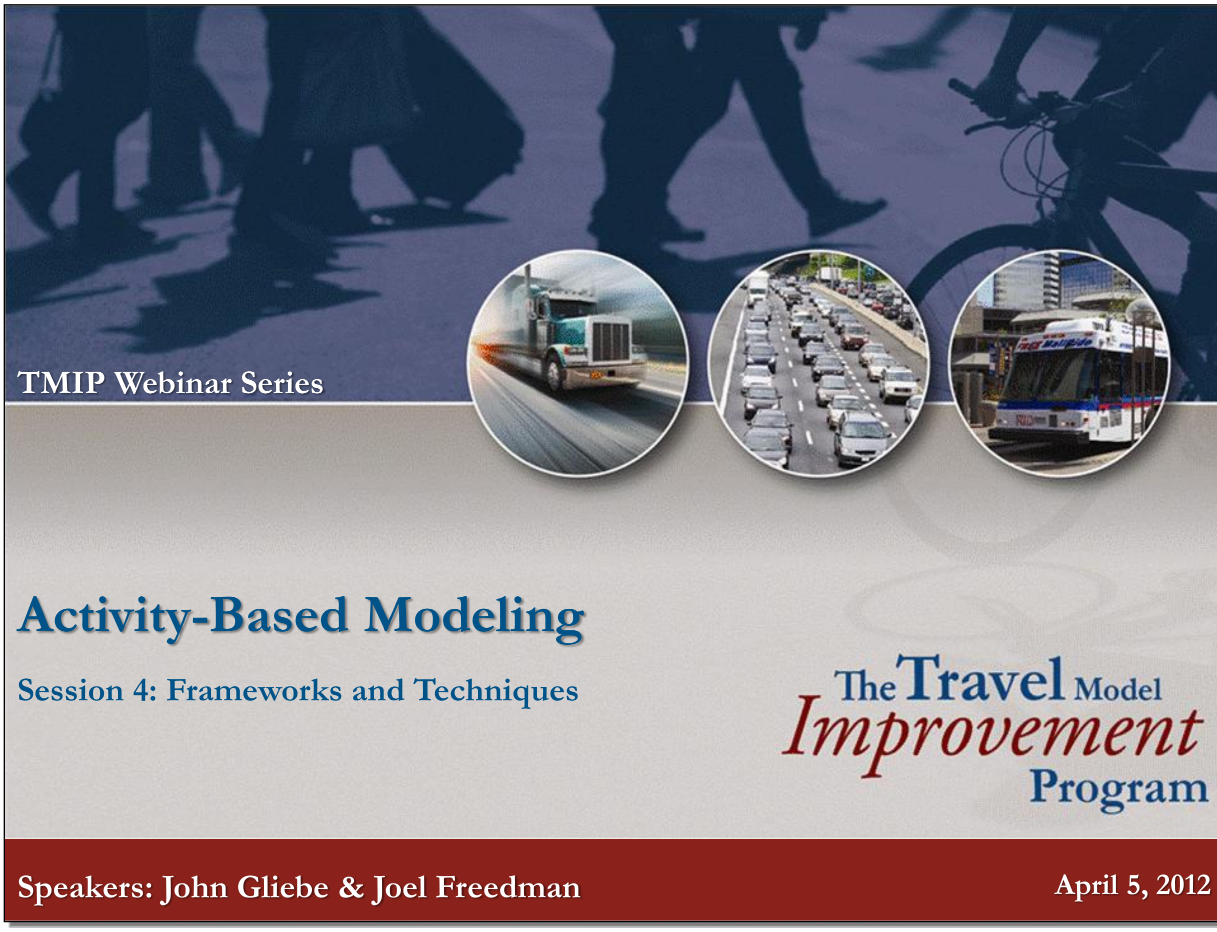

TMIP Webinar Series
Activity-Based Modeling
Session 4: Frameworks and Techniques
The Travel Model Improvement Program
Speakers: John Gliebe & Joel Freedman
April 5, 2012

Acknowledgments

This presentation was prepared through the collaborative efforts of Resource Systems Group, Inc. and Parsons Brinckerhoff.

- Presenters
  - John Gliebe, Joel Freedman
- Moderator
  - Maren Outwater
- Content Development, Review and Editing
  - Joel Freedman, John Gliebe, Rosella Picado, John Bowman, Mark Bradley
- Media Production
  - Bhargava Sana

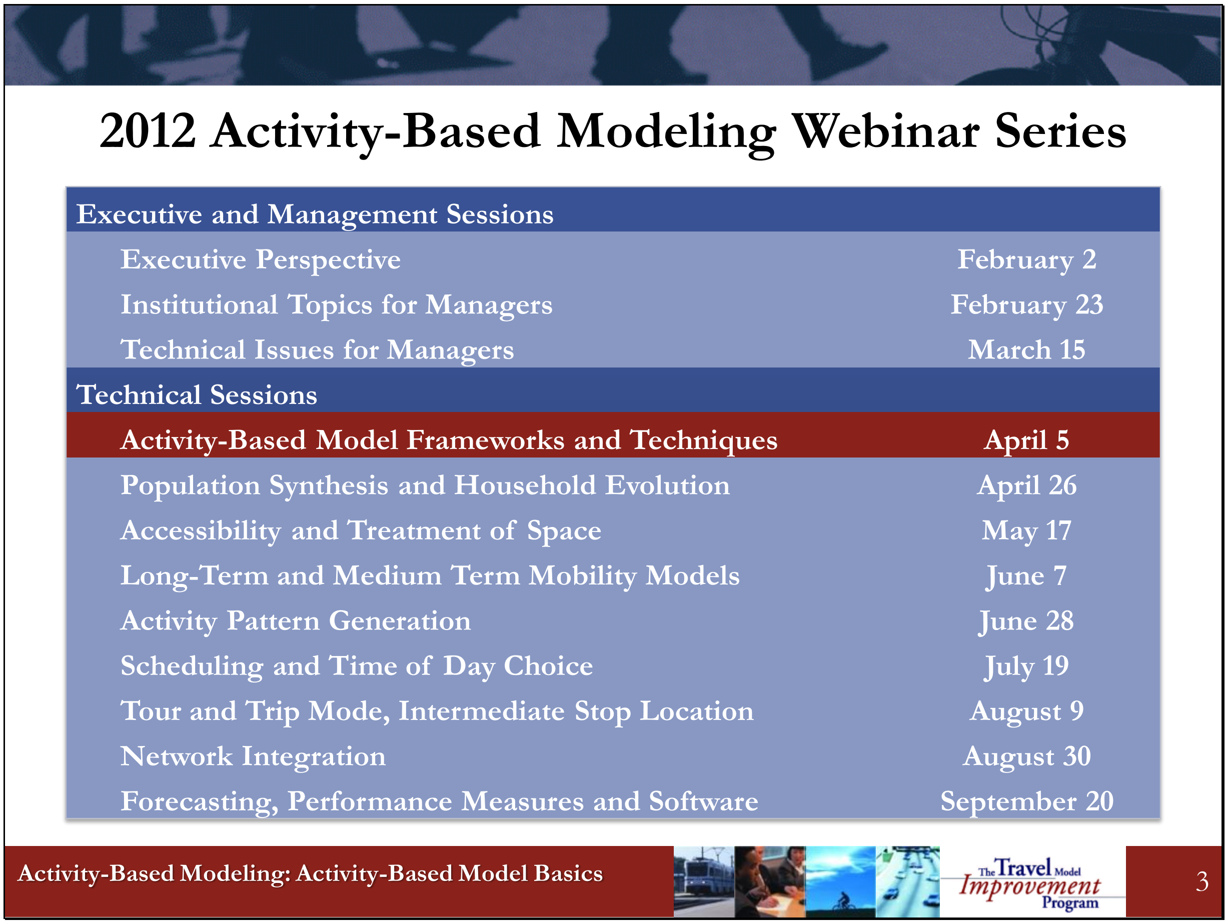
Activity-Based Modeling: Activity-Based Model Basics

The Travel Model Improvement Program

2

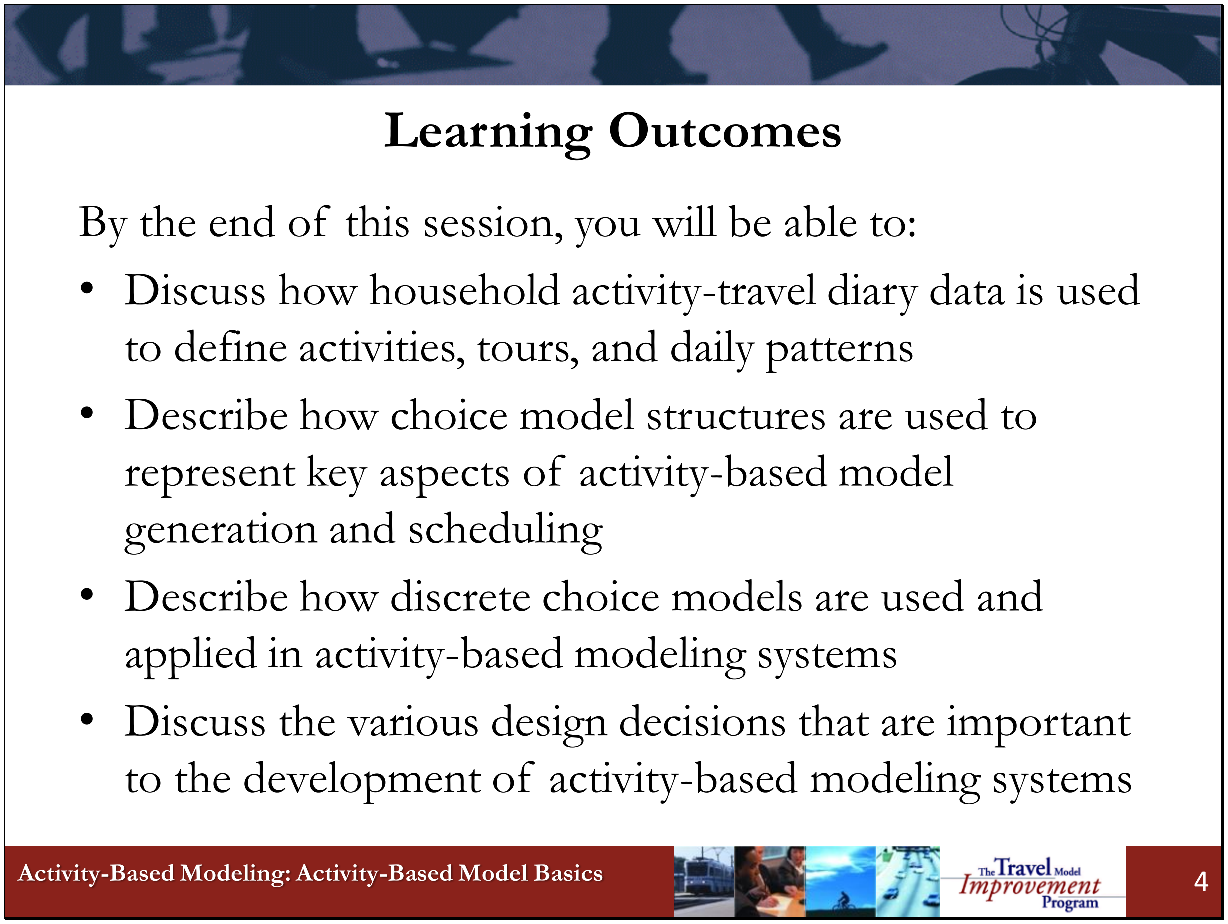
Resource Systems Group and Parsons Brinckerhoff have developed these webinars collaboratively, and we will be presenting each webinar together. Here is a list of the persons involved in producing today's session.

- John Gliebe and Joel Freedman are co-presenters. They were also primarily responsible for preparing the material presented in this session.
- Maren Outwater is the session moderator.
- Additional content was provided by Rosella Picado. John Bowman and Mark Bradley provided reviews of material
- Bhargava Sana was responsible for media production, including setting up and managing the webinar presentation.



For your reference, here is a list of all of the webinars topics and dates that have been planned. As you can see, we will be presenting a different webinar every three weeks. Three weeks ago, we covered the third topic in the series—Technical Issues for Managers. This session was designed to help managers understand the process of developing and implementing an activity-based model, some of the issues to consider when evaluating whether to move to an activity-based model, and some of the different development options available. We also provided a high-level overview of the previous two executive and management webinars.

Today's session is the first of nine technical webinars, where we will cover the details of activity-based model design and implementation. In today's session, we will prepare participants for the remaining webinars by explaining choice models, simulation, and the key components of activity-based models. In three weeks, we will cover population synthesis and household evolution models.



In today's session, we will be covering the basics of activity-based modeling. At the end of this session you should be able to:

- Discuss how household activity-travel diary data is used to define activities, tours, and daily patterns
- Describe how choice model structures are used to represent key aspects of activity-based model generation and scheduling
- Describe how discrete choice models are used and applied in activity-based modeling systems
- Discuss the various design decisions that are important to the development of activity-based modeling systems

Terminology

- Trip
- Tour
- Day pattern
- Schedule
- Discrete Choice Model
- Monte Carlo simulation

We will first cover some terminology that we will use in this webinar. Some of these terms may be familiar to you already, but they are worth mentioning so that we're all on the same page.

Trip: A movement from an origin to a destination. The trip is the core unit of travel in a travel demand model (and an activity-based travel model).

Tour: A series of trips that begin and end at an anchor location (typically either home or work).

Day Pattern: A sequence of in-home and out-of-home activities for an entire day.

Schedule: A day pattern with start and end times for tours and/or activities.

Discrete Choice Model: A probabilistic model commonly used to represent the probability of choosing one alternative from a set of mutually exclusive alternatives.

Monte Carlo simulation: A method of simulating a choice or action, by drawing a random number from a probability distribution in which the choice is one alternative outcome.

Key Concepts

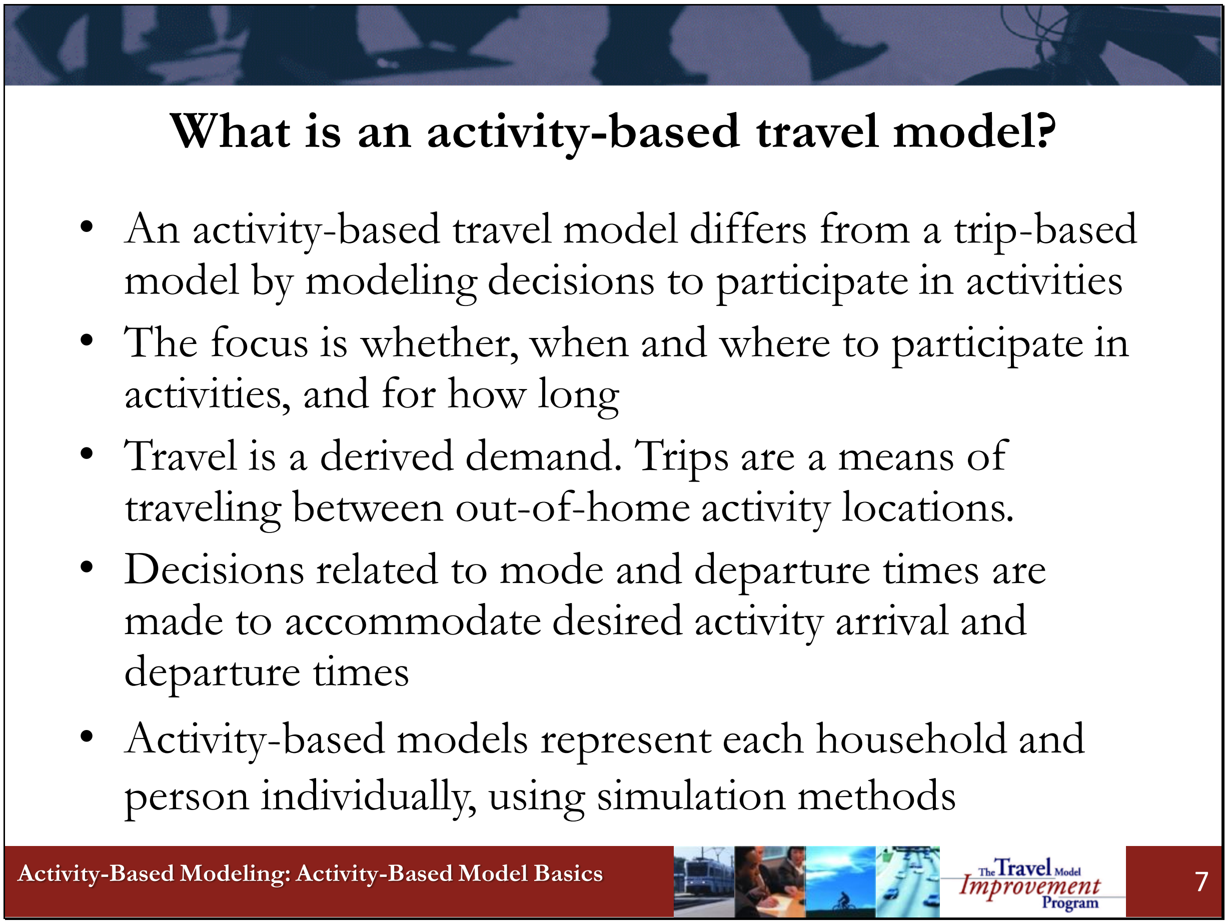
- Activity-based models attempt to model an entire daily travel pattern for each individual in a population, as affected by transportation system level of service
- While activity-based modeling systems vary, they all represent certain key aspects of the activity-travel pattern creation through integrated model components
- Discrete choice models are the most commonly used analytical formulation for model components and are applied through Monte Carlo simulation methods
- Model design involves developing structural representation of decision process and how to treat modes, space, time, and other key model parameters

Activity-based models attempt to model an entire daily travel pattern for each individual in a population, as affected by transportation system level of service. This emphasis on modeling a representative day for each person is one of the important differences between activity-based models and trip-based models.

While activity-based modeling system designs vary, they all represent certain key aspects of the activity-travel pattern creation through integrated model components. We will be describing how these key aspects are represented in survey data and their representation as model structures.

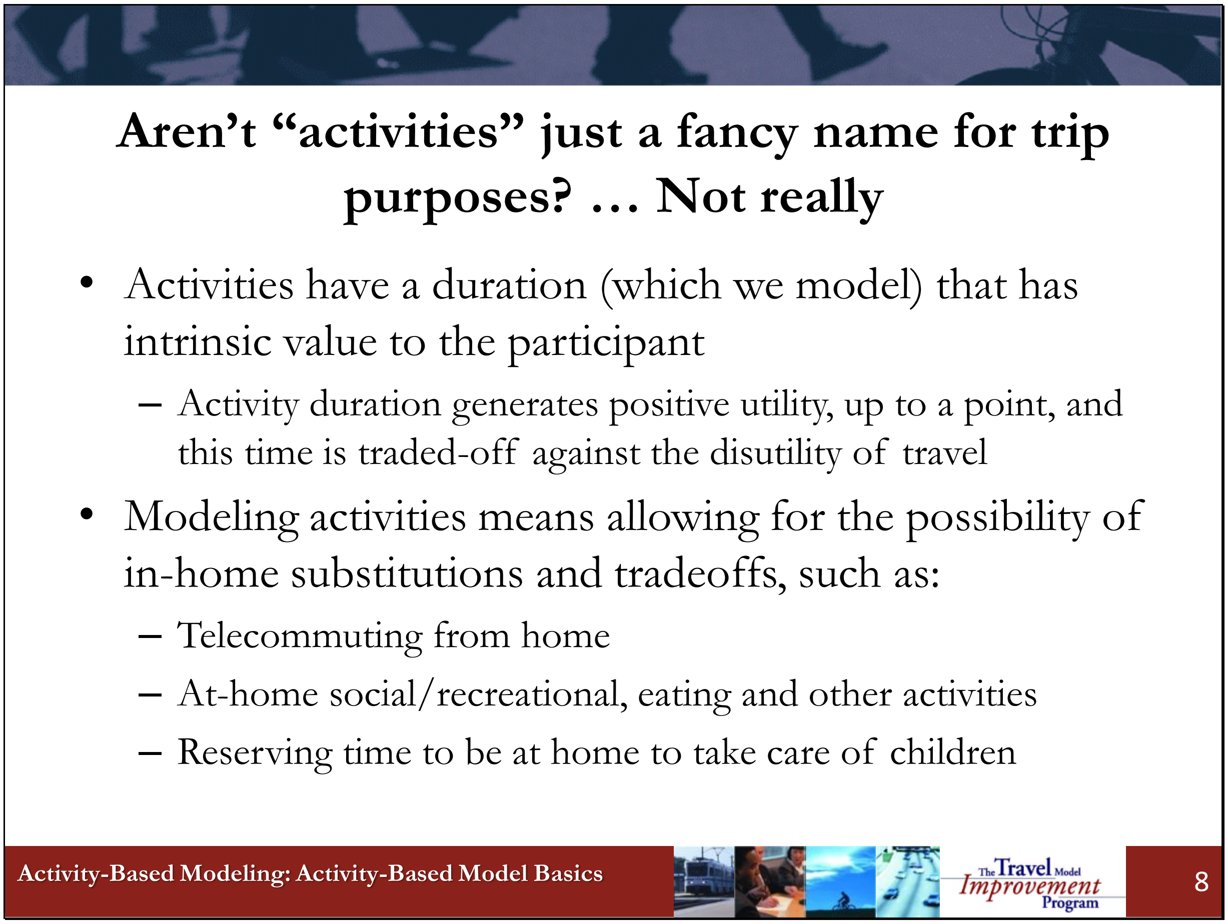
Discrete choice models are the most commonly used analytical formulation for model components and are applied in activity-based models through Monte Carlo simulation methods. Many of you may be quite familiar with discrete choice models, such as the multinomial logit model that is commonly used in mode choice modeling. We describe how discrete choice models are used throughout an activity-based modeling system to represent different structures and the implications for linkages between model components.

Model design involves developing structural representation of decision process and how to treat modes, space, time, and other key aspects of activity generation and scheduling. We will begin the model design discussion today and will flesh it out in greater detail in subsequent webinars.



In the first webinar of this series, which was aimed at an "executive" or managerial-level audience, we described what an activity-based model was. Since this is the first technically-oriented webinar on activity-base modeling, it is probably worth discussing here as well.

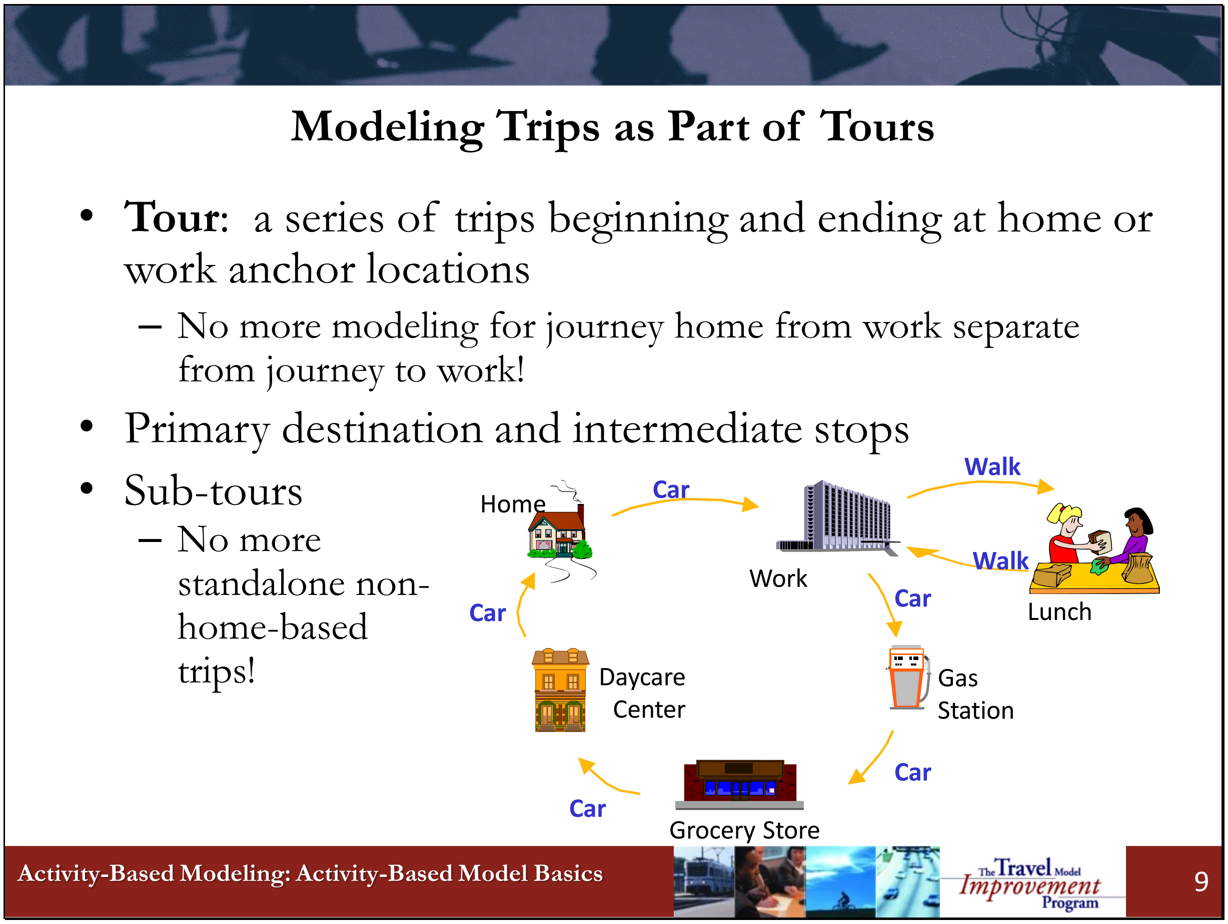
An activity-based travel model differs from a trip-based model by modeling decisions to participate in activities. The central focus of the models is whether, when and where to participate in activities, and for how long. Travel is a derived demand, resulting from the need for people to engage in activities outside the home. Trips are a means of traveling between activity locations and decisions related to trip scheduling, such as mode and departure time, are made to accommodate desired arrival and departure times from activity sites. In advanced activity-travel modeling systems, these decision are coordinated between members of the same household. Activity-based travel models are also characterized by their disaggregate representation of individuals and households, which typically using simulation methods. This enables us to track these individuals and to effectively use their demographic characteristics in analysis.



