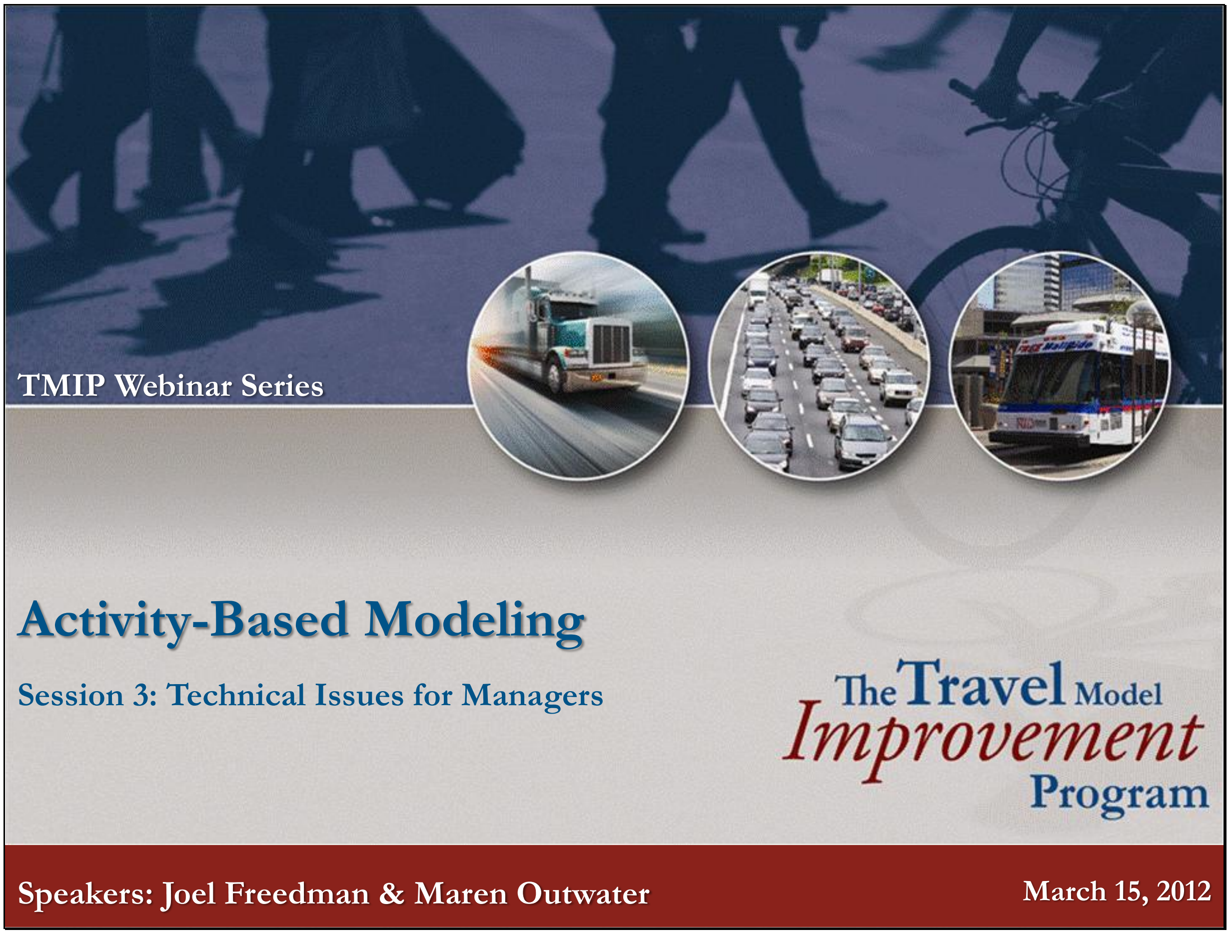

TMIP Webinar Series
Activity-Based Modeling
Session 3: Technical Issues for Managers
The Travel Model Improvement Program
Speakers: Joel Freedman & Maren Outwater
March 15, 2012

Acknowledgments

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

- Presenters
  - Joel Freedman
  - Maren Outwater
- Content Development, Review and Editing
  - Joel Freedman
  - Maren Outwater
  - John Gliebe
- Media Production
  - Bhargava Sana
  - Brian Grady

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.

- Joel Freedman and Maren Outwater are co-presenters. They were also primarily responsible for preparing the material presented in this session.
- Stephen Lawe is the session moderator.
- Content development was also provided by John Gliebe.
- Bhargava Sana and Brian Grady were responsible for media production, including setting up and managing the webinar presentation.

2012 Activity-Based Modeling Webinar Series

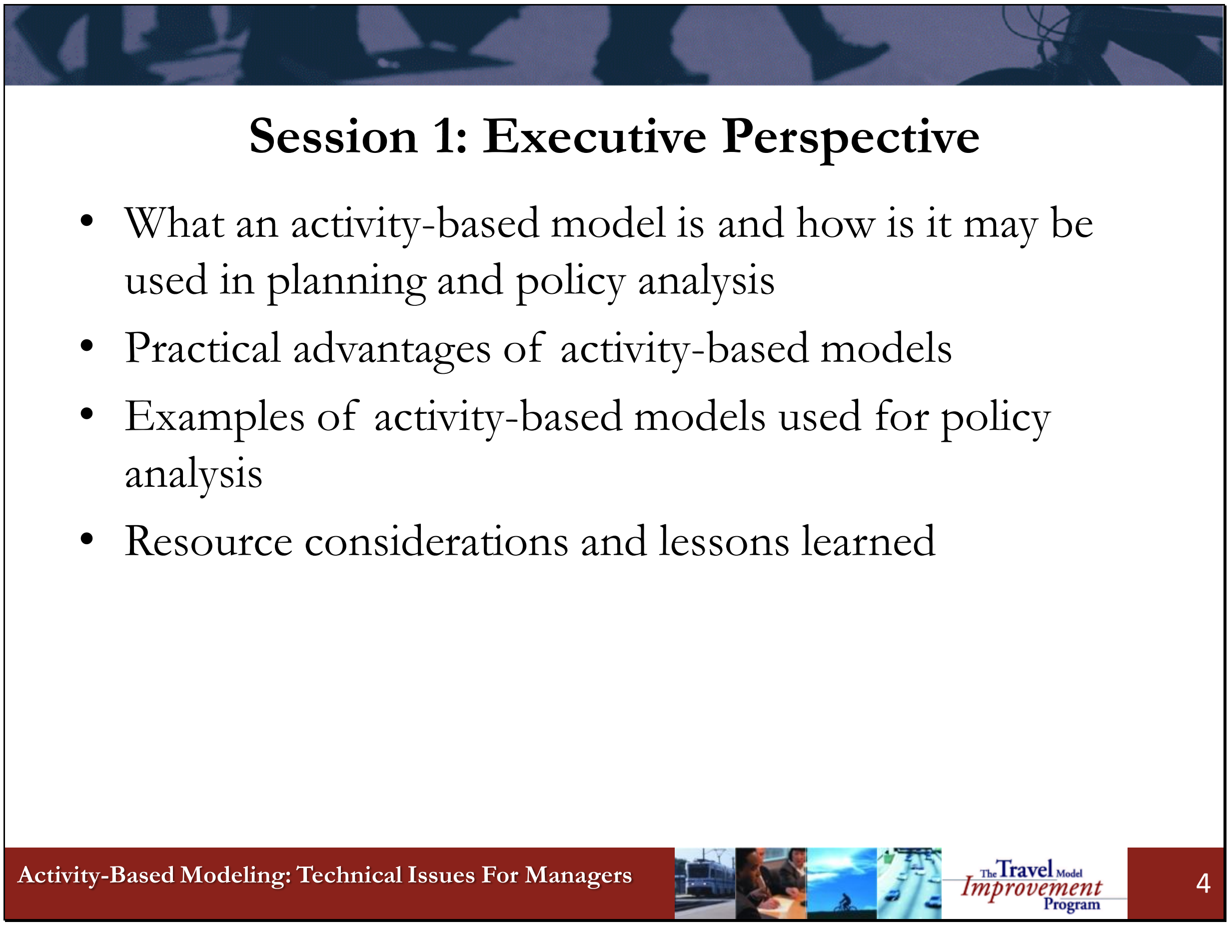
Executive and Management Sessions	
Executive Perspective	February 2
Institutional Topics for Managers	February 23
Technical Issues for Managers	**March 15**
Technical Sessions	
Activity-Based Model Basics	April 5
Population Synthesis and Household Evolution	April 26
Accessibility and Treatment of Space	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

