Transportation (2017) 44:615–640
DOI 10.1007/s11116-015-9671-5

Activity sequencing, location, and formation of individual non-mandatory tours: application to the activity-based models for Columbus, Cincinnati, and Cleveland, OH

Rajesh Paleti[1] · Peter Vovsha[2] · Gaurav Vyas[3] · Rebekah Anderson[4] · Gregory Giaimo[4]

Published online: 23 December 2015

Abstract Most of the earlier activity based models (ABMs) largely relied on a tour-based modeling paradigm which explicitly predicts tour frequency and then adds details including stop frequency, order, and location of stops within each tour. The current study is part of new tour formation design framework for an ABM in which the underlying tour structure and the stop frequency within tours emerge from temporal, sequencing, and locational preferences of activities that the traveler intends to participate during the day. In order to do this, the study developed a modified rank-ordered logit (ROL) framework that is capable of modeling sequence, locations, as well as the underlying tour structure of all activity episodes simultaneously in an integrated manner. Model estimation with the household survey data, provided several important behavioral insights into underlying choices that drive tour formation. Specifically, the study uncovered pairwise ordering preferences among episodes of different activity purposes, clustering tendencies among episodes of same activity purpose, the impact of supply side activity opportunities on the location and sequence choice dimensions, and impedance effects (including distance and

✉ Rajesh Paleti
rpaleti@odu.edu

Peter Vovsha
vovsha@pbworld.com

Gaurav Vyas
vyasg@pbworld.com

Rebekah Anderson
rebekah.anderson@dot.state.oh.us

Gregory Giaimo
greg.giaimo@dot.state.oh.us

[1] Department of Civil & Environmental Engineering, Old Dominion University, 135 Kaufman Hall, Norfolk, VA 23529, USA

[2] Parsons Brinckerhoff, 1 Penn Plaza, 3rd Floor, New York, NY 10119, USA

[3] Parsons Brinckerhoff, 1 Penn Plaza, 2nd Floor, New York, NY 10119, USA

[4] Ohio Department of Transportation, 1980 W. Broad St., Columbus, OH 43223, USA

mode and time-of-day logsums) on location and tour break dimensions. The developed models are incorporated in the operational ABM structure adopted for three major cities (Columbus, Cleveland, and Cincinnati) in Ohio.

Keywords Tour formation · Rank-ordered logit model · Activity sequence · Location · Trip chaining

Introduction

Most of the activity-based models (ABM), both in practice and research, explicitly predict tour frequency and then add details on frequency, order, and location of intermediate stops within these tours (Bowman 2012; Davidson et al. 2010; Meeks et al. 2013). This modeling framework is largely borrowed from tour-based models where the basic unit of analysis is "tour". However, this framework is not fully consistent with the central idea of ABMs that people travel to pursue activities and tours/trips are derived outcomes of this necessity to pursue different activities (Kitamura 1996; Axhausen and Gärling 1992; Goulias et al. 2011). Thus, the basic unit of analysis in an ABM should be "activity" and not "tour" or "trip". People normally do not make up-front decisions about how many tours to undertake and how many stops to make within each tour. In the real world, these tours and trips emerge from the activity participation, activity location, and activity sequencing choices made under time and space constraints imposed by activities with relatively low spatial and temporal flexibility. This paper is part of a research effort to develop an alternate behavioral framework that determines the activities first and form tours subsequently.

It can be reasonably hypothesized that an individual makes a preliminary decision on participating in a certain set of activities. The scheduling decisions are then driven by the associated temporal and spatial constraints and differential priorities. For example, in general, a worker who goes to work on the travel day will have higher priority associated with work activity compared to an individual shopping or discretionary activity planned for the same day. However, the priorities can change if, for example, the shopping activity is undertaken jointly (assuming a major shopping such as buying a car or furniture) or discretionary activity is a special "ticketed" event such as a football game. Modeling approach discussed in the paper is based on the idea that certain activities that are inflexible or less flexible can be considered as prioritized activities, when compared to other activities. The traveler plans the schedule of these prioritized activities first and then scheduling of the other activities is done around these prioritized activities. The model discussed in the paper represents an effort to better mimic the decision making process as it occurs in reality.

Earlier ABMs and tour formation

In this section, we provide a brief summary of past ABMs particularly in the context of tour formation. We restrict the discussion to representative ABMs that are closer to the econometric foundation of the current research including CEMDAP (developed by University of Texas at Austin), FAMOS (developed by Florida State University), DaySim (developed by Resource Systems Group (RSG) with the original design developed by Massachusetts Institute of Technology), TourCast (developed by Cambridge Systematics, Inc.), and CT-

RAMP (developed by Parsons Brinckerhoff). However, we acknowledge that there are several other interesting approaches beyond the traditional econometric approach including ALBATROS, MatSim, TASHA, and others where tours were generated and formed in a different way (the reader is referred to (Pinjari and Bhat 2011) for a detailed overview of different ABMs). Also, we focus our discussion on the following important feature of the model system—when and how activities and tours are generated in the model system and how the corresponding choices of activity timing, location, and tour formation are ordered.

We found that practically all ABMs in practice today (DaySim, TourCast, CT-RAMP) operate with a certain form of tour generation process as the first step. Then, additional stops are added to the tours (with less or more degree of simultaneity with the tour generation step). Activity locations and time-of-day choices are modeled later and most frequently conditional upon the primary destination of the tour. The primary destination of each tour is modeled for each tour separately and the very definition of primary destination is somewhat arbitrary except for the work and school tours. The stop frequency choice is modeled with some or no degree of simultaneity with the tour frequency choice but it is never integrated with location and timing choices. In this regard, such a basic outcome as the set of activities and their sequence for each individual during the day is never controlled explicitly. The sequence would emerge at the very last step when all stops are inserted into tours and time-of-day attributes of tours and stops are determined. In this regard, these models would be better referred to as "tour-based" rather than "activity-based". In the CEMDAP model system, activity dimension is represented in more explicit way since one of upper-level models predicts activity participation set for each individual using a utility-theory consistent multiple discrete continuous time allocation model applied at the household level (Bhat et al. 2013). However, the subsequent scheduling step includes a tour-frequency model rather than a tour-formation model although the tour-frequency choice is conditional upon the activity participation set generated at the earlier stage. Further on, this requires a stop-frequency model for each tour. CEMDAP has a similar concept of day segments that is adopted in the current paper. For workers and students, tours are generated in a specific window (before or after the mandatory activity) that adds integrity to the model system with respect to logical sequencing of activities. FAMOS represents another interesting prototype where the elements of tour formation were introduced. In FAMOS, tours are not strictly generated but emerge from the sequential process of adding stops to the tour where one of the alternatives is to end the tour and travel back home. Another strong feature of FAMOS is tracking time–space constraints in the activity-location and tour-generation process. However, FAMOS has a simpler activity generation process compared to CEMDAP and other ABMs (for example, it is not clear if this framework could accommodate joint activities and travel) and also each tour is processed independently of the other tours.

In a sense, our approach is an attempt to hybridize the best elements of activity generation and tour formation prototypes embedded in such systems as CEMDAP and FAMOS with the operational structure of applied ABMs. Also, some further generalizations to the concepts of prior activity generation (as in CEMDAP) and tour formation (as in FAMOS) were made as discussed in the subsequent sections. In particular, we adopted the concept of day segments with more added details and prioritization of activity generation as the first step before formation of travel tours and trips. In our model system, activities are first allocated to day segments and then tours are formed by sequencing these activities, finding locations for each of them, and breaking these activity sequences into tours. This is done taking into account the time–space constraints and available time windows within each day segment.