Aren’t “activities” just a fancy name for trip purposes? … Not really. Activity modeling does bear some resemblance to trip-based modeling in the sense that we generate activities, distribute them to locations, and choose modes for them. Some activity purposes—such as work and school purposes—have similar labels in the trip-based world. And in fact we do model trips within an activity-based modeling system. But modeling activities means much more than that.

Activities have a duration (which we model) that has intrinsic value to the participant. People derive satisfaction from participating in activities, and we assume that the amounts of time that we observe people participating in activities reflect the utility they derive from them. When we model the schedule of activities and travel, we take into account the expected amount of time that an individual would spend in each activity, how they prioritize their time between, work or school, and shopping and recreational activities, and how much time they devote to travel.

Modeling activities also means allowing for the possibility of in-home substitutions and tradeoffs, such as telecommuting from home, at-home leisure, eating and other activities. This is important if we’re interested in modeling future scenarios in which gasoline prices are higher or if on-line commerce were to become the norm. One response to high travel costs is to undertake activities at home. In addition, in-home activities of other household members are important. For example, many parents of young children time their work departure times and forego discretionary activities so they can be at home for their children.



Another key aspect of activity-based travel models is that travel is organized around tours. As mentioned earlier, tours are a series of trips beginning and ending at home or work anchor locations. By modeling decisions on a tour basis, there is enforced consistency between the outbound and return portions of the tour, so that a mode chosen to go to work conditions the mode available for the return home.

Common to tour-based activity modeling is the identification of a primary destination on each tour and the insertion of intermediate stops either before or after the primary destination. In addition, there may be sub-tours within a tour. In this slide’s illustration, there is a work-based sub-tour for lunch. This is a contrast from trip-based modeling protocol in which such trips would be represented as non-home-based tours.

Common Themes of Activity-Based Travel Models

- Model tours as part of a person's ***entire day***
  - Identification of a daily activity pattern
  - When conditions affect one tour they affect a person's entire day of activities and travel!
- Schedule activities consistently in ***time*** and ***space***
  - Activities occur in available time windows
  - No person can be in two places at the same time!

There are a couple of other important themes in activity-based travel modeling, which will become quite evident in this and other webinars in this series. Activity-based models attempt to model an individual's entire day of activities and travel. There are various strategies for doing this, one being the creation of daily activity pattern variables. Daily pattern variables are usually composed of some combination of tours for various purposes. For example, one pattern may be a single, simple tour in which someone goes to work and returns home. Other patterns may be more complex, such as a work tour, with a sub-tour for work-related, activities, and a separate home-based tour to pick up children from an after-school activity. The important message here is that tours and trips within the same day are to some degree interdependent. For example, if a worker has to work late, it may affect other activities in the day. So, perhaps the children have to find another way home, or a wait a bit longer. In activity-based models, scheduling is subject to time and space constraints, such that no person can be in more than one place at the same time. We will cover this in great detail in the ninth webinar in this series and will touch on it briefly later in this session.

Assembling Household Survey Diary Data for Model Development

- Household activity-travel diary
- Person attribute file
- Household attribute file

hhid	perid	dayno	actno	activity	TAZ	arrive	depart	duration	trip time	trip mode	age	female	worker	income	autos
626	1	2	0	Home	39	0	7:00	0	0	none	55	0	1	14	1
626	1	2	1	Escort	82	7:05	7:10	5	5	auto driver occ 2+	55	0	1	14	1
626	1	2	2	Work	1290	7:20	15:25	485	10	auto driver SOV	55	0	1	14	1
626	1	2	3	HH Bus	160	15:50	16:10	20	25	auto driver SOV	55	0	1	14	1
626	1	2	4	Shopping	96	16:20	17:00	40	10	auto driver SOV	55	0	1	14	1
626	1	2	5	Home	39	17:10	19:00	110	10	auto driver SOV	55	0	1	14	1
626	1	2	6	Jnt Shop/Eat	87	19:05	21:00	115	5	auto driver occ 2+	55	0	1	14	1
626	1	2	7	Home	39	21:10		50	10	auto driver occ 2+	55	0	1	14	1
626	2	2	0	Home	39	0	7:00	0	0	none	10	1	0	14	1
626	2	2	1	School	82	7:05	13:40	395	5	auto passenger	10	1	0	14	1
626	2	2	2	Home	39	14:00	19:00	300	20	school bus	10	1	0	14	1
626	2	2	3	Jnt Shop/Eat	87	19:05	21:00	115	5	auto passenger	10	1	0	14	1
626	2	2	4	Home	39	21:10		50	10	auto passenger	10	1	0	14	1

Activity-Based Modeling: Activity-Based Model Basics

The Travel Model Improvement Program

11

A good way to become familiar with the modeling of activities and travel is to look at data. This slide depicts sample record from a household activity-travel diary. Each record includes identifiers for the household, person, survey day, and activity number (sorted). Household activity-travel surveys will of course collect information on each household and person.

To this we have added information on the person’s age, gender and worker status. We have also appended information from the household records on income and number of autos. Of course, we could add other variables to these records. In this example, we’re looking at the activity-travel records for two persons in the same household for the same day. In this case, we have a 55 year-old male, and a 10 year-old female (father-daughter).

From Diary Data to Activities

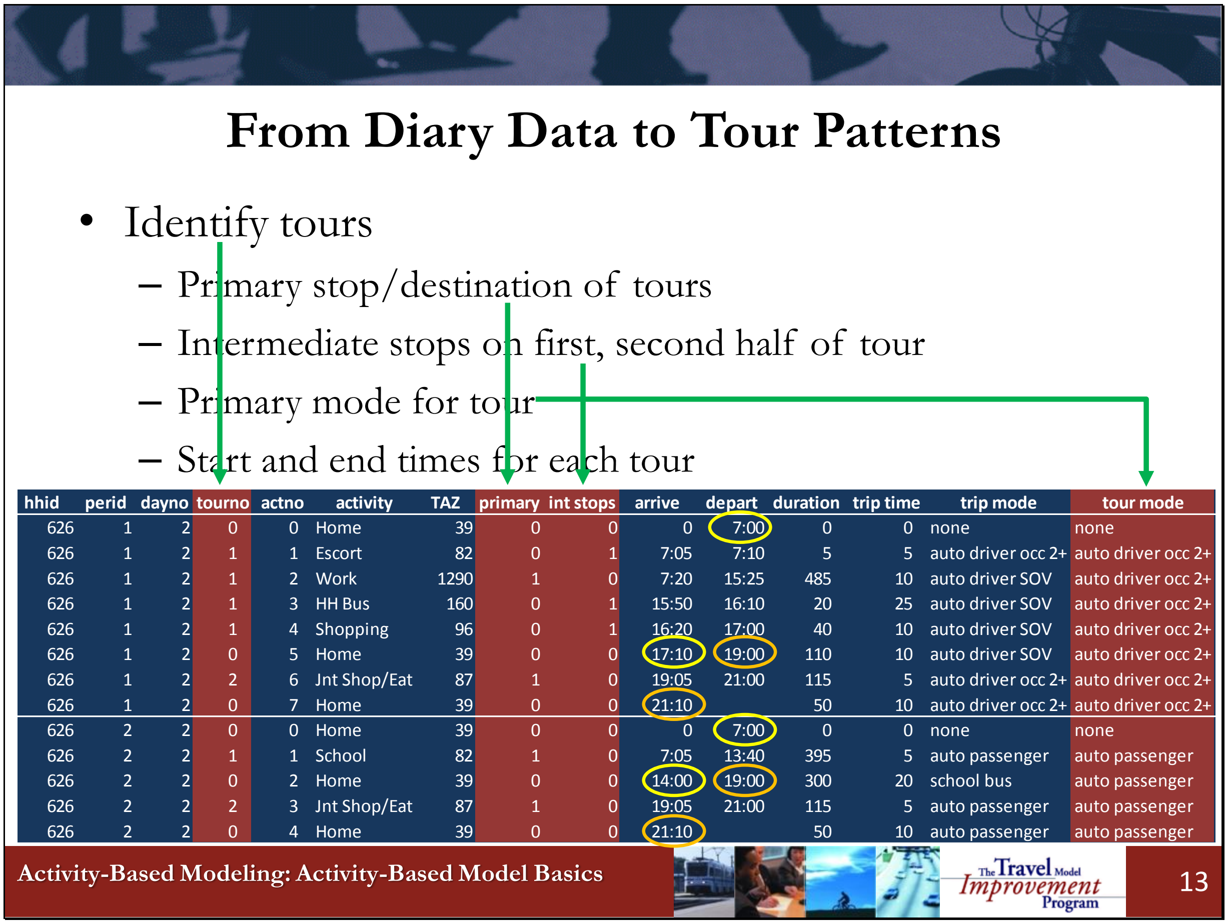
- Identify activity records
  - Place-based convention (one activity purpose per out-of-home location)
  - Some activities involve joint travel and participation
  - Activity durations are important part of scheduling activities and travel
- Familiar trip-based measures also used
  - Starting time, mode, travel time

hhid	perid	dayno	actno	activity	TAZ	arrive	depart	duration	trip time	trip mode
626	1	2	0	Home	39	0	7:00	0	0	none
626	1	2	1	Escort	82	7:05	7:10	5	5	auto driver occ 2+
626	1	2	2	Work	1290	7:20	15:25	485	10	auto driver SOV
626	1	2	3	HH Bus	160	15:50	16:10	20	25	auto driver SOV
626	1	2	4	Shopping	96	16:20	17:00	40	10	auto driver SOV
626	1	2	5	Home	39	17:10	19:00	110	10	auto driver SOV
626	1	2	6	Jnt Shop/Eat	87	19:05	21:00	115	5	auto driver occ 2+
626	1	2	7	Home	39	21:10		50	10	auto driver occ 2+
626	2	2	0	Home	39	0	7:00	0	0	none
626	2	2	1	School	82	7:05	13:40	395	5	auto passenger
626	2	2	2	Home	39	14:00	19:00	300	20	school bus
626	2	2	3	Jnt Shop/Eat	87	19:05	21:00	115	5	auto passenger
626	2	2	4	Home	39	21:10		50	10	auto passenger

Each record indicates a purpose for each activity, a transportation analysis zone (TAZ) where each activity took place, arrival and departure times, calculated activity durations and trip travel times, and a trip mode. As stated earlier, the explicit modeling of activity durations differentiates activity-based models from trip-based models. This diary example follows a "place-based" convention in which there is one activity purpose identified for each out-of-home location. Oftentimes, people have multiple purposes when going to a particular location, such as a business meeting that involves lunch, or combining social activities with meals. Nevertheless, it has become the convention, thus far in activity-based modeling to identify a primary purpose.

This particular example also includes two episodes in which these two household members interacted with one another. The most obvious one is labeled as "joint shop/eat" purpose, so maybe they went to a shopping mall together. You can see that both household members have the same departure and arrival times. For mode purposes, the older adult was the driver (multiple occupancy), and the child was identified as a passenger.

A less obvious interaction between these two is in the morning commutes. Both persons leave home at the same time, and the adult escorts the child to school on the way to his workplace. This drop off event is recorded as “escort” for the father, and simply as a “school” activity for the child, with mode of travel recorded as passenger. After the drop off, we can see the parent’s mode switches from multiple to single occupancy vehicle.



This slide depicts the identification of common tour elements. First, there is the identification of home-based tours themselves and the stops on each tour. As shown in this example, both persons have two home-based tours. The first person, the adult, goes to work on the first tour of the day, while the child goes to school. Both persons participate in a second home-based tour for either shopping and/or eating out in the evening.

The next step is to identify the primary stop and destination TAZ on each tour. Here, we have identified the work stop on the adult's first tour as the primary purpose and destination. On the child's first tour, the school stop is the only one, so it is an obvious choice. On the joint tour, there is also just one stop. Different modeling systems might use a different scheme for identifying the primary stop on a tour. Typically, the first consideration would be work, school or college stops. For tours that don't involve work, school or college, such as shopping, social/recreation and others, a few of different rules have been used. One rule is to choose the first stop on the tour; another is to choose the stop that has the longest duration; and a third is to choose the stop that is furthest from the tour origin, which will usually be home or the workplace if a sub-tour. There are pros and cons to adopting each rule.

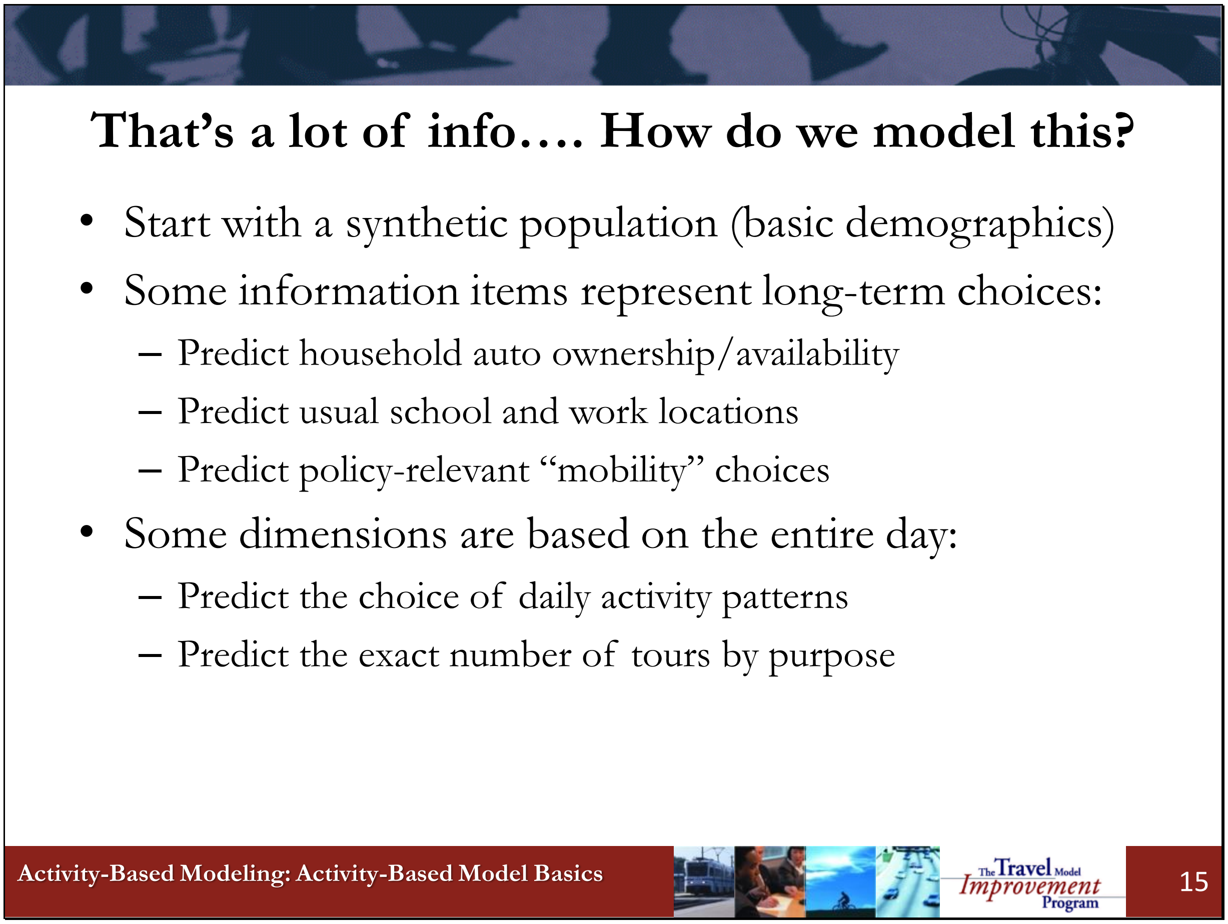
We also need to identify intermediate stops on tours. In the adult’s first tour, there is one intermediate stop (the escort) prior to the primary stop, and two more intermediate stops afterwards, for household business and shopping.

From Diary Data to Daily Activity Patterns

- Identify daily activity patterns
  - Usually defined by presence of tours by type, possibly number of tours by type
  - "Stay home all day" is a legitimate daily pattern

hhid	perid	dayno	tourno	actno	activity	TAZ	primary	arrive	depart	duration	trip time	trip mode	day pattern
626	1	2	0	0	Home	39	0	0	7:00	0	0	none	W-J
626	1	2	1	1	Escort	82	0	7:05	7:10	5	5	auto driver occ 2+	W-J
626	1	2	1	2	Work	1290	1	7:20	15:25	485	10	auto driver SOV	W-J
626	1	2	1	3	HH Bus	160	0	15:50	16:10	20	25	auto driver SOV	W-J
626	1	2	1	4	Shopping	96	0	16:20	17:00	40	10	auto driver SOV	W-J
626	1	2	0	5	Home	39	0	17:10	19:00	110	10	auto driver SOV	W-J
626	1	2	2	6	Jnt Shop/Eat	87	1	19:05	21:00	115	5	auto driver occ 2+	W-J
626	1	2	0	7	Home	39	0	21:10		50	10	auto driver occ 2+	W-J
626	2	2	0	0	Home	39	0	0	7:00	0	0	none	S-J
626	2	2	1	1	School	82	1	7:05	13:40	395	5	auto passenger	S-J
626	2	2	0	2	Home	39	0	14:00	19:00	300	20	school bus	S-J
626	2	2	2	3	Jnt Shop/Eat	87	1	19:05	21:00	115	5	auto passenger	S-J
626	2	2	0	4	Home	39	0	21:10		50	10	auto passenger	S-J

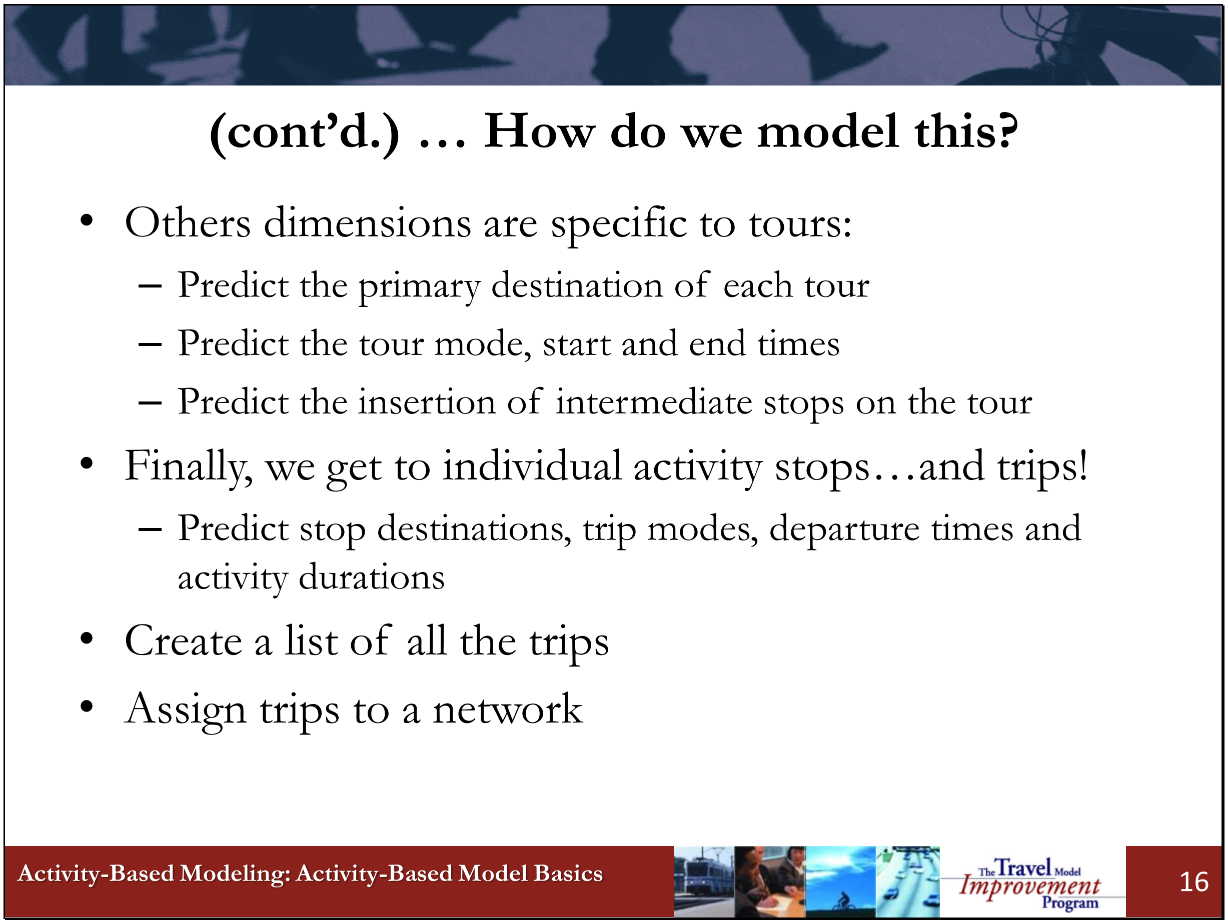
Earlier we discussed the concept of a daily activity pattern, that relates multiple tours by the same person in the same day. Daily activity patterns are usually identified by the presence of tours by types, or even the number of tours by type. It should be noted that "stay home all day" is also a legitimate daily pattern, and this happens to be observed quite commonly in diary data, particularly among certain age groups, the very young and very old. In this example, we have made up a daily pattern code, symbolized by a letter for each tour type in the pattern. In practice, different activity-based modeling systems have developed a several different daily activity pattern coding schemes. Whereas the treatment of tours and stops on tours seems to be quite similar from on activity-based modeling system to the next, there does tend to be significant differences in the ways in which daily patterns are coded, which has implications for overall model system design.



Looking at travel diary data from the standpoint of identifying daily patterns, tours, and various components of tours, there is a lot of information to consider. This goes well beyond the focus on individual trips. So, how do we model all of these activities, tours and day patterns?

The first step is to create a synthetic population, representing every person and household in your study region, with basic demographic information, such as income, gender, worker and student status, and other attribute data, as needed to support modeling. For each of these synthetic households and persons, we then predict long-term choices that are important to modeling activity-travel generation and scheduling. For persons identified as workers, we model the choice of a usual work place, and for students, a usual school or college location. Auto availability/ownership is also important and should be modeled as a household-level decision.

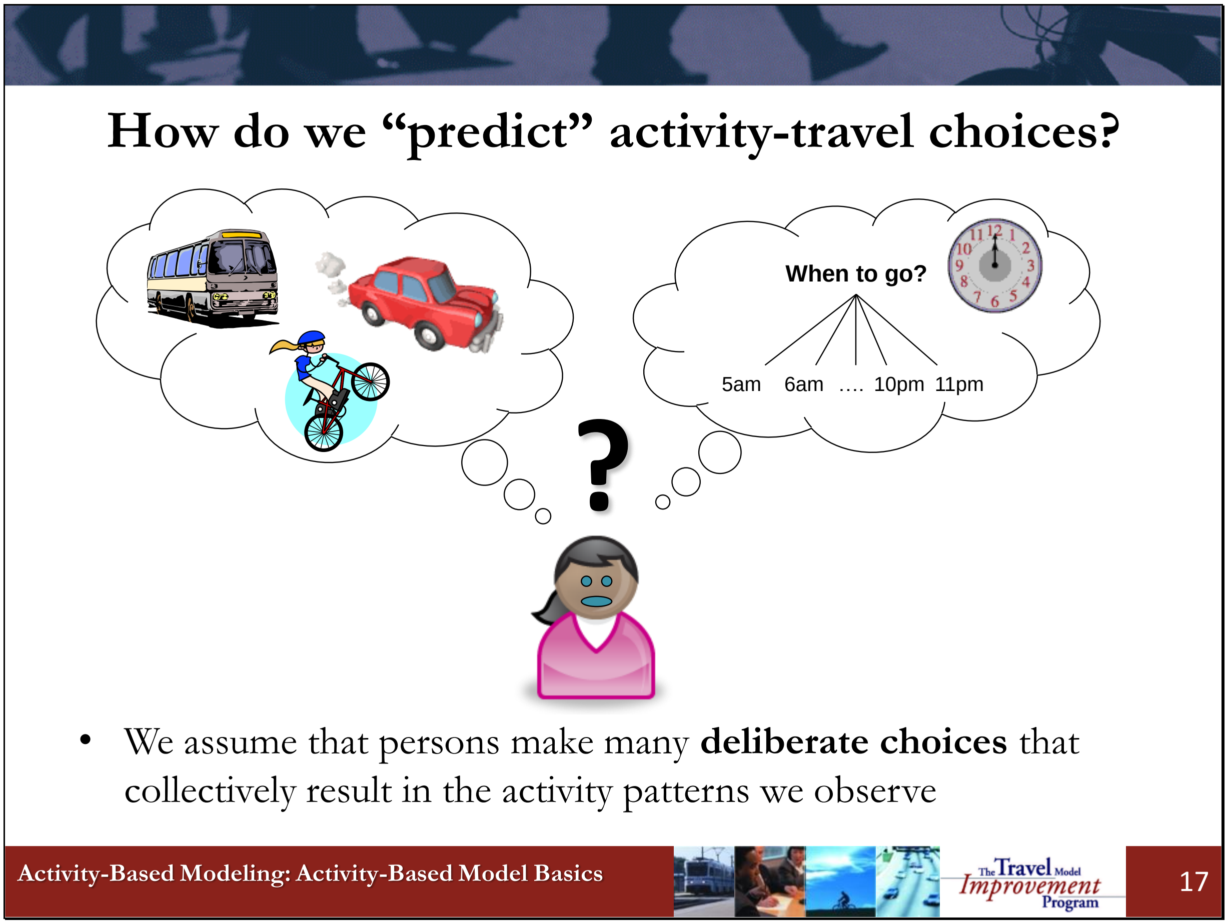
It may also be desirable to model what may be referred to as "mobility choices", such as transit pass holding, free parking at work, or participation in a travel demand management (TDM) program. These are policy-sensitive parameters that we would not expect to be available as standard inputs from Census level population inputs; therefore, we forecast them, based on models derived from our survey sample, and potentially other sources. Next, we will want to predict day-level activity patterns and the exact number of tours by purpose.



