Transportation (2008) 35:467–484
DOI 10.1007/s11116-008-9166-8

Modelling daily activity program generation considering within-day and day-to-day dynamics in activity-travel behaviour

Khandker M. N. Habib · Eric J. Miller

Published online: 15 May 2008

Abstract This paper presents a comprehensive econometric modelling framework for daily activity program generation. It is for day-specific activity program generations of a week-long time span. Activity types considered are 15 generic categories of non-skeletal and flexible activities. Under the daily time budget and non-negativity of participation rate constraints, the models predict optimal sets of frequencies of the activities under consideration (given the average duration of each activity type). The daily time budget considers at-home basic needs and night sleep activities together as a composite activity. The concept of composite activity ensures the dynamics and continuity of time allocation and activity/travel behaviour by encapsulating altogether the activity types that are not of our direct interest in travel demand modelling. Workers' total working hours (skeletal activity and not a part of the non-skeletal activity time budget) are considered as a variable in the models to accommodate the scheduling effects inside the generation model of non-skeletal activities. Incorporation of previous day's total executed activities as variables introduces day-to-day dynamics into the activity program generation models. The possibility of zero frequency of any specific activity under consideration is ensured by the Kuhn-Tucker optimality conditions used for formulating the model structure. Models use the concept of random utility maximization approach to derive activity program set. Estimations of the empirical models are done using the 2002–2003 CHASE survey data set collected in Toronto.

Keywords Activity generation · Day-to-day dynamics · Within-day dynamics · Travel behaviour · Activity utility

K. M. N. Habib (✉)
Department of Civil & Environmental Engineering, University of Alberta, , 3-004 Markin/CNRL Natural Resources Engineering Facility, Edmonton, AB, Canada T6G 2W2
e-mail: khandker.nurulhabib@ualberta.ca

E. J. Miller
Department of Civil Engineering, University of Toronto, 35 St. George Street, Toronto, ON, Canada M4S 1A4
e-mail: miller@ecf.utoronto.ca

Introduction

The paradigm shifting from trip-based to the activity-based analysis of travel behaviour in recent years is land-marked by the several operational activity-based travel demand models. ALBATROSS, CEMDAP, FAMOS, and TASHA are most cited in the literature as operational activity-based travel demand models (Arentze and Timmermans 2005; Bhat et al. 2004; Pendyala et al. 2004; Miller and Roorda 2003). All of these models can be divided into two major parts: activity program generation and activity scheduling (Bhat and Koppelman 1993; Habib and Miller 2006b, c). Another common feature is the modelling time frame, which is a 'typical day'. Here, the two most important practical issues in all activity-based travel demand models are the modelling time-span and the procedures used to generate the activity program.

Considering a typical day as the time-span for activity-based model has become a commonplace in activity–travel demand modelling. A typical day assumption refers to a hypothetical day to overcome the day-specific features of a week. Arguments in favour of typical day modelling are smaller data requirements and reduced computational burden; whereas the major argument in favour of activity-based approach against trip-based approach is capturing within-day variations in travel behaviour. In reality, our activity–travel behaviour shows significant variability across the week in addition to within-day variations (Kitamura et al. 2003; Schlich and Axhausen 2003; Doherty et al. 2004). Neglecting the daily variations within a week undermines the basic principle of the paradigm shifting to the activity-based analysis: that travel is a derived demand. The derived demand concept indicates evolution of activity–travel patterns over time (Doherty 2006). It is obvious that within-day variations of activity–travel behaviour are intricately related to the day-to-day variations. Modelling within-day variations in activity–travel behaviour in a typical day model thus becomes incomplete or adhoc if the day-to-day variations are not addressed properly.

In case of the modelling techniques, the typically used approached for the activity program generation in all operational models are either simulating from empirical distributions or a decision tree approach for generating activity episode frequency, duration, etc. The empirical distribution approach (Pendyala et al. 2004; Miller and Roorda 2003) uses observed distributions of different activity attributes to simulate activity generation. It considers the distributions to be constant over time. This approach requires some index variables to identify different market segments. Unless the relationship between index variables and the objective variables of concern are econometrically justified, this can lead to spurious results (Habib and Miller 2006a). On the other hand the decision-tree approach develops conditional decision structure for different activity attributes to simulate activity generation (Arentze and Timmermans 2005). In this case an exhaustive set of possible condition variables is defined for the homogeneous condition state and then the corresponding decision trees for the decision variables are devised. This approach can also lead to inaccurate/spurious relationships or 'relationships-by-chance' if the condition variables do not have econometric relationship with the decision variables (Habib and Miller 2006c).

This paper addresses these two critical issues of modelling time-span and activity program generation process. Contrary to arbitrary distribution or decision tree approach it presents a comprehensive econometric model of activity program generation. The target is the non-skeletal activities that are temporally and spatially more flexible compared to the skeletal activities. The skeletal activities refer to the activities that are parts of medium- to long-term household decision processes, e.g. work/school activities, etc. Unlike non-skeletal activities the planning processes of such activities do not circulate within short period of time (daily or weekly basis). So, this paper separates non-skeletal activities from the skeletal activities to devise models for activity program generation. For the explicit consideration of within-day dynamics in time-use behaviour, the worker's total number of working hours (skeleton activity) is considered as a variable in the models and for the day-to-day dynamics of time-use and activity–travel behaviour, the previous day's total executed activities are considered as variables in the models. The models exploit the concept of random utility maximization. The models represent a large scale demand system modelling approach, where the demand of all activities of concerns are modelled simultaneously and the possibility of non-participation in any specific activity is ensured by adopting Kuhn-Tucker optimality conditions. The estimations of the empirical models use 2002–2003 CHASE (Computerized Household Activity Scheduling Elicitor) survey data collected in Toronto, Canada.

The rest of the paper is organized as follows: the next section discusses the activity program generation or activity agenda formation process in details, followed by the sections discussing the concept of activity utility, econometric specification and estimation procedures of the models, description of data and interpretations of the estimated models. The paper concludes with a summary of the key findings.