Tour formation in CT-RAMP

This model is a part of the latest version of CT-RAMP (Coordinated Travel and Regional Activity Platform) adopted for the Jerusalem, Phoenix, and Ohio (Columbus, Cleveland, Cincinnati) ABMs. In CT-RAMP, for each person and household, activity generation and tour formation decisions are organized into the following four major steps:

1. *Generation of mandatory activities and formation of mandatory tour skeletons* This step includes decisions that relate to the least flexible activities. In reality, many of these activities are pre-planned before the given day and the person has to build the daily activity pattern and schedule around these activities. The process starts with participation in Special Events (for instance, a sporting event) that are characterized by a fixed location and schedule. Next step relates to generation of mandatory activities such as visiting Workplace, University, School, or any other Business-related activity. This step also includes models that predict mandatory tour skeletons with possible prioritized stops for Special Events and school escort as well as separate home-based tours for Special Events and pure escort (Paul et al. 2015).
2. *Generation of intra-household shared non-mandatory activities and formation of fully joint tours* In this step, shared non-mandatory activities by household members are generated. Although, some of these activities can be flexible in time and space, on a given day, most of them are quite constrained due to the schedule consolidation of several household members. So, these shared non-mandatory activities are considered as prioritized activities within the modeling framework. Also, the vast majority of shared non-mandatory activities are organized into fully-joint tours when all members of the travel party travel together and participate in all activities included in the tour (Vyas et al. 2015a).
3. *Allocation of individual non-mandatory activities to segments* After generation and scheduling of prioritized activities described above as well as formation of associated tours, every person can additionally have multiple individual non-mandatory activities on the given day. We assume that the modeled person has some flexibility to choose the schedule and location of these activities. Subsequent tour formation is largely derived from the way how each person groups these non-mandatory activities around the prioritized activities. In this regard, the prioritized activities and corresponding tours already formed in the first two steps play a role of important "pegs" in the day pattern. These "pegs" create distinctive segments where the modeled person may decide to allocate additional individual non-mandatory activities. The current modeling framework accommodates five different prioritized pegs: school drop-off tour (or Peg 1), primary mandatory tour (or Peg 2), school-pick-up tour (or Peg 3), fully joint tours (or Peg 4), and special event tours (or Peg 5). Pegs '0' and '6' correspond to 'Home" at beginning and end of day, respectively. The relative order of these prioritized pegs defines the pattern that serves as the skeleton that is subsequently filled by tour formation models. It is important to note that for any given individual only a subset of these prioritized pegs might be available. First segment type referred to as '*Home-Based Tour*' segment corresponds to windows between the already formed prioritized tours. So, a non-mandatory activity allocated to the 'Home-Based Tour' segment is assumed to be undertaken on a separate home-based tour. Second segment type referred to as '*Prioritized Tour*' segment corresponds to either outbound or inbound legs of the already formed prioritized tours. Third segment type referred to as '*Work Chain*' segment corresponds to the middle part of the work tour. So, any non-

mandatory activity allocated to the 'Work Chain' segment is assumed to be either undertaken on an at-work sub-tour or as an intermediate stop within the mandatory activity span including possible business meeting and other business-related locations before or after the workplace.[1] This allocation of non-mandatory activities to different segments can be thought of as a strategic scheduling decision. For example, if the modeled person has a work tour and shopping activity, this activity can be implemented before work as a separate shopping tour, as an outbound stop on the way to work, at work as a sub-tour, as an inbound stop on the way from work back home, or after return back home from work as a separate shopping tour. The model takes the form of a standard discrete choice model with all available segments as alternatives in the choice set for each non-mandatory activity. Tours can be formed only from the activities allocated to the same segment (Vyas et al. 2015b).

4. *Within-segment activity sequencing and tour formation* The final tour formation step is modeled for each day segment separately. If only one individual non-mandatory activity is allocated to a segment, it does not require any further processing except for finding the most probable location. A single additional non-mandatory activity in the 'Home Based Tour' segment generates a separate one-destination tour. A single additional non-mandatory activity in the 'Prioritized Tour' segment without previously scheduled prioritized stops (such as escorting a child to school) generates an additional stop on one of the existing tours. A single additional non-mandatory activity in the 'Work Chain' segment without prioritized stops either generates a separate at-work sub-tour or a stop in the mandatory activity span. However, when multiple non-mandatory activities are allocated to the same segment and/or if this segment already contains some prioritized stops, the tour-formation procedure involves modeling non-trivial decisions about sequencing of these activities, finding their locations, and identifying the associated tour structure. For example, if such activities as shopping and discretionary were allocated after the work tour, the following possible tour-formation alternatives should be evaluated: shopping being first and discretionary being second either on the same tour or as two separate tours, or vice versa, discretionary being first followed by shopping either on the same tour or as two separate tours. *This modeling step is the focus of the current paper.*

The rest of the paper describes details of the fourth model as well as the corresponding inputs (provided by the first three models in the ABM system application).

Data

The estimation data comprises of four different household travel surveys (HTS) conducted for the metropolitan regions of Columbus (5555 households in 1999), Dayton (1950 households in 2001), Cincinnati (2050 households in 2009–2010) and Cleveland (4540 households in 2012–2013). Each of these surveys collected detailed information regarding household composition, individual socio-demographics, residential, work, and school locations, and activity travel diary of all household members. These four data sets were

[1] An at-work sub-tour is defined as a sequence of trips that start and end at the primary workplace (for example, all lunch trips at work place constitute an at-work sub-tour).

pooled together and used for estimating this model. Only weekday records were used for this analysis. Also, only the first day survey responses were used for MPOs which had multi-day survey data. In general, the estimated models were kept generic across all regions unless it was statistically evident that some coefficients should be segmented by region. The differences between regions were captured by region-specific transportation Level-of-Service (LOS) and accessibility measures.

In the CT-RAMP modeling framework, prior to this model, all the prioritized activities and the scheduled activities within each segment are known. Among three types of segments described in the previous section, new tour (or sub-tour) formation is permitted only in 'Home-Based Tour' and 'Work Chain' segments. The remaining segments correspond to inbound and outbound directions of prioritized tours ('Prioritized Tour' segments). The activity sequencing and location decision factors differ across these three types of segments. Thus, the current sequence, location, and tour formation model is either completely segmented by major segment type or certain explanatory variables are partially segmented by the type of segment. Table 1 shows the frequency distribution of activities among the three segment types.

While the definitions of 'Home-Based Tour' and 'Prioritized Tour' segments are straightforward, 'Work Chain' segments warrant some further explanation. For a work chain such as "Home-Business (B1)-Workplace (W)-Business (B2)-Home", non-mandatory activities such as 'Lunch' can be undertaken either on 'B1-W' (segment 28) or 'W-B2' (segment 30) segments that do not result in a sub-tour or on a separate sub-tour from 'W' (segment 29). Also, please note that "H-B1" (segment 24) and "B2-H" (segment 25) which are outbound and inbound legs of commute are 'Prioritized Tour' segments. Segments 31-33 correspond to a similar second mandatory tour chain of a person with two mandatory tours. Next, consider the following daily chain "H-B1-Shop (S)-B2-H-Lunch (L)-H-B3—Maintenance (M)-B4—H". This person has two business tours "H-B1-S-B2-H" and "H-B3-M-B4-H". This type of pattern is usual for such employees as plumber and private tutor. There is no usual workplace for this person. Segment 34 relates to possible activities between the two business tours that results in an additional non-mandatory tour "H-L-H". Segment 35 relates possible activities between business episodes on the first tour "B1-S-B2". Segment 36 relates possible activities between business episodes on the second tour "B3-M-B4".

There are 35,454 individual non-mandatory activity episodes in the sample. The activity purpose at the destination end was classified into one of the following nine activity categories: escorting, shopping, household maintenance, personal maintenance, breakfast, lunch, dinner, visiting, and discretionary. Among the 35,454 activity episodes, 16,435 activity episodes correspond to individuals without any prioritized pegs (i.e., these individuals do not participate in Special Events, school escorting, mandatory, or fully joint activities during the day) and 8016 activity episodes correspond to type 1 segments. Of the remaining activity episodes, 8095 activity episodes correspond to type 2 segments and 2908 activity episodes correspond to type 3 segments. A single model was estimated for 'Home-Based Tour' and 'Prioritized Tour' segments with partial segmentation of coefficients by segment type. Another completely segmented model was estimated for activity episodes within 'Work Chain' segments.

