The Validity of Using Activity Type to Structure Tour-based Scheduling Models

Sean T. Doherty
Dept. of Geography & Environmental Studies
Wilfrid Laurier University
Waterloo, Ontario, Canada N2L 3C5
Tel : 519-884-1970; Fax: 519-725-1342
Email: sdoherty@wlu.ca

Abolfazl (Kouros) Mohammadian
Dept. of Civil & Materials Engineering
University of Illinois at Chicago
Chicago, IL 60607-7023
Tel: 312-996-9840; Fax: 312-996-2426
Email: kouros@uic.edu

Abstract

A unique set of activity scheduling data is utilized in this paper to provide much needed empirical analysis of the sequence in which activities are planned in everyday life. This is used to assess the validity of the assumption that activities are planned in accordance to a fixed hierarchy of activity *types*: mandatory activities first (work/school), followed by joint maintenance, joint discretionary, allocated maintenance, and individual discretionary activities. Such an assumption is typical of current generation of activity and tour-based travel demand models. However, the empirical results clearly do not support such assumptions. For instance, within out-of-home tours that had a mandatory activity, in less than half the cases was the mandatory activity actually planned first; remaining activity types also did not take any particular precedence in the planning sequence. Given this, a search was made for the more salient attributes of activities (beyond activity type) that would better predict how they are planned within tours. Several ordered response choice models for different tour sizes were developed for this purpose, predicting the choice order of the 1st, 2nd, 3rd, etc. planned activity in the tour as a function of activity type, activity characteristics (duration, frequency, travel time, and involved persons), and individual characteristics. Activity duration played the most significant role in the models compared to any other single variable, wherein longer duration activities tending to be planned much earlier in tours. This strongly suggests that the amount of time-use, rather than the nature of the event as indicated by activity type, is a primary driver of within-tour planning order and offers potential for a much improved and valid fit.

Keywords: activity scheduling, tours, mandatory/discretionary activities, ordered choice model.

INTRODUCTION

In the last decade researchers have been making significant advances in the development of activity-based urban travel demand forecasting models that go beyond the traditional 4-step process. The basic premise of these models is a recognition that observed travel patterns are the result of an underlying activity scheduling process. This scheduling process involves the planning and execution of peoples' activities and trips over time, space, and across individuals in a household.

Two basic approaches to modeling activity schedules include event-oriented activity scheduling process simulation models and simultaneous tour-based approaches. The scheduling process simulation models attempt to replicate the sequences in which activities and travel are "planned" (i.e. added to a schedule), modified and subsequently executed during a continuous process of schedule construction (e.g. Ettema et al. 1993; Gärling et al. 1998; Arentze and Timmermans 2000). Note especially that the planning/scheduling decision sequence may differ fundamentally from the observed sequence of executed activities/trips. Tour-based models first deconstruct a persons' day into a set of primary and secondary "tours" (defined as the travel from home to one or more activity locations and back home again), then model the number, purpose and sequence of activity stops by destination, mode of travel, and time of day based on a set sequence of discrete choice models (Shiftan 1998; Bowman and Ben-Akiva 2001; Limanond et al. 2005; Yagi and Mohammadian 2006). The most current generation of activity/tour-based regional travel demand models systems are based on a sequence of discrete choice models applied in a micro-simulation fashion (Vovsha et al. 2005; Yagi and Mohammadian 2006).

In practice, both these modeling approaches involve making assumptions concerning the sequence in which the key decisions and choices are made. This is largely a reflection of available data, rather than theoretical shortcomings, as data collection is limited to observed

activity and travel patterns collected via diary techniques which provide little to no insights into underlying decision processes. There is a persistent assumption that the highest "priority" or most "important" activity *types* are planned and added first to a schedule or tour. For example, "work", "school" and other "mandatory" or recurring activities are often assumed to be fixed in space and time and thus higher "priority", whereas more "discretionary" activities such as "leisure", "visiting friends", "entertainment" and especially "shopping", are assumed to be more flexible and thus of lesser priority.

Based on these assumptions, fixed or high priority activities are assumed to be planned/modelled first in emerging trip chain (e.g. Kitamura et al. 2000), activity scheduling (e.g. Arentze and Timmermans 2000; Miller and Roorda 2003) and tour-based models (e.g. Shiftan 1998; Bowman and Ben-Akiva 2001; Limanond et al. 2005). However, several authors (Shiftan 1998; Limanond et al. 2005; Yagi and Mohammadian 2006) recognize that activity type alone may not be sufficient in capturing the inherent flexibility and priority of activities, and instead attempt to adopt combined prioritization schemes that recognize that activities farthest from origin, or activities with longer durations, are of higher priority. The validity of these assumptions, and the limitations they may place on forecasting potential, has been scarcely addressed.

Current Generation of Tour-based Models

As background to the empirical analysis presented in this paper, the most recently developed tour-based models are reviewed in more detail. The first operational tour-based model that was intended to replace the four step model with the generation of an entire days' activity travel pattern was developed by Bowman and Ben-Akiva (2001). In the Boston prototype, a simultaneous discrete choice framework was used to model 1) a person's primary activity of the day (including option to remain at home); 2) the type of tour for the primary activity, including the number, purpose and sequence of activity stops; and 3) the number and purpose of a person's

secondary tours. The model includes the choice of time of day, destination and mode of travel for each tour, conditioned by the choice of activity pattern. A subsequent model was developed for Portland in order to look at the response to peak-hour congestion pricing (Bowman et al. 1998). Primary activities were defined as work, followed by work related, school, and all other purposes.

This tour-based modeling approach has been further enhanced and applied as regional travel demand model systems in the Metropolitan areas of Portland, San Francisco, New York, Columbus, and Atlanta (see Goulias et al. 2004; or Vovsha et al. 2005 for more details) as well as elsewhere (see Davidson et al. 2007). The modeling systems continue to be based on a sequence of discrete choice models applied in a micro-simulation fashion, but include many enhancements such as using the household for some choice dimensions and person for other dimensions .