Activity-Based Modeling: Technical Issues For Managers

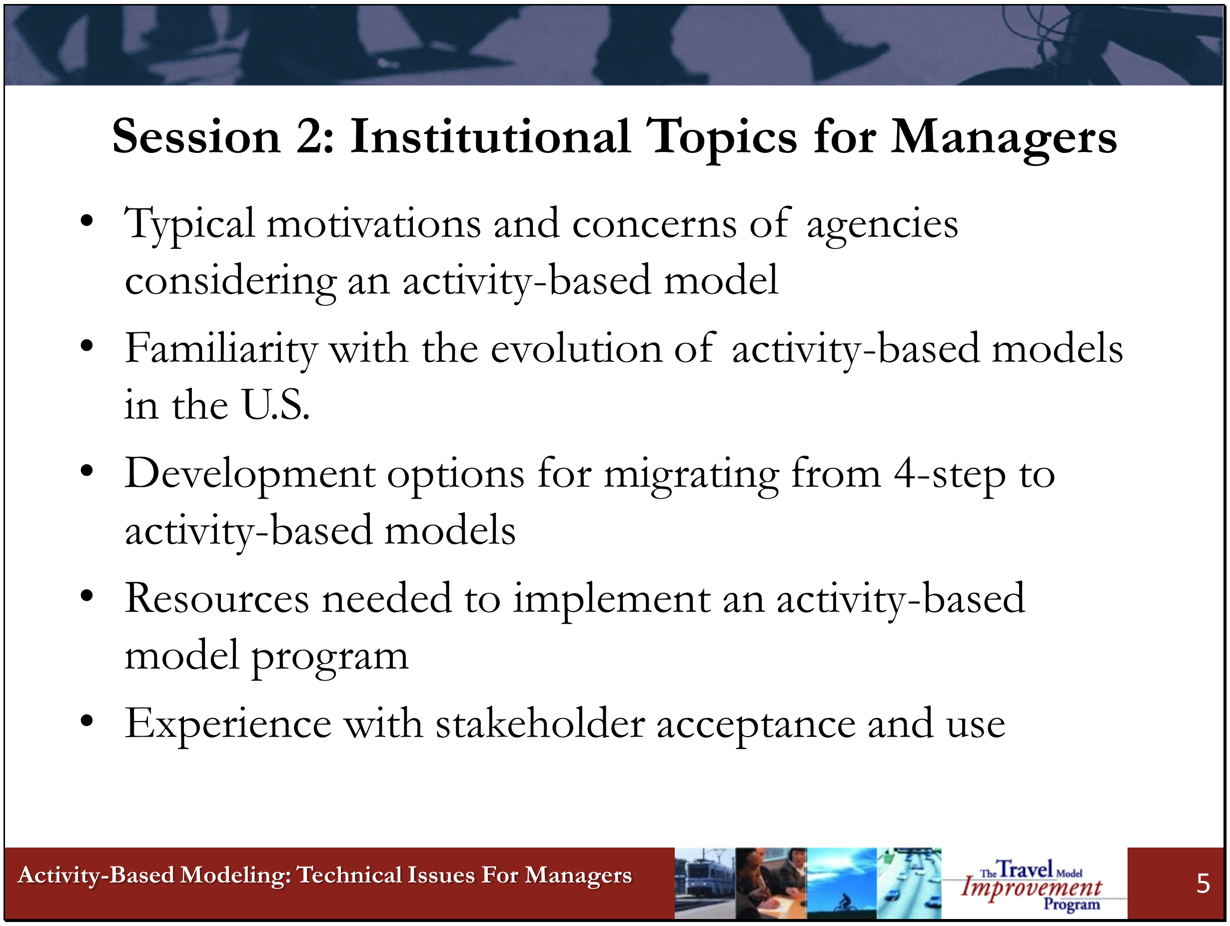
The Travel Model Improvement Program

3

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 second in topic in the series—Institutional Topics for Managers. Today we will cover technical issues for managers. Today's session provides a high-level overview of technical issues for management staff contemplating the development or implementation of activity-based models for their region. In three weeks, we will begin to cover the details of activity-based models, including choice models, simulation, and different activity-based modeling frameworks.