It is important to note that Table 1 lists all possible segments that are observed in the real-world data. However, for each person, depending on the actual prioritized activities, only a subset of specific segments would be available. For example, consider a worker with one work tour with a single work activity, no school escorting tours, and no fully joint tours. This person would have only the following five segments available for allocation of

Table 1 Observed frequency of activity episodes by segment

S. no.	Starting peg	Ending peg	Type of segment	Freq.	%
1	No prioritized pegs		–	16,435	46.4
2	Home (0)	School drop-off tour (1)	Home based tour	117	0.3
3	Home (0)	Primary mandatory tour (2)	Home based tour	664	1.9
4	Home (0)	School pick-up tour (3)	Home based tour	174	0.5
5	Home (0)	Joint tour (4)	Home based tour	1548	4.4
6	School drop-off tour (1)	Primary mandatory tour (2)	Home based tour	7	0.0
7	School drop-off tour (1)	School pick-up tour (3)	Home based tour	39	0.1
8	School drop-off tour (1)	Joint tour (4)	Home based tour	48	0.1
9	School drop-off tour (1)	Home (6)	Home based tour	167	0.5
10	Primary mandatory tour (2)	School drop-off tour (1)	Home based tour	0	0.0
11	Primary mandatory tour (2)	School pick-up tour (3)	Home based tour	24	0.1
12	Primary mandatory tour (2)	Joint tour (4)	Home based tour	179	0.5
13	Primary mandatory tour (2)	Home (6)	Home based tour	4025	11.4
14	School pick-up tour (3)	School drop-off tour (1)	Home based tour	0	0.0
15	School pick-up tour (3)	Primary mandatory tour (2)	Home based tour	0	0.0
16	School pick-up tour (3)	Joint tour (4)	Home based tour	9	0.0
17	School pick-up tour (3)	Home (6)	Home based tour	129	0.4
18	Joint tour (4)	School drop-off tour (1)	Home based tour	3	0.0
19	Joint tour (4)	Primary mandatory tour (2)	Home based tour	8	0.0
20	Joint tour (4)	School pick-up tour (3)	Home based tour	7	0.0
21	Joint tour (4)	Home (6)	Home based tour	868	2.4
22	Outbound leg of school drop-off tour (1o)		Prioritized tour	100	0.3
23	Inbound leg of school drop-off tour (1i)		Prioritized tour	185	0.5
24	Outbound leg of commute tour (2o)		Prioritized tour	2234	6.3
25	Inbound leg of commute tour (2i)		Prioritized tour	5292	14.9
26	Outbound leg of school pick-up tour (3o)		Prioritized tour	71	0.2
27	Inbound leg of school pick-up tour (3i)		Prioritized tour	213	0.6
28	Within first mandatory tour chain: Before work place (t221b)		Work chain	44	0.1
29	Within first mandatory tour chain: Work place based (t221w)		Work chain	1422	4.0
30	Within first mandatory tour chain: After work place (t221a)		Work chain	239	0.7
31	Within second mandatory tour chain: Before work place (t222b)		Work chain	1	0.0
32	Within second mandatory tour chain: Work place based (t222w)		Work chain	28	0.1
33	Within second mandatory tour chain: After work place (t222a)		Work chain	18	0.1
34	Between first & second mandatory tours (t223)		Work chain	130	0.4
35	Within first mandatory tour: Between business episodes (t221bb)		Work chain	914	2.6
36	Within second mandatory tour: Between business episodes (t222bb)		Work chain	112	0.3

individual non-mandatory activities: two ‘Home-Based Tour’ segments (‘Between Home and Work Tour’ and ‘Between Work Tour and Home’), two ‘Prioritized Tour’ segments (outbound and inbound directions of work tour), and one ‘Work Chain’ segment (‘Within first mandatory tour chain: Work place based'). All the other irrelevant segments (that are not available) are excluded from the analysis. However, it should be noted that this explanation relates not that much to the model described in the paper but to the previous one that allocates activities to different segments (Vyas et al. 2015b). The goal of this study is to develop a model that is capable of forming tours out of all individual non-mandatory activities within each segment.

Conceptual framework

As mentioned earlier, in the overall CT-RAMP system, (1) all mandatory activity participation decisions, (2) the travel pattern as determined by the relative order of prioritized tours (i.e., school escorting tours, work tours, and fully joint tours), (3) all individual non-mandatory activity participation decisions, and (4) the segments during which these non-mandatory activities are undertaken are known before the current model. The objective of the current sub-model is to package these non-mandatory activities within each segment into tours. The primary argument underlying the conceptual framework developed for this purpose is that tour formation is not a direct choice but a derived outcome of three basic choices: (1) the sequence or order in which different activities are pursued within each segment, (2) spatial location of different activities within each segment, and (3) the decision to go home after pursuing an activity (this choice is referred to as 'tour break' because it can result in the formation of a new tour to undertake remaining activities within a segment). A brief discussion of these three basic choices follows.

Sequence or order of activities

Activities of certain purpose are more likely to be pursued earlier than activities of other purposes either because of the nature of these activities or due to the opening and closing business hour constraints on the supply side. For instance, shopping is less likely to be pursued after dinner in the late evening time period but not so unlikely to be pursued after lunch in the afternoon time period. Similarly, it is possible that activities of the same purpose are either clustered together due to convenience or spread out for more variety. For instance, shopping activity episodes might be clustered back to back or spread out with one lunch activity in the middle. A simple approach to modeling sequence choice would be to develop a discrete choice model such as the multinomial logit (MNL) model with all possible sequences of activities within each segment as alternatives in the choice set. However, a segment with n activities would result in $n![= n \times (n-1) \times \cdots \times 2 \times 1]$ possible sequences thus leading to explosion in choice set size as the number of activities within a segment increases.

Spatial location of activities

For each activity within a segment, the traveler also picks a location. This choice is typically modeled using a discrete choice framework such as the MNL model in which the traveler is assumed to choose the location alternative with highest utility for pursuing an activity. The utility of each location alternative is specified as a function of the location attributes (opportunities for undertaking activities of a specific purpose), travel impedance, and traveler characteristics. The universal choice set for this decision includes all spatial units (in our case, all Traffic Analysis Zones (TAZs)) in the study region). As is the standard norm, a subset of TAZs is sampled to reduce the computational burden associated with estimating the spatial location choice model on the universal choice set.

Tour break

The decision maker has the choice of going back home (or work in case of at-work sub-tours) after pursuing an activity. A positive outcome of this choice not only breaks the current tour but can also result in the formation of a new tour if there are additional activities remaining in the same segment. For instance, consider the example where there are two activities—shopping and maintenance in the segment after work tour. Also, assume that the traveler chooses to undertake shopping before maintenance in this segment. Now, this traveler can choose to go home after shopping, take some rest, and then undertake a new home-based tour to pursue maintenance activity. Alternatively, she or he may decide to directly go to the maintenance activity location after shopping instead of going home. In the first case, the two activities in the after work segment are scheduled in two separate home-based tours whereas in the second case, these activities are scheduled within a single tour. However, it is important to note that a tour break is not possible in certain segments by definition. For instance, there cannot be tour breaks in the 'Prioritized Tour' segments corresponding to the outbound and inbound directions of work tours. So, this choice dimension is relevant only to 'Home Based Tour' and 'Work Chain' segments that can result in formation of new tours/sub-tours.

Joint nature of basic choices

Sequence or order in which different activities are pursued is a function of both activity purpose combinations and their locations (an activity can be pursued immediately after another activity simply because the locations are close). Thus, it is not behaviorally appealing to estimate a choice model for sequence based on the activity purpose alone and then determine locations conditional on the sequence of activities. A simultaneous model of sequence of activities and their locations can capture the relative trade-offs of both purposes and locations. Similarly, the chain of activities and locations along with spatial and temporal constraints determine whether these activities are chained in a single tour or multiple tours. On the other hand, it is also true that the sequence and locations for pursuing different activities depend on whether all the activities are pursued in a single tour or multiple tours.

Thus, it is not behaviorally appealing to have an independent tour frequency model and then a sequence and location choice model which determines order and location of activities conditional on tour frequency or vice versa. In this regard, it is essential to have a joint model of sequence of activities, location of activities, and tour formation to capture all the intricate interdependencies among the following three choices: (1) Sequence of activities within each segment, (2) Spatial location of activities, and (3) Tour break. In this framework, tour frequency and stop frequency are not modeled explicitly but they emerge from the above three underlying choice dimensions which is closer to how people make choices in the real world.

Model structure

In order to make this concept operational, the Rank Ordered Logit (ROL) modeling framework was adopted which was traditionally used to model ranking preferences of individuals in stated preference choice experiments (Beggs et al. 1981; Fok et al. 2012).[2] Specifically, ROL models are suited to analyze ranking data where respondents are asked to rank alternatives in the choice set instead of choosing one alternative (which is the case in traditional single discrete choice models such as the multinomial logit model). The current empirical choice context is similar to ranking data in the sense that people choose a sequence of activities in which each activity is assigned a preferential chronological order or rank. The relative ranking or order of all the activities determines their sequence. However, in a standard ROL model, every alternative in the choice set is assigned a rank. In our research, this framework was extended to simultaneously model sequence, location and tour formation choices. Specifically, the choice set comprises of alternatives defined as location and activity purpose combinations and, unlike in a traditional ROL model, only a subset of all activity and location combinations (that correspond to the chosen locations) are assigned a chronological rank in the modified framework.