In practice, the means by which household interactions are captured, and by which activities are prioritized for planning within primary and secondary tours, is by activity type. In the most recent tour-based models (Vovsha et al. 2005) and more generalize activity-based travel demand models (Davidson et al. 2007) a three-level hierarchy is used, wherein "mandatory" activities (including going to work, university, or school) are selected for planning first, followed by "maintenance" activities (including shopping, banking, visiting doctor, etc) and "discretionary" activities (including social and recreational activities, eating out, etc). To capture household effects, these are further defined according to whether they were conducted jointly, individually (no household members involved), or allocated to a particularly household member. It is then assumed that:

> "all else being equal, there is a predetermined structure of priorities in the activity
> generation and scheduling procedure along both dimensions. Mandatory
> activities take precedence over maintenance activities, while maintenance
> activities take precedence over discretionary activities. Joint activities are

considered superior to allocated activities, while allocated activities are in turn considered superior to individual activities." (Vovsha et al. 2005, p. 400)

A fixed order is thus assumed for the selection of activities to be included in tours, as follows:

1. Individual mandatory activities

2. Joint maintenance activities

3. Joint discretionary activities

4. Allocated maintenance activities

5. Individual discretionary activities

The implication for the model structure is that later decisions in the schedule building process are "conditioned" by earlier decisions and limited with respect to remaining time windows, which is a key feature since it essentially narrows the choice set down in lower levels making choice of timing, location, and mode much simpler and perhaps behaviorally realistic. The trouble is that so much depends on the above assumption – the more cases, situations, and peoples that it does not hold for, the less realistic the lower level decisions become. For example, as work (mandatory) becomes more flexible for a growing (but limited) portion of the population who telework, it will most likely decrease in priority relative to say, driving kids to school (maintenance) or going to the gym (discretionary). Thus, in reality, the structure associated with children's time-tables, or gym opening hours, play a stronger role in conditioning when work is fit in. However, under the current structure, there is no capacity to account for this, since the "mandatory" activities chosen for the primary tour are fixed. There are many other trends that will continue to erode the fixed structure (e.g. growth and increased dispersion of leisure activities), not to mention that such assumptions may not hold in all cultures.

The real concern at this point is not whether the assumption is broken, but by how much. The challenge is that there is virtually no empirical evidence available to support the above hierarchy, nor search for alternatives. This point is not missed by tour-based modelers who

recognize that "Further research is also needed to better understanding the interrelationship between activity generation and scheduling stages, and their positioning in the model system hierarchy. Similarly, relationships should be further explored between such dimensions as activity locations/durations and tour configuration, in terms of a distribution of activity episodes by tours" (Vovsha et al. 2005, p. 402).

OBJECTIVES

With respect to the latest in tour-based modelling structures used to forecast travel demand, a clear practical agenda for future work is to explore and validate the actual sequence in which activities are planned within tours. The results of this exploration will serve to: a) assess the validity of past rules concerning the fixed sequence of planning by activity type; b) lead to the suggestions of alternative rules/hierarchies based on activity type and; c) lead to suggestions of new functions/rules of within-tour planning sequence incorporating activity characteristics beyond activity type. To accomplish this, the analysis of empirical data on the sequence in which activities are planned (not just observed), in combination with the recording of activity type and involved household members, will be required.

METHODS

Data

The dataset used for this study was derived from a Computerized Household Activity Scheduling Elicitor (CHASE) survey conducted in Toronto in 2002-2003 with 271 households, including 426 adults. CHASE is unique in that it focuses on capturing multi-day information on observed activity-travel patterns whilst also tracing the underlying activity scheduling process. It does so by providing a scheduling program to respondents on a computer that provides a forum for them to self-report their on-going scheduling decisions over the course of a seven-day period. The scheduling program tracks when and why the various scheduling decisions are made, and displays back to users the results of their scheduling efforts in the form of observed

activity-travel patterns. To the user, the program appears not unlike a typical scheduler or day-planner as shown in Figure 1. Users are asked to add, modify and delete activities anywhere on their schedule as their plans evolve over a multi-day period. The final state of their activities on screen is taken as observed patterns, replicating what is captured by traditional activity diaries.

A detailed description of the design and conduct of this survey, characteristics of the sample, and data quality can be found in Doherty et al. (2004), whilst a more general description of this emerging type of survey method can be found in Doherty and Miller (2000) and Doherty (2002). Their most distinguishing feature (which is also most important to this paper) is the capturing of the sequence in which activities are planned, not just their observed sequence. This was accomplished in part, by having subjects record their preplanned activities prior to the study week, and continue to do so throughout the study week, in a prospective fashion (and thus, we know they are preplanned; but we ask for more detail on how far in-advance they were planned anyway, just to be more specific). This is followed up by regular recording of further additions, modifications and deletions over the course of week in a *prospective* and *retrospective* (after the fact) fashion. When entered (prospectively or retrospectively), subjects are immediately asked to indicate more precisely when it was planned (although we could safely assume a course level of planning without this) by asking subjects *"When did you originally plan this activity?"*, as shown in Figure 2. All options in the "Initial prompt" in Figure 2 had at least one follow-up prompt (a, b, c, d, or e in the Figure) once the user clicks "next" in the dialog, designed to more precisely assess the timing of the decision. For instance, decisions made the "same day" were recorded by subjects as being made "Just before the activity", "1-2 hours before, ">2 hours before", etc. When compared to the start-time of the activity, an accurate estimate can be made of the time-of-day the decision was made. If subjects indicated the decision was made "Before the day of the activity", they were asked what specific day this week

or last week, or if it was "Several weeks ago", "Months ago", or "Year ago". If they indicated that the activity was "part of a regular routine", they were asked to indicate when they "last thought about the planning of this activity" or "for how long have you been doing this activity in this way" – the responses to which were used similarly to estimate when it was planned. Taken together, this detailed set of prompts allowed the sequence in which activities were planned to be determined with a high level of precision – and thus form the basis of analysis in this paper. For a more detailed accounting of measurement challenges and validity of this question in the survey can be found in Doherty (2005).

A wide range of in- and out-of-home activity types were systematically probed for and organized into nine main groups (Work/School, Shopping, Services, Recreation/Entertainment, Social, Basic Needs, Household Obligations, Drop-off/Pick-up, Other) and 66 specific activity types (Grocery shopping, Internet shopping , exercise, active sports, etc.). In keeping with the classification of activities used in tour-based models, these specific activity types were re-grouped into three classes as shown in Table 1.