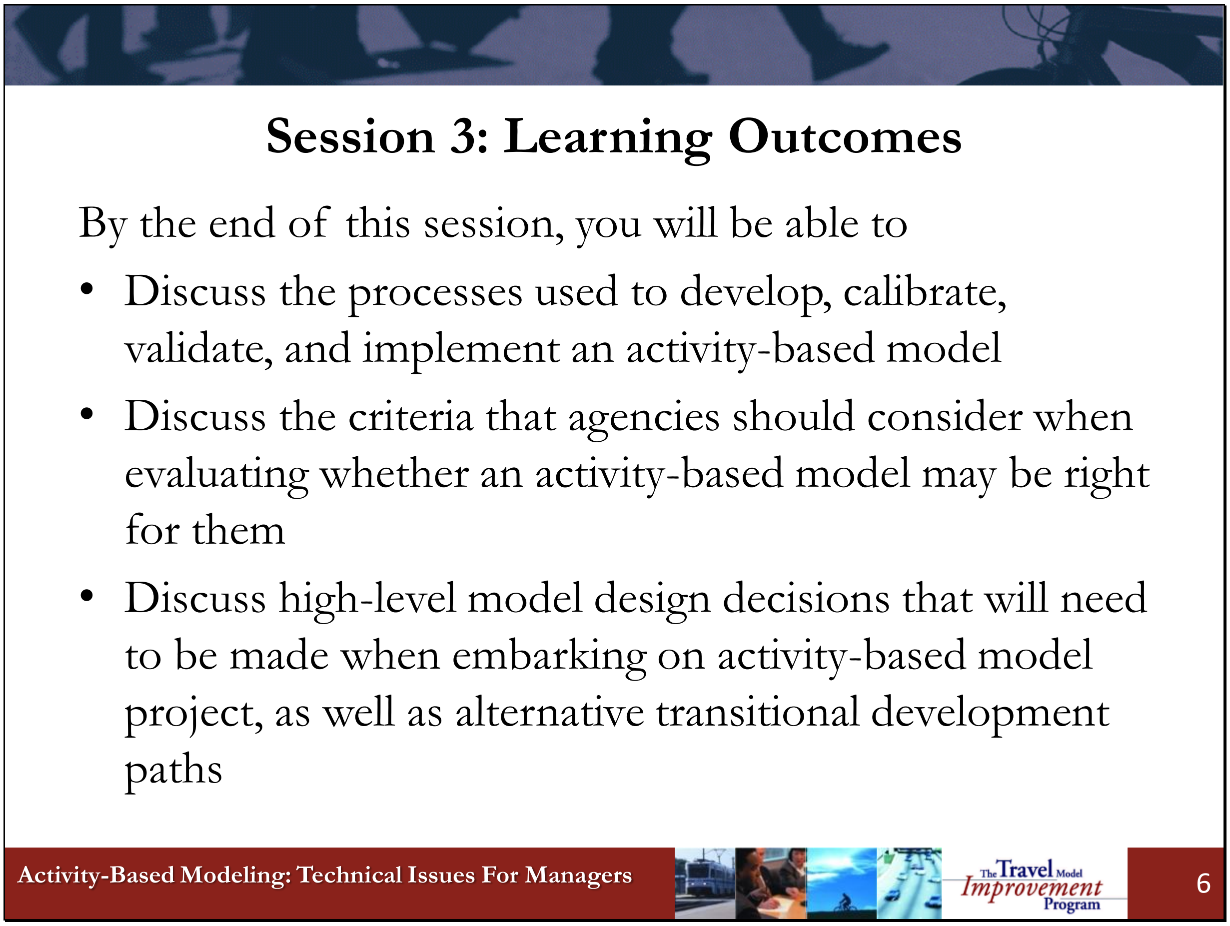


First, we'd like to summarize the topics covered in the first two sessions in the series. Session 1, the Executive Perspective, was intended to provide a high-level overview of activity-based modeling and how it could benefit an agency involved in policy studies. You may find a recording of Session 1 in the archives. In this session, we began by discussing what an activity-based model is and contrasting it with a trip-based model. We outlined some of the practical advantages and limitations of activity-based models. After that, we described a several policy examples in which activity-based models have been used to analyze some fairly complex policy scenarios, including congestion pricing, travel demand management, land use interactions, and latent demand. We concluded this first session by touching briefly on staffing, consultant, and hardware and software resource issues, and shared some lessons learned from our experience in model development projects.



The second webinar series covered institutional topics for managers. In this session, we began be describing the advent of travel modeling in the U.S., how the needs and resources of modelers in the early days dictated trip-based model design. We then describe how models began to evolve in sophistication, sometimes in response to challenges to their authenticity. We then began to describe the advantages that activity-based models provide in terms of information content and how that information can be used in policy analysis. With these advantages in mind, we then outlined the questions typically on the minds of persons who are contemplated the move to an activity-based modeling system—costs, performance, resource requirements, and data. We then listed 17 locations in the U.S. where activity-based models have been developed or are under development, and described three different approaches to model development—upfront, phased, and transferred.

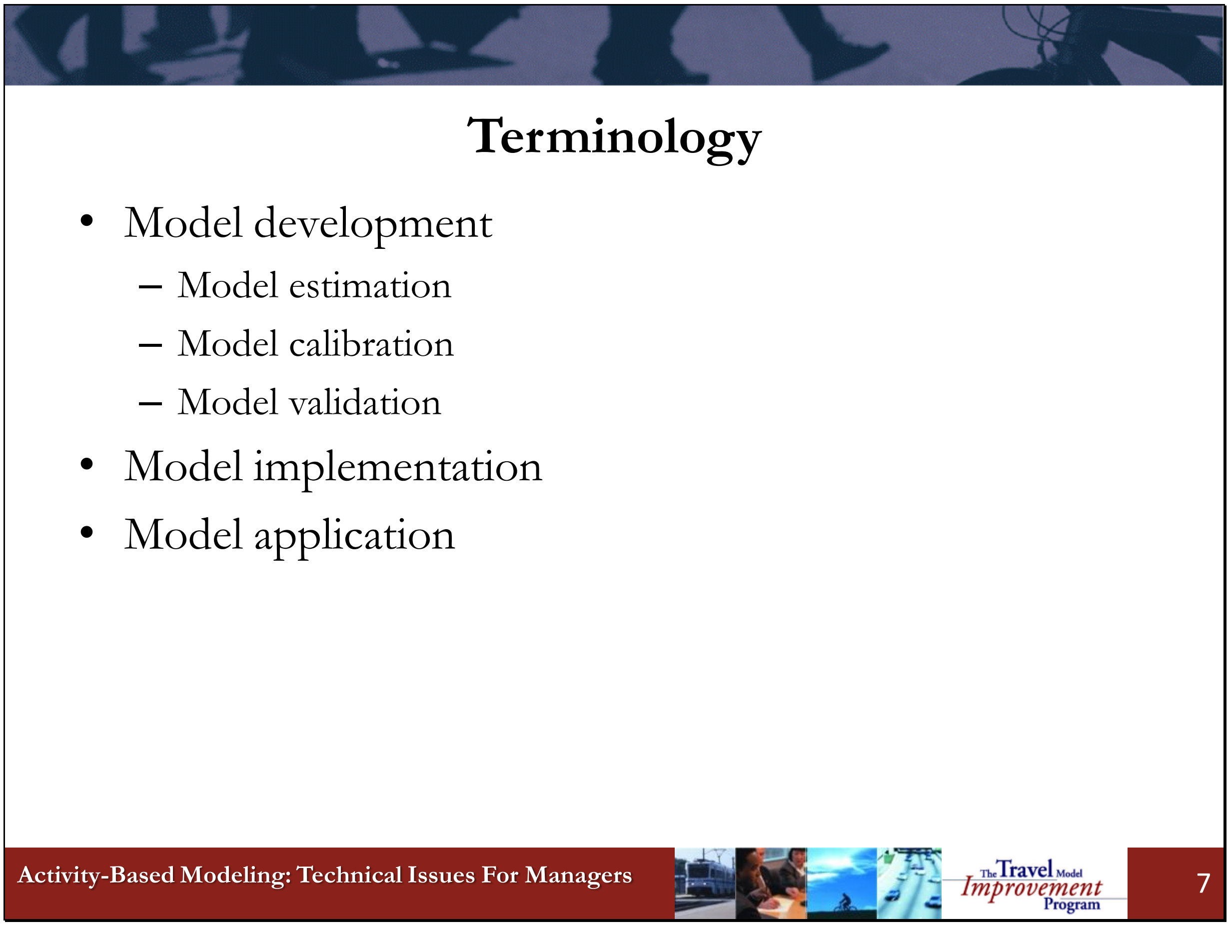
The second half of the second webinar provided several real-life examples of agencies' development costs, timelines, hardware and data requirements, and funding approaches. We then concluded with a detailed discussion of user experiences, including calibration and validation, examples of model applications, stakeholder acceptance and usage, and communicating results.



In today's session, we would like to accomplish two broad objectives. First, we would like to cover issues that we feel are important to managing the technical processes involved in activity-based model development and usage. By the end of this session, you should be able to discuss the processes used to develop, calibrate, validate and implement an activity-based modeling system.

The second learning objective is to discuss the criteria that agencies should consider when evaluating whether an activity-based modeling system is something they should pursue. This will be our attempt to "tie a bow" around the discussion of management issues related to activity-based modeling.

