 pm	71	59	2.24	41	36	0.76
7:00 pm - 2:59 am	73	66	0.68	29	40	2.79

Sample Sizes Used for Model Validation

Model Type	*N*
HB for Workers	14,494
HB for Non-Workers	2,384
NHB for Workers	11,120
NHB for Non-Workers	4,608

[a] The first time period for this model is 3:00 am to 8:59 am.

TABLE 3
Sample Estimation Results for Activity Duration Models (Weibull Distribution)

Full- or Part-time Workers		Full- or Part-time Workers		Non-Workers/Non-Students	
Return to work		*Social Recreation*		*Social Recreation*	
Variables	Estimates	Variables	Estimates	Variables	Estimates
Constant	1.29 (44.32)	Constant	0.890 (20.71)	Constant	0.804 (37.48)
D(Male)	0.027 (1.58)	D(Male)	0.076 (3.14)	D(Male)	0.131 (4.38)
D(Full-time Employ)	0.117 (5.07)	D(Full-time Employ)	0.048 (1.75)	History School	-0.0091 (-1.61)
History Work	-0.017 (-3.55)	Age	-0.0036 (-3.88)	History Social/Recn	-0.022 (-2.15)
History Return Work	-0.138 (-12.8)	History Work	-0.017 (-4.69)	D(7:00-9:00 am)	0.281 (5.42)
D(7:00-9:00am)	0.382 (5.16)	History Social/Recn	0.0065 (0.59)	D(9pm-12Midnight)	0.099 (1.08)
D(1:00-4:00pm)	-0.057 (-3.22)	D(Family;Child 5-15 yr)	-0.108 (-3.11)		
D(7:00-9:00pm)	0.248 (3.67)	D(Couple; Wife <35 yr)	-0.166 (-3.26)		
		D(5:00-7:00am)	-0.271 (-3.89)		
		D(7:00-9:00am)	-0.091 (-1.51)		
		D(7:00-9:00pm)	0.029 (0.956)		
		D(9pm-12midnight)	0.106 (2.03)		
γ	1.84 (79.37)	γ	1.15 (97.99)	γ	1.11(84.47)
Summary Statistics	*N=4,070*	*Summary Statistics*	*N=5,369*	*Summary Statistics*	*N=4,171*
Final Likelihood	-4295.327	Final Likelihood	-8997.566	Final Likelihood	-7172.690
Initial Likelihood	-6748.779	Initial Likelihood	-9330.422	Initial Likelihood	-7288.913

Note: 1. Values in the parentheses are t-ratios. D refers to dummy variable, coded as 1 or 0; brief descriptions in the parentheses identify the condition(s) for which the dummy variable is equal to 1.

2. History variables refer to the cumulative past time spent on a certain activity from the beginning of the day to the current activity.