Activity program generation

The activity program generation stage of an activity-based travel demand model generates a set of different activities that are to be scheduled within a specific time frame. A generic set of 'things to do' within a planned period of time that does not necessarily involve a scheduled sequence is called an activity-agenda. Conceptually it represents the continuous mental process, where different short-, medium- and long-term household decisions interact with each other (Litwin 2005). Practically, in an elicited survey, it is difficult to observe the individual's mental processes. We really do not observe the activity agenda formation process. We can best observe the collection of different types of activities of a particular individual prior to the scheduling process. Though the behavioural process of activity program formation is difficult to capture, we can observe the outcomes of the process and make logical hypotheses to simulate the behavioural processes (Habib and Miller 2006a). The state-of-the-art activity data collection process CHASE can collect activity-travel information prior to the scheduling process (Doherty et al. 2004; Doherty 2000, 2006) and such data can be used to develop behavioural models of the activity generation.

The hypothesis in this paper is that, for a particular person, the collection of the activity episodes (activities with duration and other attributes) that are to be scheduled within a specific time period is the optimum set among a number of possible sets he/she desires to schedule. The challenge is to simulate the behavioural process that gives the observed outcome. All of the operational activity-based travel demand models lack the explicit definition of this activity program formation in terms of its behavioural process and multiple temporal dimensions (Litwin 2005). While all models accommodate conceptually a similar level (activity program generation) of decision-making process, the lack of an explicit definition and incorporation of behaviourally plausible hypotheses may influence the policy sensitivities of the models. This is also important from the model integration points of view. In case of an integrated land-use and transportation model, it can provide sufficient scope to connect different levels of the decision-making process (short-, medium- and long-term)

(Habib and Miller 2005, 2006a; Salvini and Miller 2005). The concept of activity utility provides the opportunity to accomplish this goal in a meaningful way. Considering the multidimensional nature of interactions and trade-offs in our daily life, it is virtually inevitable that a utility-based measure is necessary to build an operational activity-based travel demand model (Miller 2005). So, the next section discusses the concept of activity utility used to derive the specification of the models.

Concept of activity utility and specification of the model

We use the utility concept of activity proposed by Winston (1987). The utility of an activity is divided into two parts: the goal utility and the process utility. The goal utility is gained by the end state accomplished by the activity and the process utility is gained by the activity execution process. The goal utility component represents the direct utility of the activity episodes and defines the activity program generation stage. Thus this utility component can be a function of activity episode duration but at the moment of activity program generation stage, the duration need not be precisely defined (Miller 2005). On the other hand, the process utility is a feature of activity itself and will generally depend upon duration, and the actual episode duration will depend on activity scheduling and re-scheduling trade-offs for different candidate activities drawn from the activity-agenda. So, for modelling the activity program generation, the goal utility of activity episode is important. Whereas the process utility component may primarily affect activity scheduling and rescheduling trade-offs. We use the specification of the goal utility used by Habib and Miller (2006a) as originally proposed by von Haefen and Phaneuf (2004). Habib and Miller (2006a) investigated the feasibility of utility theory to model activity-agenda formation and tested different alternative model specification considering a whole week as modelling span. As an extension of that work, this paper investigates the dynamics in activity-agenda formation across the week. In this specification the total utility of the activity program (activity-agenda) is divided into two parts: the utility derived from all activities under consideration and the utility derived from all other unmodelled activities. All unmodelled activities together can be referred as composite activity. The term composite activity indicates the activity types that are not of direct interest within the model but still must be accommodated in the model since we do need to account for composite time allocation within the overall time budget to maintain the continuity. Under limited time budget this activity indicates the leftover time (Z) after planning for the specific activities under consideration. Composite activity may be composed of more than one activity type, but with respect to the specific activity types under consideration, these are not quality differentiable and hence can be considered together. The total time budget (T) for the activity agenda is divided between a bundle of all specific activities and the composite activity simultaneously. The trade-off among the specific activities within the bundle is also embedded in the specification. The utility function for the specific activities (U_j) has two components: the baseline utility and the additional utility. The baseline utility fraction ($U_{baseline}$) is the exponential function that refers to the baseline preference for a specific activity compared to the composite activity. The additional utility defines the quantitative consumption of the specific activity. The additional utility fraction ($U_{additional}$) is considered to have a logarithmic functional form to ensure a decreasing rate of satisfaction with increasing consumption. It also contains a constant term, the translating parameter (θ) that ensures the possibility of zero consumption. The translating parameter provides the reference to allow zero frequency of any specific activity episode under consideration. Under the assumption of Kuhn-Tucker optimality conditions, zero consumption refers to a corner solution. So, the specification of the total utility of an activity agenda becomes:

$$\begin{aligned} U_{total} &= \sum_{j=1}^{N} U_j + U_Z \\ &= \sum_{j=1}^{N} (U_{Baseline} * U_{additional})_j + U_Z \\ &= \sum_{j=1}^{N} e^{((\beta^p X^p) + e^{\delta}\varepsilon_j)} * \ln(e^{(\beta^a X_j^a)} Y_j + e^{\theta}) + \frac{1}{1-\exp(\rho)} z^{(1-\exp(\rho))} \end{aligned} \tag{1}$$

where,

U_j is the utility of specific activity episode
$U_{baseline}$ is the baseline utility fraction of the utility of specific activity episode
$U_{additional}$ is the additional utility fraction of the utility of specific activity episode
U_z is the utility of the composite activity
X^p is the vector of variables defining the person's tastes and preferences
β^p is the parameter corresponding to X^p
δ is the scale parameter
X^a is the vector of variables corresponding to the attraction/aversion to activity j
β^a is the parameter corresponding to X^a
Y_j is the frequency of the activity j (the quantitative consumption)
N is the total number of specific activities under consideration
θ is the translating parameter that defines the slope of the indifference curve and ensures the feasibility of corner solutions.
ε_j refers to the activity specific idiosyncratic term
Z is the time allocated to the composite activity

ρ is composite activity parameter that defines the importance of composite activity in activity program formation process.