The third learning objective is to discuss the high-level model design decisions that will need to be made when embarking on activity-based model project. If an agency decides to postpone development of an activity-based model, what are some of the alternative transitional development paths that an agency can take in enhancing its modeling capabilities? It is our intent to answer this question, too.



In previous sessions, we've defined some of the terms used to describe what an activity-based model is. In future technical sessions, we discuss much more of those details. For this session, however, we'd like to define some terms related to the processes of model development and use.

Model development and analysis consists of three analytical activities: estimation, calibration and validation. Each of these activities may be characterized by different sets of methods aimed at fitting model outputs to observed data.

Estimation refers to the process of developing coefficient estimates for explanatory variables in individual model components, usually based on samples of individuals or households, such as a found in a household survey.

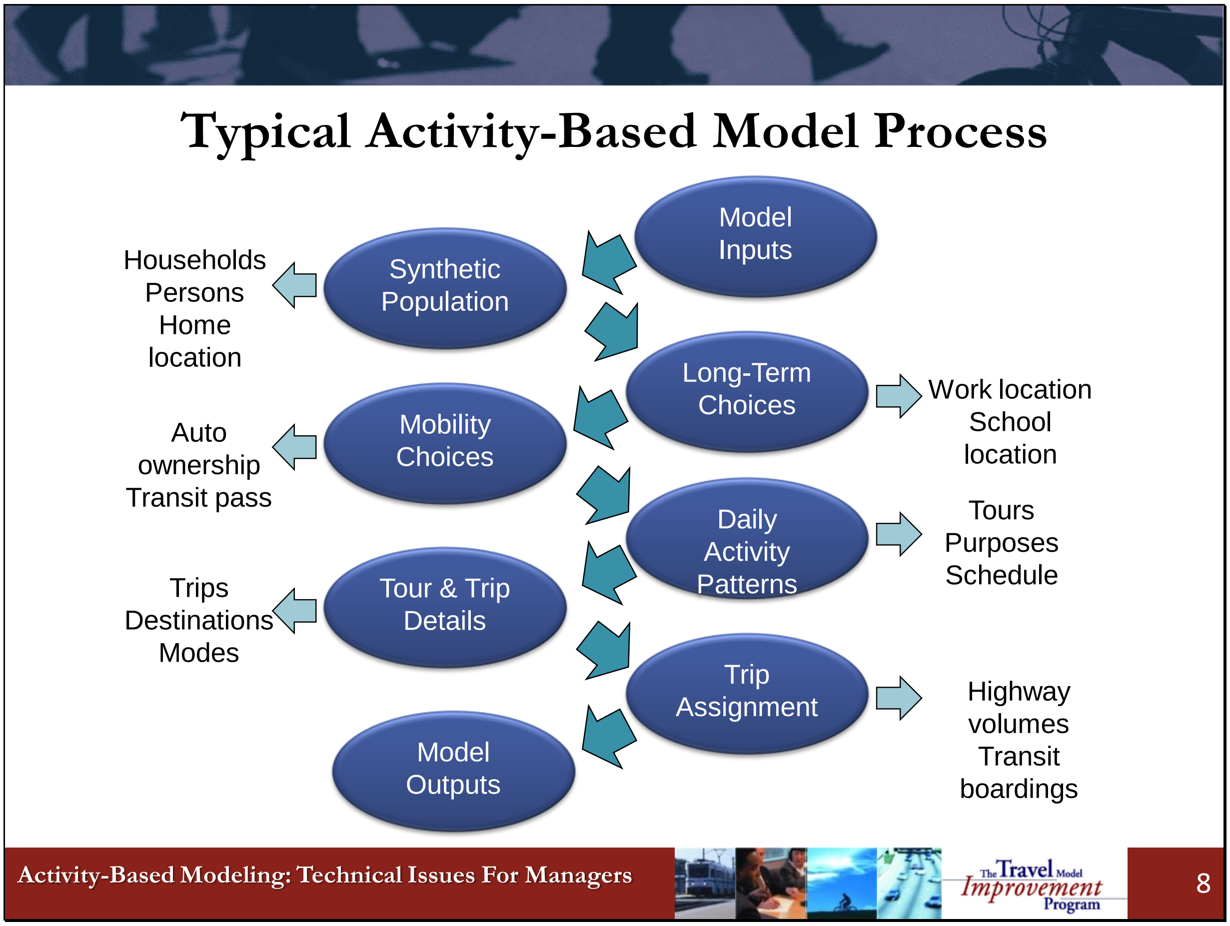
Calibration refers to applying the model to a set of inputs, usually representing the entire modeled population; comparing outputs to key benchmark values, derived from expanded survey data; and adjusting certain model components, iteratively, until a desired level of fit is attained. In some cases, it may be desirable to calibrate models to independent sources of data. A common example of this is to calibrate the means distances of workplace location choice models to Census journey-to-work data.

Validation refers to comparison of model outputs to independent observations, such as traffic counts, transit boardings, and Census-derived characteristics of the population.

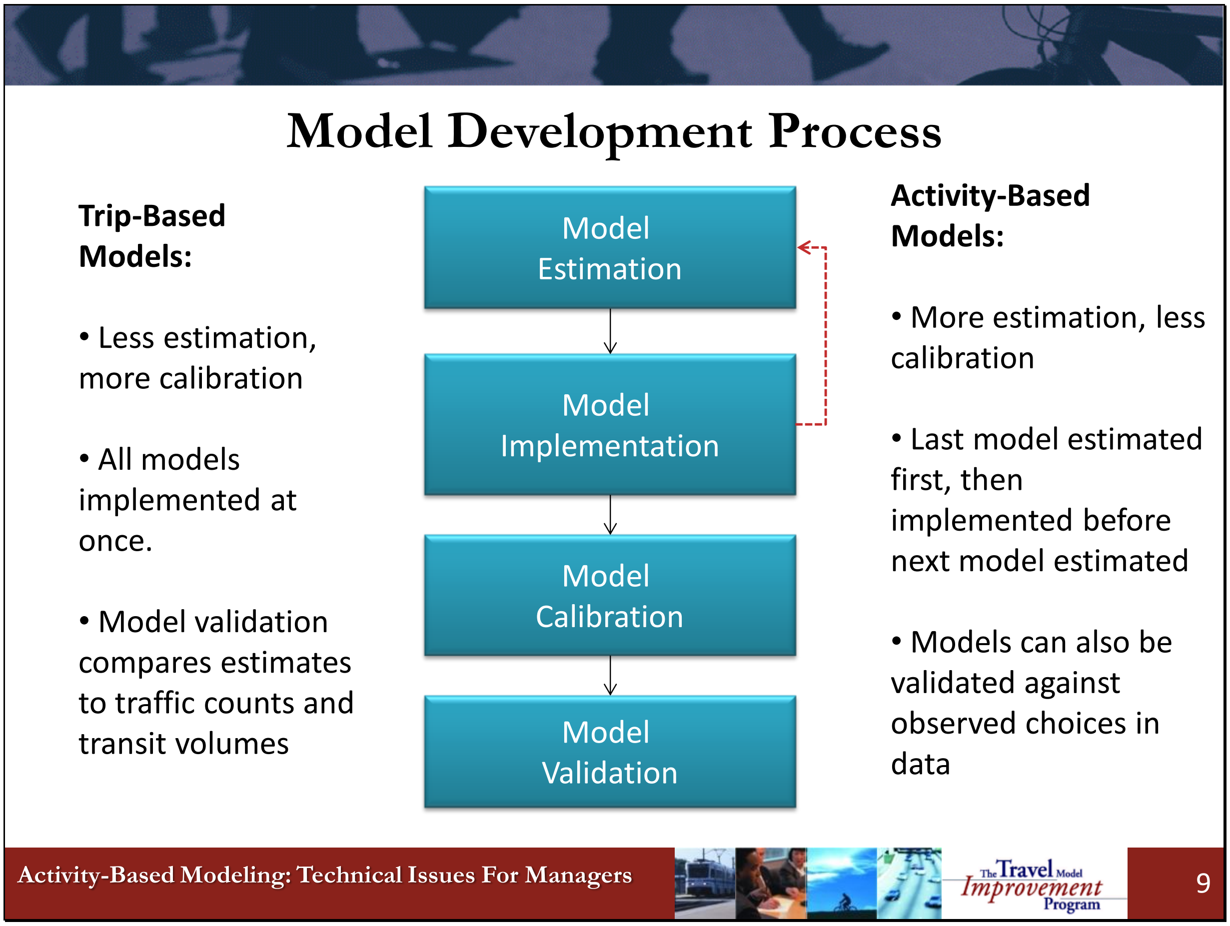
Model implementation is a term that is used in different contexts. For some, model implementation refers to the transformation of a mathematical model into an application program that can be run. In this sense, implementation is a preliminary step that allows one to begin model calibration. In another sense of the term, model implementation refers to the integration of an activity-based model into a larger forecasting system, which includes all demand model components, network models and auxiliary demand, prior to model validation. These are the two ways in which we discuss model implementation in this webinar.