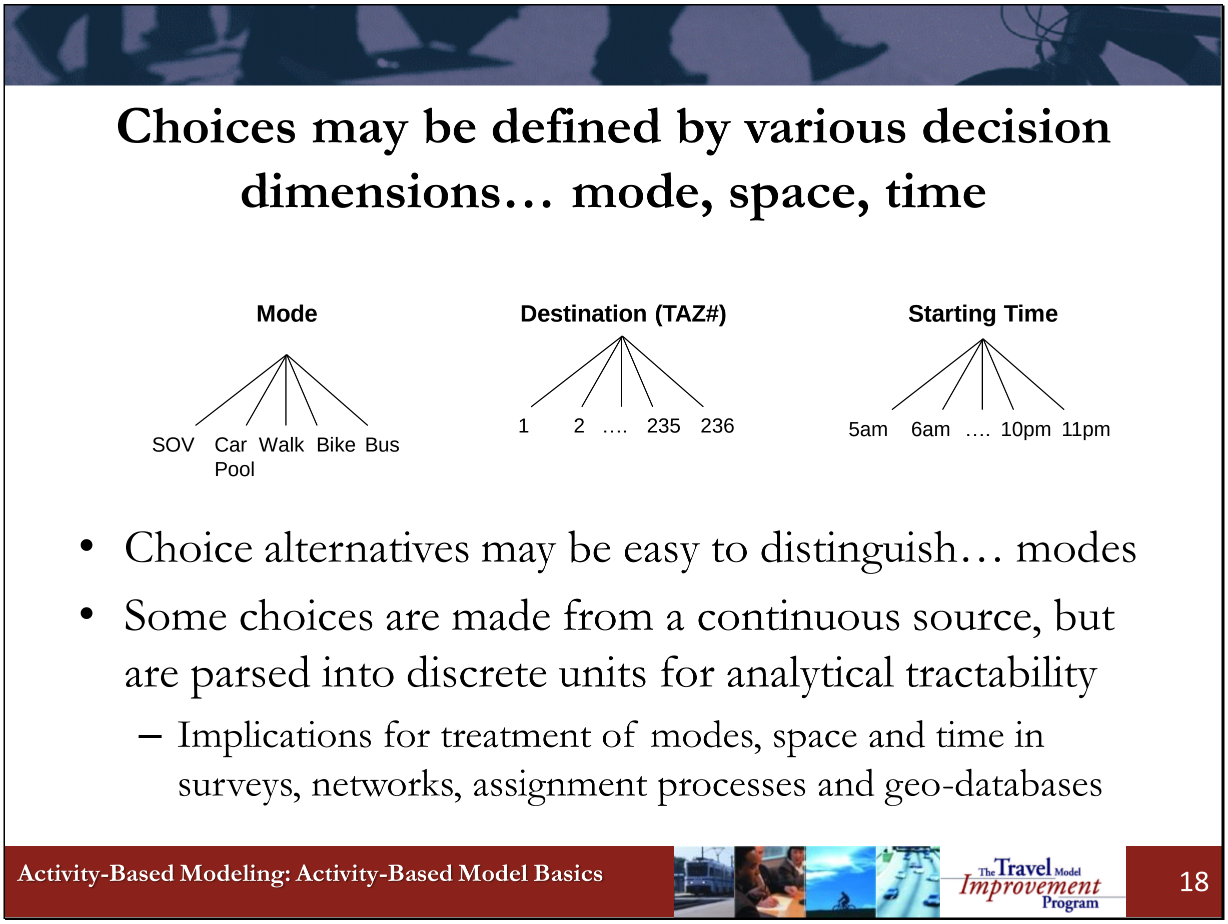
Having covered, long-term choices, mobility choices, and day-pattern choices, we now have predictions for exact number of tours by various types. We can then begin filling in the details for each tour. This would include predicting the primary destination for each tour, followed by the tour mode and start and end times. Given the primary destination and anchor, we can then predict whether there will be any intermediate stops on the tour and how many. Given a primary tour destination and mode and start and end times of day, we can then predict the details of individual trips on the tours. These details would include individual activity durations, stop destinations, trip modes and departure times. We do this for every person in our synthetic population, and we usually process entire households together, particularly in model systems with explicit intra-household interactions. We then output the results into an activity-trip list. The trip-list can then transformed into trip tables and assigned to a network.

All of these steps we will discuss in more detail below and throughout the webinar series. What we have shown here is a representative activity-based modeling process—as you shall see, different modeling systems may order the sequence of these steps somewhat differently, and some systems may describe the prediction steps somewhat differently, or even add steps.

Page 17

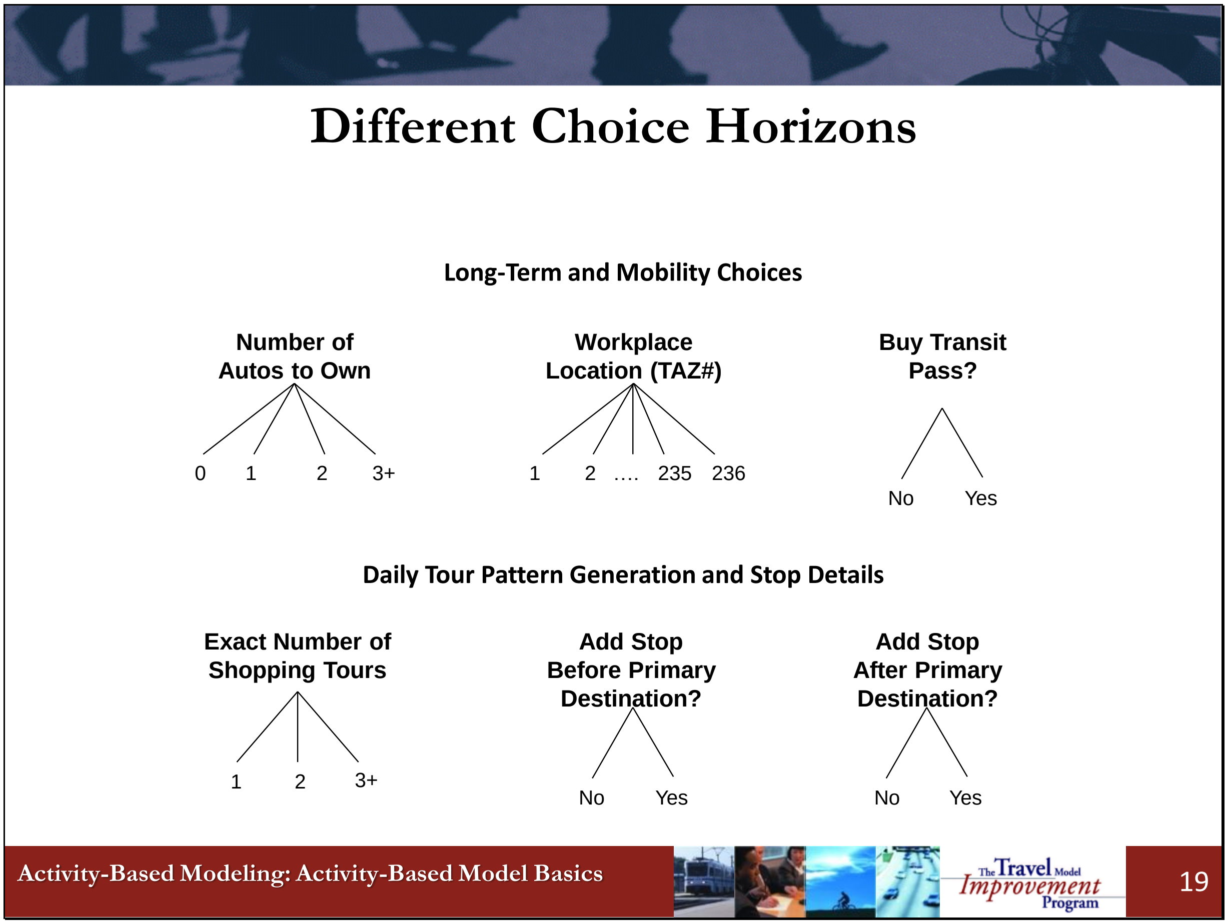


Up to this point, we've mentioned the need to predict many different elements of an activity-travel pattern, but we haven't talked about the methods for doing so. In this next section, we will touch lightly on the methods of predicting travel choices. We begin from the theoretical premise that these choices are not simply random, but rather the outcome of deliberate decision making. Further, we assume that elemental choices that we try to model collectively represent the activity-travel patterns that we observe.



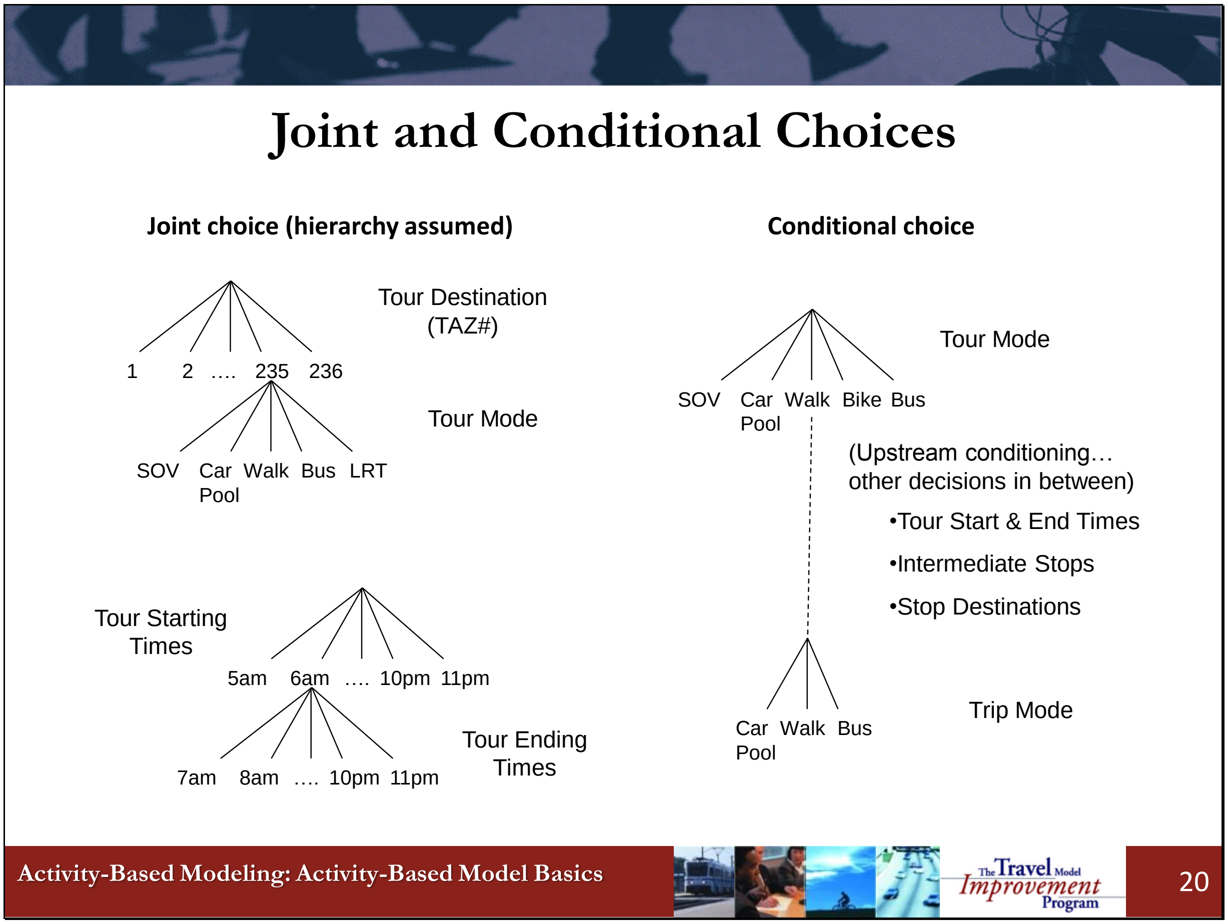
We have talked about several different choice dimensions. Anyone who has worked with travel demand models should already be familiar with mode choice models. In this diagram, we depict the choice from among five different mode choice options. However, we can model other types of choices, including the choice of destination and time of day. In these two examples, we're considering choices made in space and time respectively, which are really continuous dimensions. For analytical convenience, we can parse them into discrete units that make our job considerably easier. The fundamental unit of analysis might be a zone, or it could be something even smaller, such as a grid cell, micro-zone or even parcel. Likewise, we can parse time into intervals and choose a starting time interval for our activity. Activity-based modeling systems in use to date have used 60, 30 and even 15-minute decision intervals.

The ways in which space and time are transformed into discrete intervals has important implications for how these data are processed in surveys, for the creation of networks loading points, assignment time intervals, and the maintenance of geo-databases. We'll discuss the data implications of model design in more detail later.



In addition to different units of analysis, our models will need to consider different time horizons. In this illustration we have listed a few of the common long-term choice models that appear in activity-based modeling systems. These include the choice of workplace location (defined as a TAZ), the number of household auto to own, and whether an individual would purchase a transit pass. While workplace and transit pass are individual choices, the number of autos to own is an example of a household level choice. As you might imagine, in real life, these three decisions might be somewhat interdependent. The number of cars owned might depend on where individual household members work. In the case of some part-time workers, however, the direction of causality might be reversed. In addition, whether someone bought a transit pass might also depend on workplace and the availability of autos. The sequence in which these decisions are represented in activity-based modeling systems is part of the model design.

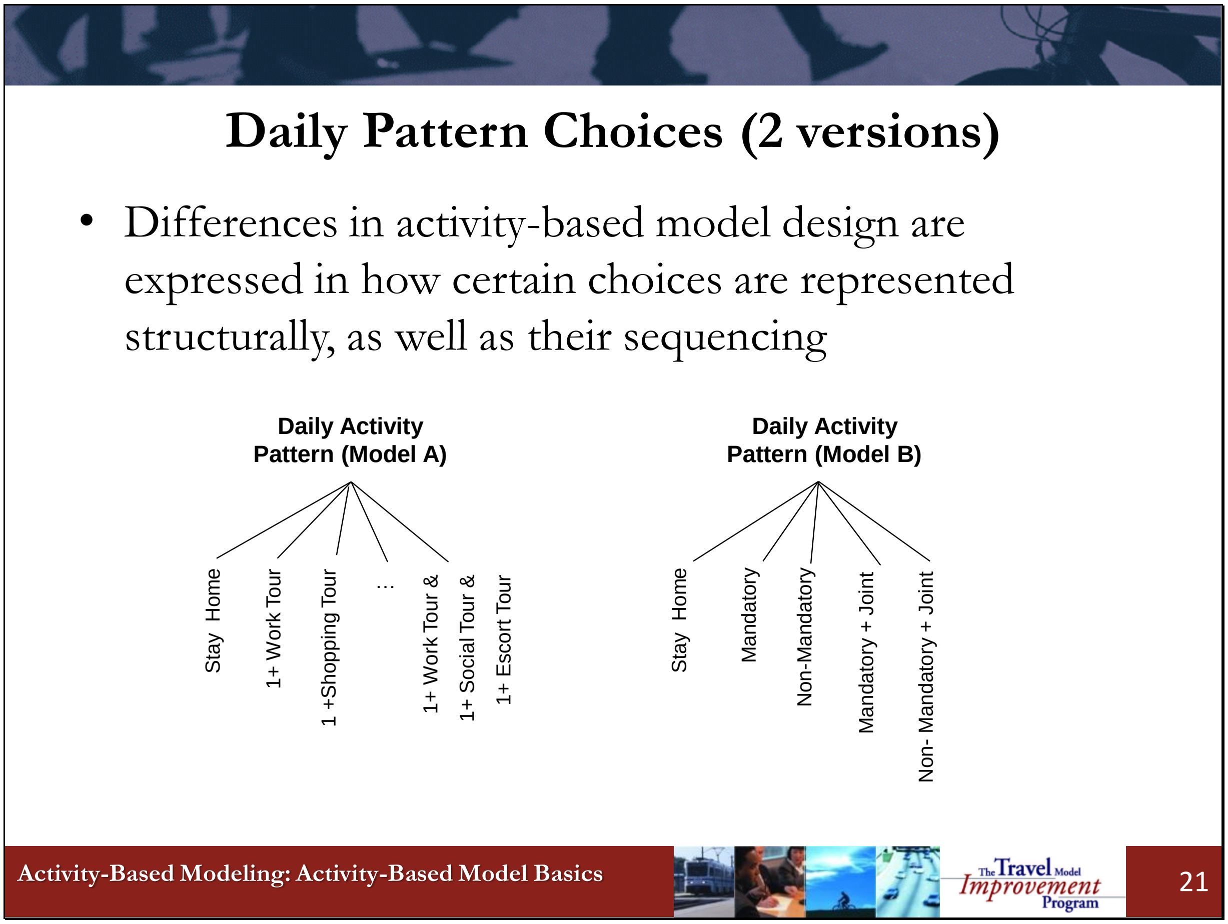
This next group of choices represents model aimed at daily tour pattern generation and certain stop details. Here, you may notice that we are actually predicting the number of shopping tours in a daily pattern, given the existence of at least one. Then we have a couple of models that predict whether to add an intermediate stop before or after the primary destination stop. So, we can use a choice mode approach to predict the frequency of occurrences of something like tours, where the number is likely to be small (say 0, 1, 2, 3+). We can also use a binary choice structure to predict a “yes-no” type of response. The add stop model might be applied multiple times. For example, applied once to predict an initial insertion of an intermediate stop, then re-evaluated to predict whether there is room in the schedule for more. As you might imagine, there are other ways that we could represent this decision process, including linked and ordered choices. In practice, we have found that some of the simplest model structures work the best over a wide range of input cases.



Because some many choices are interdependent, activity-based modeling system designs try to capture these interdependencies to the extent practical. Many people believe that what we have described as separate choices, such as mode and destination, are really bundled choices; for example, choosing between combinations of mode and destination. This first diagram represents the joint choice of tour primary destination and mode in a hierarchical manner. We could enumerate every combination of destination and mode, all on the same level, but that is not necessarily more accurate and is definitely less practical in terms of model estimation and application. In this example, we have chosen to represent the choice of destination first, and conditional upon destination, we have the choice of mode. The important take away is that the choice of destination conditions the choice of mode, and that the composite travel times and costs of the modes available to travel to each destination alternative affect the choice of the destination.

Another example of a joint or conditional choice would be the choice of tour starting and ending times. Here there is an obvious logical constraint being enforced in which the tour ending time intervals must be later than tour starting time intervals. Implicit in this choice of starting and ending times is the tour duration. The choice hierarchy seems clear due to temporal ordering of starting times before ending times; however, the utility of ending time and duration may influence the tour starting time.

Yet another example of a conditional choice would be the choice of trip mode, conditional upon tour mode. This would seem to be a rather obvious hierarchical relationship in which the mode chosen for the whole tour dictates what is available for individual trips on the tour. In this example, the person chose to walk for the tour mode, leaving SOV and bike unavailable for the subsequent trip mode choices. In an activity-base modeling system, however, there may be one or several other choice decisions that take place in between tour and trip mode choice. For example, after choosing the tour mode, there may be the choices of tour start and end times, the decision of whether and how many intermediate stops to insert, and the choices of destinations for those stops. So, while the choice of trip mode is certainly conditional upon tour mode, it is also conditional upon a handful of other choices that take place upstream.

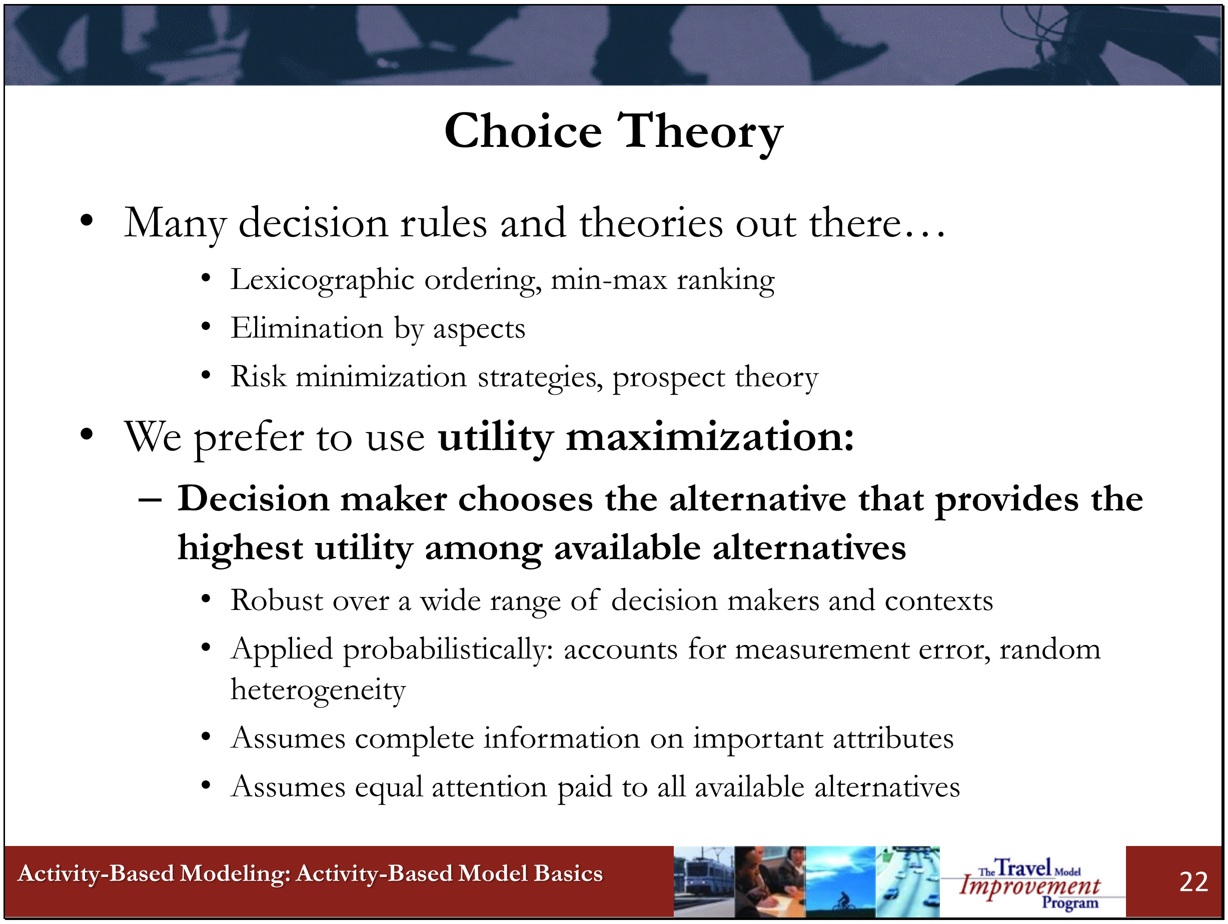


At this point, it should become apparent that there are many different ways in which these choice elements can be represented and integrated into an activity-based modeling system. Here is another example, depicting the representation of the choice of an overarching day pattern for an individual.

The first model (Model A) represents day patterns as combination of tour types. Given a large number of combinations of tours of different types, there could literally be thousands of individually defined alternatives. Although practically, this type of day pattern model, would eliminate those that are observed rarely and group certain alternatives. The exact number of tours of each type would be chosen in a subsequent series of models.

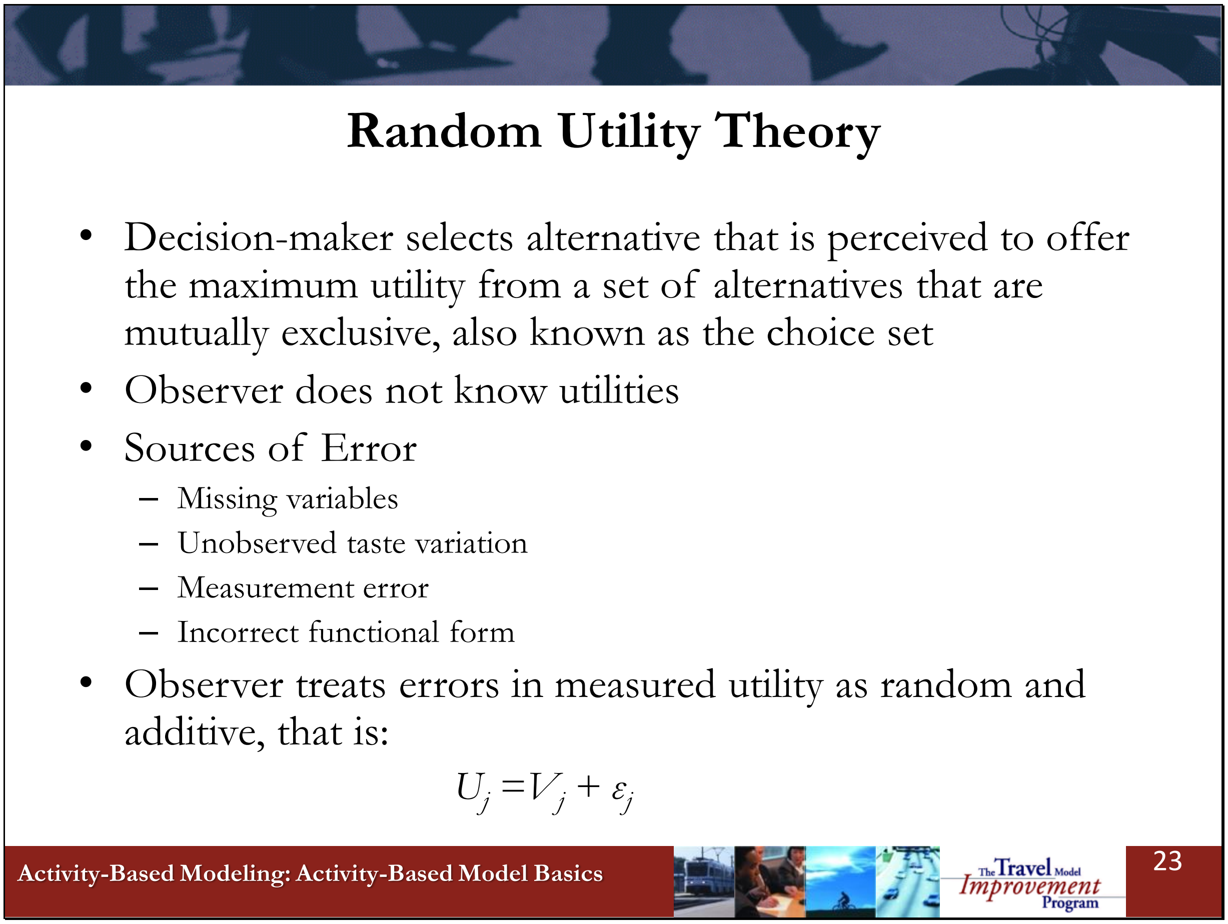
The second model depicted here defines day pattern alternatives differently by characterizing the day patterns as being either "mandatory" or non-mandatory, with a secondary choice of whether to include joint activities with other household members. Mandatory is defined as a pattern involving work, school or college activities, but may include other, discretionary activities, such as eating out, shopping, and social/recreational. A non-mandatory day pattern would include only discretionary activities. Joint activities with other household members are an extra dimension that could be added to either a mandatory or non-mandatory day pattern. The exact number of mandatory, non-mandatory activities and tours, as well as joint activity participation would be determined in downstream models. What is shown here is a simplification.