Equally important to this paper is the detail on the involved person for each activity that was captured in the survey. For each activity, subjects were asked to select the names of people that were directly involved with the activity from a pull-down list as shown in Figure 1. This list was initially constructed during an upfront interview, wherein the first name and relationship of each person was noted (son, daughter, spouse, relative, friend, etc.). During the survey week, subjects could add to this list by specifying new names and relationships. In addition, parents with young children were asked to specify which of their children were "under their care" during the activity, also shown in Figure 1. This information was used to further differentiate the three main activity types from above into five categories used in the most recent versions of tour models (Vovsha et al. 2005), including:

1. Individual mandatory activities (i.e. not conducted with other household members, but may be with other non-household members);

2. Joint maintenance activities (i.e. conducted with other household members)

3. Joint discretionary activities (i.e. conducted with other household members)

4. Allocated maintenance activities (i.e. not conducted with other household members, but may be with other non-household members)

5. Individual discretionary activities (i.e. not conducted with other household members, but may be with other non-household members).

Almost all activities fell into this categorization, except a portion of individual mandatory activities conducted with other household members.

Sample

The 271 household sample was drawn from a randomly selected list of household telephone numbers and addresses from the telephone directory for the study area. The study area includes the central and suburban areas in the City of Toronto. Out of a total of 1935 numbers in the sample, 1637 households were contacted successfully. Given that 271 agreed to participate, the effective response rate was 16.6%. Considering the level of commitment expected of the respondent, this is considered to be an acceptable response rate. In 75% of households, all adult members completed the CHASE survey, with an average of 1.6 adults participating per household.

Respondents recruited for the survey were found to be reasonably representative of the total population. The proportion of responding households in each main area of the city (central city, mature suburbs, and newer suburbs) closely matched the proportion of actual households in these areas as derived from the 2001 Census. When compared against 2001 data from the Transportation Tomorrow Survey (TTS), a travel survey conducted on 5% of the households in the Greater Toronto Area and weighted to reflect census household counts (Data Management Group 2003), household size, number of vehicles and dwelling type distribution were all found to be a close match to the population. The one difference was a slightly over-sampling of single

family homes and townhouses versus apartment dwellers. Full details on these comparisons and sample characteristics can be found in Doherty et al. (2004).

From the 271 households, a total of 453 individuals participated in the CHASE survey. After an extensive data quality screening process, 354 of these were chosen for the analysis in this paper. Selection was based on a combined rule requiring a minimum number of login days (3) and a maximum amount of missing activity time for the week (generally less than 7 hours). Full details on data quality assessment can be found in Doherty (2004). The 354 subjects reported a total of 28,907 observed activities. Of these, the timing of when they were originally planned could not be recalled (or was unknown) in 16% of the cases. This left a final sample of 24,238 activities.

Analytical Techniques

Descriptive analysis of the frequency to which the five categories of activities are planned 1st, 2nd, etc. within activity tours is first presented. An out-of-home tour consists of a sequence of activities located outside the home and bounded by in-home activities (e.g. work-shopping-visiting). An in-home tour simply consists of a sequence of activities that occur exclusively within the home location, and are bounded by out-of-home activities (e.g. Meal-Cleaning-TV). More than one of each type of tour may occur in a day.

The descriptive results provide a basis to assess the validity of assuming that activities are planned within tours in a fixed sequence by activity type. In light of these results, an alternative approach is developed wherein planning order is modelled as a function of activity characteristics and activity type in an attempt improve predictive power, flexibility and behavioural validity. This reflects the search for the more salient attributes of activities that drive how they are scheduled, versus the often arbitrary assignment by activity type.

Ordered Response Models

The most common framework used for choice behavior modeling in recent years has been the multinomial logit (MNL) model. However, MNL model has the "independence of irrelevant alternatives" (IIA) property, which requires that the random utility components of the utilities of choice alternatives be independent. As a result, the MNL model assumes categories of dependent variables are independent and therefore, MNL model have difficulty dealing with an ordered classification of the dependent variable.

In the current study, the observed activity planning sequences in tours form an ordered structure in which activities can be the 1st activity planned in the tour, the 2nd, etc. (an ordinal categorical variable). Given this, an ordered response choice model was considered. The structure of an ordered choice model involves a latent regression that uses thresholds to separate the successive activity sequence orders on the underlying latent scale. Since the ordered response model is capable of replicating all activity planning sequences, it offers an attractive alternative to the standard multinomial logit model.

There are numerous studies in the literature that use ordered response approach to model a variety of transportation related problems. Bhat and Pulugurta (1998) developed an ordered choice model to analyze automobile ownership decision behavior of households. Stinson and Bhat (2004) used an ordered response structure and developed a model to predict frequency of bicycle commute. Kockelman et al. (2001) used an ordered probit approach to model the perception of sidewalk-section crossing difficulty. Kemperman et al. (2003) used an ordered logit model to predict the duration of theme park visitors' activities.

The ordered choice model is a flexible model which can show the probability of an event as a function of its attributes. For example, consider the case when examining activity scheduling order choice in a tour. The probability that a decision-maker will add an activity to his tour is dependent on the duration of the tour and the time already assigned to other activities

in the tour, conditional on the fact that the person still has not planned this activity. The

underlying model in this example can be based on either an ordered probit, ordered logit, or

other ordered functional forms in which an unknown parameter can be estimated for each

ordered interval. Two most commonly used functional forms of ordered response models are

ordered logit and ordered probit models developed by Zavoina and McElevy (Zaviona and

McElevy 1975). It can be shown that an ordered response model is built around a latent

regression in the same manner as the binomial probit model (Greene 2003):

$$y_i^* = \beta'x_i + \varepsilon_i \qquad\qquad [1]$$

where x_i is the vector of explanatory variables and β' is a corresponding vector of coefficients to

be estimated. The observation mechanism results from a complete censoring of the latent

dependent variable as follows:

$$
\begin{aligned}
y_i &= 0 & \text{if} \quad & y_i \le \alpha_0 \\
&= 1 & \text{if} \quad & \alpha_0 < y_i \le \alpha_1 \\
&= 2 & \text{if} \quad & \alpha_1 < y_i \le \alpha_2 \qquad\qquad [2]\\
&\ldots \\
&= J & \text{if} \quad & y_i > \alpha_{J-1}
\end{aligned}
$$

In these equations, y_i is the observed counterpart to y_i^*. The αs are unknown parameters

defining the threshold bounds and are estimated with β. The ordered probit model is based on

the normal distribution of the error term (Zaviona and McElevy 1975). The variance of ε_i is

assumed to be one since as long as y_i^*, β, and ε_i are unobserved, no scaling of the underlying

model can be assumed from the data, although it is also possible to assume heteroscedasticity

for the variance of ε_i and develop a heteroscedastic model. Since αs are free parameters, there

is no significance to the unit distance between the set of observed values of y (Greene 2002).