This specification ensures that the overall utility function of an activity-agenda is a strictly increasing function of each individual activity type and the composite activity enters as additive to the collective sub-function of individual activities. It ensures the scope of modelling the trade-off between planned and unplanned activities (composite activity) at the first stage. This specification also ensures that if an activity is not chosen (i.e. $Y_j = 0$) the changes in the corresponding X_a variables do not have any influence on the total utility of the activity program (activity-agenda). This property indicates that the addition of any activity type in the choice set of activity-agenda does not have any effect on the agenda utility unless that specific activity is chosen. Thus it allows considering a very large choice sets in the demand system, which is advantageous for activity program generation modelling. Another key point is that this functional form is a direct utility function with an endogenous idiosyncratic error component that gives rise to an endogenous regime switching type of model under time budget and non-negativity of activity frequency constraints. The functional form is strictly differentiable, strictly increasing and quasi-concave, which can give 2^N combinations of possible activity-agenda sets, where N is the total number of activities in the choice set (2 represents 1 for interior and 1 for corner solution).

Considering a specific time budget, the proposed utility maximization model for activity-agenda formation becomes:

$$
\begin{aligned}
&\textit{Maximize}:\\
&U(X^p, X^a, \varepsilon) = \sum_{j=1}^{N} e^{((\beta^p X^p)+e^{\delta}\varepsilon_j)} \ln(e^{(\beta^a X_j^a)} Y_j = +e^{\theta}) + \frac{1}{1-\exp(\rho)} z^{(1\,\exp(\rho))}\\
&\textit{Subject to}:\\
&\sum_j d_j^T . Y_j + z = T, \quad d_j \textit{ indicates the average duration of activity } j\\
&\quad T \textit{ indicates the total time budget}\\
&Y_j \geq 0, \quad j = 1, 2, \ldots, N
\end{aligned}
$$

To ensure the possibility of corner solutions we exploit the Kuhn-Tucker optimality condition. For this optimization problem the Kuhn-Tucker first order optimality conditions are (Kuhn and Tucker 1951):

$$
\frac{\partial U}{\partial Y_j} - d_j * \lambda \leq 0 \text{ and } \frac{\partial U}{\partial z} - \lambda \geq 0 \tag{2}
$$

where λ is the Lagrange Multiplier.

The equality condition of Eq. 2 holds when the Y_j is non-zero. Considering the equality for z, the constant λ can be expressed as:

$$
\lambda = \frac{\partial U}{\partial z}
$$

Substituting the value of λ, we get:

$$
\frac{\partial U}{\partial Y_j} \leq d_j * \frac{\partial U}{\partial z} \quad j = 1,\ 2,\ 3, \ldots N \tag{3}
$$

$$
\begin{aligned}
&\varepsilon_j \leq f_j(X^p, \beta^p,\ \delta,\ X^a, \beta^a, Y_j, \theta, \rho):\\
&\qquad \varepsilon_j = f_j(X^p, \beta^p,\ \delta,\ X^a, \beta^a, Y_j, \theta, \rho) \quad \textit{if } Y_j > 0\\
&\qquad \varepsilon_j < f_j(X^p, \beta^p,\ \delta,\ X^a, \beta^a, Y_j, \theta, \rho) \quad \textit{if } Y_j = 0
\end{aligned} \tag{4}
$$

Putting the exact expressions of the total utility function, we get:

$$
\varepsilon_j \leq \frac{1}{\exp(\delta)} \left(\begin{array}{r} -\left(\exp((\beta_0 + \beta^P X^P)\right) + \ln \dfrac{d_j}{\exp(\beta^a X^a)} \\ + \ln(\exp(\beta^a X^a) * Y_j + \exp(\theta)) \\ + \left(-\exp(\rho) \ln(T - \displaystyle\sum_j d_j Y_j) \right) \end{array} \right), \forall j \tag{5}
$$

Intuitively the above conditions indicate that for positive-frequency activities the marginal utilities are the same across the optimal allocation. Assuming a distributional assumption of ε_j, we can easily derive the likelihood function for estimating the structural parameters of the proposed model. Let us consider an individual, i with the activity frequencies $Y_j > 0$ for $j \leq K$ and $Y_j = 0$ for $j > K$ (where $j = 1, 2, 3\ldots N$ and K is the total number of non-zero activity frequencies). Using the 'transformation of variable theorem', the contribution of this person i to the overall likelihood function of the sample is:

$$l = \int_{-\infty}^{f_{K+1}} \ldots \int_{-\infty}^{f_N} G(f_1, f_2, ..f_K, \varepsilon_{K+1}, ...\varepsilon_N)|J|d\varepsilon_{K+1}, ...d\varepsilon_N \quad (6)$$

where, $G(\cdot)$ indicates a generating function.

$|J|$ is the determinant of the Jacobian of transformation from (ε_1, ε_2, ε_3,…ε_N) to (Y_1, Y_2, Y_3,…Y_K, ε_{K+1},…ε_N).

Here the Jacobian, $J = \partial\varepsilon_j/\partial S_j$, where $S_j = Y_K$ if $j \leq K$ otherwise $S_j = \varepsilon_K$. This gives the Jacobian a partitioned structure whose lower right component is identity and the upper right component is zero. So, the determinant of the Jacobian becomes just the determinant of a $K \times K$ matrix of $\partial\varepsilon_j/\partial Y_j$ (von Haefen and Phaneuf 2004).

Considering that the ε_j has a normalized type I extreme value (IID) distribution, we get the likelihood function:

$$l(Y|\beta_0, \beta^a, \delta, \beta^p, \theta, \rho) = \prod_j \left([\exp(-f_j(\cdot)) \big/ \exp(\delta)]^{1_{Y_j>0}} |J| \exp[-\exp(-f_j(\cdot))] \right)$$

where $1_{Yj>0}$ is the indicator function for the chosen activity.

So, if the total number of individual observations is P in the sample, the total likelihood function becomes

$$L(Y|\beta_0, \beta^a, \delta, \beta^p, \theta, \rho) = \prod_P \left(\prod_j \left([\exp(-f(\cdot))/\exp(\delta)]^{1_{Y_j>0}} \cdot |J| \cdot \exp[-\exp(-f(\cdot))] \right) \right)$$

This is a closed form likelihood function. For a fixed parameter model, the structural parameters of the above likelihood function can be estimated by using a numerical gradient-based search method for fairly large number of activities without any difficulties. For this paper, the loglikelihood function is estimated using the code written in GAUSS using BFGS algorithm (Aptech Systems 2006; von Haefen and Phaneuf 2004).