Again, these are model design decisions. We will begin to discuss different approaches to model design in the second half of this webinar. We will get into the finer points of model design options in the next several webinars in this series. For example, Webinar 8 is devoted to activity pattern generation.



We have discussed the various ways in which activity-travel choices are represented structurally. Next, we will discuss the mathematical formulations for these choice models. There are a number of competing rules and theories that have been proposed by behavioral psychologists and economists to describe the ways in which people make choices. Some of these include lexicographic ordering and ranking strategies, elimination by aspects, risk minimization strategies, and prospect theory. These all have merit in describing certain decision makers in certain contexts.

In activity-based travel modeling, we tend to use utility maximization as a theoretical underpinning. The assumption is that people choose the alternative that provides them with the highest utility among available alternatives. This has been found to be robust over a wide range of decision makers and choice contexts. While it carries with it certain assumptions, it is applied probabilistically in model formulations, which allows us to account for measurement error and random heterogeneity in the population. Some of its somewhat less realistic assumptions include that the decision maker has full knowledge of the attributes of each alternative and pays equal attention to all available alternatives.



Many of you have been exposed to discrete choice models either through your academic training or on-the-job experience in travel demand modeling, although I am sure there are a few persons in the audience who less familiar. Most travel demand modeling professionals are at least familiar with mode choice models. So, this is not intended to be an extensive tutorial on discrete choice models. Our purpose here is to highlight certain important aspects of discrete choice models that are central to their use in activity-based travel modeling. We want to make sure that you are familiar with important terminology. In particular, we want to discuss the roles of choice sets, composite utility or log sums, and how models are applied in a simulation environment.

Starting with Random Utility Theory, we assume that the decision-maker selects alternative that is perceived to offer the maximum utility from a set of alternatives that are mutually exclusive, which we call the choice set. The observer does not know utilities; however, they may be inferred from the choices made.

Sources of error include: missing variables, unobserved taste variation (preferences), measurement error (actual versus perceived travel time), and using the incorrect functional form (linear, non-linear, hierarchical, etc.). We treat these errors and random and additive, resulting in the utility formulation shown here. Total utility is composed of a systematic portion V_j, which we represent through the variables in our model, and a random component symbolized by the epsilon error term.

Choice Probabilities

- Probability of choosing alternative *i* from a set of choice alternatives *C*

$$P(i:C) = \text{Prob}\left(U_i \geq U_j\,, \quad \forall j \subset C\right)$$
$$= \text{Prob}\left(V_i + \varepsilon_i \geq V_j + \varepsilon_j\,, \quad \forall j \subset C\right)$$

- Under general assumptions, the model is…

$$P(i:C) = \frac{\exp\left(V_i\right)}{\sum_{\forall j} \exp\left(V_j\right)}$$

- Probability is based on the difference in utility between alternatives

$$\frac{\exp(V_i)}{\exp(V_j)} = \exp(V_i - V_j)$$

The probability of choosing an alternative “i” from a set of choice alternatives “C” may be expressed probabilistically in this formula. General assumptions for the distribution of the error term, following a Gumbel distribution, lead to the familiar multinomial logit model shown here. This is the model we see so often representing mode choices. It is also worth noting that probabilities in a logit model are based on the difference in utilities between alternatives, not their ratios.

Utility Expressions: Mode Example

$$
\begin{aligned}
\text{Utility}_{transit} = \; & a * \text{in-vehicle time} \\
+ \; & b * \text{fare} \\
+ \; & c * \text{access time} + \text{egress time} \\
+ \; & d * \text{wait time} \\
+ \; & \text{mode-specific constant}
\end{aligned}
$$

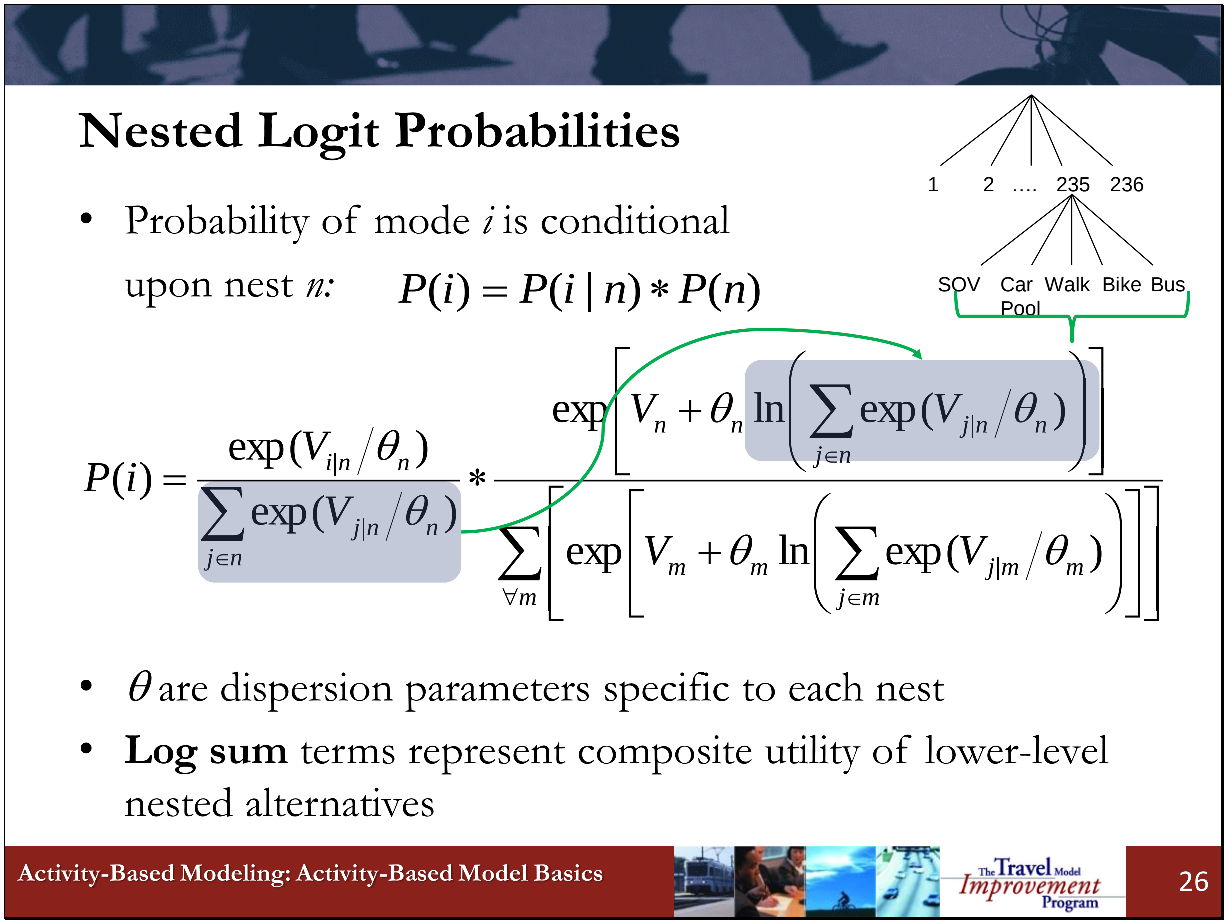
- V_i = Systematic Utility -- the weighted sum of the attributes
- a, b, c, d are the weights, or parameters, in the model
- Parameters are estimated from survey data or borrowed/asserted
- They convert the times and costs to *utiles*
- They are negative if multiplied by time/cost (*disutility*)
- The mode-specific constant is the value of the “non-included” attributes

Activity-Based Modeling: Activity-Based Model Basics

The Travel Model Improvement Program

25

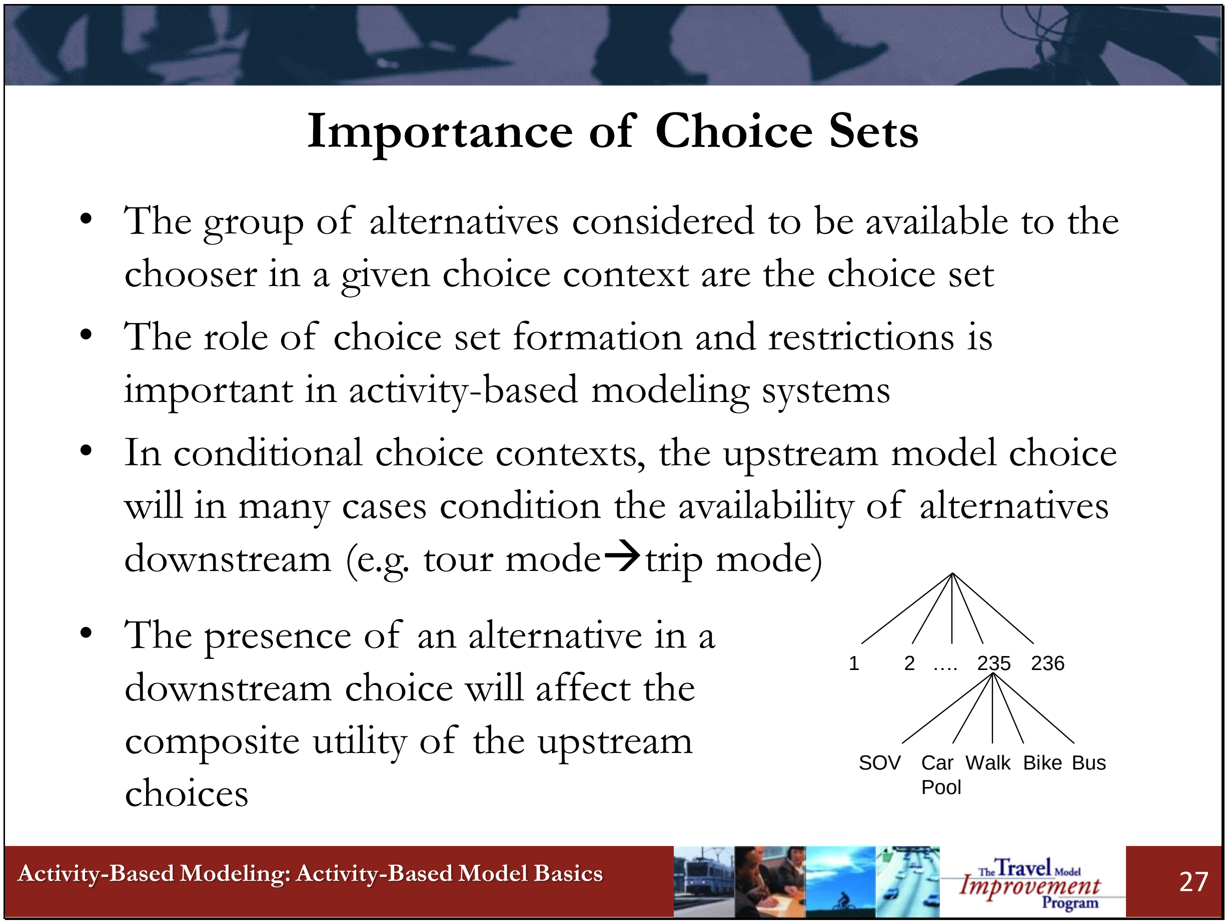
Let’s use a mode choice example to illustrate how a utility function is formulated. The utility equation for transit is shown on this slide. Utility equals the weighted sum of the attributes of the alternative. The weights in the model are known as model parameters, shown here as a, b, c, and d. These parameters can be estimated from survey data, borrowed from another model, or asserted based on experience. The parameters convert the modal attributes in various units such as minutes and cents to a general value called a utile (since they measure utility). This has important implications for how the weights can be compared to one another. Note that there is also a term called a mode-specific (or alternative-specific) constant. This represents the value (in utiles) of all of the attributes of the alternative that are not explicitly listed in the utility equation. In the case of transit, this could include difficult-to-measure factors such as transit reliability, transit safety, and the influence of weather on the choice of transit.



Earlier, we were discussing several examples of joint and conditional choices. Activity-based modeling systems make extensive use of such hierarchical or "nested choices". For example, we looked at the nested choice of model conditional upon destination. We also considered the nested choice of tour ending time, conditional upon tour starting time. We can represent the conditional probability of a choice that appears in a lower-level nest upon the choice made in the upper-level nest as follows in these formulas. The theta parameters are dispersion terms that reflect the correlation between alternatives in the same nest. In order to be consistent with utility maximization parameters, theta must have values greater than zero and less than or equal to one.

The term highlighted here represents the composite utility of the nested alternative. Notice that the denominator for the lower-level choice (mode) appears in the utility expression of the upper-level choice of destination zone. Because we take the natural log of this sum, this term is commonly referred to as the "log sum". It represents the maximum expected utility that may be derived from the lower-level choice, which in this case is mode. In the choice of a destination shown here, the log-sum term represents the mode-weighted accessibility for travel to each zone alternative. Another portion of the utility of the zone alternatives V_m is shown here, and represents other attributes of the zone, such as attraction variables.

Thus, it is common in activity-based models to use composite accessibilities, such as mode choice log-sums to account for travel times and costs by all available modes when choosing a destination. The assumption is of course that the destination is chosen first. We will talk more about how log-sums are used in activity-travel model components in the second half of this webinar and in future webinars.



The choice set is the group of alternatives considered to be available to the chooser in a given choice context. The roles of choice set formation and restrictions are important in activity-based modeling systems. This is particularly true for nested choices and conditional choice relationships, as discussed above. In conditional choice contexts, the upstream model choice will in many cases condition the availability of alternatives downstream. For examples, as we showed earlier, the choice of tour mode conditions the availability of certain trip modes. Generally speaking, if a person does not choose to drive for the tour mode, then we would not expect drive to be available for any trip on the tour. The same would be true with bicycle.

In addition, the presence or absence of an alternative in a lower-level choice may greatly affect the composite utility of the upper level choice. So, in a policy context, if we were to add a new transit service to the region that would greatly improve travel time by certain zones, then this addition of a new alternative to serve those zone pairs would make those destinations more attractive. This change in accessibility would be reflected in the mode choice log sums that would be used by an upstream destination choice model and possibly even long-term workplace and auto ownership choice models.

Prediction: Allocation vs. Simulation

- Classical 4-Step Trip-Based Model—Mode Choice
  1. For each market segment, defined by trip purpose and household demographic group, predict the probability of each mode for each O-D pair.
  2. Allocate the number of trips for each segment and O-D pair to modes in proportion to their predicted probabilities.
  3. Sum over market segments to form trip tables.
- Activity/Tour-Based/Simulation—Mode Choice
  1. Predict probability of each simulated chooser selecting each mode for a specific O-D pair and purpose.
  2. Use Monte Carlo random draws to predict mode choice.
  3. Sum over choosers and purposes, grouped by O-D pair, to form trip tables for network assignment.

Prediction using simulation methods is another important difference between activity-base modeling systems and trip –based models. Let's use mode choice models as an example, since this is the one place where discrete choice models are consistently used in trip-based modeling systems. In a trip-based model, we define market segments by trip purpose and household demographic group, and predict the probability of each mode for each O-D pair. We then allocate the number of trips for each segment and O-D pair to modes in proportion to their predicted probabilities. This is an aggregate prediction, which we then sum over all market segments to form trip tables.

In an activity-based model using simulation, we predict the probability of each simulated chooser selecting a mode for a specific O-D pair and purpose. We then use Monte Carlo random draws to predict a single mode choice. To form trip tables for network assignment, we aggregate over individual trip records, grouped by O-D pair.

Monte Carlo Prediction

1. Predict the probability and cumulative probability for each alternative outcome

	SOV	HOV	Bus	LRT	Walk	Bike
Probability	0.56	0.28	0.03	0.08	0.01	0.04
Cum. Prob.	0.56	0.84	0.87	0.95	0.96	1.00

2. Draw a random number from a uniform distribution on the unit interval (0…1): **e.g. Rand() = 0.76**
3. Select the alternative with the range on the cumulative probability array that includes the random draw

	SOV	HOV	Bus	LRT	Walk	Bike
Lower Bound	0.00	0.57	0.85	0.88	0.96	0.97
Upper Bound	0.56	0.84	0.87	0.95	0.96	1.00

There are three basic steps in Monte Carlo prediction. First, predict the probability of each choice for each household or person making a choice. Here we use mode choice to illustrate an example, but the same applies to any of the choice models we have discussed in this webinar. Next, calculate the cumulative probability of the array of choices, as shown here, such that they add up to 1.0. These values represent the upper bound of prediction bins.

The second step is to draw a random number from a uniform distribution on the unit interval.

The third step is to select the range on the cumulative probability array that includes the random draw. In this particular example, we drew .76, which falls into the bin range for the second alternative, Auto HOV. Had we drawn a different number, say 0.33, we'd be choosing Auto SOV.

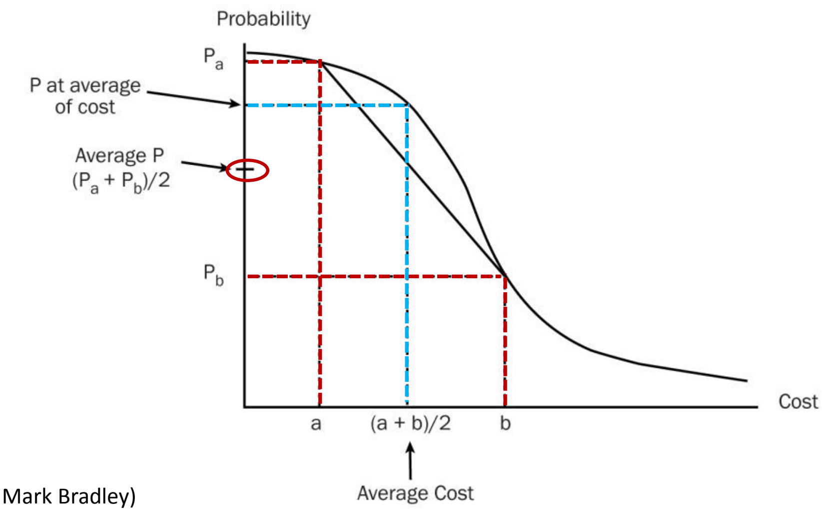
Monte Carlo Simulation Advantages

- Advantages
  - Computational efficiency (when number of segments exceeds number of agents)
    - Full availability of all variables
    - Outcomes of previous model components can be used as explanatory variables in subsequent components
  - Provides distribution of results
  - Avoids aggregation bias

Monte Carlo simulation has advantages and disadvantages compared to expected values used in trip-based models. The key advantage of Monte Carlo simulation is that explanatory variables can be included in models with little computational overhead (as opposed to aggregate models, in which each market segment increases the number of calculations exponentially). Monte Carlo simulation also can provide a distribution of results, though in practice this has not been fully taken advantage of. The main reason for not taking advantage of the ability to forecast a distribution of outcomes is that it would require excessive total run time to do, say, 100 runs. Monte Carlo simulation also helps us avoid aggregation bias in prediction.

Logit Models and Aggregation Bias (I)

- Average probability is not equal to the probably at the average of explanatory variables.



(Graphic source: Mark Bradley)

Activity-Based Modeling: Activity-Based Model Basics

The Travel Model Improvement Program

31

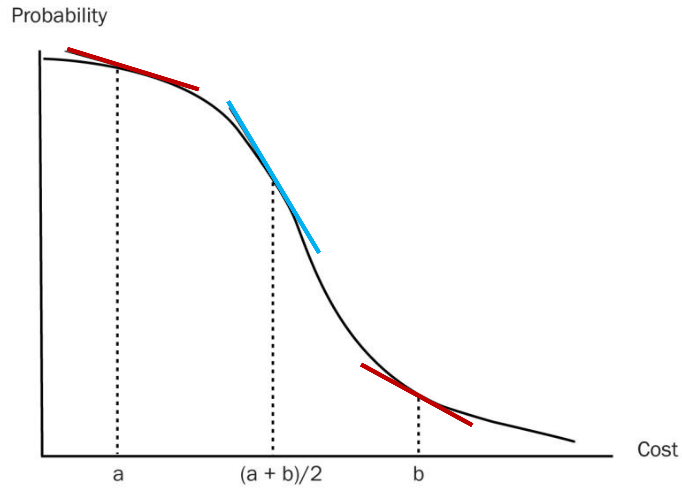
This slide is illustrates the issue of aggregation bias using logit models. On the horizontal axis you have the cost of a choice for two different decision makers, A and B. Follow the red lines. Note that the probabilities for each individual, as predicted by the model are quite different.

Now, if we were to aggregate these two individuals and take their average cost, we'd obtain a different probability that is somewhere in between. This is shown by the blue lines.

However, notice that the probability of this average cost is different than the average probability we obtain when we calculate each person's probability

Logit Models and Aggregation Bias (II)

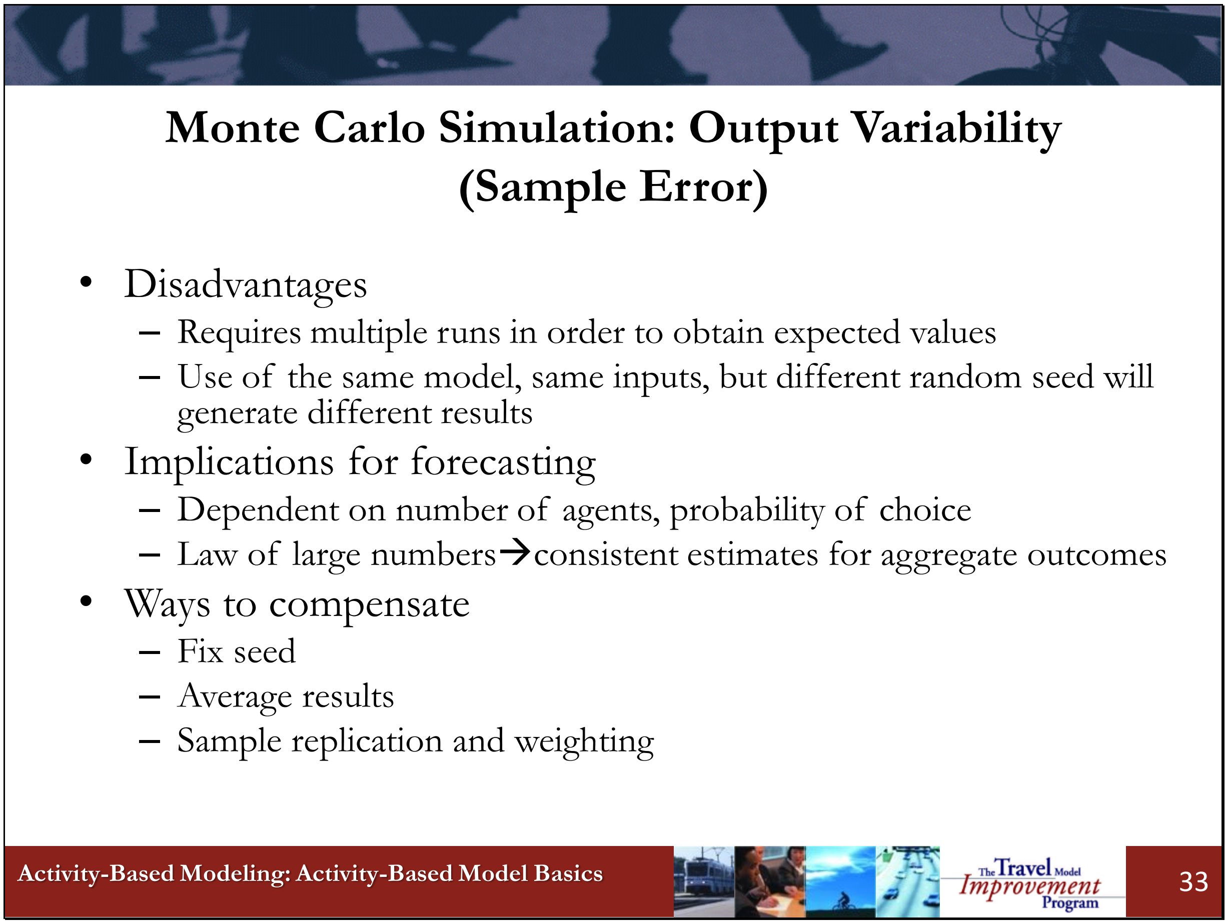
- The average impact of a change (average of slopes at a and b) is not equal to the impact calculated at the average of the explanatory variables.



(Graphic source: Mark Bradley)

This slide illustrates another aspect of aggregation bias. The average impact of a change is not equal to the impact calculated at the average of the explanatory variables. This is symbolized by the tangents to the curve, representing slopes at each point. Again, the red lines represent the individual outcomes, and the blue line represents the slope corresponding to the averaged outcome. Due to the sigmoid (S-shape of the curve), the logit model is most sensitive (elastic) to change in inputs at its center region, and is relatively less sensitive (inelastic) to changes in inputs at its top and bottom ends. This is one reason why some aggregate models predict larger shifts in response to scenario inputs changes than disaggregate models.