Therefore, even though in the case of modeling the planning sequences of activities, the

dependent variable is not continuous, application of an ordered response model can be useful.

In the case of ordered probit, the probability associated with each term in Equation 2 is determined by the threshold bounds and the latent response y_i^*. This can be written as:

$$\text{Prob}\,(y_i = j \mid x) = \Phi(\alpha_j - \beta'x_i) - \Phi(\alpha_{j-1} - \beta'x_i)$$ [3]

where Φ is the cumulative standard normal distribution function. Model estimates are obtained by maximizing the likelihood function given by:

$$\ln L = \sum_i \ln L_i = \sum_i \ln[P(y_i = j)]$$ [4]

The model can be estimated either with individual data, with y_i = 0, 1, 2,… or with grouped data. Other assumptions for the underlying distribution function of ε_i will result in various forms of ordered choice models. The ordered logit model results from the assumption that ε_i has a standard logistic distribution instead of standard normal.

Model Variables and Structure

The ordered response modeling structure used in this study was based on the hypothesis that a single discrete variable represents the latent activity planning sequence in the tour. Therefore, the dependent variable in the model is the sequence order in which the activity has been planned in a tour. This is a discrete variable and can take on values 1, 2, 3, etc. Models were developed only for the 3587 out-of-home tours involving 6647 activities. It is assumed that the order of each activity in the tour is treated independently; therefore no more than one activity can be planned *first*, *second*, *etc*. It should be noted that an alternative approach to model within tour activity planning sequence would be to use a series of MNL models, where the choice is made between all of the remaining activities for which one to plan, rather than independent ordered models for each activity. For example, for a four activity tour, modeler could first run an MNL to pick which activity is planned first, then run the model again to pick the activity to plan second, etc. until all activities are planned. In this way, the activities in the tour are competing

with each other for each planning slot rather than remaining independent. Developing such a model remains as a future research task.

Three groups of explanatory variables were considered. The first group consisted of binary variables for the five typical activity type labels. The second group consisted of the following observed activity characteristics:

- duration (minutes)
- frequency (observed weekly)
- travel time *to* the activity (minutes)
- involved persons indicator (0 alone; 1 one+ persons involved)

The third group consisted of the following individual and household characteristics:

- Gender
- Age (years)
- Income (Canadian dollars)
- Household size (adults plus children aged up to 19)
- Total # difference activities conducted during the week

Various combinations of different variables were considered within each group and interaction effects of variables across the various groups were also explored. Using intuitive considerations and utilizing results learned from earlier research (Mohammadian and Doherty 2005), a systematic process was used to select suitable variables in the model and eliminate statistically insignificant variables. In order to interpret the results, one should note that a positive (negative) coefficient on a variable indicates that an increase in the variable has the effect of increasing (decreasing) the planning sequence order – i.e. being planned later (earlier).

RESULTS

Descriptive Analysis of Within-tour Planning Sequence

An initial breakdown of frequency of the 24,238 activities by activity type is provided in Table 2 by location. Note that only the first five categories are normally considered in emerging tour-based models, whereas for completeness, two additional categories are listed ("joint" mandatory and night sleep). Looking only at the 6,647 activities occurring out of home, about 1/3 are

allocated maintenance activities, 1/4 are individual mandatory, 1/5 individual discretionary

activities, 1/10 joint discretionary, and 8% joint maintenance. A small portion of *out-of-home*

mandatory activities (2.5%) were found to be conducted jointly with other household members.

In subsequent analysis, these are merged into the first category, as it is largely assumed they

retained their "mandatory" nature despite the involvement of other household members.

Compared to out-of-home, the proportion of mandatory activities in- home goes down

substantially (3%), whereas individual discretionary and night sleep go up.

In terms of planning sequence of activities, higher numbers of daily activities (and hence

higher planning sequence values) were evident for in-home tours versus out-of-home. In-home,

planning sequences were as high as 30 (meaning, someone had 30 in-home activities in a day,

planned in sequence from 1, 2, ... 30), although 99% of them were below 19, and 95% below

13. This largely reflects a high level of detail on in-home activities captured by the survey. The

highest number of out-of-home activities in a day was 24 (one case), although 99% were below

10 and 95% below 7.

The more revealing analysis is of the planning order/sequence within tours by activity

type, as shown in Table 3 for out-of-home tours of varying sizes (up to 5 activity tours; sample

sizes dwindle thereafter). The most significant trend concerns the frequency that mandatory

activities are planned first, which is only 60% for two-activity tours, and falls steadily to as few as

30% in larger tours. On the whole, mandatory activities are planned first less than half the time

(or about 45% of tours with 2-5 activities). Perhaps equally interesting is the more uniform

distribution of planning order amounts other activity types with few significantly high or low

values – never fewer than 10% and rarely over 30%. An even more uniform distribution of

results was found for in-home activity tours. Overall, these results do not reflect any particular

fixed order or hierarchy, and are certainly far from always having mandatory activities planned

first and individual discretionary activities planned last as popularly assumed.

Ordered Response Models of Within-tour Planning Sequence

Recognizing the questionable validity of assigning the within-tour planning order statically by activity type, the planning sequence was modelled as a function of the activity characteristics including duration, frequency, travel time and involved persons. As shown in Table 3, the percentage of observed out of home tours with 5 or more activities is small, therefore only tours with up to 4 planning order categories are modeled in this study. Each activity in the tour is treated independently and 3 different models are estimated for different tour sizes (i.e., 4, 3, or 2 activities in the tour). Therefore, tour sequences are not estimated using the same model.

Overall, use of activity characteristic variables led to an overall good fit as shown in part (a) of Table 4. The overall good model fit, strong parameter estimates, and highly significant chi-squared statistic suggest that the model is capable of estimating the order and planning sequence of all activities in the tour (1st, 2nd, etc.), not just the first one. Among activity characteristics, only duration and travel time were found to be significant in all 3 models. The sign of the parameters are in the expected direction – wherein activities with higher durations and higher travel times tend to be planned earlier in the tour formation.