Once the structural parameters of the model are estimated, the model becomes simply a constrained optimization problem. This means, for given structural parameters, constrained optimization technique should be used to estimate optimum values of activity frequencies (Y_j) and composite activity time (Z). To obtain the optimum activity-agenda of an individual, the objective function should be integrated over the possible values of unknown error terms (a similar approach is taken by Bhat 2005). In other words, being consistent with the behavioural assumption that people are not always global optimizers, one can generate a set of random numbers as a candidate vector for the unknown error terms in the model and plug those values in the objective function to generate a random activity-agenda for that person (Habib and Miller 2006a). However, it should be mentioned that this optimality does not refer to the global optimization of activity behaviour, which is strongly denied by behavioural psychologists (Gärling et al. 1989; Gärling and Garvill 1993; Gärling et al. 1994; Gärling et al. 1999). This refers to the optimality of a sub-module of overall activity decision framework that is behaviourally plausible (Habib and Miller 2006a).

A similar modelling structure is proposed by Bhat (2005), which is referred as MDCEV (Multiple Discrete Continuous Extreme Value) model. Although the MDCEV model appears to have similar features, there is a significant difference between our approach and the modelling approach proposed by Bhat (2006). The major difference is the consideration of left-over time after spending into specific activity types given a specific time budget. This left-over time is defined in our approach as 'composite activity' and Bhat (2006) defines it as 'outside good'. According to Hick's theory of composite good, this item can be a bundle of similar types of undefined alternatives, (Hicks 1946). It represents the time left over in the time budget after time spent on the specific alternatives under consideration is accounted for. Given a fixed time budget, if the time spent on specific alternatives under consideration is assumed to be random, the composite item function becomes implicitly random too (deduction of the sum of random variables from a fixed variable is also random). For this reason, unlike MDCEV model, our modelling approach does not consider any explicit random element in this composite element. However, this explicit inclusion of a random term in the outside good imposes certain restrictions on the model estimation process to ensure identification. As per Bhat (2006), the outside good cannot be composed of more than one good. In the MDCEV modelling framework, considering more than one good as a one common outside good (the left over component within the total budget after allocating to the specific goods) requires either explicitly distinguishing each of the outside goods separately and considering only one as the outside good or considering an individual error term with each outside good item. Considering separate error terms for each outside good item would not lead to the same formulations as of MDCEV model.

Data and variable specification

CHASE survey data from the first wave of the Toronto Travel-Activity Panel survey (TAPS) are used in this paper. (Roorda and Miller 2004). The survey was conducted in Toronto, in 2002–2003 with 426 individuals in 271 households. The detail description of the methods and general data summary are available in Doherty et al. (2004) and Doherty and Miller (2000). This is a multi-day self-reporting survey. The participant in this survey lists the processes of activity planning and scheduling over a period of 7 days. The participant first adds the activities before starting of the day and may or may not modify before the activities are performed. The CHASE program tracks activity episodes that are added first, then modified or deleted over time. The implementation state of an activity episode in terms that whether the episode got modified before implementation is also available. The 'added' attributes of the episode indicate the first entry and are thereby assumed to be the output of the mental process of the activity program generation. To simulate the activity program generation process, the choice set of the activity-agenda is defined by 15 individual activity types. The generic activity classification is devised to ensure the homogeneity in terms of basic type and average duration of the activities within the same group. The basic activity types are:

1. Activity01: Basic needs like lunch, coffee etc but not at home
2. Activity02: Household obligations: cleaning, maintenance, meal preparation, attending kids, attending pets etc.
3. Activity03: Drop off/Pick up goods like video/CD/DVD rental, dry cleaning, postal mails etc.
4. Activity04: Shopping like Major grocery shopping (at least 10 items)
5. Activity05: Shopping like Minor grocery shopping (less than 10 items), convenience store, drug store etc.
6. Activity06: Shopping like Personal clothing, houseware etc.
7. Activity07: Service appointments like doctors/medical appointment etc.
8. Activity08: Personal services like saloon/barber/beauty, banking, auto servicing etc.
9. Activity09: Recreations like hobbies, workout, sports, theatre, other outdoor recreational activities

10. Activity10: Recreations like at home TV program, video, napping, relaxing etc.
11. Activity11: Social activities like visiting/hosting, planned social events, bars/special clubs etc.
12. Activity12: Social activities like religious and cultural activities
13. Activity13: Other all types of activities that fall as both social as well as recreational type activities but do not belong to above types
14. Activity14: ICT use like telephone (more than 10 min), internet shopping, browsing etc.
15. Activity15: Volunteer activities

The above-mentioned activities do not include work/school, night sleep and at-home basic need type activities because these are considered skeletal (most regular) activities and are to be modelled separately (Habib and Miller 2006c). It should be mentioned that among the skeletal activities work/school is considered as fixed activity that is often defined by external factors (job type, office schedule) and hence is deducted from the time budget. On the other hand, night sleep and at home basic type activities are quite regular (skeletal) but highly flexible too. Hence these are considered as the composite activity that provides the slackness in daily time allocations.

A typical week is considered as the modelling time-span. Monday (Day0) is taken as the start of the week. Saturday (Day5) and Sunday (Day6) are the weekends. After cleaning for missing information, a total 405 out of 426 individuals remain in the data set for model estimation. Figure 1 presents weekly total and maximum daily frequencies of activities in the data set. Considerable variations are clear across the individual activity types. It is clear that people participate in some specific type of activities repeatedly throughout the week. Household obligation, basic need, personal shopping, out-of-home and at-home recreation, and ICT use activities have higher value of total weekly and maximum daily frequency. Major grocery shopping, service appointment and volunteer activities have lower values of total weekly and maximum daily frequency. These results provide some assurance that the sample represents a typical week of the individuals under consideration.