In yet another sense, agencies may say that their model has been implemented, indicating that they have reached the milestone event of being able to use that model in a production context, which often implies that it has also been validated.

Model application is used a little less ambiguously. Typically, this refers to running the model to produce some type of analysis.

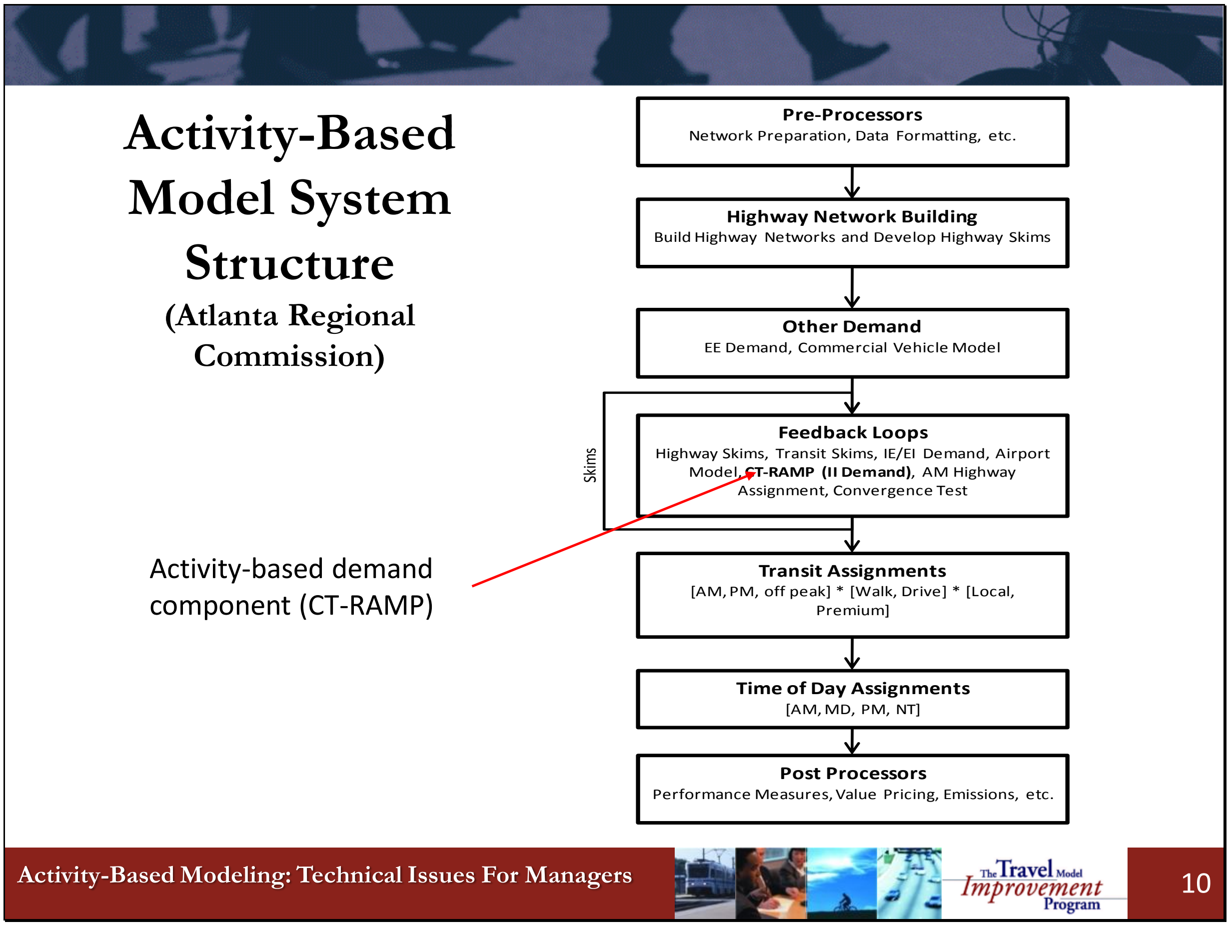


For purposes of discussion in the remainder of this webinar, it might be useful to review the basic components of an activity-based modeling system. While activity-based models can vary in structure, this diagram shows the location of tour and trip detail choices (tour mode, primary destination, intermediate stop location and trip mode) in a typical model stream. The text on the outside shows the types of outcomes predicted by each model stage. When we are ready to predict tour and trip details, we already have a synthetic population of households and persons with their home locations; we have predicted the primary work and school locations, auto ownership and other mobility decisions; and we have generated and scheduled tours using a daily activity pattern model. We do not yet know the primary destination of any non-work and non-school tours, the tour mode, the location of intermediate stops, or the trip mode. Once we are able to fill in these details, we are ready to convert the simulation data into trip tables that can be assigned to the network.



Trip-based models are typically developed in a linear fashion. First models are estimated using statistical techniques. For example, trip rates are calculated for household segments. Then models are implemented in application software. Then each model is calibrated, or adjusted, to better match expanded survey data. For example, friction factors are adjusted in gravity models to better match trip length frequency distributions. Mode-specific constants are adjusted to better match trips by mode. Finally, model outputs, such as link volumes and/or transit boardings, are compared to independent data to judge the validity of the system.

The process of developing activity-based models is very similar. However, because choices models are the foundation of most activity-based model systems, there is a greater reliance on model estimation rather than aggregate calibration. Also, since models are more inter-connected, often estimation begins with the last model in the sequence, such as trip mode choice. Then the model is implemented and the information from that model is used to estimate the next higher level in the model chain. Activity-based models can be validated in traditional ways; however, it is also possible to run the data that was used to estimate the models through the system to check the model predictions against the survey outcomes.



This diagram is a typical example of where an activity-based demand model would be integrated into a larger model system run. This is an example from the Atlanta Regional Commission. It might be useful for you to keep this diagram in mind as we discuss model calibration and validation.