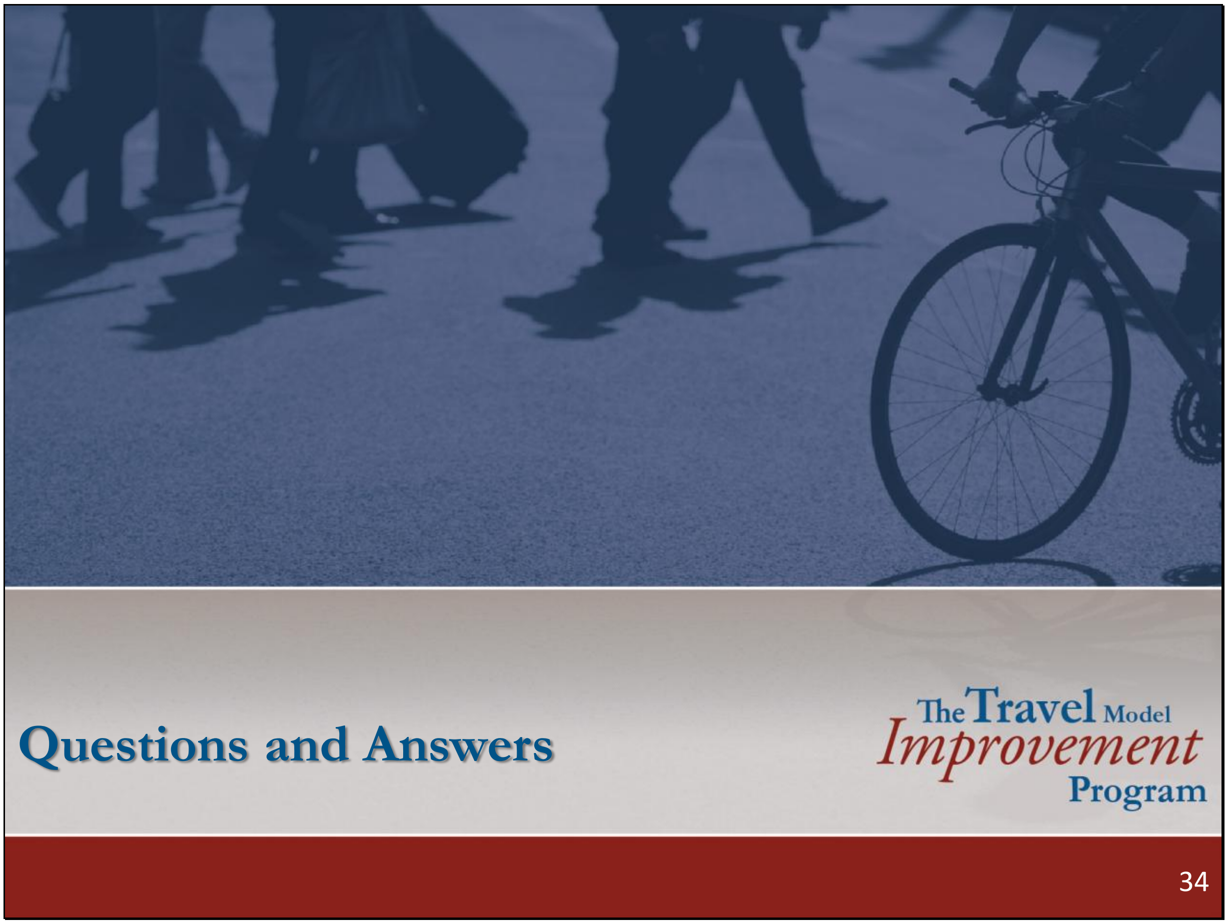
An example of this might be a mode shift in response to a new toll charge. Imagine the perceived cost of the toll being affected by personal values of time, where Person A has a high willingness to pay (so perceived cost is not so onerous) and Person B has a low willingness to pay (so perceived cost is considered to be very onerous). Because both persons are already at the far ends of the distribution, they are less likely to react to a cost change by changing their baseline choices. By grouping travelers under a single average value of time, however, the perceived cost represents an average condition, the blue slope, and has the potential to overestimate the elasticity of response to the toll.



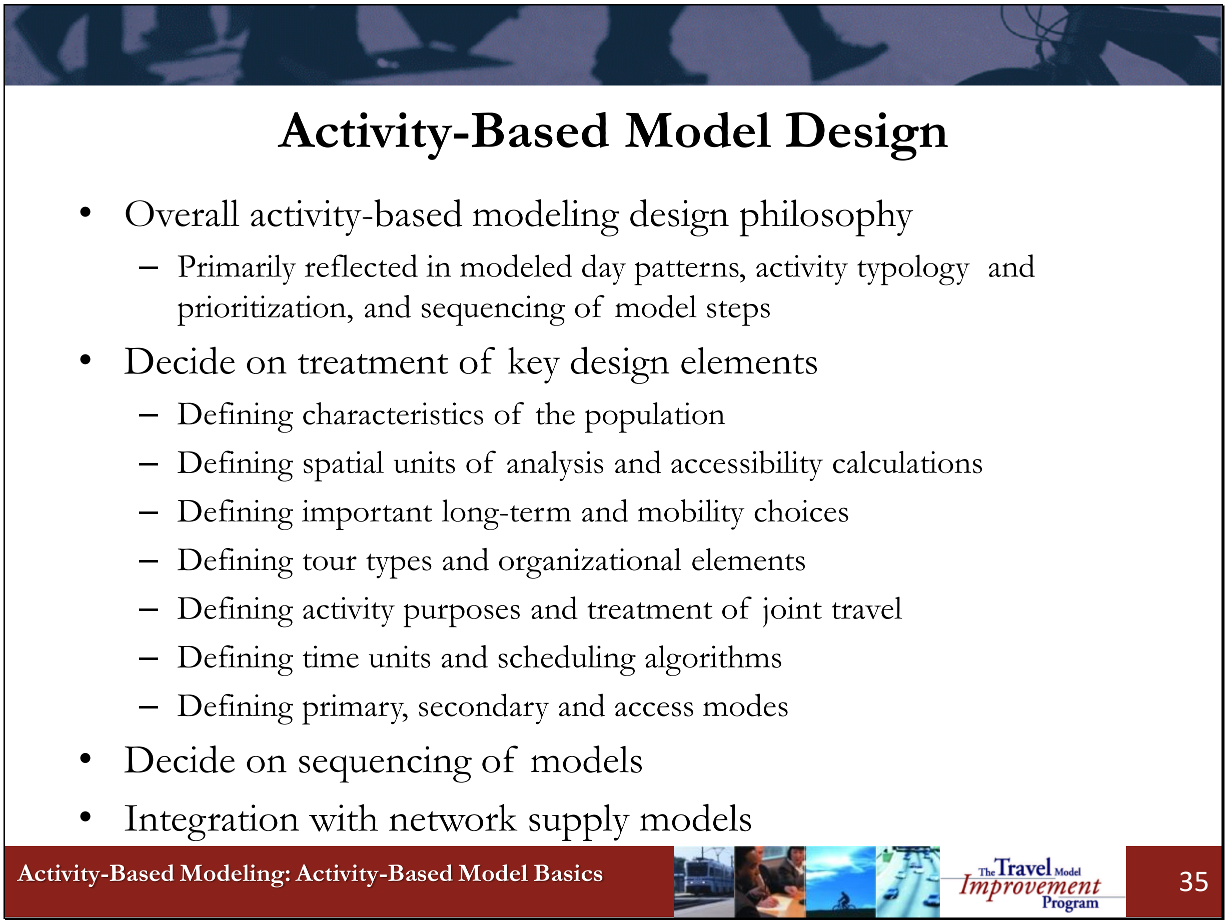
The key disadvantage of Monte Carlo simulation is that multiple runs are required in order to determine the expected values, or average results, for certain model outputs. This has implications for forecasting, but most practical activity-based models compensate for this.

The amount of variability in model results is dependent on the number of decision makers in the choice decision, and the size of the probability of the choice. For example, lower probability choices have more variability in their outcomes than higher probability choices. Because most outputs from activity-based models are aggregations of choices, one run of the model can be a sufficient indication of the expected outcome from a policy. For example, regional VMT, VHT, district-level tour flows, tours and trips by mode, and higher facility-type link estimates and transit line boardings are very stable from run to run. However, more disaggregate analysis, such as TAZ-level origins and destinations, lower facility type link loadings, and lower ridership transit routes, can have more variation and therefore multiple runs of the model system may be required, where results are averaged across the runs.

One way to compensate for Monte Carlo variability is to fix the random number seed, such that the results will only vary according to changes in inputs. This ensures stability from run to run, but at the cost of representing only one possible outcome. Also, results can be averaged to obtain an expected value. Finally, the synthetic population can be replicated, and the model can be run once where each sample can be weighted to obtain an average or expected outcome.


Questions and Answers
The Travel Model
Improvement
Program
34

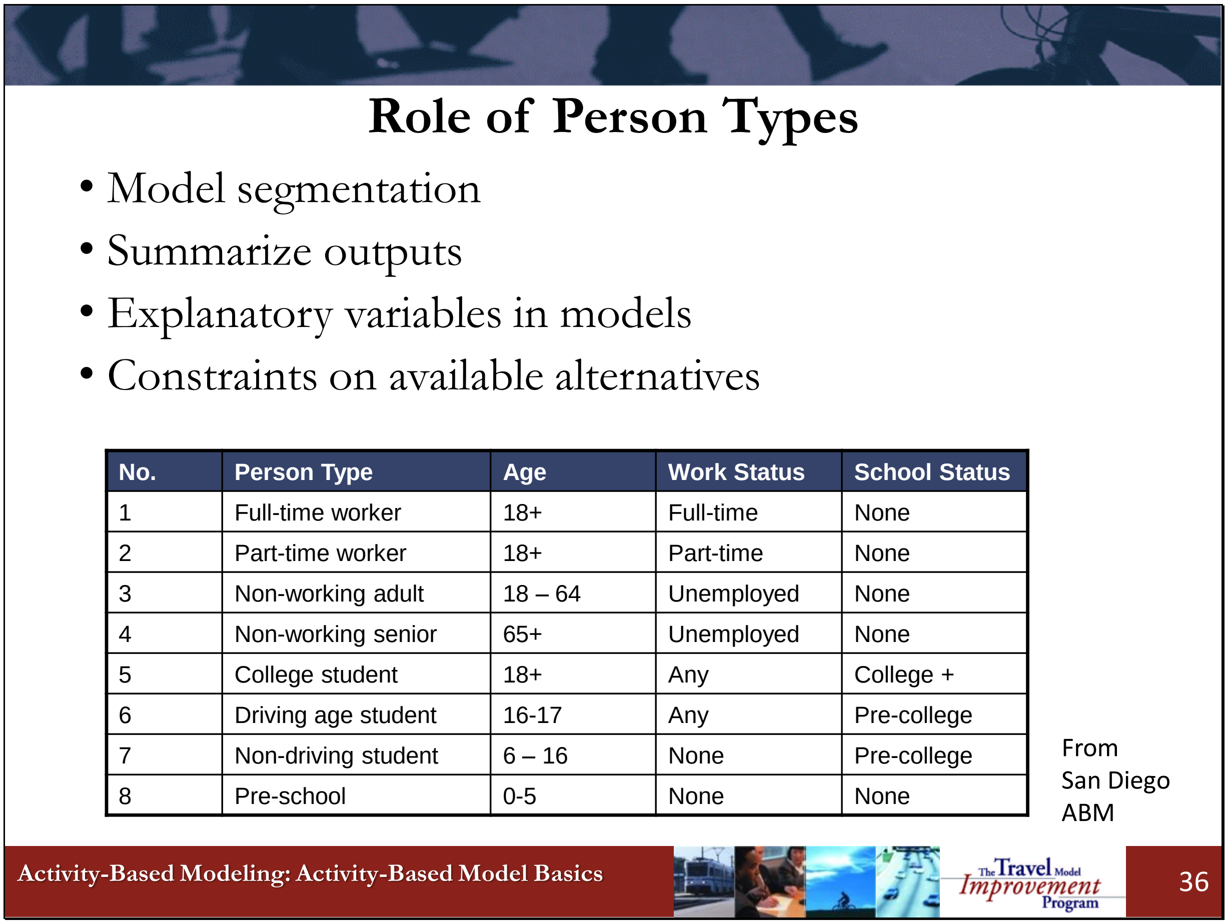
Page 35



Having examined how activity modeling choices are represented in both data and model structures, we can now talk about model system design. If you study the different model designs in detail, you will find different philosophies expressed in terms of how to define daily activity patterns, how activities themselves are defined, and the sequencing of model steps. In addition, there are several key design elements that must be decided upon. These include:

- Defining characteristics of the population
- Defining important long-term and mobility choices
- Defining spatial units of analysis and accessibility calculations
- Defining tour types and organizational elements
- Defining activity purposes and treatment of joint travel
- Defining time units and scheduling algorithms
- Defining primary, secondary and access modes

In addition, we need to decide on the sequencing of model steps and how the model system will be integrated with network supply models.



Role of Person Types

- Model segmentation
- Summarize outputs
- Explanatory variables in models
- Constraints on available alternatives

No.	Person Type	Age	Work Status	School Status
1	Full-time worker	18+	Full-time	None
2	Part-time worker	18+	Part-time	None
3	Non-working adult	18 – 64	Unemployed	None
4	Non-working senior	65+	Unemployed	None
5	College student	18+	Any	College +
6	Driving age student	16-17	Any	Pre-college
7	Non-driving student	6 – 16	None	Pre-college
8	Pre-school	0-5	None	None

From San Diego ABM

Activity-Based Modeling: Activity-Based Model Basics

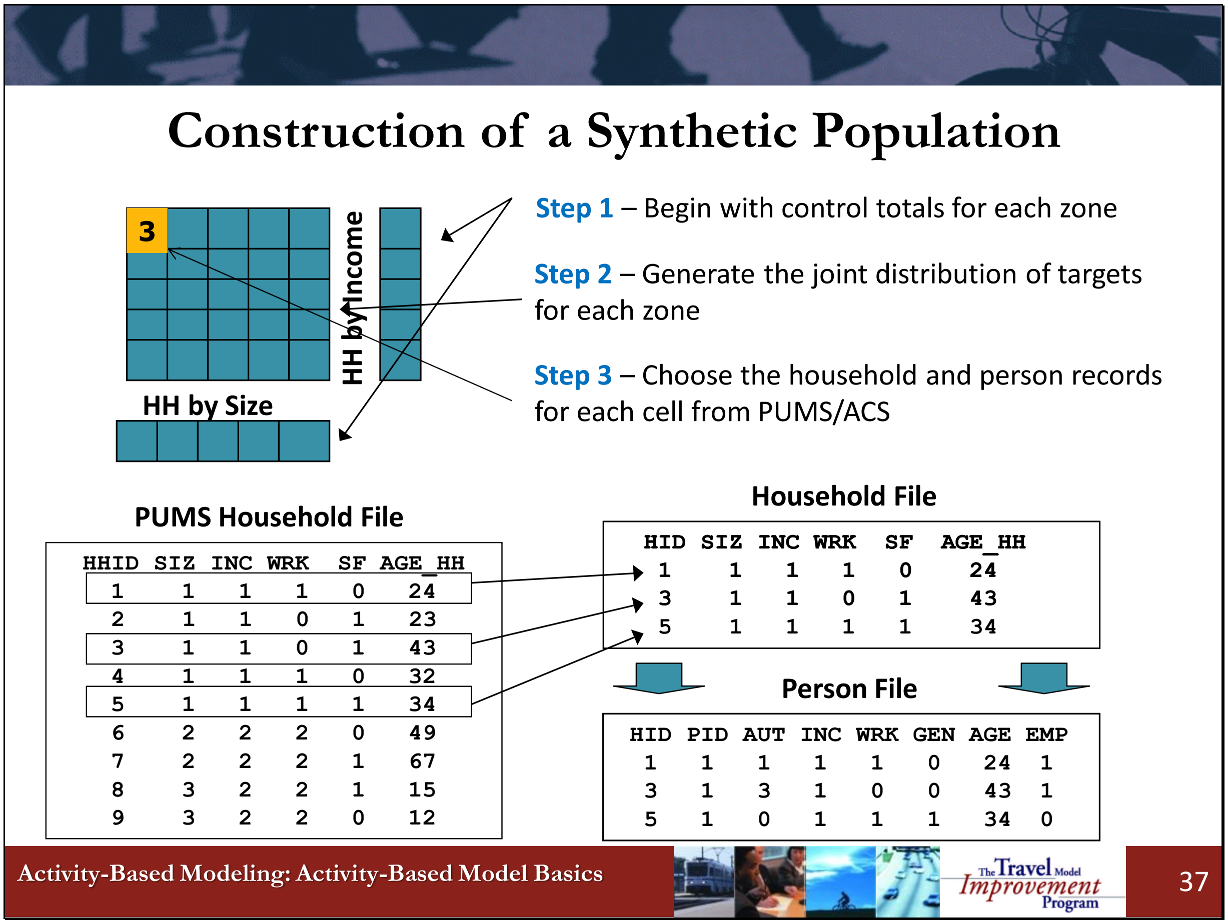
The Travel Model Improvement Program

36

One of the first design elements to be considered is the synthetic population and the role that person and household attributes will play in the overall model design. Although we have been telling you throughout this series how beneficial it is to model persons in disaggregate form, it is often useful to create person-type categories for the sake of modeling convenience. The table shows an example of person-type categories and definitions. Person-type categories can be used for a number of purposes:

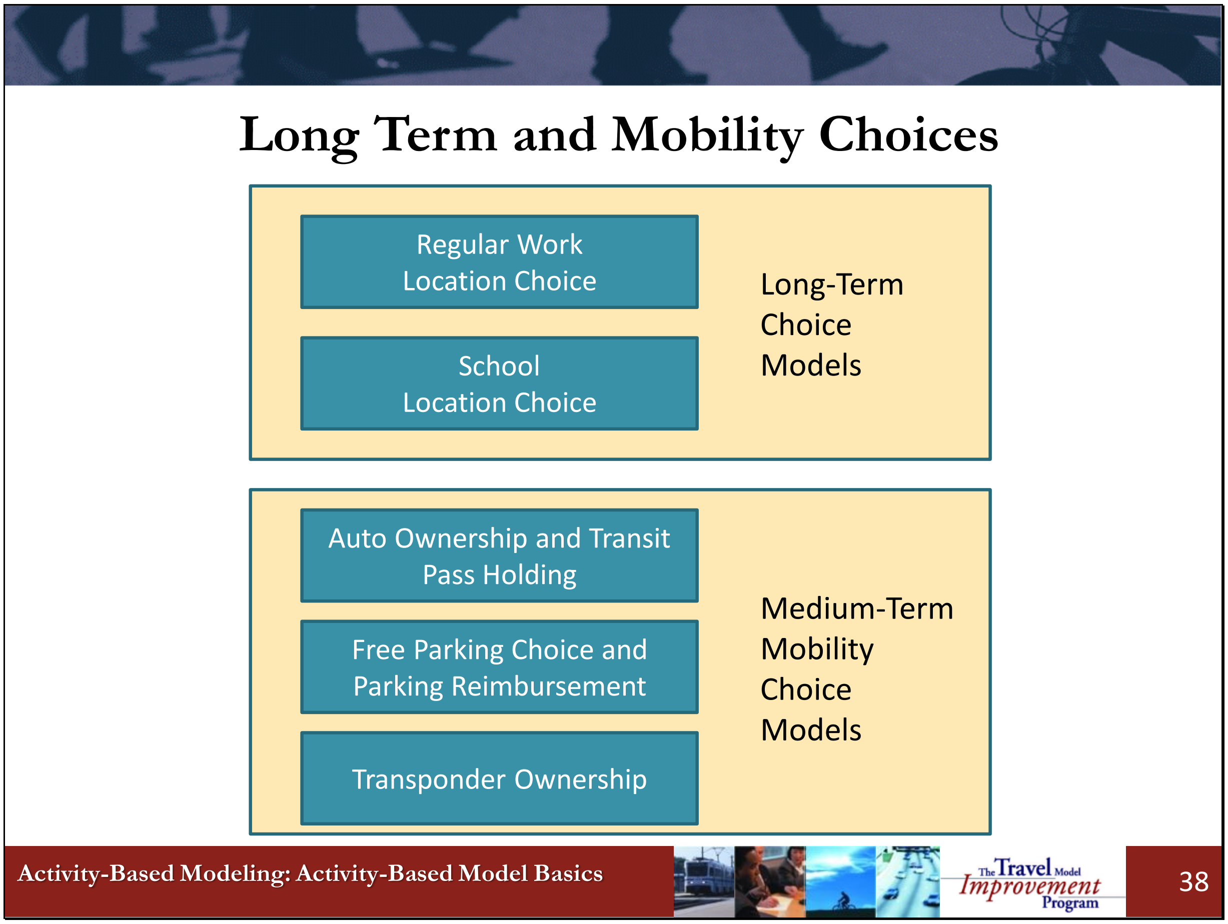
- As a basic segmentation for certain models, such as daily activity pattern models.
- To summarize and compare observed versus estimated data and calibrate models.
- As explanatory variables in models.
- As constraints on alternatives that are available; for example, mandatory activities are only available to workers and students.

Some categories, such as work and school status, are there because they make logical sense. Determining the proper cutoff points for categorical variables related to age and income are usually derived in the model development process empirically through descriptive statistical work. For example, a community with a large population of retirees might use different age groupings than a college town. Not having proper cutoff points could result in important market segments being under-represented in model specifications and estimated models not explaining as much behavioral variation in the population as they could.



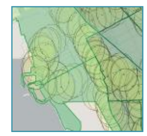
The attributes and attribute levels specified in the synthetic population should be consistent with the person types actually needed in the model. This slide shows one possible approach for construction of a synthetic population. In the first step, control totals for each zone are generated. In the base-year, these can be obtained from Census. In the future year, these control totals can be based upon growth factor models, allocation methods, and/or land-use models. In the second step, the aggregate characteristics of the synthetic population are defined, in the form of a joint distribution of household attributes for each zone. In the third step, individual person and household records are drawn at random from a sample of households, such that combined they meet the joint distribution defined for the zone. There are other approaches, such as weighting disaggregate PUMS records directly according to the control totals, at both a household and a person level, and then selecting households according to those weights.

These techniques will be described in more detail in the Webinar 5. Some of the decisions that need to be made will include which household and person attributes are most important to control and at what level of spatial resolution, given the available data, and what other household or person attributes can be added as uncontrolled attributes.



Long-term choice models include models of regular work location for workers, and school location choice for students. These have become fairly standard in all activity-based modeling systems. Possible medium-term mobility models include auto ownership, transit pass holding, free parking reimbursement (for workers who work in parking-constrained areas such as downtowns), and transponder ownership. The reason that these models are placed first in the model chain is that these choices are made on a long-term temporal scale, and they condition later choices that are made on a daily or even hourly time scale. Some model systems provide the ability to “turn off” these models, keeping their choices fixed from a baseline model run, so that only short-term or daily decisions can be analyzed. One motivation for this feature is the ability to measure “ramp-up” periods. For example when a new toll road is first opened, workers have not yet adjusted their work location choice based on the level-of-service offered by the facility, and many households have not acquired transponder units. In the longer-term, workers might be expected to adjust to the level-of-service provided by the facility and households who are willing to pay for the toll facility would be expected to acquire transponder units. In webinar 7, we will cover both long-term choice models and medium-term mobility models and discuss these issues in more detail.

Treatment of Space

Spatial Representation	Diagram
Zones •Already exists for most MPOs •The most aggregation error, particularly for non-motorized and transit modes	
Sub-zones • Created by buffering around transit lines, stops • Improved representation of walk-transit, but may not be consistent with skims • Doesn't help with non-motorized representation (intra-zonal walk and bike)	
Micro-zones • Created by sub-dividing zones (7-10:1) • Best representation of transit accessibility when coupled with stop-stop skims • Improved representation of non-motorized time	
Parcels • Created via parcel database • Improves representation of walk-transit, but need to make consistent w/ skims • Best representation of non-motorized time	

Activity-Based Modeling: Activity-Based Model Basics

The Travel Model Improvement Program

39

As discussed earlier, another important design element is the treatment of space and the calculation of accessibilities. This slide shows commonly-used spatial systems in activity-based models, and describes some trade-offs between them. There are generally four different treatments of space that can and have been used in activity-based models.

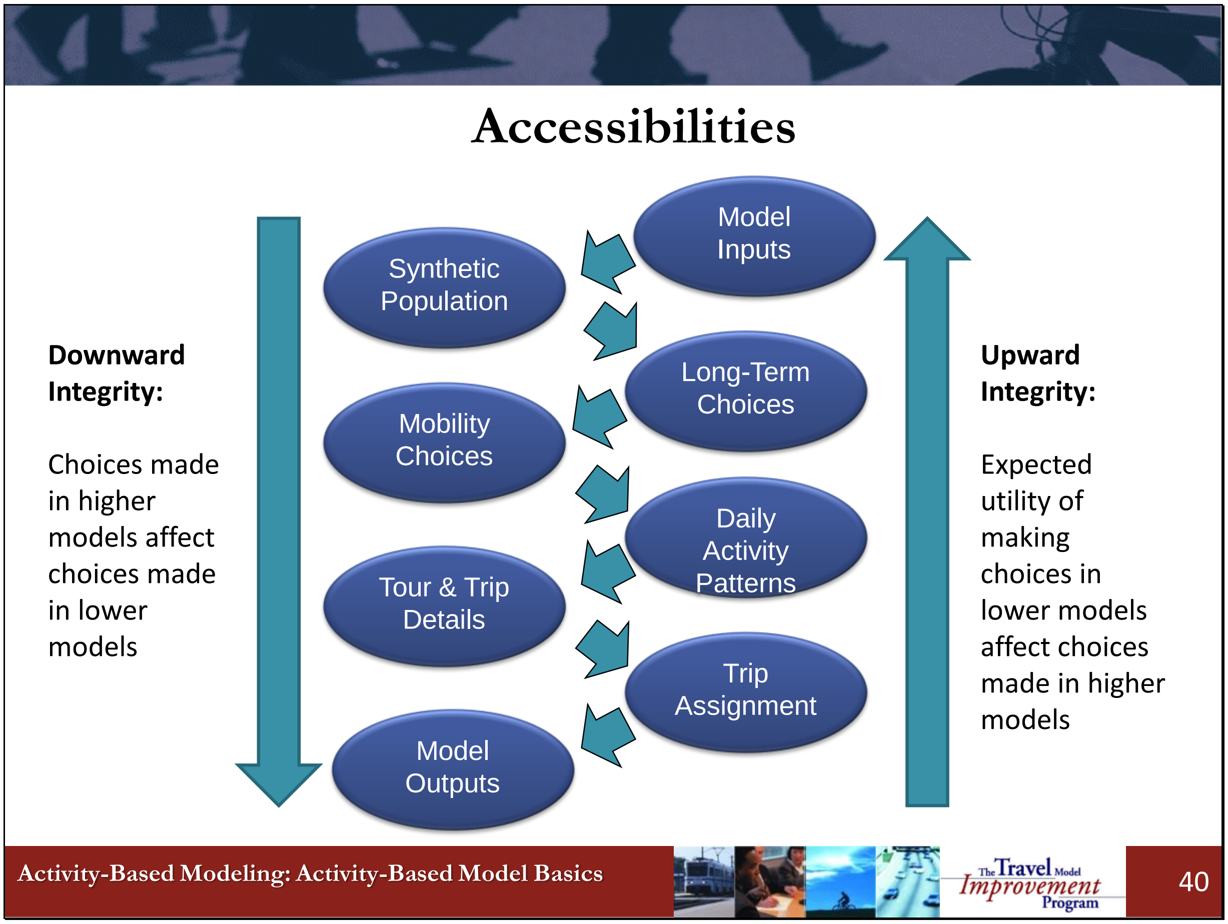
Transportation analysis zones, or TAZs, were used in earlier versions of models, and are still used for smaller regions where TAZs tend to be quite small. The advantages of TAZs are that they are readily available and it is generally easy to estimate land-uses at the zone level. The disadvantages are that they typically have some level of aggregation bias with respect to intra-zonal and close-in travel, particularly for transit access/egress and non-motorized travel.