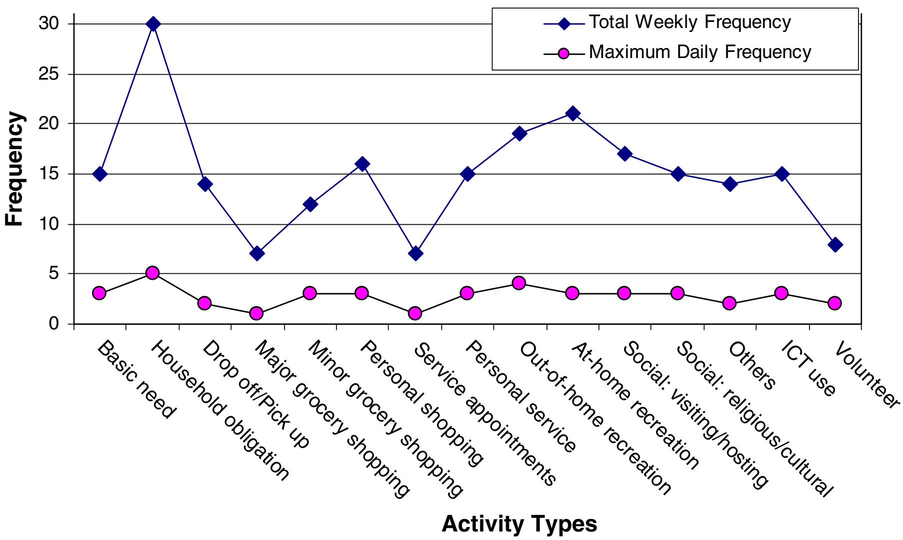

Fig. 1 Observed maximum frequencies of the activities

For modelling the activity program generation process, the time budget (T) for each day is considered as 24 h and for the whole week time budget it is 7 times 24 h. Average durations (d_j) of individual activities are considered to be the weekly average values. Several socio-economic variables are considered as covariates (X^p). These are: the individual's age, gender, total yearly income in Canadian Dollars, an indicator of having driving license, total work/school hours in the week, size of the person's household and total number of autos in the household. Two activity-specific variables (X^a) are considered: the 'number of possible activity locations' and 'travel ratio'. The travel ratio is the ratio of the summation of activity duration and travel time to the activity duration (Ettema 2005). It refers to the price of activity time in terms of travel time. By definition it is greater than or equal to 1. The longer the travel time required to get to an activity location, the higher the value of the 'travel ratio'. This variable brings trip information inside the activity program generation model in a normalized way, rather than simply using travel time. This variable makes the model sensitive to transportation system performance. The 'number of possible locations' variable represents the spatial distribution of activity spaces for the person. This variable is collected during the survey by the EWR (End of Week Review) questions. This is also a key variable that defines the spatial dispersion of activity places and its influence on activity/travel behaviour. This variable makes the model sensitive to various land-use policies.

Worker's total number of daily working hours (skeletal activity time) is considered as a variable for the baseline utility component of day-specific models. For the week-long model, this indicates total weekly working hours. Work is considered as a skeletal activity and the assumption is that planning of such activity is the part of medium- to long-term decision processes. For short-term activity planning (daily or weekly) this activity type is not considered as part of activity-program formation. Considering the total time allocation to this skeletal activity in the non-skeletal activity-agenda formation model ensures the interactions between these two basic activity types and helps to address the within-day dynamics of time allocation and activity planning behaviour. Assuming the weekly cycle as the rhythm of life, day-to-day variations and trade-offs are obvious in our activity/travel behaviour. The issues of repeated participation and serial non-participation in specific activities are also clear in the data. In order to accommodate these issues another variable is considered in the model: the previous day's total executed activity frequencies. This variable is included in the additional utility component; so that the previous day's total executed frequencies will influence the present day's activity program formation. This variable accommodates the day-to-day variations in activity program formation and it also introduces an activity scheduling element inside the activity generation process.

Interpretations of the empirical models

Models are developed for each day of the week, considering Monday as the starting day. For comparison purposes, a weekly model is also developed, considering whole week as the planning period. The models are presented in Table 1. Considering the relatively small data set compared to the large number of parameters to be estimated, the estimated coefficients are considered statistically significant if the corresponding two-tailed 't' statistics satisfy the 90% confidence interval, ($t \geq 1.64$). However, some variables with statistically insignificant parameters are also retained in the models because they provide considerable insight into the behavioural process. Retaining of some of the insignificant

Table 1 Kuhn-Tucker models for activity program generation

Parameters	Monday Values	Tuesday Values	Wednesday Values	Thursday Values	Friday Values	Saturday Values	Sunday Values	Whole Week Values
Scale Parameter	−0.084*	−0.031*	−0.060*	−0.044	−0.018	−0.040	0.002	−0.160*
Baseline utility component								
Constant	−11.664*	−13.233*	−12.290*	−12.435*	−13.681*	−13.090*	−14.082*	−11.810*
Age in 10 years	0.005*	0.007*	0.007*	0.002	0.007*	0.005*	0.003	0.003
Male (Dummy)	−0.080	0.093	0.079	0.065	0.080	0.109	−0.108	−0.070
Income in 1000 CAD/Year	0.001	0.000	0.000	0.001	−0.002*	−0.001	0.000	8E-4
Driving License (Dummy)	0.105	0.126	0.161	0.227	0.288*	0.271*	0.083	0.140*
Household size	0.120	−0.043	0.054	0.085	−0.035	−0.142	0.116	−0.020
Household Autos>2 (Dummy)	0.029	0.076	0.342*	−0.016	0.176	0.227*	0.092	0.020
Total working hours	−0.035*	−0.041*	−0.033*	−0.038*	−0.026*	−0.021*	−0.032*	0.002
Activity specific dummy								
Basic need	1.444*	1.072*	0.995*	1.309*	1.184*	1.014*	1.070*	1.150*
Household obligation	2.621*	2.487*	2.382*	2.532*	2.428*	2.496*	2.459*	2.50*
Drop off/Pick up	−0.318	−0.175	−0.099	−0.258	−0.298	−0.198	−0.201	−0.240*
Major grocery shopping	0.614*	0.513*	0.476*	0.208	0.463	0.188	0.344	0.620
Minor grocery shopping	0.366*	−0.060	0.289*	−0.039	−0.299	−0.007	−0.232	0.140
Personal shopping	1.065*	0.863*	1.076*	1.017*	0.771*	0.952*	0.797*	1.250*
Service appointments	0.186	−0.192	0.305	−0.439	0.060	−0.040	0.259	0.080
Personal service	0.191	0.258	0.234	0.249	−0.084	0.270	0.012	0.430
Out-of-home recreation	2.271*	2.218*	1.963*	2.307*	2.275*	2.070*	2.075*	2.060*
At-home recreation	3.130*	3.025*	2.927*	3.081*	2.929*	2.949*	3.046*	3.190*
Social: visiting/hosting	2.099*	2.341*	2.045*	2.206*	2.092*	2.257*	1.886*	2.330*
Social: religious/cultural	1.054*	0.931*	1.328*	1.181*	1.085*	1.072*	0.888*	2.330*