As you can see, many of the steps you see here are the same or quite similar to what you would have in a trip-based modeling system. This includes the steps typically found in a model run stream, including: pre-processing network and other data, network and skim building, bringing in static auxiliary demand tables from other sources, and then entering the central model loop. Within the feedback, the activity-based demand model is run, along with any other models that might be sensitive to congestion effects. The feedback loop would include intermediate highway (and sometimes) transit assignments and would be run iteratively until convergence. Once the feedback loop has converged, final transit and highway assignments are run, and any post-processing modules are applied.

Data for Estimation & Calibration: Surveys

- Household survey
  - Provides complete daily trip diaries for entire household
- On-Board survey
  - Difficult to use for ABM because only trip data is reported
  - Special questions required to construct tours
- Special market surveys
  - Origin-destination
  - Downtown parking
  - Visitors
  - Special events
  - Air passengers
  - Other special markets/populations
- Census data
  - Calibration summaries
  - Synthetic population controls

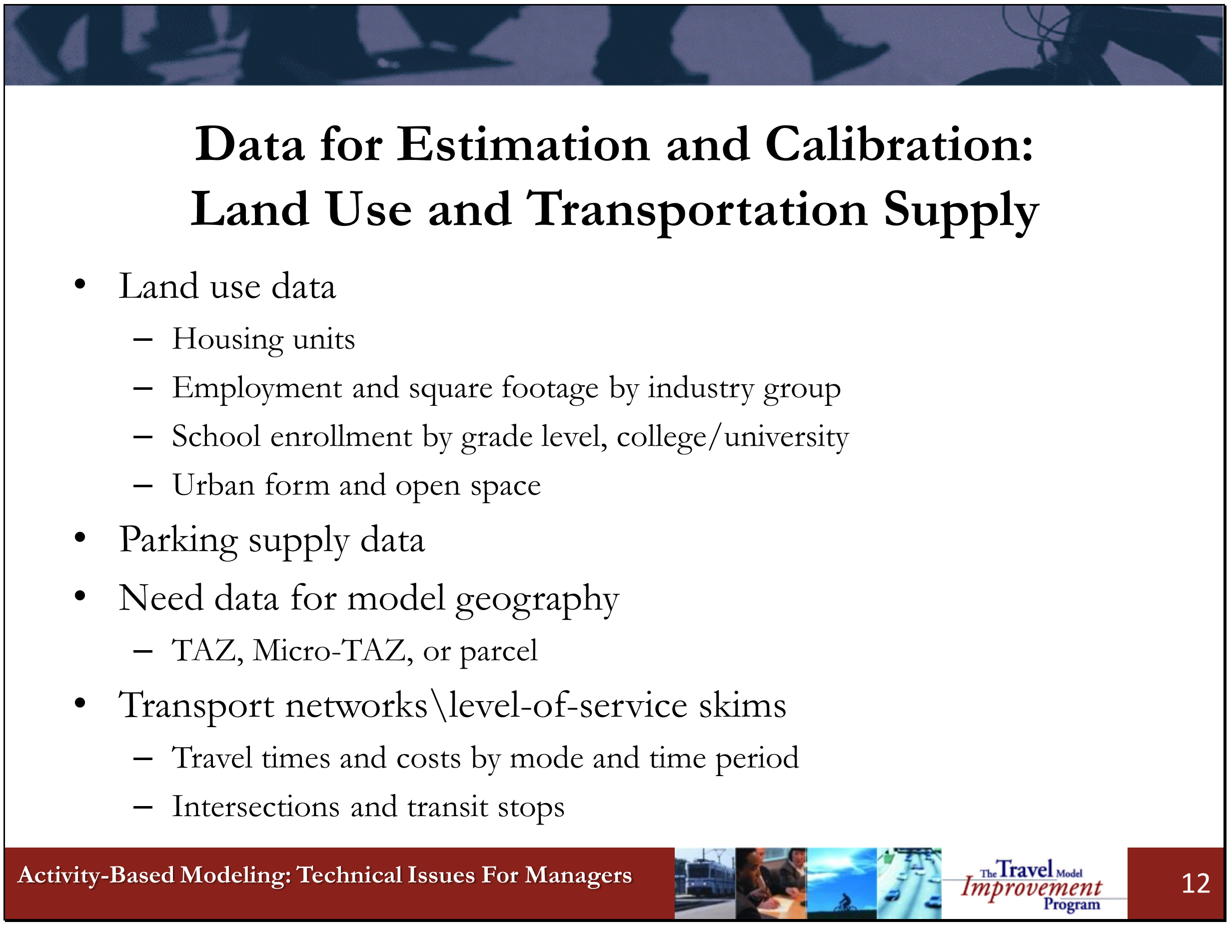
A household travel survey provides the information required to estimate an activity-based model. Several other types of surveys provide information useful for estimating and calibrating the model.

Household survey: The household survey ideally records all travel by all household members that occurs on a designated travel day. Required information for each trip includes location, party composition, start and end times, purpose and mode. In addition to the trip information, the survey compiles information about the household and its members, such as income, number of vehicles owned, usual workplace and school locations, and availability of free or subsidized parking at the workplace. Care should be taken to perform, in real-time, adequate consistency checks on the information provided, to ensure high quality data for constructing complete tours, joint tours, and activity scheduling and geo-coding.

On-Board survey: On-board survey data is most often used in activity-based model calibration. On-board surveys are difficult to use for estimating activity-based models because typically only one trip out of the entire tour is reported. Questions can be added to the survey form to elicit information that allows inferring the entire tour. An alternative is to over-sample transit users when conducting the household survey (though this likely provides only one transit user in the household).