Another approach involves developing sub-zones by buffering around transit stops, and using the sub-zone definition to over-ride skimmed walk access/egress times to/from transit. This approach is also commonly used in trip-based models. Advantages of this approach are that it is easy to create the buffers using simple GIS procedures, and it offers an improved representation of transit access/egress. However, there is still some aggregation bias with respect to transit skims, and the approach is not helpful for non-motorized travel.

Spatial disaggregation, to the extent that it can be done realistically, improves the ability to measure activity attractiveness and travel impedance. And differences in attractiveness and impedance, which vary a LOT within traditional zones, have VERY LARGE impacts on people's travel choices. As activity-based models become spatially more disaggregate, it is becoming more and more feasible to realistically incorporate non-motorized modes into the models.

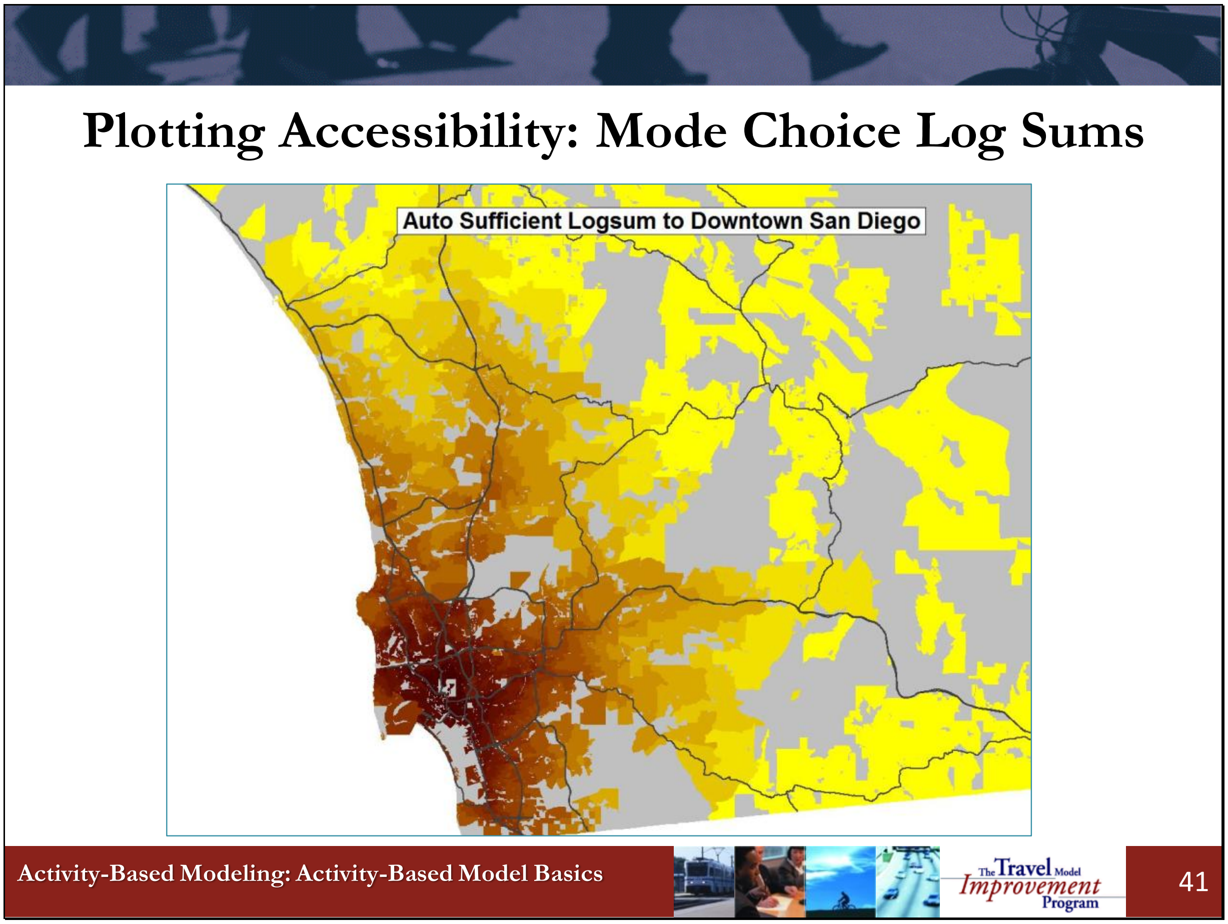
The third approach is to code a set of micro-zones by sub-dividing zones, using grids, or census blocks. Advantages include the ease of creating micro-zones, and the approach offers a very precise measurement of walk access/egress to/from transit, especially when coupled with stop-level transit skims. However, employment data can be more difficult to allocate to micro-zones than to TAZs, depending on the availability of good employment data.

Finally, parcels can be used in activity-based models. They are sometimes available from existing sources, and are the most disaggregate with respect to non-motorized distances and walk times. However, parcels are not necessarily stable across time, and allocating employment data to parcels can be challenging. We will cover this topic in even more detail in Webinar 6.



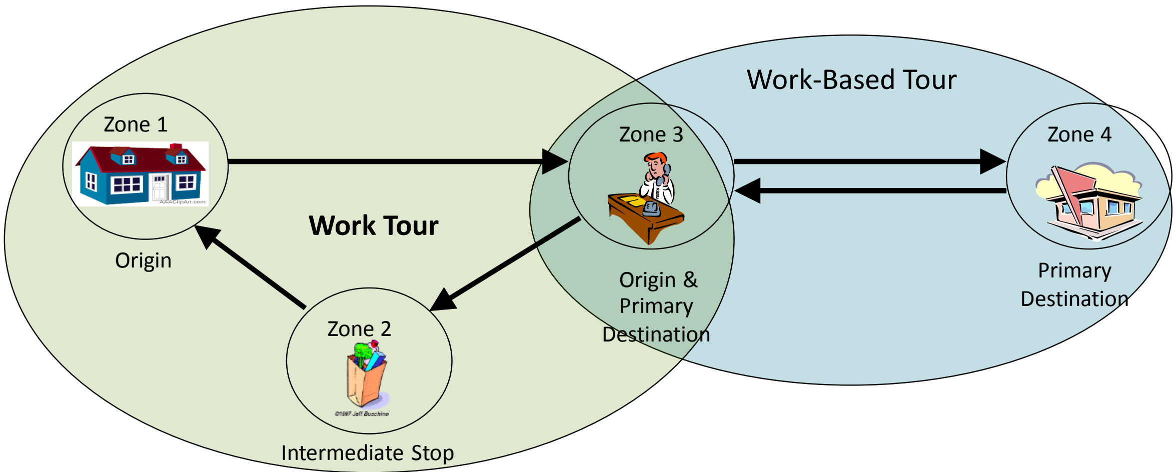
Webinar 6 will also discuss the calculation and use of accessibilities in the activity-based model system. While activity-based models can vary in structure, this diagram shows the location of tour and trip detail choices (tour mode, primary destination, intermediate stop location and trip mode) in a typical model stream. As the model system progresses, travelers make decisions: whether to travel, where to go, how many stops to make, what mode to choose, and so on. Earlier decisions influence and constrain the decisions made later; for example, the number of vehicles owned, as modeled in the auto ownership (mobility) model, influences the number of tours and the mode used on each tour. The mode used for the tour then influences the location of stops on the tour, and so on. This is referred to as 'downward vertical integrity'.

Activity-based models also use information from models that are lower in the model chain to inform the choices made by decision-makers in upper-level models. This information typically takes the form of accessibilities that are based upon all of the information that is relevant for a lower level choice. For example, a mode choice log-sum, which reflects accessibility by all modes of transport, can be used to inform the choice of destination for the tour or stop. This is referred to as ‘upward vertical integrity’.



Log-sums can be plotted. This a map of the mode choice log-sum from all micro-zones to downtown San Diego, specifically for travelers who reside in households with at least as many cars as adult drivers. The darker colors indicate more accessible zones and the lighter colors indicate less accessible zones from downtown, consider the time and cost of travel by all modes of travel, but specifically for auto sufficient households. Log-sum variables may be created for specific market segments and for mode choice are typically segmented by household auto availability or sufficiency (which relates workers or drivers to available autos). Market segmentation of log-sums is yet another key design element.

Treatment of Tours and Trips



Tour Data

HH #	Per #	Tour #	Purpose	Origin TAZ	Destin. TAZ	Outbound Stop1 TAZ	Return Stop1 TAZ	Mode	Outbound Time	Return Time
1023	1	1	Work	1	3	0	2	Transit	7 :30AM	5:00 PM
1023	1	2	Work-Based	3	4	0	0	Walk	12:00 PM	1:00 PM

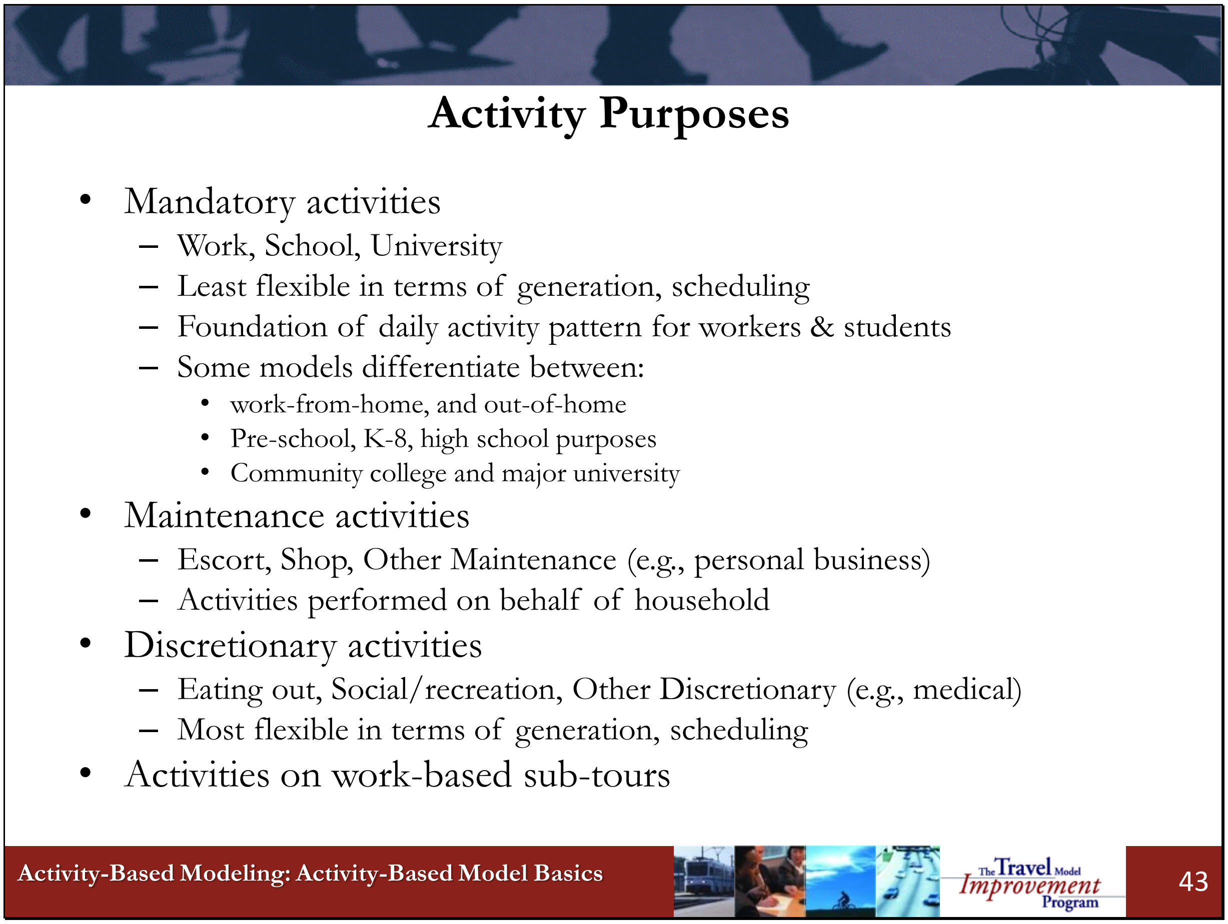
Activity-Based Modeling: Activity-Based Model Basics — The Travel Model Improvement Program — 42

At beginning of this webinar, we looked at some household diary data from a survey. Here is a graphic depiction of a day's tour pattern for one person, with a table that combines all of the activities into two tour records. A worker leaves home in the morning and travels to work in zone 3. The worker then goes to lunch in zone 4 and returns back to work. At the end of the day, the worker returns home, but stops on the way in zone 2 for some groceries.

The definition of tours and components of tours is fundamental to activity-based modeling. This travel pattern consists of two tours. One tour is a work tour which consists of the set of trips from home (zone 1) to work (zone 3) to shop (zone 2) back to home (zone 1). In this example, the home location is the *tour anchor or origin* (the start and end of the sequence of trips for the tour). The *primary destination* is the key location on the tour that defines the tour purpose. It is often the main reason for travel. For a work tour, it is the workplace.

How to determine which stop in the tours is the primary destination is one key design decision. While it is possible in recent tour-based household surveys to ask the primary purpose of the tour, this has not always been the case and is certainly not true in all surveys, particularly older ones. Using a hierarchical typology based on activity purposes is one method, which works well for work, school and college purposes, but for other purposes primacy is less clear. Other tie-breaking rules include the first stop on the tour, the stop furthest from the home anchor point, and the stop with the longest duration. This has important implications for the construction of tour schedules, since time window availability criteria for the insertion of intermediate stops would be influenced by both activity duration and travel time to the primary destination.

There can be zero or more *intermediate stops* on the tour, which are stops made between the *anchor location* and the *primary destination*. Some activity-based modeling systems refer to sequence of one or more stops between the anchor location and the primary destination as the first half of the tour, or outbound half; and the sequence of one or more stops between the primary destination and the anchor location as the second half of the tour, or return half. On this tour, there is one intermediate stop on the *return half* of the tour, between work and home. There are no stops on the *outbound half*, which is between home and work. Whether to model stops on tours using this "half tour" schema, or a more sequential method, is another design decision. The sequence of trips between work and lunch is referred to as a work-based tour (or sub-tour). In this case, the anchor location for the tour is the workplace, and the primary destination is lunch. There are no intermediate stops on this tour. Another design decisions is whether to allow non-work locations to be anchors for sub-tours, such as school.

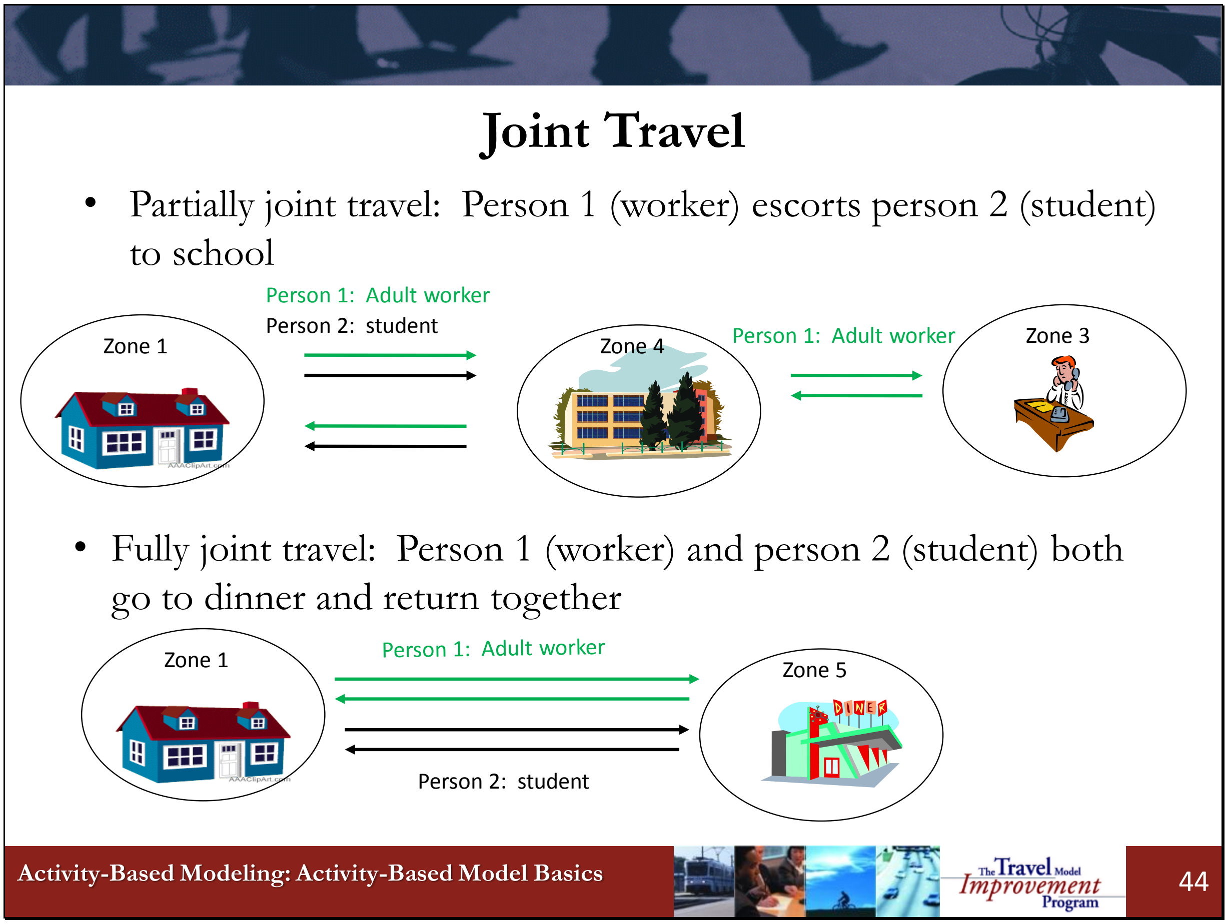


In general, disaggregation of travel purposes by activity types makes activity-based models more sensitive to variations in travel behavior than trip-based models and allows them to be more accurate when matching person types with activity locations and times of day.

Activities are often grouped into three key categories according to priority in the daily activity pattern schedule: Mandatory, Maintenance, and Discretionary. Mandatory activities consist of work and school. They are the least flexible in terms of generation and scheduling, and are the basic building blocks of activity schedules for workers and students. Some model systems differentiate between work-at-home (telecommuting) and work-out-of-home activities. Some models also categorize school activities by grade level.

Maintenance activities include escort, shopping, and other maintenance which can include doctor's visits. Some modeling systems model certain purposes explicitly, while others combine them into more general categories, like "other". This is a design decision that should depend on local modeling needs. For example, in areas with a large contingent of senior citizens, explicit modeling of a medical activity purpose may be desirable.

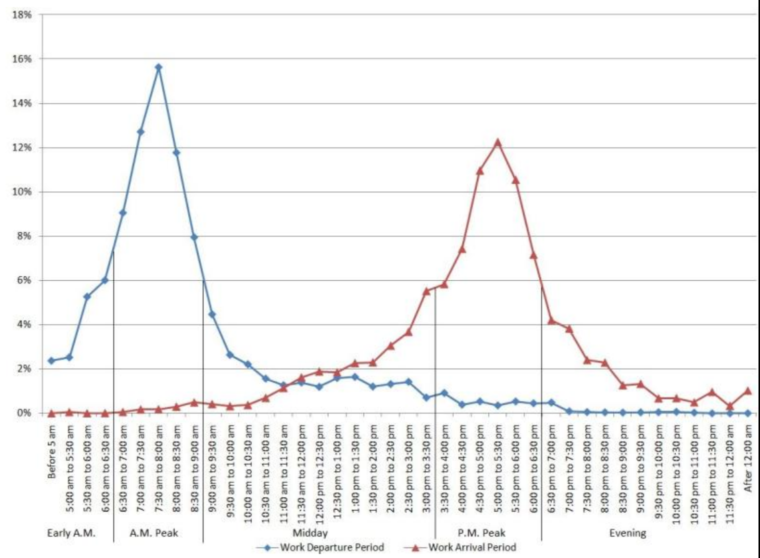
Many of these activities are performed on behalf of the household, such as picking up or dropping off household members, or going grocery shopping. In model systems that represent joint travel explicitly, the escort purpose may be replaced by more detailed descriptions, as we shall see in the next slide. Finally, discretionary activities include eating out, visiting, and other recreational activities. They are the most flexible in terms of generation and scheduling, and often substituted for in-home activities, particularly for households with poor accessibilities to recreational opportunities. In addition, some modeling systems differentiate activities on work-based sub-tours from those belonging to the main home-based tour. One reason for this is because sub-tours tend to be more constrained in terms of time; therefore, activities on work-based sub-tours are likely to have significantly shorter average durations and travel distances.



Joint travel refers to tours and trips that consist of at least two persons from the same household. Research shows that over 80% of shared-ride trips are made by members of the same household. Depending on the model system design, it may be represented implicitly or explicitly. Explicit representation of joint travel is useful for modeling high-occupancy vehicle demand, mode choice elasticity, and vehicle allocation among household members. There are two types of joint travel. “Partially joint” travel refers to tours where picking up or dropping off passengers occurs on the tour, as shown above. Fully joint travel refers to tours where all members travel to all stops on the tour together, shown below. We will discuss different ways to generate joint travel in Webinar 8, and will revisit joint travel in Webinar 10, where we discuss tour and trip mode choice.

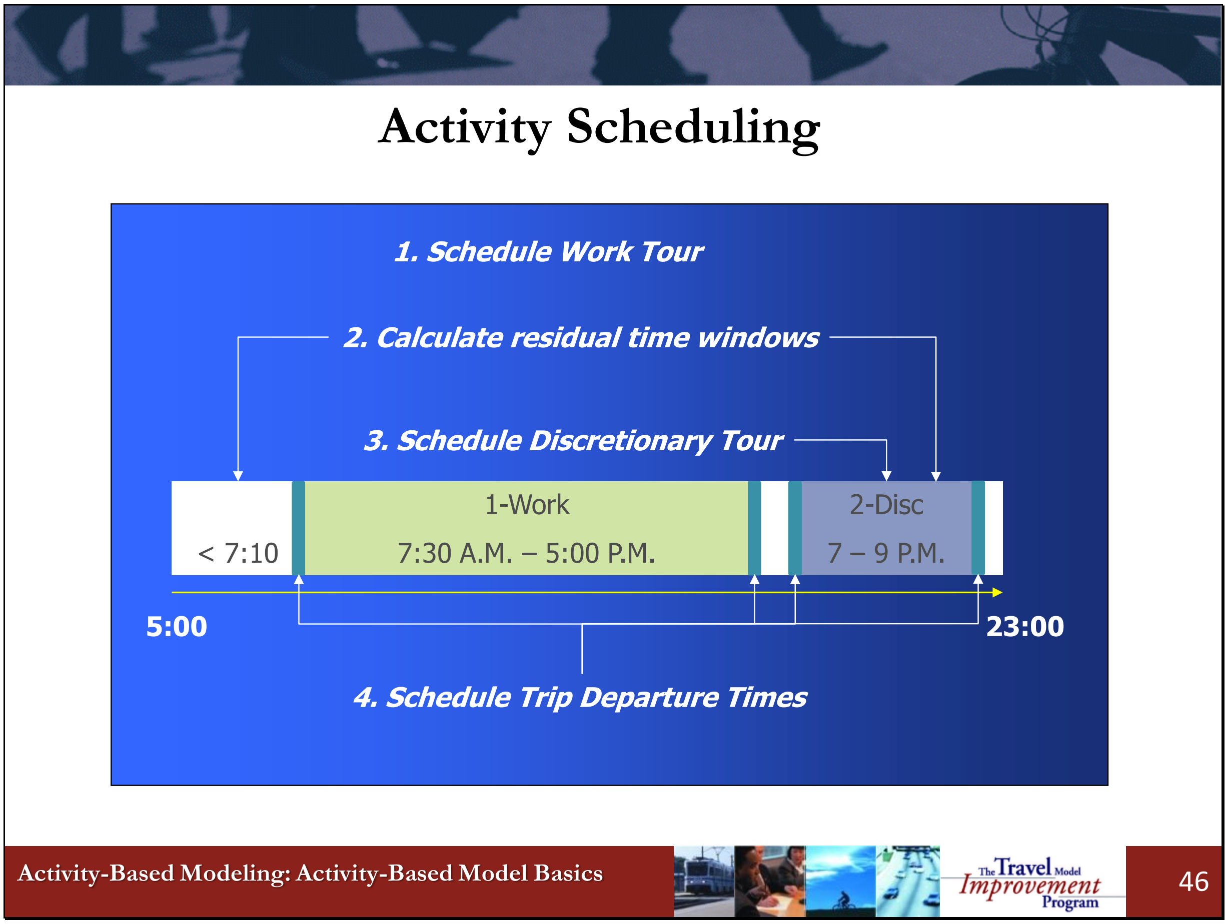
Treatment of Time

- Different temporal systems used
  - 5 time periods, hourly, half-hourly, continuous
- Many models have more aggregate time periods for skims and disaggregate time periods for activity scheduling
  - Example: 5 periods for level of service matrices, half-hourly periods for scheduling activities



Disaggregated treatment of time in activity-based models makes them more sensitive than trip-based models to changes in level of service as well as the idiosyncratic preferences of individuals doing things at certain times of day (such as meals) and business hours. This makes many activity-based models a more effective tool for modeling peak spreading behavior, congestion pricing, and TDM policies that aimed at shifting travel demand across time.