The observed sequence, location, and tour frequency choices within each segment are assumed to be outcomes of several sequential choice instances in which an individual chooses the highest ranked combination of activity purpose and location alternative. Please note that although the model is applied sequentially over multiple choice occasions, the overall framework relates to an integrated formulation of choices of choices for activity location, sequencing, and tour breaks. This is because the model is applied sequentially *but not in a predetermined order*. The order itself is an outcome of the sequencing decision that keeps the model integrated. The choice set at each choice occasion is a union of location alternatives for all the activities that have not been yet chosen prior to the current choice occasion. Table 2 shows the choice alternatives for an individual with two activities: shopping (S) and maintenance (M) in 'Home-Based Tour' segment that allows formation of new tours, as an example. The only difference in the construction of choice alternatives for segments that do not allow new tour formation (for example, 'Prioritized Tour' segments) is that the "Go through Home" alternatives would be unavailable for activities in these segments.

Several important observations can be made from Table 2:

1. The set of location alternatives varies across different activities. So, different locations may be sampled for different activities, i.e. [$L_D(S,1) \neq L_D(M,1)$].
2. The location alternatives may change across choice occasions, i.e., location alternatives in second choice occasion for maintenance activity in Table 2 can be different from the location alternatives in choice occasion 1, i.e. [$L_D(M,1) \neq L^*_D(M,1)$].
3. Once a certain activity is chosen, the location alternatives corresponding to that activity are made unavailable in all subsequent choice occasions, i.e., none of the shopping alternatives are available in choice occasion 2.
4. Each location alternative for each activity is further duplicated to produce two sub-alternatives with respect to the tour structure. The first one correspond to "going

[2] Please note that the Rank Ordered Logit (ROL) model used in this study is different from the ordered logit model used for analyzing ordinal variables (Paleti and Bhat 2013).

Table 2 Illustration of choice data for a sample individual

Choice occasion	Go shopping directly		Go shopping through home		Go maintenance directly		Go maintenance through home		Chosen alternative
	1	2	3	4	5	6	7	8	
1	$L_D(S,1)$	$L_D(S,2)$	–	–	$L_D(M,1)$	$L_D(M,2)$	–	–	$L_D(S,2)$
2	–	–	–	–	$L^*_D(M,1)$	$L^*_D(M,2)$	$L^*_H(M,1)$	$L^*_H(M,2)$	$L^*_H(M,2)$

directly to activity location" [$L^*_D(M,1)$, $L^*_D(M,2)$] whereas the second one corresponds to "going to activity location through a stop at home" [$L^*_H(M,1)$, $L^*_H(M,2)$].

5. Alternatives corresponding to "going to activity location through home" are unavailable for the first activity in the segment. For instance, the "going to activity location through home" alternatives, i.e., [$L_H(S,1)$, $L_H(S,2)$] and [$L_H(M,1)$, $L_H(M,2)$], are unavailable for the first choice occasion.
6. Tour and stop frequencies emerge from the sequence, location, and tour formation choices. For instance, the chosen alternatives for the example shown in Table 2 results in two tours—shopping tour followed by maintenance tour with one stop each.
7. The total number of choice occasions in a segment with N activities is equal to N.

Let i denote the index for activity, j for location i.e., Traffic Analysis Zone (TAZ), and k denote whether the person travels to the location directly ($k = 1$) or goes to the activity location through home ($k = 2$), and t indicate the index for choice occasion. The observed part of the utility function for each alternative in the choice set at choice occasion t can be written as:

$$V^t_{i,j,k} = V^t_i + V^t_j + V^t_k, \tag{1}$$

where V^t_i is the utility component specific to activity purposes, V^t_j is the utility component specific to location alternatives, V^t_k is the utility component specific to tour formation decision, $i \in C^t_A$ is the set of all activities in the segment at choice occasion t, and $C^t_A = C^{t-1}_A - \{A^{t-1}\},\ \forall\ t > 1$ where A^{t-1} is the chosen activity at choice occasion $t - 1$ and C^1_A is the set of all activities in the segment $j \in C^t_{L.i}$ is the set of all location alternatives for activity i at choice occasion t,

$$k = \begin{cases} 1 = \text{go directly to the activity location} \\ 2 = \text{go to the activity location through home} \end{cases}$$

V^t_i can be further split into two parts as follows:

$$V^t_i = V^t_{i,\text{sequence}} + V^t_{i,\text{cluster}}, \tag{2}$$

where $V^t_{i,\text{sequence}}$ is the utility component specific to activity sequence effects (i.e., the relative order in which activities of different purposes are pursued), and $V^t_{i,\text{cluster}}$ is the utility component that captures the clustering effects (i.e., whether activities of the same purpose are pursued together). This "clustering" component in the sequencing utility constitutes the difference between the case with sequencing of two different activities and case where two successive activities are the same. If this utility component is positive it means a higher propensity of the two episodes of this activity (say, shopping) to be implemented back to back (all else being equal) (but not necessarily at the same location). If this utility component is negative, it means that most probably the two episodes of this activity will have a space (other activities) between them. V_j^t also can be split into two parts as follows:

$$V_j^t = V_{j,\text{size}}^t + V_{j,\text{impedance}}^t, \tag{3}$$

where $V_{j,\text{size}}^t$ captures the activity-specific zonal attraction effects for location j, $V_{j,\text{impedance}}^t$ captures all the impedance effects that determine location choices

We assume that the unobserved factors that govern the sequence, location, and tour formation choices are independent and identically distributed across individuals and alternatives implying that the choice probability at each choice occasion t is given by:

$$P^t(i,j,k) = \frac{\exp\left(V_{i,j,k}^t\right)}{\sum_{i' \in C_A^t} \sum_{j' \in C_{L,i'}^t} \sum_{k'=1}^{2} \exp\left(V_{i',j',k'}^t\right)}. \tag{4}$$

So, the complete log-likelihood function across all choice occasions can be written as:

$$P = \prod_t P_c^t(i,j,k), \tag{5}$$

where each of the $P_c^t(i,j,k)$'s correspond to the probability of observed activity, location, and tour break combination at choice occasion t in the estimation data. Given that we pooled multiple data sources for estimating this model, we tried estimating different scale parameters but these attempts did not yield any statistically significant results.[3] Also, it is possible to explore the correlations across choice occasions of the same person. Essentially it would replace the current MNL setting with a panel mixed logit structure as was introduced in (Srinivasan et al. 2006). Our first attempts with this approach yielded inconclusive results. Also, it should be also noted that due to a large number of destination alternatives it is necessary to employ a sampling procedure. This makes the proposed model a "rank ordered logit model with sampling". Until very recently, sampling and subsequent calculation of correction terms is difficult to justify for any error term structure beyond a simple MNL. However, (Guevara and Ben-Akiva 2013) proved that standard sampling methods produce consistent estimates in logit mixture models. Exploring different mixing specifications across different choice occasions with sampling is an avenue for future research.

Sampling location alternatives

As mentioned earlier, it is computationally not feasible to consider all location alternatives. Moreover, this problem is more pronounced in this model because there can be multiple activity episodes within each segment and considering all locations for each activity

[3] To test scale heterogeneity, we used a nested logit specification where the choice set was duplicated for each sub-sample and placed under a separate nest with a linkage across all model coefficients except for the scale parameters. Polydoropoulou and Ben-Akiva (2001) and Paleti et al. (2014) used similar techniques for uncovering scale heterogeneity in revealed and stated preference datasets. However, both unconstrained optimization of the resulting log-likelihood function and grid search over an extended range of scale parameters did not indicate scale parameters significantly different from 1.