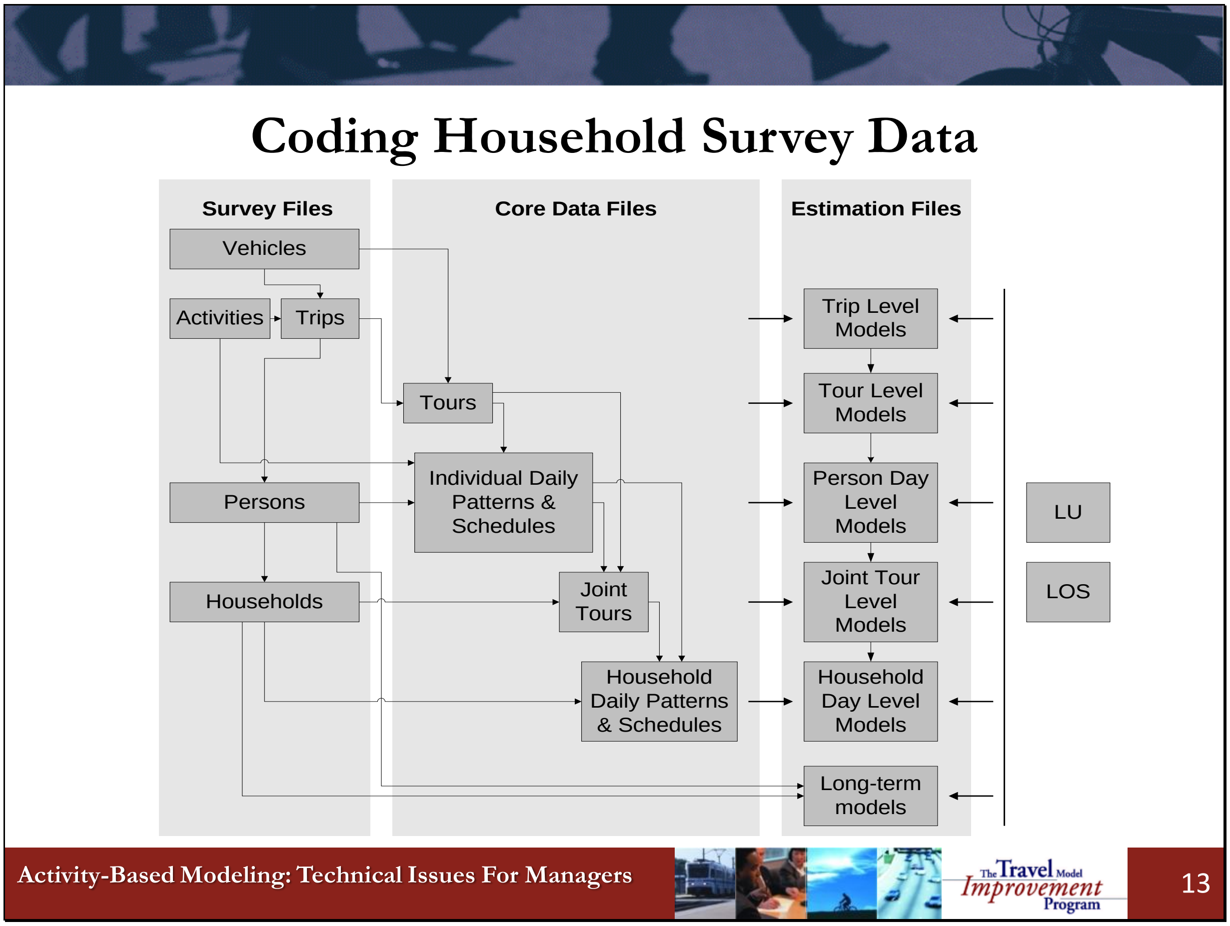
Special market surveys: Various targeted surveys can be conducted to understand and gather information about special markets of significance in the model region, such as visitors, air passengers, downtown parking users, and others. The importance of these markets varies from region to region and may be relevant only for particular transportation projects.

Census data: The Decennial Census and American Community Survey provide information critical for developing the synthetic population controls and for calibrating the auto ownership and workplace location models.

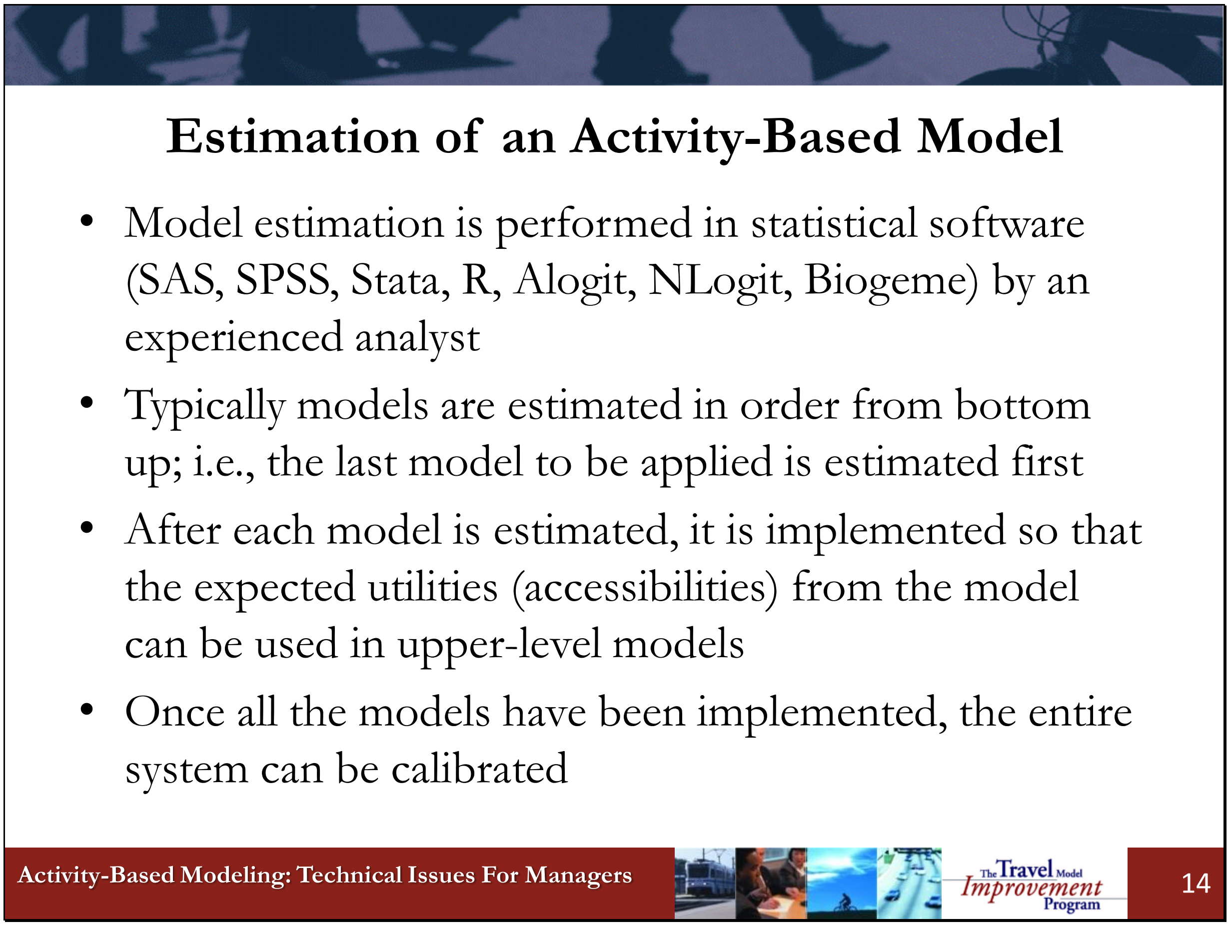


Land use data: Land use data characterize the location and type of activities in the model region. They may be gathered from many different sources, including census, general plans, real estate appraiser's databases, departments of education, payroll tax databases, and other sources of employment data. Similar data is used in trip-based models. Activity-based models that have been designed to work on parcel or micro-zones require that the land use data be expressed at these levels of geographic aggregation, instead of the larger traffic analysis zones.