Although real time is continuous, many activity-based models often treat time as intervals, or time periods. As we saw earlier in this presentation, we can parse time into intervals in a way that allows us to choose starting and ending times for activities using discrete choice models. These intervals range from half-hourly, to more aggregate. Sometimes multiple temporal systems are used in the same model; for example, network skims are created for 4-5 time periods but activities are scheduled into half-hourly intervals (or even assigned continuous start and end times). Some activity models schedule activities in quasi-continuous time, either by using continuous functions to predict activity duration, given a starting time. Moving towards finer temporal resolution in both activity-decisions and network assignment is an area of growing research. In Webinar 9, we will cover the treatment of time in activity-based models, and ways to schedule tours, trips, and activities.



In addition to decisions regarding temporal resolution, one of the other fundamental design decisions in activity-based modeling is the design of scheduling algorithms. This slide shows an example of an activity scheduling process. Activities are scheduled consistently, such that they fit into available time windows, taking into account expected travel times. For mandatory activities in which arrival times are important, departure times may be "backed out" using the expected travel time from skims (by the expected travel mode). Higher-priority activities are typically scheduled before lower priority activities. "Residual time windows" refers to the time left over after activities have been scheduled, and provide opportunities for further activity engagement. Different modeling systems use alternative scheduling processes, but the outcome (that activities are scheduled consistently and realistically with no overlaps) should be observed across all systems. No person can be in more than one place at one time.

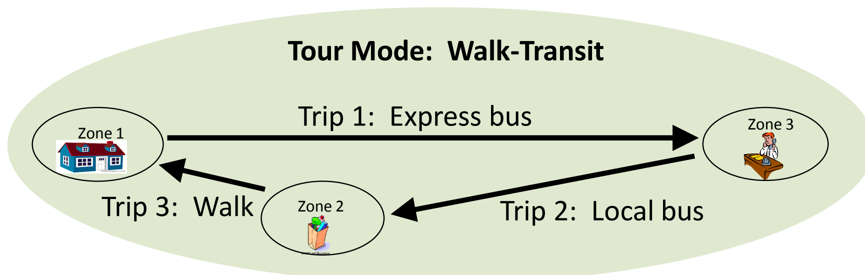
We will discuss the treatment of modes in Webinar 10. Different modes imply different model designs. This depends on how precisely one needs to model access modes and sub-modes. These range from simpler definitions (like auto, transit) to more precise descriptions such as High-Occupancy vehicle lanes, light-rail transit, bus rapid transit, heavy rail and commuter rail. Examples of each mode are shown on this slide.

- **High-Occupancy Vehicle (HOV) Lanes:** Vehicles are restricted based on occupancy; typically 2 or more persons are required, although some lanes require 3 or more persons.
- **Bus Rapid Transit:** A type of transit mode featuring buses that offers operating characteristics similar to rail, such as separate rights-of-way, rail-like station amenities such as covered platforms and/or rider information, and low-board vehicles for easy boarding and alighting.
- **Light-Rail Transit:** A type of rail transit that can operate in either separate right-of-way or in mixed-flow with auto traffic.

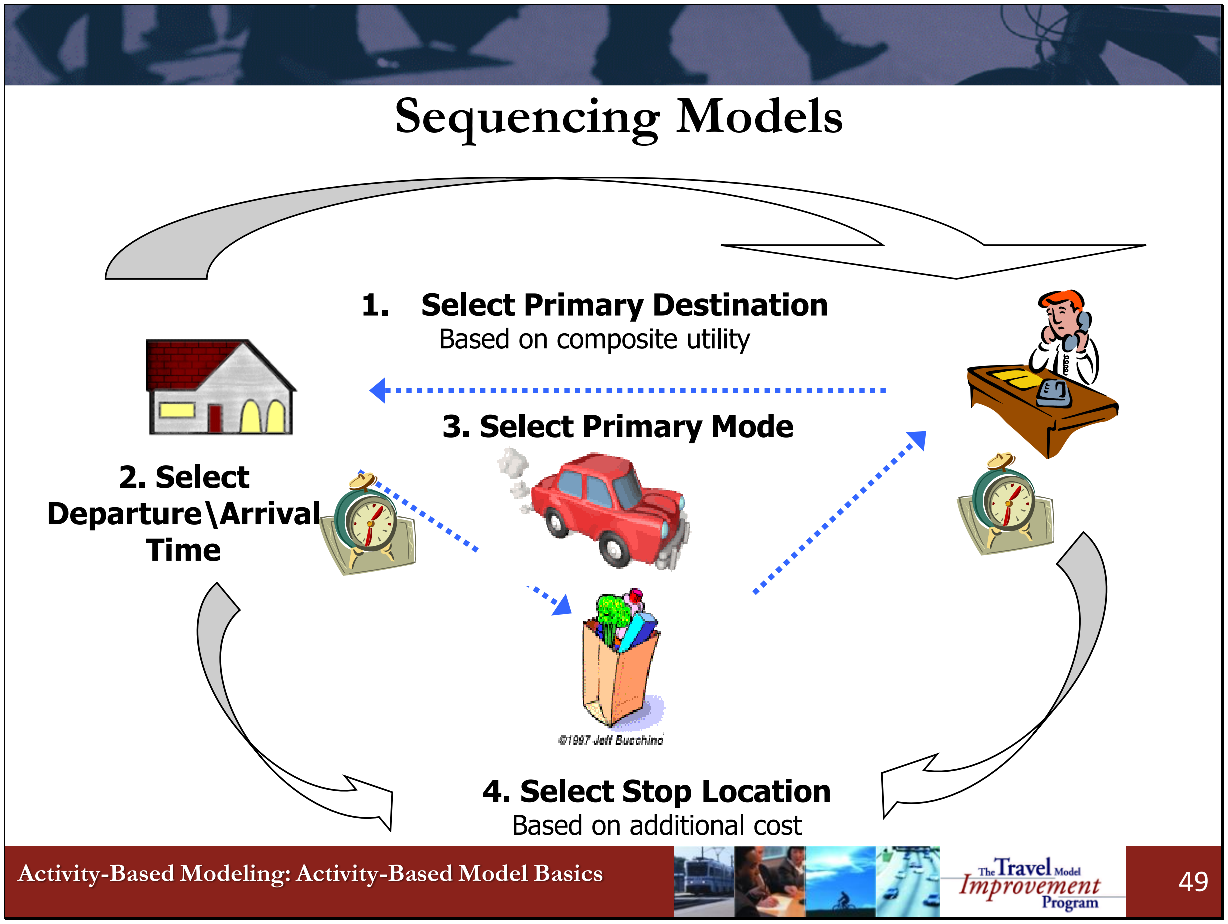
- **Heavy rail:** Higher-speed passenger rail cars operating singly or in trains of two or more cars, on fixed rails in exclusive rights-of-way excluding other vehicular and foot traffic.
- **Commuter rail** is passenger train service operating between a central city, its suburbs, and/or another central city.

Treatment of modes

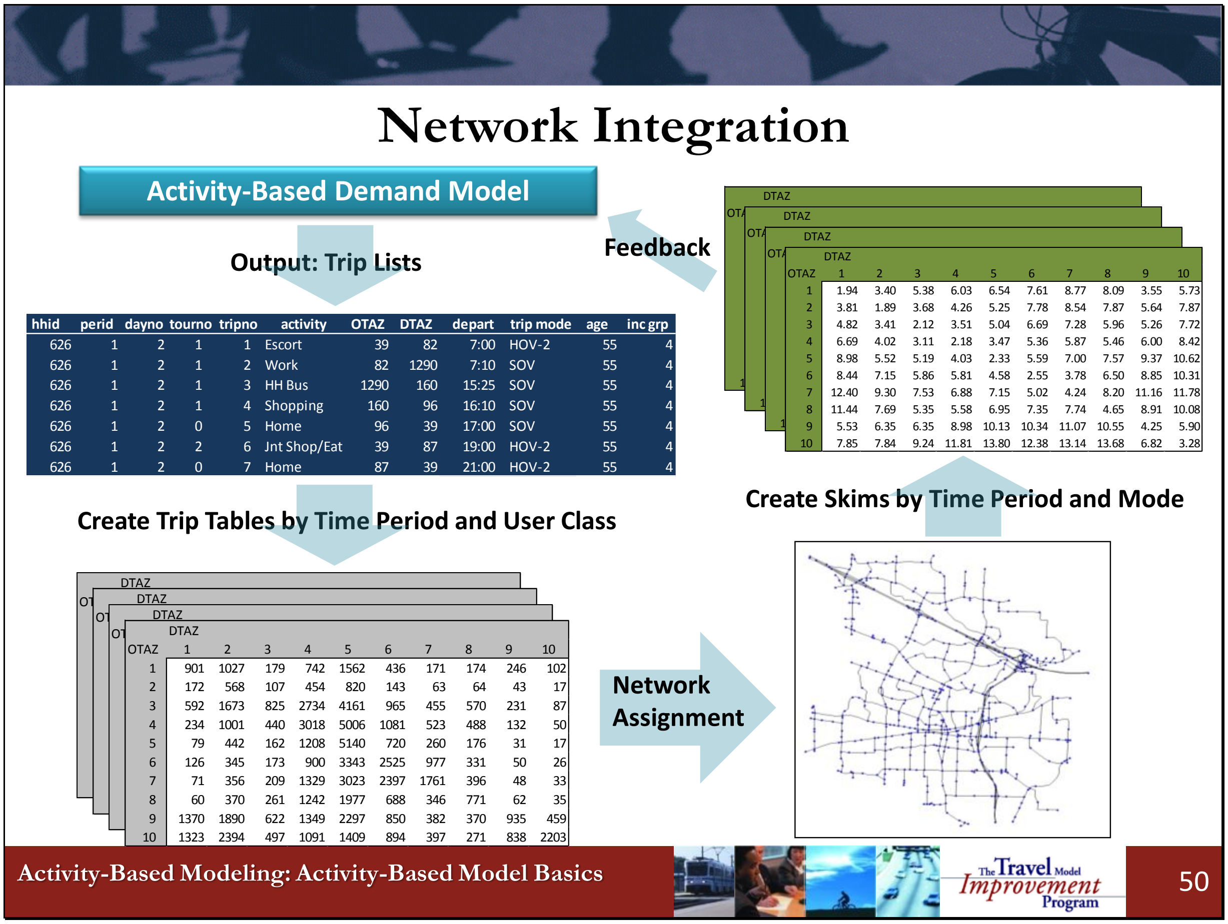
- Tour mode versus trip mode
  - Tour mode is the overall mode for the tour
  - Trip mode is the actual mode chosen for each trip on the tour
- Other considerations
  - Auto driver versus passenger?
  - Explicit line-haul mode (local versus express versus rail) choice?
  - Park-and-ride versus kiss-and-ride?
  - Explicit toll\HOV choice?



Activity-based models can treat modes simply or can address modes with a great amount of detail, as described on this slide. A key aspect of activity-based models is the difference between tour mode and trip mode. The trip mode refers to the mode used for each trip on a tour. This is observed in household survey data when a respondent reports their travel for the day. Ultimately, trips are assigned to transport networks based upon the trip mode. The tour mode is an abstract concept used to classify the mode for the tour as a whole. Typically tour mode is based upon the modes used for each trip. For example, the diagram shows the same work tour previously shown, with each trip mode labeled. The tour mode is walk-transit, even though one of the trips (the last one) is made by walking. The tour mode influences the location of stops on the tour and the modes of the trips used for the tour.



One of the key differences in activity-based model system designs is the use of different algorithms for sequencing models. This slide shows one way of sequencing tour destination, mode, and stop location choice. In this example, the primary destination for the tour is selected first. This selection is based upon the composite utility, or mode choice log-sum, for the tour mode choice model. Once the primary destination is chosen, the departure time from home, and arrival time back at home, is chosen. Next, the primary mode is chosen for the tour. Finally, the locations of stops on the tour are determined, taking into account the additional travel time and cost required to access the stop based upon the location of the tour origin and primary destination. We will talk more about model sequencing in Webinar 10, and the pros and cons of various sequencing options.

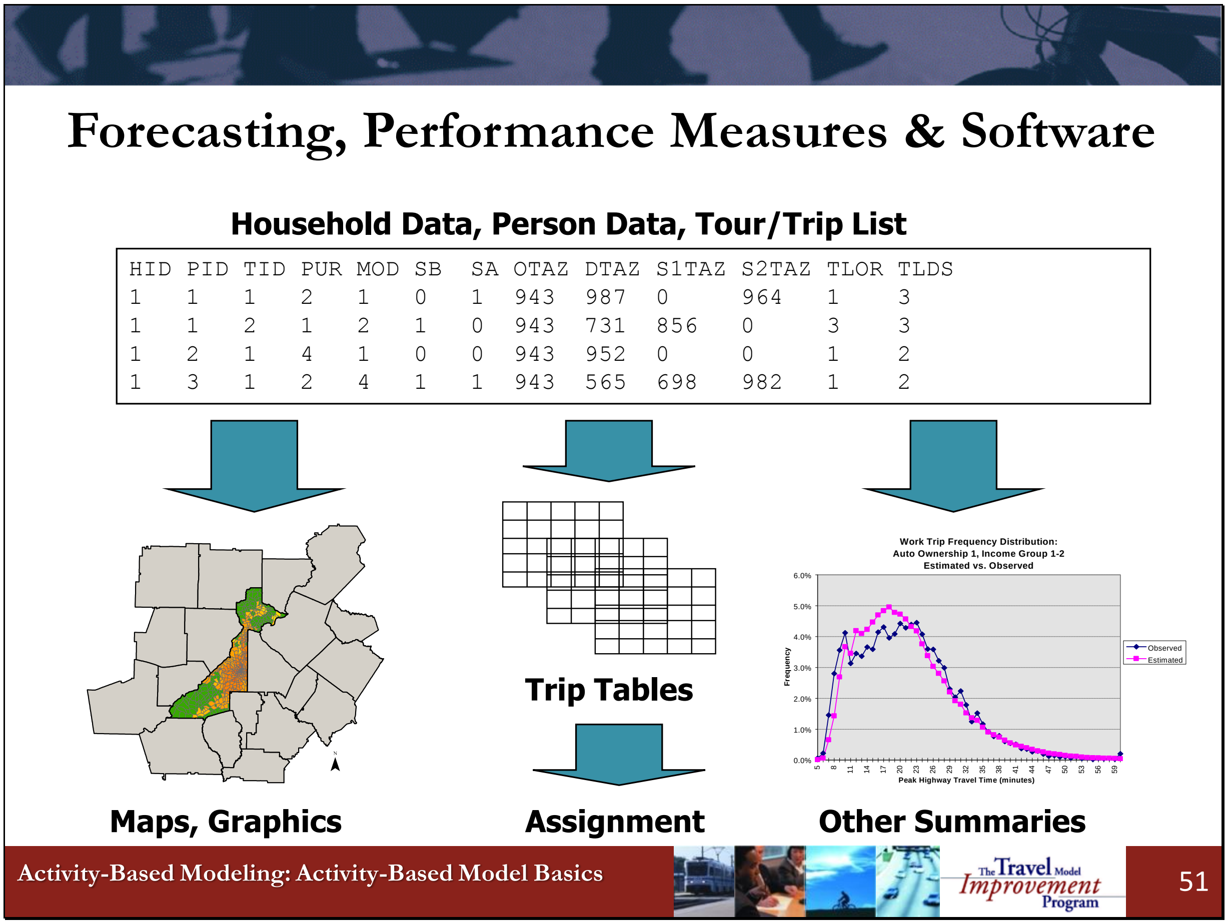


Among the current generation of activity-based models now in use, network integration is fairly straightforward and not too dissimilar from trip-based models. The activity-based demand model will create a list of activity-travel events that looks very much like a travel diary!

From these activity-travel records, a trip list is formed, representing vehicle trips for persons traveling in autos. This trip list includes origins and destinations as well as very specific (minute-by-minute) trip departure times. The trip list may also include contextual or person-attribute information that might be used for classifying the traveler. For example, in addition to mode, the trip might be identified by the activity purpose at the destination end and possibly by the driver's income group, both of which could be used to assign the trip to a "value of time" grouping. This is not a feature of the early activity-based modeling systems, but is becoming an increasingly common design feature of more recent model designs for the purposes of modeling heterogeneous responses to tolls and fares.

This user class information, along with time of day and origin-destination identifiers, would then be used to group trips from the trip list into trip tables. Yes, these are the same style of trip tables that should be familiar to all travel demand modelers. One difference, of course, may be that activity-based models are trending towards more assignment time periods user classes based on value of time.

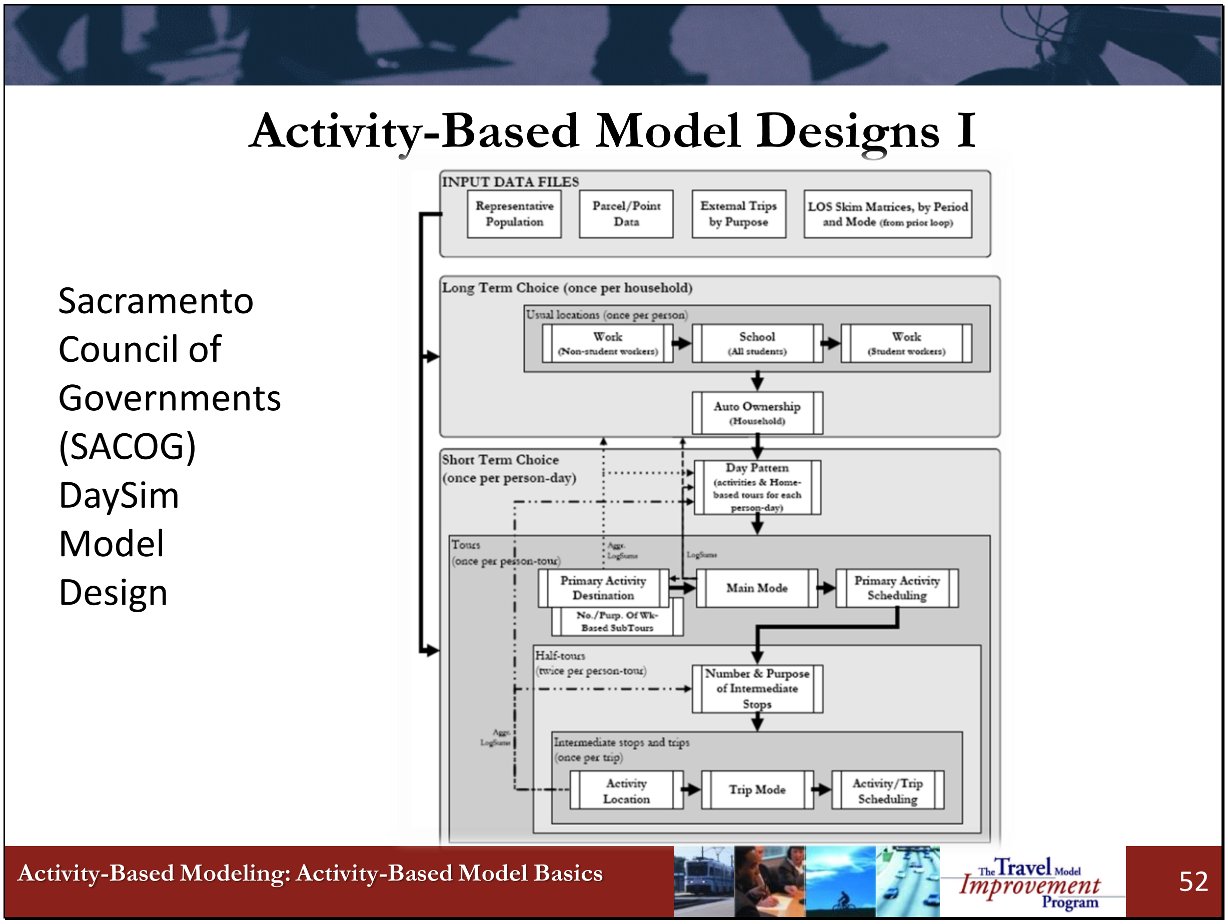
Once highway and transit assignments are run, travel time and cost "skim" matrices are produced, and fed back to the activity-based model. Current activity-based models use speed-feedback-loop systems very much like trip-based models and use the same network assignment software. We will cover network integration in greater detail in Webinar 11. It is also worth mentioning that there are a couple of active research projects going on right now on the integration of dynamic traffic assignment (DTA) and activity-based models. This includes the SHRP2 C10 program, the SimTravel project being conducted by Arizona State, University of Arizona, and U.C. Berkeley, and independent investigations by the San Francisco County Transportation Authority.



In addition, the design of the simulation software itself is an important consideration. Since activity-based models forecast using Monte Carlo simulation, developers have created mechanism by which the user can fix random number seeds and, in some places, freeze certain model components to support New Starts analysis or other analyses in which it is desirable to hold certain model outcomes constant while varying just one, like mode choice.

In Webinar 12, we will cover the software used to implement activity-based models, and the use of models in forecasting. This slide illustrates how disaggregate data produced by the model, which includes household, person, tour and trip level information, can be summarized and analyzed in a variety of ways. Familiarity with the data produced, and the underlying models that produce that data, is essential for properly interpreting model output and using it to evaluate policy scenarios. Because activity-based models produce such a large amount of output, it may be difficult to know what to expect as “normal,” particularly when looking at various tour-based measures, activity durations, and other measures that may be new to trip-based modelers.

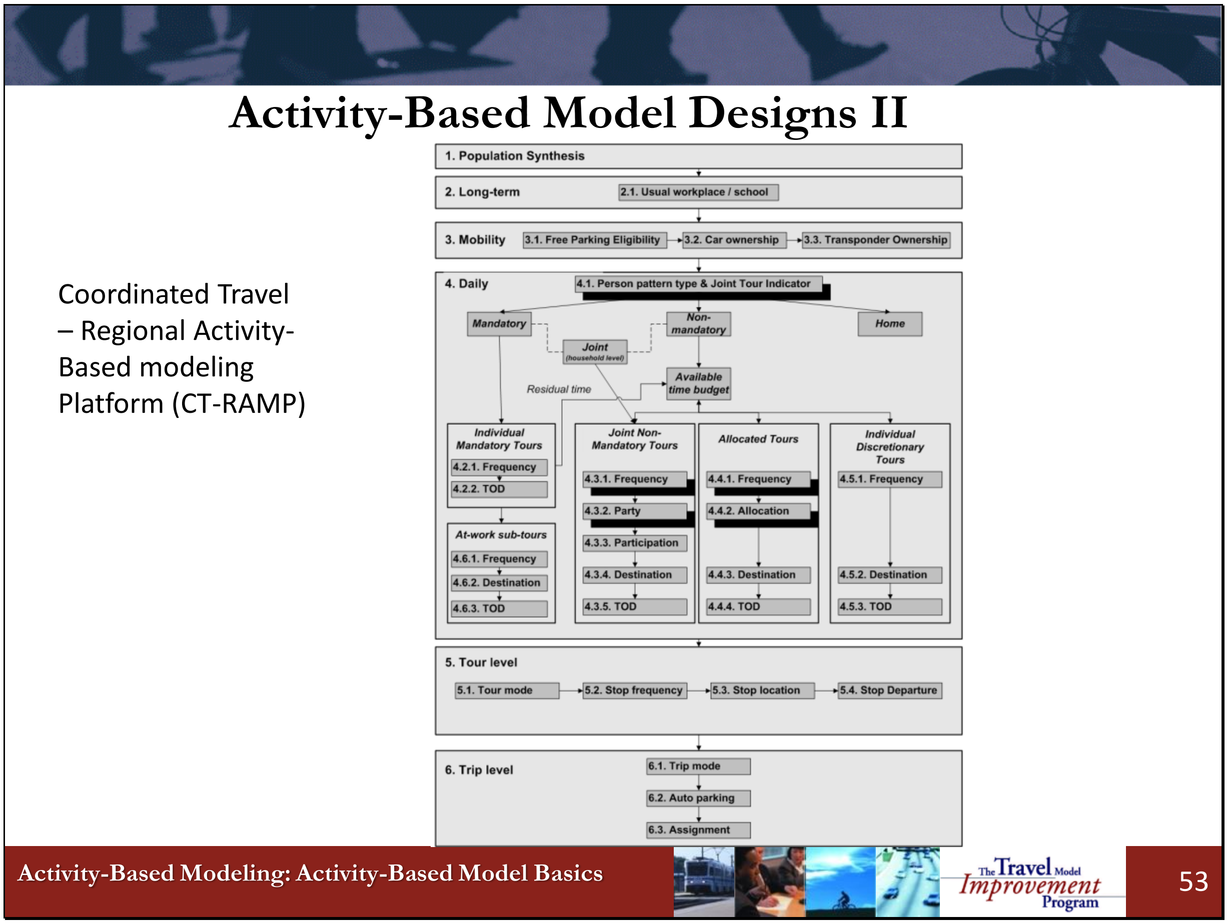
Visualization tools, combined with training, documentation, and experience examining results are keys to success.



So, let's see what it looks like when we begin to put these different design elements together. We're going to look at three different model designs, two that are currently in use and one that is now under development. This slide shows the DaySim model design for Sacramento Association of Governments (SACOG). Similar models have been implemented for San Francisco County Transportation Authority (SFCTA) and Denver Council of Governments (DRCOG), with models under development for Puget Sound Regional Council (PSRC - Seattle MPO), in San Joaquin and Shasta Counties, California and in Jacksonville and Tampa, Florida.