Table 1 continued

Parameters	Monday Values	Tuesday Values	Wednesday Values	Thursday Values	Friday Values	Saturday Values	Sunday Values	Whole Week Values
Others	1.081*	1.203*	1.123*	1.169*	0.963*	1.155*	0.922*	1.200*
ICT use	1.266*	1.313*	1.241*	1.473*	1.367*	1.338*	1.366*	1.040*
Volunteer	0.211	−0.084	0.372	0.514*	0.329	0.339	0.172	−0.070
Translating parameter								
Theta	−0.406*	−0.373*	0.082	0.104	0.140	−0.141	0.306*	−0.320*
Additional utility component								
Travel ratio	−0.226*	−0.481*	−0.218*	−0.223*	−0.199*	−0.328*	−0.028	−0.340*
Number of possible locations for the activity	−0.026*	−0.019	−0.011	−0.008	0.005	−0.034*	−0.008	−0.010
Previous day number of executed activities		0.690*	0.878*	0.951*	0.828*	0.812*	0.946*	
Composite good parameter								
Rho	0.691*	0.817*	0.735*	0.755*	0.819*	0.787*	0.870*	0.550*
Pseudo-R square	0.177	0.220	0.200	0.225	0.204	0.217	0.197	0.177

* indicates *t*-statistics greater than or equal to 1.64 (corresponding 90% confidence limit)

variables is also due to the expectation that if a larger data set were available these parameters might show statistical significance. The goodness of fit of the overall model is measured by the pseudo-R square value, which is 1 minus the ratio of log-likelihood value of the full model and the loglikelihood value of the null model (constant only model). The pseudo-R square value closer to 1 represents better fit to the observed data.

Individually, the goodness of fit values of the day-specific models is better than that of the aggregate weekly model. This is because of the within-day and day-to-day variations in time-use and activity/travel behaviour. The aggregate weekly model suppresses these variations, whereas in the day-specific models these are addressed explicitly. Among the day-specific models, the Monday model, which represents the beginning of the week, gives lower goodness of fit value. This is because of the fact that this model does not consider the variable 'previous day's total executed activities'. The CHASE data set used in this paper considered Monday as the beginning of the week. Hence the data are left censored for Monday and we lack the information of previous day's executed activities. The best goodness of fit is seen for Thursday model. Also the maximum number of statistical significant parameters is for the Thursday model. Thursday is the middle of the week. So, the modelling approach considering previous day's executed activities together with dynamic linkage between skeletal and non-skeletal activity information enables to capture maximum behavioural elements for Thursday. This is also indicating that for a single-day modelling many of the behavioural elements remain censored.

In terms of the components of individual models, the socio-economic variables in the baseline utility components are considered generic because it reduces the number of parameters to be estimated. At the same time, it is justified in a way that the baseline utility component mainly defines the trade-offs in time allocation to specific activities and the composite activity, but the activity specific dummy variables in the baseline utility component introduce the individual activities under consideration explicitly. So, the socio-economic variables in the baseline utility component in general indicate the activity participation tendencies and the activity-specific dummy variables indicate the baseline relative importance of individual activities in activity program generation. The variables in the additional utility component indicate individual activity's potential/attraction to derive the optimum set of activities. The translating parameter θ gives the reference for the additional utility of zero frequency activities and the composite activity parameter (ρ) indicates the effects of the time budget in activity program formation. For a coherent preference specification, the value of $1 - \exp(\rho)$ should be less than +1. A lower value of $1 - \exp(\rho)$ indicates the lower effect of time budget or the composite activity on the activity program generation.

The specification of the models is in general a large-scale demand system model specification, where under a given time budget constraint and the given average duration of the activities, the model determine whether to participate (discrete) in an activity and, if participating, then how many times (continuous) to participate. In such a specification the scale parameter of the extreme value error plays a key role in linking the discrete and continuous part of the model (Arora et al. 1998) and also according to the Eq. 1, the higher the value of the scale parameter (exponential of the parameter), the higher the randomness in activity program generation. According to Table 1, the randomness in activity program generation is the highest for Sunday, which is consistent with our expectation that weekend activities are more random than weekday activities. Comparing the day-specific models and the weeklong model, the randomness reduces when the planning period is the whole week and behaviourally it indicates that people become more definitive when thinking in the long term, while randomness/uncertainty increases with the reduction of the planning period (time budget).

The constant term in the baseline utility component indicates the reference baseline utility (if all variables in the baseline utility component are considered to be zero). This component is the highest for Sunday compared to the other days of the week. The high reference utility indicates the lack of sufficient information for Sunday compared to the other days.

The translating parameter, θ gives the non-zero reference for the additional utility component and ensures the possibility of zero participation in any activity types. The value of this parameter is the lowest for Monday and increases with the weekdays and again reduces at the end (Sunday). The higher value of this parameter basically indicates the higher possibility of zero participation in some activities but higher number of total activity participations. The higher value for the days at the middle of the week indicates that we usually participate in a higher number of activities in mid-week, but these tend to be concentrated in certain activity types, while other activities tend not to be engaged in at all in mid-week. It also indicates serial non-participation in some activities during the middle of the week. Whereas, the lower values at the beginning or at the end of the week indicate that we tend to participate in variety of activities.

The ρ parameter indicates the sensitivity of the effect of the composite activity on activity program generation. The value of $1 - \exp(\rho)$ is higher for the starting day of the week compared to the other days. It indicates that we are more cautious for allocating time to the composite activity at the beginning of the week but eventually the time pressure increases and we become less concerned about it. This finding is very important and it also reinforces our argument in favour of a week-long modelling time frame as opposed to the typical-day modelling time frame.