episode leads to choice set explosion. However, a completely random sampling approach can lead to illogical zig-zag tour paths with very long intermediate trips which is usually not how people make activity location decisions. In the real world, individuals optimize location alternatives and put them in a sequence that avoid large deviations from their current tour trajectory. To capture this behavior, the location alternatives at each choice occasion are sampled based on two pivot locations—origin TAZ (i.e., chosen location TAZ of the previous choice occasion) and final destination TAZ (which is home TAZ for all 'Home-Based Tour' segments and one of the prioritized locations including work, school drop-off or pick-up locations for 'Prioritized Tour' and 'Work Chain' segments). Such a sampling-by-importance mechanism ensures that potential location alternatives for subsequent activities do not deviate substantially from the tour path. A sampling-by-importance model with TAZ activity-specific size terms and a coefficient of -0.1 for both "Distance from origin TAZ" and "Distance to final destination TAZ" variables was applied. Five TAZs were sampled for each activity in the segment and up to 8 activities were considered within each segment resulting in a maximum of 40 activity and location combinations. These 40 alternatives were duplicated to create 40 additional alternatives that correspond to the option of going to the activity location through home. Thus, the maximum possible number of alternatives in the model is 80. Also, all activity records in each segment were duplicated twice before sampling and weight of each record set was set to 0.5. All location alternatives were sampled independently for these two sets of duplicate records. This approach is roughly equivalent to sampling 10 location alternatives for each activity in the segment. In model estimation, a correction term equal to $\ln\left(\frac{n_i}{N \times q(i)}\right)$ was added to the utility function to account for the difference in the sampling probability and the frequency of the alternative in the sample. The sampling correction term represents natural logarithm of the ratio of the sampling frequency to selection probability for each alternative as was substantiated in the theory (McFadden 1978; Ben-Akiva and Lerman 1985). In this correction term, $q(i)$ is the selection probability (probability to be drawn) which is a function of size variable and simplified distance-based impedances, n_i is the selection frequency in the sample or the number of times an alternative is chosen, and N is the sample size (which is 5 because we sample five locations for each activity). During model application, the same sampling strategy was adopted to reduce computational complexity. Specifically, if there are N activity episodes in a segment, then N choice occasions were created and a subset of location alternatives was sampled for each activity episode at first choice occasion using the exact same sampling-by-importance model used during model estimation. For the first choice occasion, the origin TAZ and final destination TAZ serve as the pivot locations whereas for all subsequent choice occasions, the chosen TAZ in previous choice occasion and the final destination TAZ serve as pivot locations. So, the model is applied sequentially in a microsimulation framework where different location alternatives may be sampled for the same activity episode in different choice occasions depending on chosen locations in prior choice occasions.

correction terms

Modifications for 'Prioritized Tour' and 'Work Chain' segments

Tour breaks are not allowed in 'Prioritized Tour' segments. Thus, "Going to activity location through home" location alternatives are made unavailable for activities within these segments. However, tour breaks (at workplace) are allowed in 'Work Chain' segments and these tour breaks result in the formation of workplace based sub-tours. Table 3 shows the type 3 day segments where workplace based sub-tours are allowed. Moreover,

Table 3 Tour breaks in 'work chain' segments

Work tour structure	Sub-segment type for tour break rules	Sub-segment as listed in Table 1 for activity allocation	Tour break (visiting workplace)
Single workplace visit (can be one or two such tours)	Chain starting and ending at workplace	T221 W, T222 W	Before each activity except for the 1st
	Business chain before visiting workplace	T221b, T222b	No
	Business chain after visiting workplace	T221a, T222a	No
Two workplace visits (can be only one such tour)	From 1st workplace to 1st business between workplaces	T221w	Before each activity including 1st business except for the 1st
	From last business between workplaces to 2nd workplace	T221w	Before each activity except for the 1st (i.e., last business)
	From business to business	T221bb	No
	Business chain before visiting 1st workplace	T221b	No
	Business chain after visiting 2nd workplace	T221a	No

'Prioritized Tour' segments may include school escorting stops as part of outbound and inbound commute tours. Also, there might be multiple episodes of school escorting stops on outbound and inbound directions of school escorting half tours. Similarly, there might be several pre-determined business and work locations in the mandatory tour chain for 'Work Chain' segments. These stops serve as spatial anchors in addition to the left and right end of the segment because the location and the sequence of these stops are already known. However, it is still possible to undertake non-mandatory individual activity stops before or after these stops with pre-determined locations. To accommodate stops with a fixed location within the same modeling framework, we do not sample any locations for these activities in our sampling mechanism. For example, if there is one shopping activity and one school-pick-up stop in the inbound direction of a commute tour, five locations for shopping activity will be sampled but only one known location will be used for the school-escorting activity.

Enforcing sequence of activities with known order

The relative order of certain activities is fixed by definition. Specifically, all eating-out activities must follow a sequential order: breakfast first, then lunch, and then dinner. Similarly, school drop-off and pick-up locations and their sequence are known prior to this model. So, this sequence must be enforced appropriately during model estimation. This is achieved by an appropriate exclusion of later activities from the choice set. For instance, if both breakfast and lunch are in the choice set at any given choice occasion, lunch is made unavailable if breakfast is also in the choice set. This ensures that lunch is never chosen prior to breakfast. Moreover, based on data, several decisions were made regarding the presence and sequence of school escorting and other non-mandatory activities within

'Prioritized Tour' segments. These decisions and assumptions were guided by the observed data and are presented in Table 4.

Main explanatory variables

As noted earlier, the observed part of the utility function can be split into 5 parts:

1) activity sequence,
2) activity clustering,
3) activity-specific zone size (attraction) variable,
4) travel impedance, and
5) tour formation represented by additional stops at home for 'Home-Based Tour' segments or at work for 'Work Chain' segments ("tour breaks").

The *activity sequence* component captures the chronological order or sequence-related preferences of different activity purposes. Unlike a simple MNL model, where just alternative-specific constants would capture the mean preference for an alternative, in the current model, alternate-specific constants are segmented by the presence of activities of other purposes in the choice set. The underlying idea is that the chronological order in which activities are pursued within a segment is most likely dependent on the type of activities that have been already scheduled during that segment. This implies that the "independent of irrelevant alternatives" property does not hold for the estimated model. The ratio of choice probabilities for two alternatives depends on other alternatives in the choice set. Furthermore, it was also tested if these relative preference effects vary across different types of segments (0-X: 'Home-Based Tour' segment sandwiched between home and prioritized tour, X-X: 'Home-Based Tour' segment sandwiched on both sides by prioritized tours, X-6: 'Home-Based Tour' segment sandwiched between prioritized tour

Table 4 Presence and sequence of activities in 'Prioritized tour' segments

'Prioritized tour' segment sub-type	Assumption		
	Presence of school escort stops	Presence of other stops	Ordering rule
Outbound leg of school DO tour (1o)	Yes	No	Not required
Inbound leg of school DO tour (1i)	No	Yes	Not required
Outbound leg of commute tour (2o)	Yes	Yes	School drop off always occurs before other stops
Inbound leg of commute tour (2i)	Yes	Yes	None (All possible sequences are allowed)
Outbound leg of school PU tour (3o)	No	Yes	Not required
Inbound leg of school PU tour (3i)	Yes	No	Not required
Outbound leg of school DO tour (3i)	No	Yes	Not required

and home, and 0–6: full-day segment without any prioritized tours) and depending on whether the segment lies before or after the primary mandatory tour.

While greater size variables are associated with higher likelihood of choosing a destination TAZ over other TAZs, the corresponding effect on the sequence of activities is not readily apparent. In fact, we hypothesize that the impact of size variable on activity sequence component might be negative, i.e., locations with higher size variable will increase the likelihood of pursuing the corresponding activities rather later than earlier (big shopping malls being a good example for the last rather than first activity on a tour). This is because bigger establishments may provide more flexibility to pursue activities later in the day compared to smaller establishments with constrained hours of operation and resources. So, for every pair of activity episodes with distinct activity purposes (i,i') in a segment, we constructed the following measure to capture the impact of size variable on activity sequence component: $R_{ii'} = LN\left[\frac{S_{ijk}}{\sum_{j'\in C^t_{L,i'}} S_{i'j'k}}\right]$. This measure represents the relative value of size variable for activity and location combination (i,j) and other competing activities i' across all locations $j' \in C^t_{L,i'}$. For the example shown in Table 2, for the first and second shopping alternatives, the value of this measure would be: $LN\left[\frac{S_{\text{Shop, L1}}}{S_{\text{Maintenance, L1}} + S_{\text{Maintenance, L2}}}\right]$ and $LN\left[\frac{S_{\text{Shop,L2}}}{S_{\text{Maintenance,L1}} + S_{\text{Maintenance,L2}}}\right]$, respectively. We used a ratio measure instead of some weighted average of difference in size terms of (i,j) and (i',j') across all $j' \in C^t_{L,i'}$ to bring different activities (such as shopping and discretionary) with different size terms to a common denominator in terms of sequencing. "Big shop" in terms of employees is not directly comparable to "Big gym". With this scaling, we make "Big" and "Small" relative to the total stock of corresponding attractions and more comparable across different activities.

The *clustering* component captures the tendency to undertake activity episodes of the same purpose together. All clustering variables are also segmented by the type of segment to test if there is a variation in clustering effects across different types of segments.

The *zone size (attraction) variable* captures the zonal attraction power for different location alternatives as a function of different zonal population and employment factors. In the current study, the size variables were pre-calculated using coefficients obtained from regressions of weighted number of trip destinations in each zone by purpose on different zonal population and employment characteristics.