The results of adding in activity type variables to capture unmeasured attributes of activities had the effect of slight increase in model goodness-of-fit, as shown in part (b) of the Table 4. These models retain Duration and Travel Time from the previous model, plus activity type variables. As expected, in tours with 3 or 4 activities, mandatory activities tend to be planned earlier in tours versus other activity types. However, in the tours with only two activities, in the absence of mandatory activity variable in the model, discretionary activities seem to be planned earlier. This may be as a result of the inclusion of duration, as mandatory activities tend to being normally long in duration.

It is interesting to note that there is no fixed planning order for other activity types. The observed variation in the order of activities in tours with different sizes suggest that activities are

not planned according to a fixed hierarchy of activity types, rather the activity characteristics play more important role in identifying the order of that activity in the tour.

On a final note, additional sets of models were estimated at every stage in the analysis incorporating individual and household characteristics. However, in all cases their inclusion did not lead to an improved model fit over the above variables, and so results are not presented.

DISCUSSION AND CONCLUSIONS

In this paper, a unique set of activity scheduling data was utilized to provide much needed empirical analysis of the sequence in which activities are planned in everyday life. This analysis was closely tied to the activity types used in emerging tour-based activity scheduling models. This was done to test the validity of the assumption that activities are planned according to a fixed hierarchy of activity types including mandatory, maintenance, and discretionary activities.

The descriptive analysis was meant to demonstrate as clearly as possible the relationship between planning sequence and activity type. As a result, the message is clear and simple. Within out-of-home tours that had a mandatory activity, in less than half the cases was the mandatory activity actually planned first; the distribution of remaining planned activities was roughly uniform, with barely a hint of any particular activity type taking precedence in planning sequence.

These results provide little support to the assumption that activity types are systematically planned in the order prescribed in tour-based models. In particular, assumptions such as those adopted by Vovsha et al. (2005), as in most other emerging activity-based models (Davidson et al. 2007), that "maintenance activities take precedence over discretionary activities", "Joint activities are considered superior to allocated activities", and "allocated activities are in turn considered superior to individual activities" are clearly not supported by

these results. Even the assumption that mandatory activities are always planned first would likely not hold in about one-half of the cases – a significant concern.

Taken together, and considering the typical fixed planning sequence assumption of past models (Individual mandatory planned first, then Joint maintenance, Joint discretionary, Allocated maintenance, and Individual discretionary), Table 3 suggests that for 2-activity tours, only 60% of first planned activities would be correct (assuming that the generation of activities within the tour is known without error). For tours with 3 activities, the % correct would be much less; 42% correct on first <u>and</u> about a 30% chance at best of correctly predicting among the remaining non-mandatory activity types; taken together, less than about a 13% correct rate. The % correct on tours with 4 or more activities would be even less. Overall, the % correct for all tours would likely be no more than 30% at best.

Overall, the results provide evidence that activity type is at best, weakly associated with the actual planning sequencing of activities within tours, and that the validity of any assumptions of a static fixed planning sequences by activity type are likely to be weak in reality, and a potential hindrance to the ability of tour-based models to make behaviorally valid forecasts. Given this, a search was made for the more salient attributes of activities that would better explain how they are planned and sequenced. Several ordered response choice models for different tour sizes were developed for this purpose, predicting the choice order of the 1st, 2nd, 3rd, etc. planned activity in the tour as a function of activity characteristics and activity type.   It is worth noting that attempts were made to keep the model as simple as possible, otherwise the explanatory power of the model can be improved even further by incorporating other detailed variables. Activity duration played the most significant role in the models compared to any other single variable, wherein longer duration activities tending to be planned much earlier in tours. In fact, adding additional variables beyond duration improved the overall fit only slightly, including activity type variables and travel time. This strongly suggests that the amount of time-used,

rather than the nature of the event as indicated by activity type, is a primary driver of within-tour planning order and offers potential for a much improved and valid fit.

An alternative, but less parsimonious approach would be to use a combination of frequency, travel time and activity type indicator variables (Individual Mandatory, Allocated Maintenance, and Individual discretionary activity binary variables) to achieve a somewhat similar model fit. Such an approach could be useful if the duration variable is unavailable or if additional variables are desired for specific policy analysis scenarios. But even though it has a similar model fit, using duration would have several additional compelling advantage including: it is a continuous variable; the model has fewer variables; behaviorally, duration it is a more *valid* "driver" of the decision process, since increased duration can be more naturally associated with increased value of time, and hence increased "priority" (whereas activity type is really just a surrogate for these factors); measuring duration on surveys is a straightforward matter of comparing start and end times (rather than for example, complex check-boxes); duration, being time-based, is more universally understood regardless of cultural or socio-demographic background (the same cannot be said for interpretation of activity categories, not to mention language and cultural-interpretation concerns); perhaps most importantly, it allows any variety of non-standard activity types (like volunteer activities, non-worker activities, etc.) to become high priority for planning, and thus more amendable to a variety of population segments. Interestingly, involved persons, measured as either the total # or as a binary variable (alone versus with others), did not have a significant effect on any of the tested or presented models.

At a minimum, if activity types are to be used in isolation as a basis for planning order, the order should not be fixed – instead, allow a variety of activity types to be planned first/earlier in a more probabilistic fashion, using frequencies such as those provided in this paper as a guide. However, the problem with this approach that remains is that activity types are merely a label given to an activity, and it is left to the researcher to assume that a given activity type

alone is sufficient to capture how it is planned by all people in all situations – which is clearly not the case, as even discretionary activities are often planned first.

What is needed is consideration of other activity characteristics (beyond type label) that alone or in combination, affect activity planning. Results presented here point to the use of activity duration for this purpose, rather than involved persons as adopted in several emerging tour-based models (i.e. through definition of joint versus individual activity types). For instance, a new set of ad-hoc rules could be developed that would more realistically replicate the variety of activities chosen earliest for planning within tours (not just mandatory ones), such as assuming the longest duration activity is planned first, followed by other mandatory activities, etc. A more elaborate approach would be to adopt a choice-based multivariate function for predicting the ordering of activities based on a combination of activity characteristics (not just type) including at least duration – this would provide much more flexibility to handle a variety of activity and tour situations people face, rather than just fixed sequences. It would also improve the usefulness of the model for assessing the impacts of policies or other changes that inherently affect activity characteristics and not necessarily their generic labels. For instance, incorporation of duration in a within-tour planning function/rule would make the model sensitive to the impacts of shortened work hours or increased leisure time – effects that would be difficult to account for using an activity type only decision rule. The subsequent effects this would have on tour formation and subsequent observed patterns could be significant, such as an increased tendency for leisure to be planned earlier in tours compared to work, thereby affecting the very nature and available time windows of work.