In terms of the variables in the baseline utility component: age of the person is significant for earlier days of the week, indicating that older people plan for higher numbers of activities than younger people. Although the dummy variable representing sex do not show a significant parameter, the values and signs indicate that males plan for higher numbers of activities in the middle of the week but females plan for higher numbers of activities in the last day of the weekend (Sunday) and the first day of the weekdays (Monday). Income becomes significant for the last day of the weekdays (Friday) and it shows negative effect. It indicates higher income people plan for lower number of activities on Friday compared to lower income people. Higher household automobile ownership and possession of driving license give people higher modal accessibility and influence them to plan more activities. These variables help to link the activity program generation model with the other components of an integrated model of short-, medium- and long-term household decisions (e.g. an auto ownership model) in a meaningful way. The total working hours of the workers, which is a skeleton activity component (Habib and Miller 2006c) is highly significant in every day-specific model. It shows similar negative effects in every model except for the whole week model, where the sign is opposite but statistically insignificant. This implies that this variable captures the within-day dynamics in time allocation behaviour. The work activity is fundamentally a different type of activity. People earn money by participating in work activity, whereas they usually spend money for participation in other activities. The negative sign of this variable indicates that the work activity acts as a temporal peg, not only in the case of activity scheduling, but also at the activity program generation level. If people spend more time in working, they allocate less time to other activities, rather than reducing time from the composite activity. Whereas in the case of the whole week aggregate model, the interpretation of the positive sign of this variable is that people who plan larger amounts of time for work throughout the week allocate lower amounts of time for composite activities rather than reducing time from the other activities. It also implies that the aggregate weekly model fails to capture the within-day dynamics in time-use and activity planning behaviour.

The activity specific dummy variable in the baseline utility component gives significant insight into the behavioural process. In this case the absolute values and signs of the coefficients do not have any practical significance, but the relative values indicate the corresponding importance of individual activities in agenda formation. In other words, these variables can be interpreted as relative direct utility gain or relative direct satisfaction or relative preference. It is clear that the least preferred activity is the drop off/pick up goods activity. It is intuitive that the drop off/pick up goods type of activities may not give direct utility *per se* but these activities provide necessary support for other activities. Volunteer activity also seems to be less preferred in comparison to other activities. The most preferred activity is at-home recreation type activities. This type of activity does not require travel, and people enjoy the process of the activities too. Household obligation type activity also has higher relative importance and it seems that people get high satisfaction by performing this type of activities. Among the three shopping activities the personal shopping gives higher relative satisfaction compared to the other two shopping types, an intuitively justified result. Between the two types of social activities, the visiting and hosting type social activities show higher satisfaction than the religious or cultural type social activities. The relative direct utility from activities involving ICT (Information and Communication Technology) has lower direct utility than direct interactive social activities and recreation activities, but online shopping (ICT) has higher direct utility than major grocery type shopping activity. It should be noted that these behavioural findings are consistent in all models shown in Table 1.

The variables in the additional utility component are key variables and they define the attraction/detraction potential of the activities under consideration. The variable 'travel ratio', which can be considered as the time cost of the activity has a negative effect in all models. Intuitively it is sensible that an activity that involves longer travel time reduces the relative interest in that activity compared to other activity types involving shorter travel time. It is interesting to note that this variable becomes statistically insignificant for the case of Sunday. The justification is that the travel time often becomes an insignificant factor for activity planning in the weekends. The difference between Sunday and Saturday is that Saturday is the beginning of the weekend and it may carry the time pressure developed through out the week and hence, travel time at that case does affect significantly in activity planning. The variable 'number of possible activity locations' also shows a negative effect. It indicates that the higher the number of possible locations the person has for any activity; the lower the frequency for that activity.

The variable indicating the total number of executed activities of each activity type under consideration is highly significant in every day-specific model. This is a key variable that defines the day-to-day dynamics in time-use and activity/travel behaviour. It also introduces the dynamic linkage between activity scheduling and activity program generation. The sign of this variable is positive. It appears that the positive sign captures the repeated participation in specific activities or serial non-participation in certain activities. As indicated in Fig. 1, total observed frequencies of some activities are very high compared to some other activities (for example household obligation versus major grocery shopping), and in practical life our all activities are not uniformly distributed throughout the week.

Conclusions

This paper describes a random utility maximizing model for day-specific activity program generation for week-long activity/travel scheduling. Considering the daily time budget, the model uses the Kuhn-Tucker optimality conditions to derive the optimum set of activities (including unspecified composite activities) within the budgeted time period. The model can capture the within-day dynamics of time-use and activity/travel behaviour by incorporating 'total working hours' in the model and day-to-day dynamics are accommodated considering 'previous day's total executed activities' as variables. These two types of variables also make the activity program generation models sensitive to the activity scheduling process. The incorporation of the variable, 'travel ratio' introduces the transportation system performance inside the activity program generation model. A number of socio-economic and activity-specific variables give insight into the behavioural process of time-use and activity/travel behaviour. The models are estimated using 2002–2003 CHASE survey data collected in Toronto. The empirical models show significant goodness of fit to the observed data. Significant variations in behaviour are obvious throughout the week. The starting (Monday), middle (Thursday) and ending (Sunday) of the week are distinctively different from the other days in terms of goodness of fit to the observed data, number of statistically significant variables, the effects of travel time in activity/travel planning etc. The models also replicate our repeated participation and serially non-participation in some activities.

In terms of application, the models described in this paper are part of the continuing effort to make the activity-based travel demand model more behavioural and more sensitive to various land-use and transportation policies. This paper compliments other work of the authors, in which a daily activity skeleton was derived (Habib and Miller 2006c). All of this research is leading to the development of 'second generation' version of the existing activity-based travel model TASHA (Miller and Roorda 2003).

Acknowledgement This research was supported by a Transport Canada TPMI grant, as well as financial contributions by the Neptis Foundation, EJLB Foundation, McLean Foundation, Salamander Foundation, and the Greater Toronto Airports Authority. The data used in the study were collected as part of a Canada SSHRC MCRI grant.