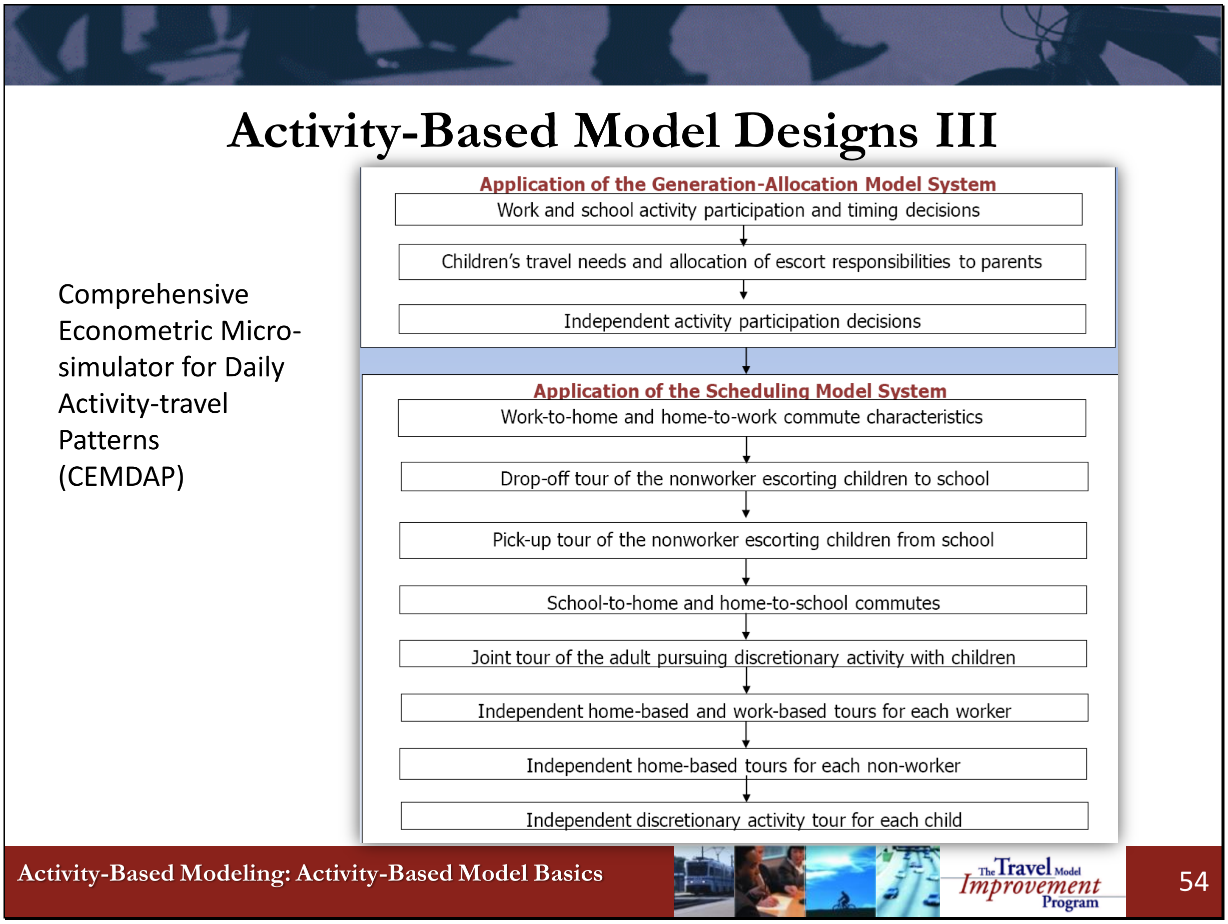
In this model system long-term and mobility choices appear at the top of the diagram. This would seem to be a fairly standard treatment of these modeling steps across different platforms. After this we have activity generation and scheduling steps. This model system is characterized by an enumeration of possible combinations of tours of different types, which define individual day patterns. We looked at an example of this earlier. There are literally hundreds of combinations represented in the model system, which provide a strong conditioning for the models to follow. Downstream models include models to determine the exact number of tours of each type, destinations, mode and starting and ending times of tours. Given the establishment of tours and their primary destinations, modes and start and end times, the next model to be applied determines whether there is room left over the in schedule for intermediate stops on the tour. For intermediate stops, we then choose locations, modes and departure times.

Page 53



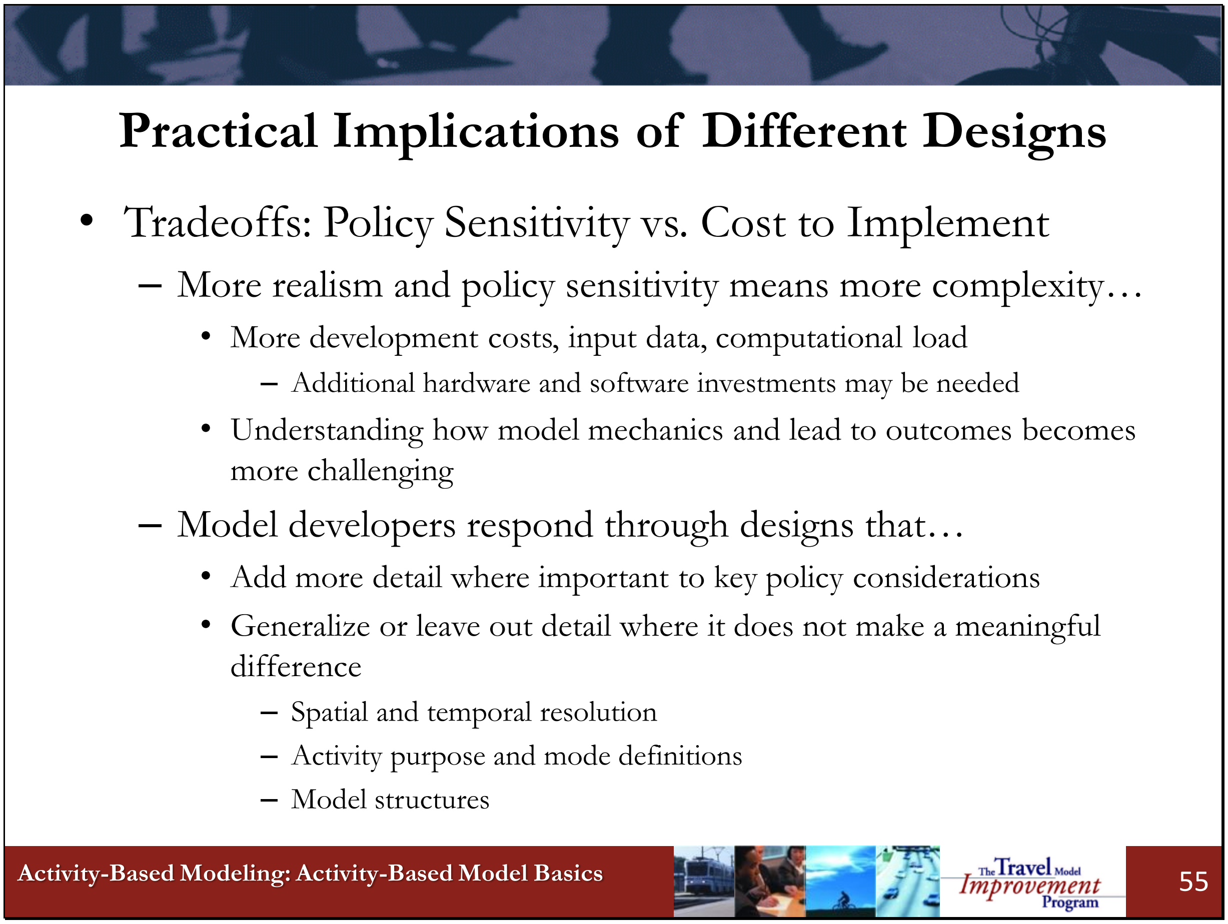
This slide shows the CT-RAMP model diagram for San Diego Association of Governments (SANDAG). Similar models have been developed for Mid-Ohio Regional Planning Commission (MORPC), Tahoe Regional Planning Agency (TRPC), Atlanta Regional Commission (ARC), Metropolitan Transportation Commission (MTC), and under development for Maricopa Association of Governments (MAG), the Miami region, and Jerusalem, Israel.

In this model system, we also see population synthesis and long-term and mobility choices at the top of the structure. This model system is characterized by the identification of mandatory activities (work and school), which are prioritized in schedule creation. Persons are assigned daily pattern types, which are defined rather simply as being mandatory, non-mandatory or home. It is also characterized by explicit modeling of joint activities, through a series of models that determine joint tour frequency, party size, and individual participation. Time windows play an important role in this model system as well, with mandatory activities scheduled first, followed by joint activities, and finally independent discretionary activities. Frequencies, time of day, and destination decisions are made at this level. This is followed by tour modes, stop, frequency, stop location and departure times. Trip modes, auto parking and network assignment are the final steps.



This slide shows the Comprehensive Econometric Micro-simulator for Daily Activity-travel Patterns, by Dr. Chandra Bhat of University of Austin-Texas. The model has been applied to the Dallas/Ft-Worth region, and a version of the model is currently under development for the Southern California Association of Governments (SCAG).

This model system is characterized by an activity generation and allocation step the predetermines the amount of activity participation, expressed in terms of time spent in an activity. This is done through a discrete-continuous model formulation. Household-level responsibilities for ridesharing are also determined at this stage. The system then proceeds through a series of scheduling steps, which prioritize certain types of tours or partial tours. For example, round trip work-commute characteristics are identified first, followed by any drop-off and, then, pick up tours for non-workers who are escorting children to school. CEMDAP's scheduler then proceeds through a number of other tour types, ending with independent discretionary tours for children. Although not shown here, within each tour scheduling group, there are the usual activity stop generation, location, mode and timing decisions. In addition, population synthesis and long-term and mobility choices are determined through a separate module, not shown here.

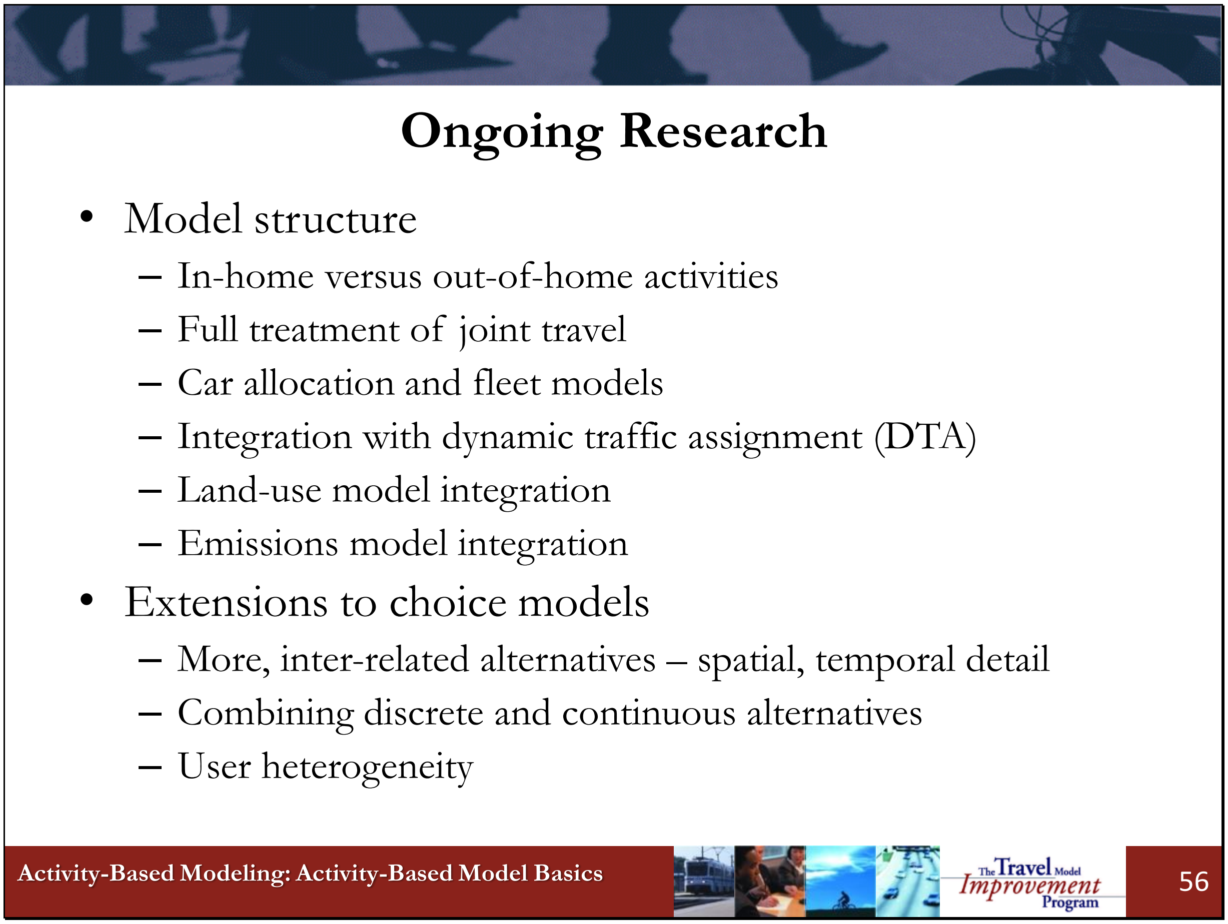


The practical implications of different model system designs come down to a tradeoff between sensitivity to certain policy inputs versus the cost of running and maintaining the modeling system. By cost, we're talking about both time and money. The more complexity that is built into a modeling system, the more it costs to develop, the more input data needs to be developed and maintained, and the greater the computational load. More complex models will take longer to run and, accordingly, will require investments in hardware and software to make it feasible to run them in a production context. In addition, a more complex model could be more difficult to interpret, although, this can be overcome by careful design of output reporting capabilities and features, and of course user training and documentation.

The developers of these modeling systems are well aware of these tradeoffs. They have carefully thought through their model designs, choosing to add more detail where they felt it was most important, and leaving out certain details where they felt it was less important. These design decisions tend to be fairly nuanced. For example, one decision is whether to model at a fine level of spatial resolution, such as parcels, or to use a more aggregate measure of spatial units such as

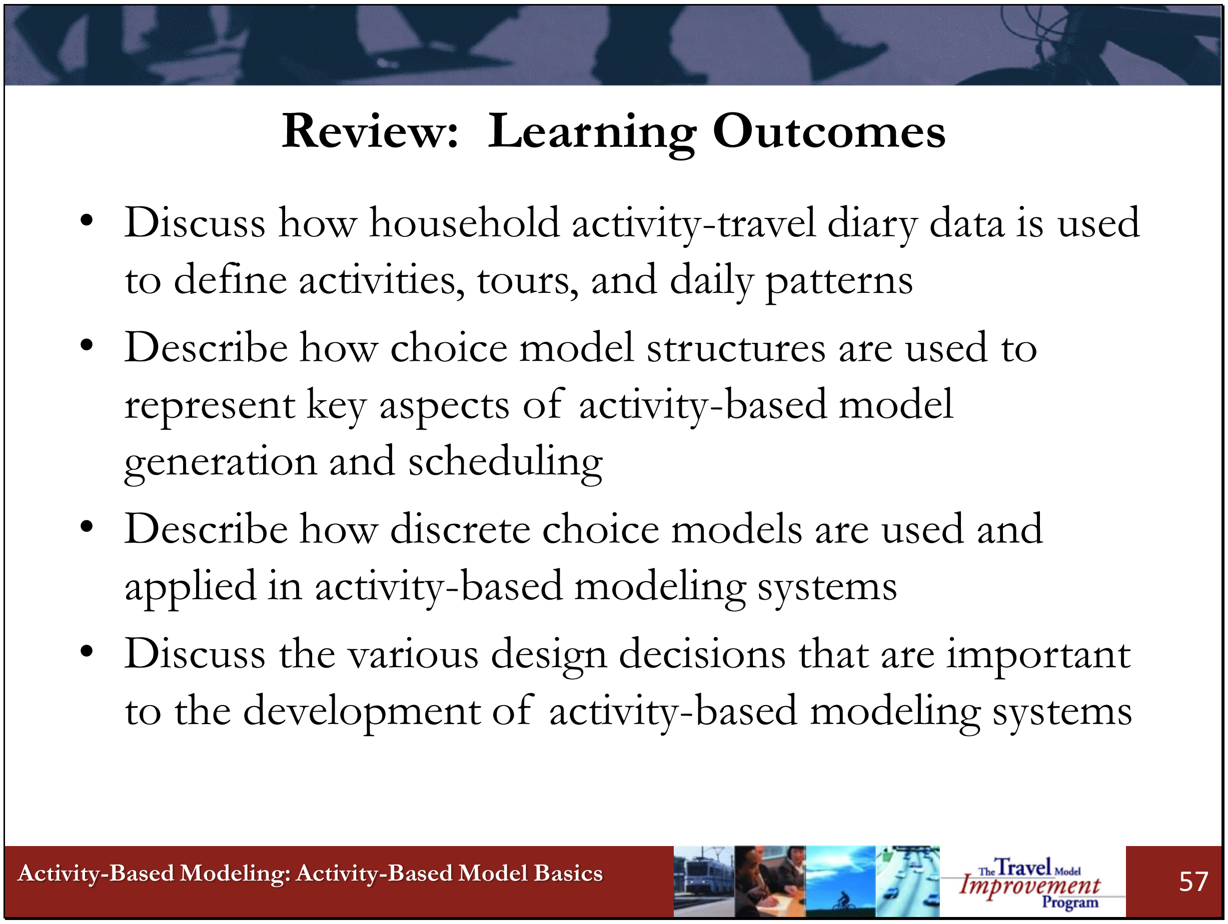
TAZs. The use of parcels has the potential to provide the best estimates of non-motorized travel times, but adds complexity to network loading, calculations of travel times and impedances, and use of skims.

Another example might be whether to model intra-household interactions explicitly—joint activity participation and pick up and drop off events—or to assume that these events happen, but that it is sufficient to treat each person’s experience independently. Modeling intra-household interactions explicitly is more realistic, but adds a layer of model components due to the need to generate and coordinate patterns between individuals. In both of these examples, the question is whether the added complexity produces enough of a meaningful difference in model results to merit the extra costs. It is very important for agencies to on their modeling needs and priorities when developing a new model system.



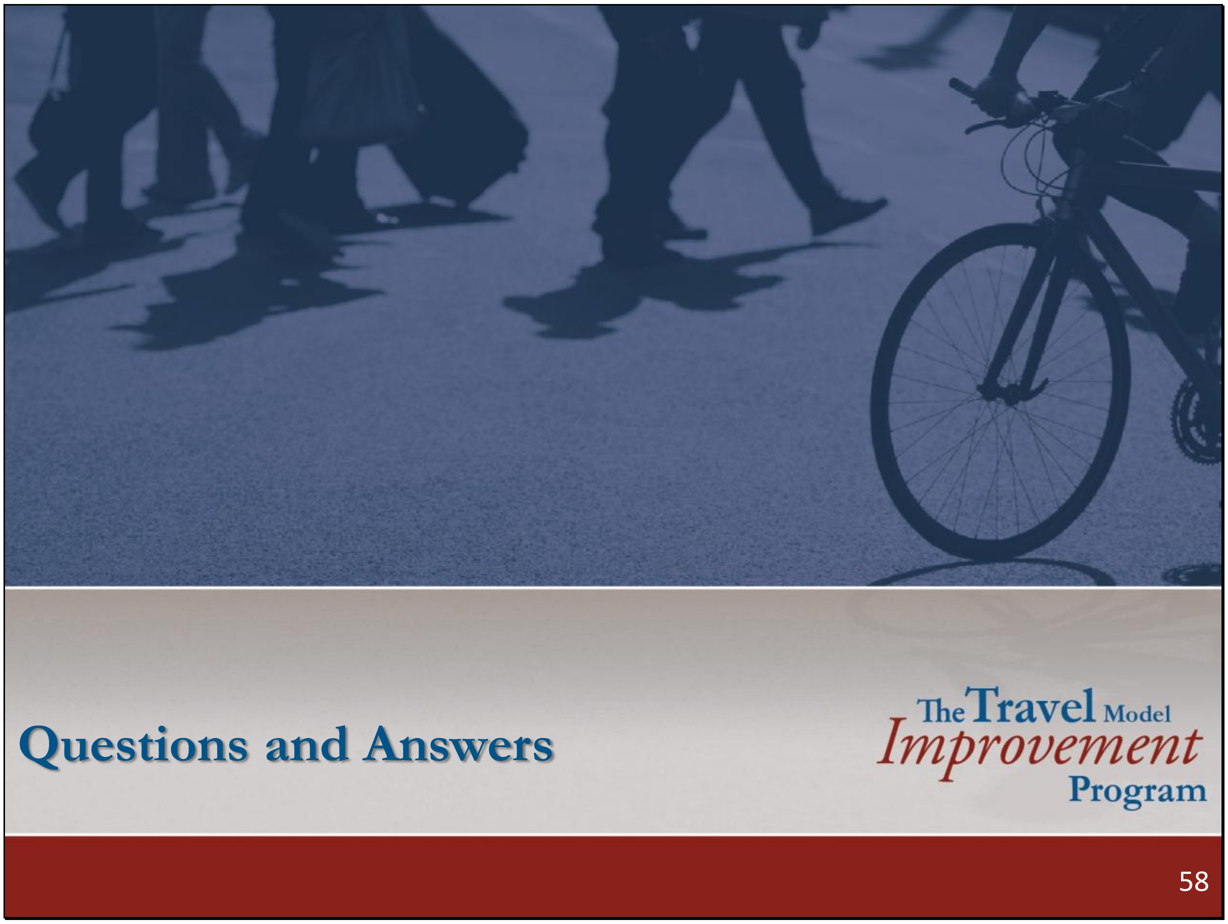
Practical activity-based travel modeling is relatively new, and there are a number of areas of ongoing research. Activity-based model structures are being extended in novel ways, to explicitly address in-home versus out-of-home activities, full treatment of joint travel, household car allocation and fleet models, integration with dynamic traffic assignment (DTA), land-use models, and emissions models.

There is also a very active body of research in the realm of discrete choice models. Model structures are being explored that address more, inter-related alternatives. Models have been developed that simultaneously model both discrete and continuous alternatives, such as number of vehicles owned and amount of miles driven on each. There are many models that address the issue of user heterogeneity, such as cost sensitivities that are expressed as a distribution of cost parameters rather than as an average parameter that applies to everyone. Finally, some models explicitly model user preferences for packages of alternatives or attributes of alternatives, rather than relying solely on socio-economic variables to explain different sensitivities. We will explore some of these topics in subsequent webinars.



You should now be able to…

- Discuss how household activity-travel diary data is used to define activities, tours, and daily patterns
- Describe how choice model structures are used to represent key aspects of activity-based model generation and scheduling
- Describe how discrete choice models are used and applied in activity-based modeling systems
- Discuss the various design decisions that are important to the development of activity-based modeling systems


Questions and Answers
The Travel Model Improvement Program
58

2012 Activity-Based Modeling Webinar Series

Session	Date
Executive and Management Sessions	
Executive Perspective	February 2
Institutional Topics for Managers	February 23
Technical Issues for Managers	March 15
Technical Sessions	
Activity-Based Model Frameworks and Techniques	April 5
Population Synthesis and Household Evolution	**April 26**
Accessibility and Treatment of Space	May 17
Long-Term and Medium Term Mobility Models	June 7
Activity Pattern Generation	June 28
Scheduling and Time of Day Choice	July 19
Tour and Trip Mode, Intermediate Stop Location	August 9
Network Integration	August 30
Forecasting, Performance Measures and Software	September 20

Continue the discussion online…

The new TMIP Online Community of Practice includes a Discussion Forum where members can post messages, create forums and communicate directly with other members. Simply sign-up as a new member, navigate to http://tmiponline.org/Community/Discussion-Forums.aspx?g=posts&t=523 and begin interacting with other participants from today's webinar session on Activity-Based Modeling.

Session 4 Questions and Answers

Have multi-day GPS-based household surveys been used to develop activity-based models?

John: It's possible that has been done in a research context. Kay Axhausen at ETH Zurich has done a lot of work with GPS data in Zurich.

Joel: There are multiple regions in the US that have done GPS surveys across multiple days even if the paper diaries are only taken for one day. We're still learning to use the data. One of the benefits of this data is more accurate departure and arrival times than what is reported in the paper diary.

Has any method other than the Monte Carlo method been incorporated in an activity based model in practice?

Joel: One of the first activity-based models was developed in Portland, where aggregate probabilities were calculated across a number of dimensions and applied to groups of persons. The problem with this approach is that there is not enough memory space to cover all of the dimensions. Monte Carlo simulation allows us to more efficiently process all of these dimensions. If you consider destinations, times of day, modes, number of tours and stops, the probability tree becomes huge, and you have to use Monte Carlo simulation to address all the choices.

Where do the probabilities for destination choice come from? Are they based on TAZ employment or retail employment, or what?

John: We will cover destination choice models in more detail in later webinars, but we have observations of what destinations people chose, and we will estimate logit models where the utility function includes not only impedance between the origin and possible destinations, but also a term which estimates the attractiveness of the alternative, similar to the attraction variable in a gravity model. In estimation, alternative destinations are sampled to reduce the number of computations compared to using all zones.

You talked about different ways to overcome the disadvantages of Monte Carlo simulation. To address concerns about variability in project evaluation, can you recommend preferred methods for fixing and holding constant the Monte Carlo variation?

Joel: If you're talking about FTA's requirements that trip tables be held constant between the baseline and build alternative, or other cases where holding certain choices constant between alternatives makes sense; we'll run a baseline alternative, hold early choices such as destination choice fixed from the baseline alternative, and write the choices to disk. Then we'll read the skims from the build case in and re-simulate only later choices such as mode choice. The issue with that is that in ABMs everything is connecting to everything else, so e.g. intermediate stop locations in the baseline alternative might not make sense given the location of service in the build alternative. Different model platforms have different techniques for addressing such issues. Another case where you might hold some choices constant is when you want to know short-term responses to policies. You might hold long term choices like auto ownership and work location constant and only re-simulate mode and route choices to evaluate short-term impacts.

Is the Poisson probability distribution used in activity-based models?

John: The Poisson distribution is usually used to model a random arrival process. It is occasionally used to model trip generation processes, as a counting process. It isn't very commonly used. Not many of the other choices in an activity-based model are amenable to a counting process.

Is the mode for a tour determined comparing utilities to the primary destination only, or is travel to intermediate stops factored into the choice of tour mode? And are trips between intermediate stops constrained depending on the choice of tour mode?

Joel: Yes, we are typically choosing a tour mode based on round trip travel times and costs to the primary destination, though we can use information from the stop location choice model to influence the tour mode. For example the accessibility to stops (the stop location choice log-sum) can influence the number of intermediate stops, and this information can also be fed back to influence the tour mode. It is typically found that the mode for the tour is more influenced by round trip times and costs than characteristics of the traveler or the tour. The second part of the question is about mode switching. Most models do allow for mode switching. The tour mode is an abstract concept. The trip mode is what is actually observed in the activity diary. A lot of mode switching occurs in surveys, for example switching between shared ride and drive alone auto modes on the tour, so we do need to allow for that switching to happen.