Overall, at this stage, it would seem prudent to go beyond ever more detailed activity type hierarchies that we feel better explain planning sequence (such as differentiating the current five-category classification further by two or more categories of duration), as there will invariably be frequent exceptions to every rule like this (if these results are any indicator),

whether because of a change in location, person, or situation, or as a result of policy changes. Further testing is needed of a variety of functions/rules of within-tour planning sequence that are sensitive to a wide variety of activity characteristics beyond type, as well as to persons, household and situational factors. Perhaps it is also time we take seriously the revision of our larger scale travel survey to provide a continuous source of such data, rather than focusing so intently on capture activity types alone.

ACKNOWLEDGEMENTS

REFERENCES

Arentze TA and HJP Timmermans (2000). *ALBATROSS: A learning based transportation oriented simulation system*. Eindhoven, The Netherlands, The European Institute of Retailing and Services Studies.

Bhat CR and V Pulugurta (1998). A Comparison of Two Alternative Behavioral Choice Mechanisms for Household Auto Ownership Decisions. *Transportation Research B* 32: 61-77.

Bowman JL and ME Ben-Akiva (2001). Activity-based Disaggregate Travel Demand Model System With Activity Schedules. *Transportation Research A* 35(1): 1-28.

Bowman JL, M Bradley, Y Shiftan, KT Lawton and M Ben-Akiva (1998). Demonstration of an Activity-Based Model System for Portland. 8th World Conference on Transport Research, Antwerp, Belgium.

Data Management Group (2003). Transportation Tomorrow Survey 2001: Design and Conduct of the Survey. Toronto, Joint Program in Transportation, University of Toronto.

Davidson W, R Donnelly, P Vovsha, J Freedman, S Ruegg, J Hicks, J Castiglione and R Picado (2007). Synthesis of first practices and operational research approaches in activity-based travel demand modeling. *Transportation Research A* 41(5): 464-488.

Doherty ST (2002). Interactive Methods for Activity Scheduling Processes. In *Transportation systems planning: methods and applications*. K Goulias, Ed. New York, CRC Press**:** 7-1 to 7-25.

Doherty ST (2004). Rules for assessing activity scheduling survey respondents' data quality. *Journal of the Transportation Research Board*(1870): 109-115.

Doherty, S. T. (2005). How far in advance are activities planned? Measurement challenges and analysis. *Journal of the Transportation Research Board* 1926: 41-49.

Doherty ST and EJ Miller (2000). A Computerized Household Activity Scheduling Survey. *Transportation* 27(1): 75-97.

Doherty ST, E Nemeth, M Roorda and EJ Miller (2004). Design and Assessment of the Toronto Area Computerized Household Activity Scheduling Survey. *Journal of the Transportation Research Board* 1894: 140-149.

Ettema D, A Borgers and H Timmermans (1993). Simulation model of activity scheduling behavior. *Transportation Research Record* 1413: 1-11.

Gärling T, T Kalén, J Romanus and M Selart (1998). Computer simulation of household activity scheduling. *Environment and Planning A* 30: 665-679.

Goulias KG, M Bradley, V Noronha, R Golledge and PS Vovsha (2004). Data needs for innovative modelling workshop. National Household Travel Survey Conference: Understanding our Nation's Travel, Washington D.C., November 1-2, Available at http://trb.org/Conferences/NHTS/Workshop-DataNeeds.pdf.

Greene W (2002). *Limdep Version 8.0, Econometric Modeling Guide*. Plainview, NY, Econometric Software, Inc.

Greene W (2003). *Econometric Analysis*. London, Prentice Hall.

Kemperman A, A. Borgers, H. Oppewal and H Timmermans. (2003). Predicting the Duration of Theme Park Visitors' Activities: An Ordered Logit Model Using Conjoint Choice Data. *Journal of Travel Research* 41(4): 375-384.

Kitamura R, C Chen, RM Pendyala and R Narayanan (2000). Micro-simulation of daily activity-travel patterns for travel demand forecasting. *Transportation* 27(1): 25 - 51.

Kockelman K, Y. Zhao and C Blanchard-Zimmerman. (2001). Meeting the Intent of ADA in Sidewalk Cross-Slope Design. *Journal of Rehabilitation Research and Development* 38(1): 101-110.

Limanond T, DA Niemeier and PL Mokhtarian (2005). Specification of a tour-based neighborhood shopping model. *Transportation* 32: 105-134.

Miller EJ and MJ Roorda (2003). A Prototype Model of Household Activity/Travel Scheduling. *Journal of the Transportation Research Board* 1831: 114-121.

Mohammadian A and ST Doherty (2005). A Mixed Logit Model of Activity Scheduling Time Horizon incorporating spatial-temporal variables. *Journal of the Transportation Research Board* 1926: 33-40.

Mohammadian A and ST Doherty (2006). Modeling Activity Scheduling Time Horizon: Duration of Time between Planning and Execution of Pre-Planned Activities. *Transportation Research A* 40(6): 475-490.

Shiftan Y (1998). Practical approach to model trip chaining. *Transportation Research Record* 1645: 17-23.

Stinson MA and C Bhat (2004). Frequency of Bicycle Commuting: Internet-Based Survey Analysis. *Transportation Research Record* 1878: 122-130.

Vovsha P, M Bradley and J Bowman (2005). Activity-based travel forecasting models in the United States: Progress since 1995 and Prospects for the Future. In *Progress in Activity-Based Analysis*. H Timmermans, Ed. Oxford, Elsevier**:** 389-414.

Yagi S and A Mohammadian (2006). An Activity-Based Microsimulation Model of Travel Demand in the Jakarta Metropolitan Area. 11th International Conference on Travel Behaviour Research. Kyoto.