References

Aptech Systems: GAUSS User's Manual. Maple Valley, CA (2006)

Arentze, T., Timmermans, H.J.P.: ALBATROSS Version 2.0 A Learning Based Transportation Oriented Simulated System, EIRASS (2005)

Arora, N., Allenby, G.M., Ginter, J.L.: A hierarchical Bayes model of primary and secondary demand. Market Sci **17**(1), 29–44 (1998)

Bhat, C.R., Koppelman, F.S.: A conceptual framework of individual activity program generation. Transp Res A **27**(6), 433–446 (1993)

Bhat, C.R.: A multiple discrete-continuous extreme value model: formulation and application to discretionary time-use decisions. Transp Res Record Part B **39**, 679–707 (2005)

Bhat, C.R.: The multiple discrete-continuous extremem value model: role of utility function parameters, identification considerations and model extension. Paper presented at the 11th IATBR Conference, Kyoto, Japan 2006

Bhat, C.R., Guo, J.Y., Srinivasan, S., Sivakumar, A.: A comprehensive econometric micro-simulator for daily activity travel patterns (CEMDAP). Paper presented at 83rd Annual Meeting of the Transportation Research Board, Washington, DC 2004

Doherty, S.T., Miller, E.J.: A computerized household activity scheduling survey. Transportation **21**(1), 75–97 (2000)

Doherty, S.T.: Should we abandon activity type analysis? Redefining activities by their salient attributes. Transportation **36**(6), 517–536 (2006)

Doherty, S.T., Nemeth, E., Roorda, M.J., Miller, E.J.: Design and assessment of Toronto area computerized household activity scheduling survey. In: Transportation Research Record: Journal of the Transportation Research Board **1894**, TRB, National Research Council, Washington DC, pp. 140–149 (2004)

Ettema, D.: Latent activities: modelling the relationship between travel times and activity participations. In: Transportation Research Record: Journal of the Transportation Research Board **1926**, TRB, National Research Council, Washington DC, pp. 171–181 (2005)

Gärling, T., Brännäs, K., Garvill, J., Golledge, R.G., Gopal, S., Holm, E., Lindberg, E.: Household activity scheduling. In: Transportation policy, management and technology towards 2001: selected Proceedings of Fifth World Conference of Transportation Research vol. 4, pp. 235–248. Western Periodical, Ventura, CA (1989)

Gärling, T., Gillholm, R., Montgomery, W.: The role of anticipated time pressure in activity scheduling. Transportation **26**, 173–191 (1999)

Gärling, T., Kwan, M-P., Golledge, R.G.: Computation-process modelling of household activity scheduling. Transp Res B **28**(5), 355–364 (1994)

Gärling T., Garvill J.: Psychological explanation of participation in everyday activities. In: Gärling, T., Golledge R.G. (ed.) Behaviour and Environment: Psychological and Geographical Approach, Amsterdam, New York, Elsevier Science Publisher (1993)

Habib, K.M.N., Miller, E.J.: A conceptual model of activity-based travel demand model. In: Proceedings of PROCESSUS Second International Colloquium on the Behavioural Foundation of Land-Use Transportation Model: Frameworks, Models and Applications, Toronto, 12–15 June (2005)

Habib, K.M.N., Miller, E.J.: Modelling activity generation: a utility based model for activity agenda formation. Paper presented at the 11th international conference on travel behaviour research, Kyoto (2006a)

Habib, K.M.N., Miller, E.J.: Modelling individuals' frequency and time allocation behaviour for shopping activities considering household level random effects. In: Transportation Research Record: Journal of the Transportation Research Board 1985, TRB, National Research Council, Washington DC, pp. 78–87 (2006b)

Habib, K.M.N., Miller, E.J.: Modelling skeleton components of workers' daily activity schedules. In Transportation Research Record: Journal of the Transportation Research Board 1985, TRB, National Research Council, Washington DC, pp. 88–97 (2006c)

Hicks J.R.: Value and Capital. 2nd edn Oxford University Press, London (1946)

Kitamura, R., Yamamato, T., Susilo, Y.O., Axhausen, K.W.: On the day-to-day variability of prism vertex location. Paper presented at 82nd annual meeting of the transportation research board, Washington, DC 2003

Kuhn, H.W., Tucker, A.W.: Nonlinear programming. In: Neyman, J. (ed.) Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, pp. 481–492, University of California Press, Berkeley, CA (1951)

Litwin, M.: Dynamic Household Activity Scheduling Process. Ph.D Thesis, Department of Civil Engineering, University of Toronto (2005)

Miller, E.J., Roorda, M.J.: Prototype model of household activity/travel scheduling (TASHA). In: Transportation Research Record: Journal of the Transportation Research Board, No. 1831, TRB, National Research Council, Washington, DC, pp. 114–121 (2003)

Miller, E.J.: Project-based activity scheduling for household and person agents. In: Mahmassani, H. (ed.) Flow, dynamics and human interactions-Proceedings of the 16th International Symposium on Transportation and Traffic Theory, Elsevier, New York (2005)

Pendyala, R., Kitamura, R., Kikuchi, A., Yamamato, T., Fujii, S.: FAMOS: the Florida activity mobility simulator. Paper presented at 83rd Annual Meeting of the Transportation Research Board, Washington, DC 2004

Roorda, M.J., Miller, E.J.: Toronto area panel survey: demonstrating the benefits of a multiple instrument panel survey. Preprint CD of Seventh International Conference on Travel Methods, Costa Rica, 1–2 August 2004

Salvini, P., Miller, E.J.: ILUTE: an operational prototype of comprehensive microsimulation model of urban system. Network Spatial Econ **5**, 217–234 (2005)

Schlich, R., Axhausen, K.W.: Habitual travel behaviour: evidence from a six-week travel diary. Transportation **30**, 13–36 (2003) von Haefen R.H., Phaneuf D.J.: Kuhn-Tucker demand system approach to non-market valuation. In: Scrapa, R., Alberini, A. (eds) Applications of Simulation Methods in Environment and Resource Economics, Kluwer Academic Publishers, Boston/Dordetch/London (2004)

Winston, G.C.: Activity choice: a new approach to economic behaviour. J Econom Behav Organ **8**, 567–588 (1987)

Author Biographies

Khandker M. Nurul Habib is an assistant professor of Civil & Environmental Engineering at the University of Alberta. He completed his PhD from the University of Toronto in 2007 and was a Post Doctoral Fellow during 2007–2008. His research interest includes econometric modelling, activity-based travel demand modelling, microsimulation, land use-transportation interactions, environmental impact of transportation, transportation policy analysis, transportation demand-supply interaction and advanced statistical theory.

Eric J. Miller is Behen-Tanenbaum Professor of Civil Engineering at the University of Toronto where he is also Director of the Urban Transportation Research and Advancement Centre (UTRAC). His research interests include integrated land-use/transportation modelling, activity-based travel demand modelling, microsimulation and sustainable transportation planning.