The *impedance* component captures the zonal impedance factors that govern any location decision. Specifically, this component includes origin–destination joint mode and time-of-day logsums, several linear and non-linear distance effects, and their interaction with key demographic factors such as gender, person type (eight categories including full-time worker, part-time worker, university student, non-worker, retiree, driving age school child, pre-driving age school child, and pre-school child), household income, and presence of children. For each location alternative, two locations serve as anchor points from which all impedance measures were computed: (1) the chosen location in the previous choice occasion or the start anchor of the segment, and (2) right anchor with known location which is either home (for all 'Home-Based Tour' segments) or one of the prioritized pegs (for 'Prioritized Tour' segments). While the left anchor is always moving (as determined by the chosen location of the previous choice occasion), the right anchor for most of the individuals is fixed. However, for 'Prioritized Tour' segments with school escorting stops, the next school escorting stop will serve as the moving right anchor. For instance, consider an outbound commute tour with two school-escorting stops and one maintenance stop in

Table 5 Moving anchors in 'Prioritized tour' segments

Prioritized tour segment type	Initial start anchor	Moving start anchor	Initial end anchor	Moving end anchor
Inbound leg of school DO tour	Last drop-off	Previous activity location	Home	
Outbound leg of commute tour	Home, last drop-off	Previous activity location	Workplace, University, Business	
Inbound leg of commute tour	Workplace, University Business	Previous activity location	Home	Next pick-up location
Outbound leg of school PU tour	Home	Previous activity location	First pick-up	
Inbound leg of school PU tour	Last pick-up	Previous activity location	Home	

between these two escorting stops (as the modeled alternative). The first escorting location will serve as right anchor for all maintenance location alternatives in the first choice occasion but the second escorting location will serve as the right anchor for all maintenance location alternatives in the second choice occasion. Table 5 summarizes the start and end anchors of activities in 'Prioritized Tour' segments.

The *tour break* component captures the factors that impact whether an individual goes directly to the next activity location or go to the next location through home resulting in the formation of a new tour. This component includes factors such as activity purpose of the previous stop, activity purpose at the destination of the current stop, deviation computed as additional distance the person has to travel if she or he goes through home instead of going directly, length of the segment, number of tours already scheduled in the segment, number of remaining activities in the segment, and number of stops in the current tour.[4]

Estimation results

Tables 6, 7, 8, 9, 10 presents the estimation results for sequence, location, and tour formation for activities in 'Home-Based Tour' and 'Prioritized Tour' segments. Missing cells indicate that the corresponding parameter estimates were not statistically significant. Some coefficients were fixed during model estimation based on logical considerations (particularly those with values +9, −9). Table 6 shows the impact of size variables on the activity sequence component of the utility function. It can be seen that most of the size effects are negative supporting our hypothesis that activity locations with more opportunities are more likely to be pursued later than activity locations with fewer opportunities.

Tables 7 show the parameter estimates that indicate the mean preferred pair-wise ordering of different activities for different types of segments. It can be seen from these tables, that the parameter estimates of the sequence component form an upper triangle matrix in which each cell indicates the relative preference of undertaking activities in the corresponding column prior to activities in the corresponding row. For instance, negative coefficient for the "shopping-maintenance" cell in Table 7 indicates that, on average,

[4] Deviation is calculated as follows: $(OH + HD - OD)$ where OH is Origin-to-Home distance, HD is Home-to Destination distance, and OD is Origin-to-Destination distance.

Table 6 Activity sequence component (size effect)

	Escorting	Shopping	HH maintenance	Personal maintenance	Breakfast	Lunch	Dinner	Visiting	Discretionary
Activity sequence component (size effect)									
Escorting			−0.0706				−0.2020		−0.3388
Shopping					0.1497		−0.1388		−0.2717
HH Main.							−0.5012		−0.3508
Pers. Main.									−0.4180
Breakfast									
Lunch									−0.2259
Dinner									−0.1948
Visiting									−0.4061
Discretionary									
Activity sequence component (base effect)									
Escorting							0.6899	0.1984	
Shopping			−0.1913	−0.5959	−0.9187		0.5021		
HH Main.							0.8072		
Pers. Main.							1.2557	0.4541	
Breakfast									
Lunch							0.7714		
Dinner									−0.2179
Visiting									
Discretionary									

Table 7 Additional activity sequence component results

Activity-sequence pair	Segment	Parameter estimate
Shopping—discretionary	'Home Based Tour' segment between Home & Prioritized Tour (0-X)	−0.5663
Personal maintenance—dinner	'Home Based Tour' segment between Home & Prioritized Tour (0-X)	9.0000
Personal maintenance—lunch	'Home Based Tour' segment between Prioritized Tour & Home (X-6)	9.0000
Personal maintenance—lunch	Any segment that is after work	9.0000
Personal maintenance—discretionary	Any segment that is after work	9.0000

Table 8 Clustering component

Activity	Parameter
Escorting	
All segments	−0.4235
'Home Based Tour' segments between prioritized tour and home (X-6)	−0.6603
Visiting	
All segments	−0.5455
'Home Based Tour' segments between two prioritized tours (X-X)	−9.0000

shopping activities are less likely to be pursued before maintenance activities. Other parameter estimates may be similarly interpreted. Also, some of these pairwise sequence effects are different across different segments as indicated by the statistically significant parameters in Table 7.

Table 8 shows the results corresponding to the clustering component. The parameter estimates in this component are the diagonal entries in the results shown in Table 7, i.e., the tendency to undertake activities of the same purpose together. So, the results indicate that escorting and visiting activities are less likely to be clustered together in any segment. The clustering effect is much stronger for X-6 type of segments (i.e., 'Home Based Tour' segments that are sandwiched between prioritized peg on the left and home on the right) for escorting activities. The stronger −9 coefficient relates to visiting activity for the X-X segments (i.e., 'Home Based Tour' segments between two prioritized tours).

Table 9 presents the parameter estimates from the impedance component. Several non-linear distance terms (including linear, logged, squared and cubed distance terms) both from the left and the right anchors were tested. Among all these distance impedance effects, logarithmic specification of distance gave the best data fit. Furthermore, several interactions between the distance and traveler socio-demographics also came significant. Figures 1, 2 show the marginal effects of distance on probability for different population groups (by worker status, gender, person type, and income group) for 'Home Based Tour' and 'Prioritized Tour' segments. One common observation across all segments is that the

Table 9 Impedance component

Variable	Parameter	T-Stat
Impedance measures from left peg		
Distance effects for 'Home Based Tour' segments		
Logarithm	−0.1479	−6.02
Logarithm × Female with kid	−0.0662	−1.46
Logarithm × Income >75 K	−0.1857	−6.99
Logarithm × Part-time Worker	−0.0972	−2.41
Logarithm × Retiree	−0.0842	−2.69
Distance effects for 'Prioritized Tour' segments		
Logarithm	−0.4068	−10.33
Logarithm × Female	−0.1029	−1.96
Mode & TOD Choice LogSum for 'Home Based Tour' segments	0.8164	28.12
Mode & TOD Choice Logsum for 'Prioritized Tour' segments	1.0000	
Destination zone is same as Origin TAZ	0.5331	11.53
Impedance measures from right peg		
Distance effects for 'Home Based Tour' segments		
Logarithm	−0.1479	−6.02
Logarithm × Female with kid	−0.0662	−1.46
Logarithm × Income >75 K	−0.1857	−6.99
Logarithm × Part-time Worker	−0.0972	−2.41
Logarithm × Retiree	−0.0842	−2.69
Distance effects for 'Prioritized Tour' segments		
Logarithm	−0.4068	−10.33
Logarithm × Female	−0.1029	−1.96
Mode & TOD choice LogSum for 'Home Based Tour' segments	0.8164	28.12
Mode & TOD choice Logsum for 'Prioritized Tour' segments	1.0000	
Other variables		
Logarithm of distance to pick-up location in ride sharing	−1.1079	−1.80
Destination zone is same as Origin TAZ	0.5331	11.53

distance effect is always negative and the marginal effect of distance is gradually decreasing with increasing distance. Logically, the distance effects corresponding to activities in 'Prioritized Tour' segments are stronger than the corresponding effects for 'Home Based Tour' segments suggesting that 'Prioritized Tour' segments are more constrained spatially and temporally. Also, distance to the next ride sharing location (i.e., school escorting stop in commute tours) has a much stronger negative impact on sequence and location choice indicating considerable time pressure that the driver is subjected to in these ride sharing tours. The coefficient on the composite mode & time-of-day choice logsum is positive and in the unit interval as required by theory. The log-sum variable used in this model was computed assuming a normalized nested logit framework with destination choice at the upper level, time-of-day choice at the intermediate level, and mode choice at the lower level. Similar to the findings in the distance impedance effects, the coefficient on the logsum variable was higher for 'Prioritized Tour' segments than 'Home