Zaviona W and RD McElevy (1975). A Statistical Model for the Analysis of Ordinal Level Dependent Variables. *Journal of Mathematical Sociology* 4: 103-120.

Table 1 CHASE survey activity types categorized by their tour-based equivalent

Mandatory activities	Maintenance activities	Discretionary activities
At work	Wash/dress/pack/snacks	In-home meals
Telework	Cleaning/Maintenance	Bagged lunch
Volunteer work	Meal preparation	Restaurants
At School	Attending to children	Coffee/snack shop
Schoolwork	Attending to pets	Religious activities
Training/special classes	Other household obligations	Hobbies
Other work/school	Picking-up/dropping-off people	Exercise or active sports
	Picking-up/dropping-off meals	Spectator Events/Theatre
	Picking-up/dropping-off Snacks/drinks	Playing/parks
	Picking-up/dropping-off Other items	Regular TV programs
	Groceries shopping (<10 items)	Unspecific TV
	Groceries shopping (10+ items)	Watching videos
	Houseware shopping	Relaxing/napping/reading
	Clothing/personal items shopping	Email/internet
	Drug store shopping	Other recreation/entertainment
	Other shopping	Hosting visitors
	Medical/professional services	Visiting people
	Barber/salon/beauty	Planned social events
	Banking	Special clubs
	Other services	Telephone >10 minutes
		Other social activities
		Pleasure driving

Table 2 Activity frequency by Activity Type and Location

	Location					
	At-home		Out-of-home		All	
Activity Type	Count	(%)	Count	(%)	Count	(%)
Individual mandatory activities	523	(3.0)	1620	(24.4)	2143	(8.8)
Joint maintenance activities	664	(3.8)	508	(7.6)	1172	(4.8)
Joint discretionary activities	1256	(7.1)	705	(10.6)	1961	(8.1)
Allocated maintenance activities	5647	(32.1)	2218	(33.4)	7865	(32.4)
Individual discretionary activities	6398	(36.4)	1372	(20.6)	7770	(32.1)
Joint mandatory activities*	25	(0.1)	164	(2.5)	189	(0.8)
Night sleep*	3078	(17.5)	60	(0.9)	3138	(12.9)
TOTAL	17591	(100.0)	6647	(100.0)	24238	(100.0)

* These classes not normally included in emerging tour-based models.

Table 3 Planning order of activities within out-of-home tours

Tour Type	Planned Order	Mandatory		Joint maintenance		Joint discretionary		Allocated maintenance		Individual discretionary	
		Count	(%)	Count	(%)	Count	(%)	Count	(%)	Count	(%)
One Activity Tours	1st	592	(100.0)	128	(100.0)	208	(100.0)	539	(100.0)	464	(100.0)
	TOTAL	592	(100.0)	128	(100.0)	208	(100.0)	539	(100.0)	464	(100.0)
Two Activity Tours	1st	221	(59.7)	34	(30.9)	61	(35.1)	93	(19.2)	84	(31.7)
	2nd	149	(40.3)	76	(69.1)	113	(64.9)	392	(80.8)	181	(68.3)
	TOTAL	370	(100.0)	110	(100.0)	174	(100.0)	485	(100.0)	265	(100.0)
Three Activity Tours	1st	121	(41.9)	22	(25.9)	29	(26.4)	87	(20.3)	43	(20.8)
	2nd	78	(27.0)	20	(23.5)	33	(30.0)	106	(24.8)	68	(32.9)
	3rd	90	(31.1)	43	(50.6)	48	(43.6)	235	(54.9)	96	(46.4)
	TOTAL	289	(100.0)	85	(100.0)	110	(100.0)	428	(100.0)	207	(100.0)
Four Activity Tours	1st	76	(35.3)	9	(12.7)	16	(23.5)	56	(17.5)	37	(21.8)
	2nd	59	(27.4)	14	(19.7)	9	(13.2)	66	(20.6)	33	(19.4)
	3rd	41	(19.1)	18	(25.4)	18	(26.5)	77	(24.1)	36	(21.2)
	4th	39	(18.1)	30	(42.3)	25	(36.8)	121	(37.8)	64	(37.6)
	TOTAL	215	(100.0)	71	(100.0)	68	(100.0)	320	(100.0)	170	(100.0)
Five Activity Tours	1st	43	(29.3)	5	(10.6)	7	(13.7)	21	(11.5)	10	(13.5)
	2nd	34	(23.1)	12	(25.5)	9	(17.6)	34	(18.6)	9	(12.2)
	3rd	28	(19.0)	10	(21.3)	10	(19.6)	34	(18.6)	14	(18.9)
	4th	19	(12.9)	9	(19.1)	8	(15.7)	37	(20.2)	19	(25.7)
	5th	23	(15.6)	11	(23.4)	17	(33.3)	57	(31.1)	22	(29.7)
	TOTAL	147	(100.0)	47	(100.0)	51	(100.0)	183	(100.0)	74	(100.0)

* Note of Interpretation: read values top-to-bottom within each activity type and tour type grouping. The % values are calculated such that they correspond to the % number of occurrences the given activity was planned 1st, 2nd, etc. when the given activity was present in the tour. For instance, in two-activity tours involving "Individual/joint mandatory" activities, "Individual/joint mandatory" activities were planned first 59.7% of the occurrence (221 cases) and planned second 40.3% of the occurrences (149 cases).