Table 10 Tour break component

Variable	Parameter	T-Stat
Deviation (OH + HD − OD)	−0.0455	−12.34
Number of tours already in the segment		
Number of tours already in the segment = 1: Number of pegs = 0	−8.4123	−5.65
Number of tours already in the segment = 2: Number of pegs = 0	−8.1692	−5.46
Number of tours already in the segment ≥3: Number of pegs = 0	−8.5295	−5.54
Number of tours already in the segment = 1: X-X	−8.1773	−5.94
Number of tours already in the segment = 2: X-X	−7.1410	−3.75
Number of tours already in the segment ≥3: X-X	−9.0000	
Number of tours already in the segment = 1: 0-X	−8.6441	−5.80
Number of tours already in the segment = 2: 0-X	−8.7126	−5.51
Number of tours already in the segment ≥3: 0-X	−10.2515	−4.11
Number of tours already in the segment = 1: X-6	−7.9700	−5.93
Number of tours already in the segment = 2: X-6	−8.0915	−5.67
Number of tours already in the segment ≥3: X-6	−8.2571	−3.46
Time window		
LN(1 + time window)	0.7496	3.48
Number of remaining activities in day segment (zero is base category)		
One		
Two	−0.4930	−4.09
Three	−0.7383	−4.32
Four or more	−0.9661	−3.90
Purpose at destination TAZ	−1.1957	−4.36
Escorting		
HH maintenance	0.9938	6.42
Personal maintenance	0.4085	2.95
Breakfast	−2.4044	−2.32
Dinner	0.2271	0.99
Visiting	0.1424	0.91
Discretionary	1.1294	7.53
Previous purpose at origin TAZ		
HH Maintenance	−0.1876	−1.62
Dinner	−1.0149	−2.97
Visiting	−0.7329	−4.39
Number of existing stops in the tour (zero is base category)		
Two	−0.2210	−1.72
Three	−0.2887	−1.44
Four or more	−0.5088	−2.11

This component is applicable only to "Going through Home" alternatives

Based Tour' segments. Moreover, the coefficient on logsum variable for 'Prioritized Tour' segments was not significantly different from 1.

Table 10 presents the estimation results for the tour break component of the utility function. One of the key determinants of the decision to go to the activity destination directly or through home is the additional distance an individual needs to travel if she/he

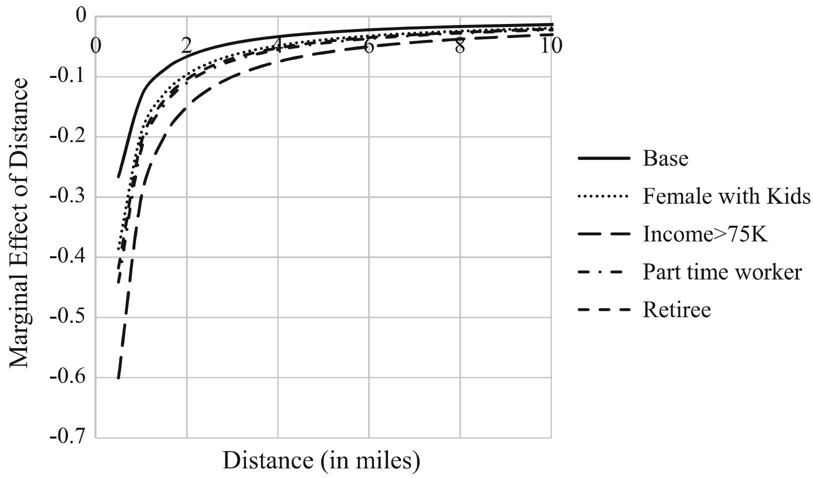

Fig. 1 Marginal effect of distance on probability in 'Home Based Tour' segments

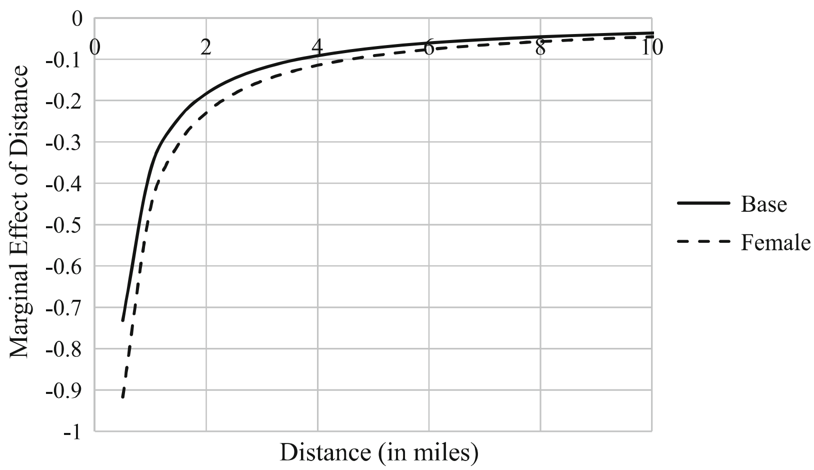

Fig. 2 Marginal effect of distance on probability in 'Prioritized Tour' segments

goes to home before going to the activity location. As expected, this effect is strong and negative indicating larger deviations are not conducive to (additional) tour formation. The next set of variables in Table 10 capture the impact of tours already present in the segment. This effect is further segmented by the type of segment. As expected, it is less likely to have an additional tour break in a segment if there is already a tour in that segment. Length of the time window of the segment is positively related to a higher likelihood of taking a tour break in that segment. Length of a day segment is defined as the duration between end time of previous prioritized activity and start time of next prioritized activity discounted by a buffer time for travel time from home to prioritized activity locations. If at any choice occasion there are more activities that are yet to be scheduled, then the likelihood of taking a tour break becomes lower. Also, if there are already many activities scheduled in a tour by the current choice occasion, then a tour break becomes less likely. These two results are probably indicative of the general tendency to chain activities in the same tour. Also, tour breaking tendency is found to depend on the activity purpose at the origin and destination ends of the trip. For instance, the traveler is more likely to take a tour break if the destination activity purpose is "discretionary" whereas if the activity purpose at the origin end is "dinner" then the traveler is less likely to go home before pursuing other activities in the segment.

Conclusions

Explicit tour frequency models with subsequent addition of details on stop frequency and location applied sequentially have been the norm for modeling activities in most of the applied Activity-Based Models (ABMs). However, in reality, activity location decisions are not always conditional on tour frequency decisions. On the contrary, in most cases, tour and stop frequencies emerge out of sequence and location decisions associated with each of the activities in the segment. By definition, "activity" is the basic unit of analysis in ABMs and all travel should be a derived outcome of the necessity to participate in different activities with varying spatial and temporal constraints and preferences.

However, until now there has been a very limited research that focused on tour formation that works bottom up from activities. Tour formation is a complex behavior phenomenon with many intricate relationships among activities and their preferred locations and time periods (relative to prioritized activities with generally tighter spatial and temporal constraints). So, it necessitates a new modeling paradigm that accommodates main behavioral aspects involved in the formation of tours. The current study is part of a bigger research effort to develop such a comprehensive modeling framework that can account for the possible decision parameters involved in the tour formation process. Specifically, the models developed in this paper constitute the last important step to determine the sequence, location, and underlying tour structure of all activities within each segment. It is based on the known structure of the segments for each individual and known allocation of individual non-mandatory activities to each segment (that are modeled in the ABM system prior to the tour formation step). To our knowledge, this is the first study which simultaneously models all these three choices in an integrated manner and we believe that it is a significant advancement in the design of behaviorally realistic ABMs.

Moreover, the proposed approach does not add much complexity to the entire model system. It rather moves some inevitable complexity from one sub-model to another sub-model. In the end, any travel model should generate tours with intermediate stops. In many activity-based models in practice, this is achieved by a set of sub-models that predict tour frequency, stop frequency for each tour and half-tour, and then find locations for each stop. Overall, the complexity and number of associated choices in this sequence of sub-models is comparable to the proposed approach. We replace tour frequency, stop frequency, and stop location set of sub-models with the tour-formation model where tours emerge from activity sequences and locations.

In addition to research interests and desire to mimic the real-world decision-making of individuals as close as possible, it is expected that a better behavioral framework would make the ABM more useful for policy studies that involve potential changes in activity scheduling and trip chaining. In this regard, the proposed model framework can be a useful tool to analyze the consequences of congestion pricing and other policies with a greater level of behavioral impacts compared to the existing travel models in practice. For instance, consider a global mileage-based pricing scenario. When a tour frequency model is applied up front as part of the individual daily pattern, the trip chaining response to pricing can be only captured through aggregate accessibility measures. Subsequently, impact of congestion pricing on stop frequency and location will be limited to each tour. In this highly sequential structure, the trade-offs between number of tours, number of stops, and stop locations will be largely lost and the only hope will be that each of the sub-models will be driven by the same accessibility measures that would guarantee some implicit integrity. In the proposed model structure, effect of global pricing will propagate through all steps of tour formation explicitly. First, people will look for shorter travel due to a greater cost per mile that would affect choices of destinations and sequences of activities. Secondly, due to shorter travel from home there can be more opportunities for tour breaks and some avoidance of trip chaining. However, on the other hand, additional detours for stopping at home might be themselves onerous leading to more trip chaining. Ultimately, the number of tours can increase or decrease depending on the density of attractions in the area.

However, like any other model, the proposed model has inherent limitations. The most basic one is that, in the current framework, we assume that individuals optimize location alternatives and put them in a sequence that avoid large deviations from their current tour trajectory. To capture this behavior, the location alternatives at each choice occasion are sampled based on two pivot locations—origin TAZ and final destination TAZ. However, this is not necessarily true and it is possible that travelers directly evaluate chains of destinations for different activities in a single step instead of updating the location alternatives for activities based on the two pivot locations over multiple choice occasions. Modeling chained destination choices poses several interesting challenges that serve as possible avenues for future research including effective chain sampling mechanisms and possible strong correlations among chains of locations. The next possible extension of this study is to model correlation across different choice occasions in the ROL framework for the same traveler by building upon the recent advances in sampling methods for mixed logit models (Guevara and Ben-Akiva 2013).

Acknowledgments The authors would like to thank four anonymous reviewers whose comments helped improve an earlier version of the paper considerably.

References

Axhausen, K.W., Gärling, T.: Activity-based approaches to travel analysis: conceptual frameworks, models, and research problems. Transp. Rev. **12**(4), 323–341 (1992)

Beggs, S., Cardell, S., Hausman, J.: Assessing the potential demand for electric cars. J. Econom. **17**(1), 1–19 (1981)

Ben-Akiva, M.E., Lerman SR, : Discrete choice analysis: theory and application to travel demand, vol. 9. MIT press, Cambridge (1985)

Bhat, C.R., Goulias, K.G., Pendyala, R.M., Paleti, R., Sidharthan, R., Schmitt, L., Hu, H.-H.: A household-level activity pattern generation model with an application for Southern California. Transportation **40**(5), 1063–1086 (2013)

Bowman, J.L., Bradley, M.: DAYSIM Activity and Travel Schedule Simulator Technical Documentation, Resource Systems Group, Inc. (2012)

Davidson, W., Vovsha, P., Freedman, J., Donnelly, R.: CT-RAMP family of activity-based models. In: Proceedings of the 33rd Australasian Transport Research Forum (ATRF) 2010. Citeseer

Fok, D., Paap, R., Van Dijk, B.: A rank-ordered logit model with unobserved heterogeneity in ranking capabilities. J. Appl. Econom. **27**(5), 831–846 (2012)

Goulias, K.G., Bhat, C.R., Pendyala, R.M., Chen, Y., Paleti, R., Konduri, K.C., Huang, G., Hu, H.-H.: Simulator of activities, greenhouse emissions, networks, and travel (SimAGENT) in Southern California: Design, implementation, preliminary findings, and integration plans. In: Integrated and Sustainable Transportation System (FISTS), 2011 IEEE Forum on 2011, pp. 164–169. IEEE

Guevara, C.A., Ben-Akiva, M.E.: Sampling of alternatives in Logit Mixture models. Transp. Res. Part B: Methodol. **58**, 185–198 (2013)

Kitamura, R.: Applications of models of activity behavior for activity based demand forecasting. In: Activity-based travel forecasting conference, New Orleans, Louisiana 1996
McFadden, D.: Modelling the Choice of Residential Location. University of California, Institute of Transportation Studies (1978)
Meeks, W.S., Sabina, E.E., Childress, S., van Slyke, C.D., Mullins III, J.A., Ziering, E.A., Rossi, T.F.: TourCast: Trade-offs in Activity-Based Model Implementations. In: Transportation Research Board 92nd Annual Meeting 2013, vol. 13-3712
Paleti, R., Bhat, C.R.: The composite marginal likelihood (CML) estimation of panel ordered-response models. J. Choice Model. **7**, 24–43 (2013)
Paleti, R., Vovsha, P., Givon, D., Birotker, Y.: Joint modeling of trip mode and departure time choices using revealed and stated preference data. Transp. Res. Record **2429**, 67–78 (2014)
Paul, B.M., Vovsha, P., Hicks, J.E., Vyas, G., Livshits, V., Jeon, K.: Generation of Mandatory Activities and Formation of Mandatory Tours: Application to the Activity-Based Model for Phoenix, AZ. In: Transportation Research Board 94t^{h} Annual Meeting 2015, vol. 15-5149
Pinjari, A.R., Bhat, C.R.: Activity-based travel demand analysis. A Handbook of Transport Economics **10**, 213–248 (2011)
Polydoropoulou, A., Ben-Akiva, M.: Combined revealed and stated preference nested logit access and mode choice model for multiple mass transit technologies. Transp. Res. Record **1771**, 38–45 (2001)
Srinivasan, S., Bhat, C., Holguin-Veras, J.: Empirical analysis of the impact of security perception on intercity mode choice: a panel rank-ordered mixed logit model. Transp. Res. Record **1942**, 9–15 (2006)
Vyas, G., Vovsha, P., Paleti, R., Givon, D., Birotker, Y.: Investigation of Alternative Methods for Modeling Joint Activity Participation. In: Transportation Research Board 94th Annual Meeting 2015a, vol. 15-1908
Vyas, G., Vovsha, P., Paul, B., Givon, D., Livshits, V.: Allocation of Individual Non-Mandatory Activities to Day Segments in the Tour Formation Procedure: Application to the Activity-Based Models for Jerusalem, Israel and Phoenix, AZ. In: Transportation Research Board 94th Annual Meeting 2015b, vol. 15-4462

Rajesh Paleti is an Assistant Professor of transportation in the Civil and Environmental Engineering Department at Old Dominion University (ODU) specializing in developing econometric models for analyzing travel behavior. His primary research interests include activity-based travel demand modeling, integrated land-use and transportation modeling, and traffic safety. Prior to joining ODU, he worked as a Transportation Systems Modeler in Parsons Brinckerhoff, NY. He earned his doctoral degree in Civil Engineering (Transportation) from the University of Texas at Austin.

Peter Vovsha Ph.D., is a Principal with Parsons Brinckerhoff. He has numerous publications on discrete choice models, regional model system design, and behavioral research. He is internationally recognized for development of original modeling constructs such as the cross-nested logit model. He led pioneering applications of advanced activity-based travel models and integrated agent-based models of travel demand and network simulations in the US. He managed large-scale research projects for the National Cooperative Highway Research program (NCHRP) and Strategic Highway Research Program 2 (SHRP 2) on impacts of road congestion and pricing on travel demand.

Gaurav Vyas has joined Parsons Brinckerhoff in 2012 after completing Master in Civil Engineering from University of Texas in Austin in December 2011. The focus of his work is primarily on activity-based (ABMs) travel models, with experience in trip-based model development as well. He has worked on the development of advanced ABMs in such metropolitan regions as Los Angles, Columbus, and Jerusalem.

Rebekah Anderson P.E. is a Transportation Engineer for the Ohio Department of Transportation. She has managed several model development and travel survey projects including ABMs for Columbus, Cincinnati/Dayton and Cleveland, and GPS-Based and Smartphone-Based Household Travel Surveys and Tablet-Based Transit On-Board Surveys as well as other research projects.

Gregory Giaimo P.E. is a Transportation Engineer for the Ohio Department of Transportation. He manages development and use of travel demand models and forecasting methods in Ohio. In this role he converted the travel demand models from their original mainframe versions to microcomputers, reinitiated travel survey and other model related data collection programs, oversaw development of Ohio's first statewide travel demand model, developed Ohio's first modern version of the four step modeling paradigm for smaller urban

areas, developed Ohio's congestion management system, it's software for implementing USEPA's MOVES program, an automated implemented of the Highway Capacity Manual for travel demand modeling, it's automated forecasting system for producing HPMS and minor project forecasts, it's traffic estimation system for safety analysis and it's enhancements to the NCHRP 255 procedures which were adapted in NCHRP Report 765.