Table 4 Ordered Response Models for Different Tour Sizes

	Variables	Parameter	t-statistic	Variables	Parameter	t-statistic
	(a) Activity Characteristics Only			**(b) Activity Characteristics and Activity Types**		
Tour Size = 4	Constant	0.881	17.04	Individual Mandatory act	0.555	4.61
	Duration (hr)	-0.113	-6.45	Joint discretionary act	0.891	6.49
	Travel Time (hr)	-0.004	1.29	Individual discretionary act	0.877	9.29
	$\alpha 1$	0.697	18.93	Joint Maintenance act	1.051	8.27
	$\alpha 2$	1.391	29.03	Allocated Maintenance act	0.863	13.17
				Duration (hr)	-0.071	-3.19
				Travel Time (hr)	0.046	1.32
				$\alpha 1$	0.699	18.93
				$\alpha 2$	1.398	29.04
	No. of Observations	816		No. of Observations	816	
	$L(\beta)$	-1110.16		$L(\beta)$	-1105.828	
	$L(C)$	-1131.22		$L(C)$	-1131.216	
	$-2[L(C)-L(\beta)]$	42.10 (2 df)		$-2[L(C)-L(\beta)]$	50.78 (6 df)	
Tour Size = 3	Constant	0.564	11.72	Individual Mandatory act	0.391	3.95
	Duration (hr)	-0.075	-5.63	Joint discretionary act	0.433	3.69
	Travel Time (hr)	-0.011	-1.93	Individual discretionary act	0.572	6.67
	$\alpha 1$	0.879	22.35	Joint Maintenance act	0.527	4.14
				Allocated Maintenance act	0.588	9.87
				Duration (hr)	-0.054	-3.26
				Travel Time (hr)	-0.013	-2.23
				$\alpha 1$	0.875	22.37
	No. of Observations	1131		No. of Observations	1131	
	$L(\beta)$	-1224.81		$L(\beta)$	-1225.222	
	$L(C)$	-1242.53		$L(C)$	-1242.53	
	$-2[L(C)-L(\beta)]$	35.44 (2 df)		$-2[L(C)-L(\beta)]$	34.617 (6 df)	
Tour Size = 2	Constant	0.217	5.02	Joint discretionary act	0.135	1.34
	Duration (hr)	-0.093	-8.45	Individual discretionary act	0.149	1.86
	Travel Time (hr)	-0.010	-1.61	Allocated Maintenance act	0.338	5.84
				Duration (hr)	-0.074	-5.34
				Travel Time (hr)	-0.011	-1.80
	No. of Observations	1504		No. of Observations	1504	
	$L(\beta)$	-1003.17		$L(\beta)$	-996.28	
	$L(C)$	-1042.49		$L(C)$	-1042.49	
	$-2[L(C)-L(\beta)]$	78.65 (2 df)		$-2[L(C)-L(\beta)]$	92.43 (6 df)	

Notes:
1) $L(\beta)$ is the loglikelihood at convergence (estimated model) and $L(C)$) is the loglikelihood of the restricted model.
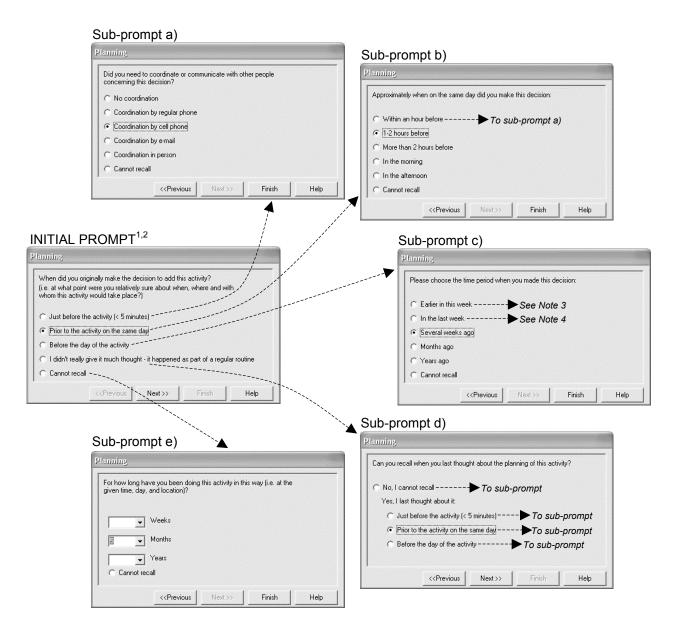2) The likelihood ratio test statistic $-2[L(C)-L(\beta)]$ estimated here is asymptotically chi-square distributed with degrees of freedom equal to the number of free parameters in the model.

Figure 1   CHASE (Computerized Household Activity Scheduling Elicitor) example main screen as it may appear on Monday, showing Add Entry dialog box for activities and their attributes

Note: Original screen is in color.

Figure 2 Planning time horizon queries following addition of an activity

Sub-prompt a)

Planning

Did you need to coordinate or communicate with other people concerning this decision?

○ No coordination
○ Coordination by regular phone
◉ Coordination by cell phone
○ Coordination by e-mail
○ Coordination in person
○ Cannot recall

<<Previous | Next >> | Finish | Help

Sub-prompt b)

Planning

Approximately when on the same day did you make this decision:

○ Within an hour before ------▶ *To sub-prompt a)*
◉ 1-2 hours before
○ More than 2 hours before
○ In the morning
○ In the afternoon
○ Cannot recall

<<Previous | Next >> | Finish | Help

INITIAL PROMPT[1,2]

Planning

When did you originally make the decision to add this activity?
(i.e. at what point were you relatively sure about when, where and with whom this activity would take place?)

○ Just before the activity (< 5 minutes)
◉ Prior to the activity on the same day
○ Before the day of the activity
○ I didn't really give it much thought - it happened as part of a regular routine
○ Cannot recall

<<Previous | Next >> | Finish | Help

Sub-prompt c)

Planning

Please choose the time period when you made this decision:

○ Earlier in this week ------▶ *See Note 3*
○ In the last week --------▶ *See Note 4*
◉ Several weeks ago
○ Months ago
○ Years ago
○ Cannot recall

<<Previous | Next >> | Finish | Help

Sub-prompt e)

Planning

For how long have you been doing this activity in this way (i.e. at the given time, day, and location)?

[▼] Weeks
[6   ▼] Months
[▼] Years
○ Cannot recall

<<Previous | Next >> | Finish | Help

Sub-prompt d)

Planning

Can you recall when you last thought about the planning of this activity?

○ No, I cannot recall ------▶ *To sub-prompt*
Yes, I last thought about it:
  ○ Just before the activity (< 5 minutes) -----▶ *To sub-prompt*
  ◉ Prior to the activity on the same day -----▶ *To sub-prompt*
  ○ Before the day of the activity -------▶ *To sub-prompt*

<<Previous | Next >> | Finish | Help

Notes:
1) Following this initial question, the user would select "Next" leading to at least one additional sub-prompt as shown, before finishing.
2) The word "add" in the main question "When did you originally make the decision to add this activity" was changed to "modify" or "delete" to accommodate modification and deletion decisions.
3) This choice followed by an additional sub-prompt asking "Please choose the day earlier in this week when you made this decision"
4) This choice followed by an additional sub-prompt asking "Please choose the day last week when you made this decision"