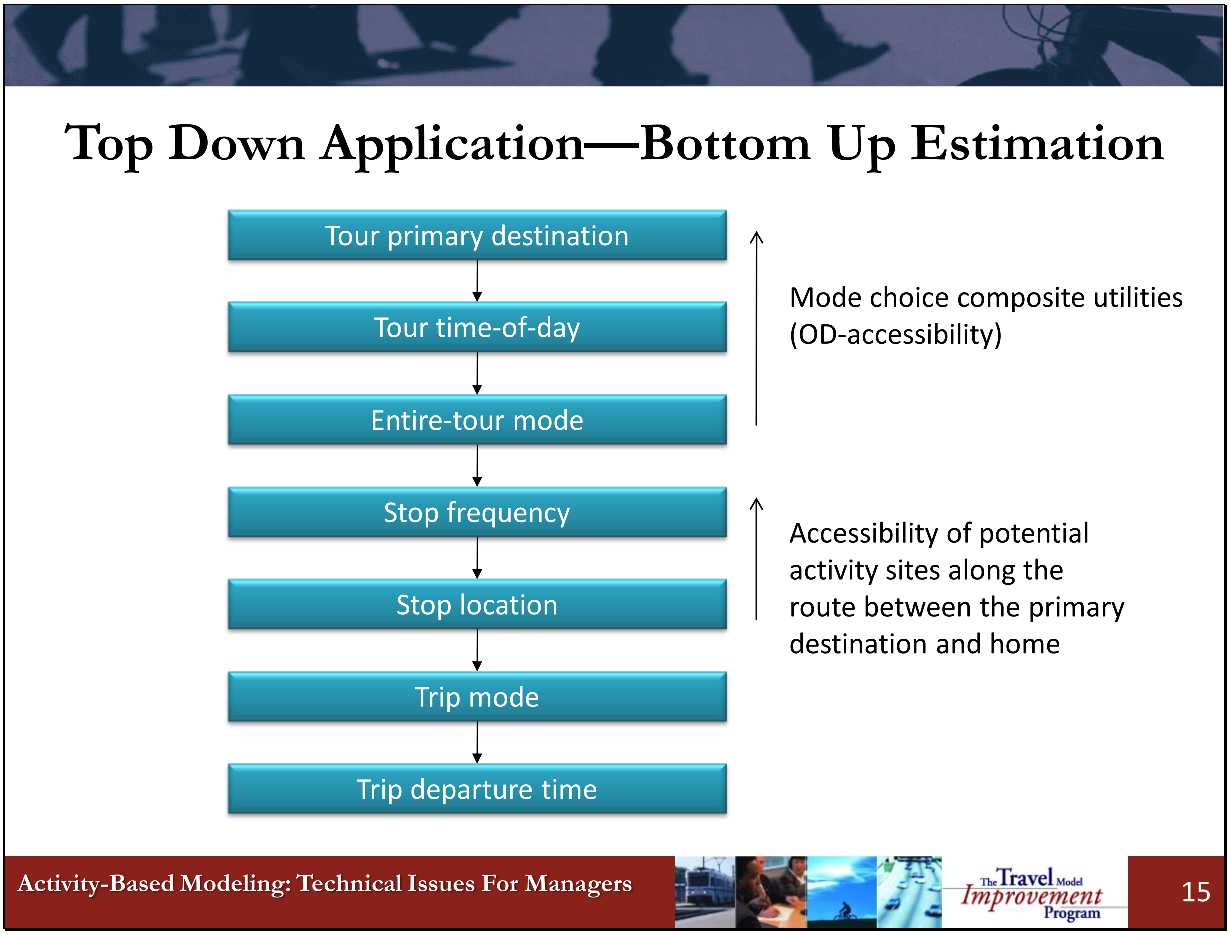
Parking supply data: Like trip-based models, activity-based models require an inventory of parking lot locations and prices for parking-constrained areas.



Household survey data is the foundation of activity-based models. Survey data typically consists of household, person, and trip (or activity) files, plus a vehicle file. These survey data files, are shown on the left. They are assembled into core data files, as shown in the middle of the graphic. They include tour files and daily activity pattern files that describe what each person did on the surveyed day. These are the observed data files that are used to calibrate the system. A more comprehensive set of data files, as shown on the right, is required for model estimation. This involves appending land-use and level-of-service data to the core data files, typically with one record per choice with descriptions of each alternative that was available for that choice. These files are read by estimation software.

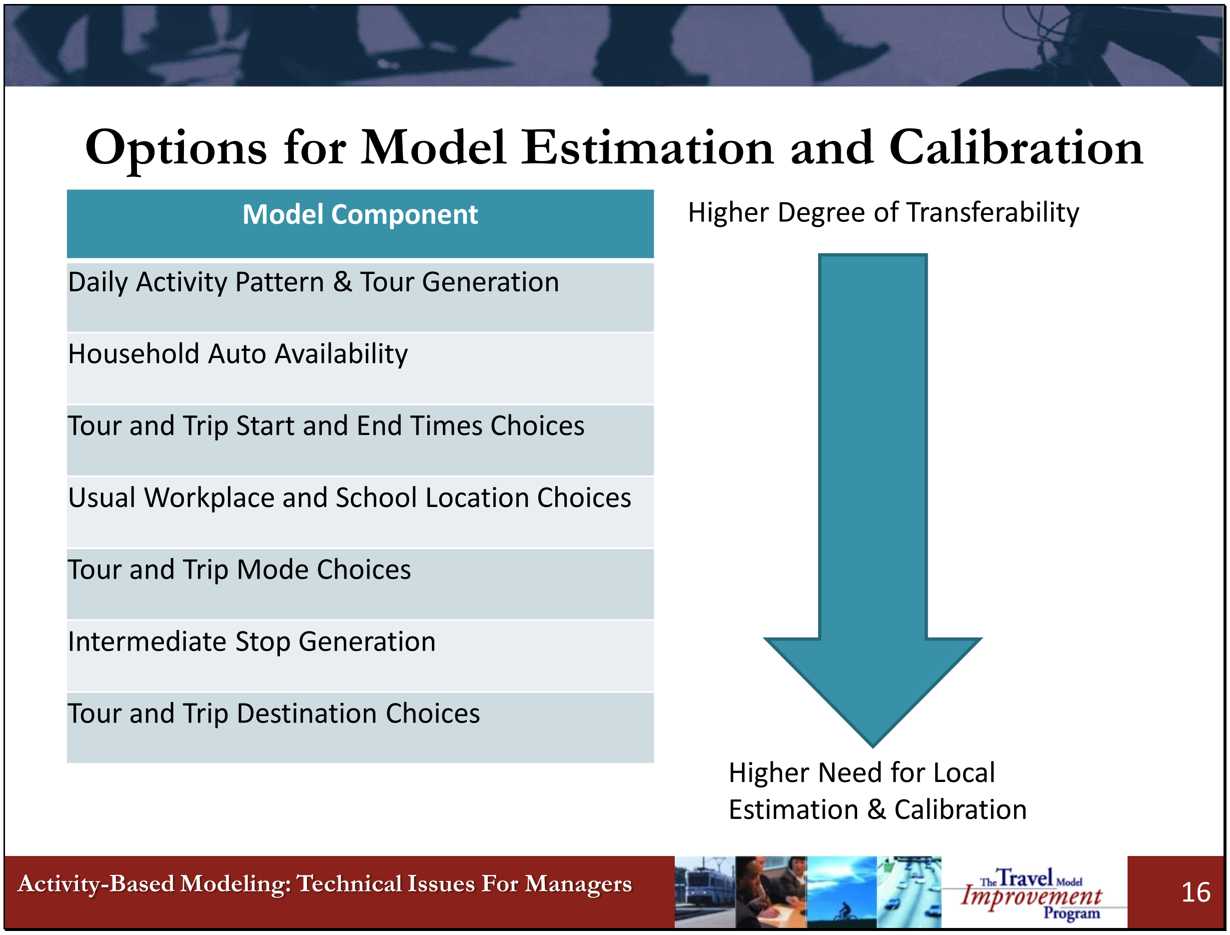


Model estimation is performed in statistical software (SAS, SPSS, Stata, R, Alogit, NLogit, Biogeme) by an experienced analyst. Typically models are estimated in order from bottom up; that is, the last model to be applied is estimated first. After each model is estimated, it is implemented so that the expected utilities (accessibilities) from the model can be used in upper-level models. Once all the models have been implemented, the entire system can be calibrated.

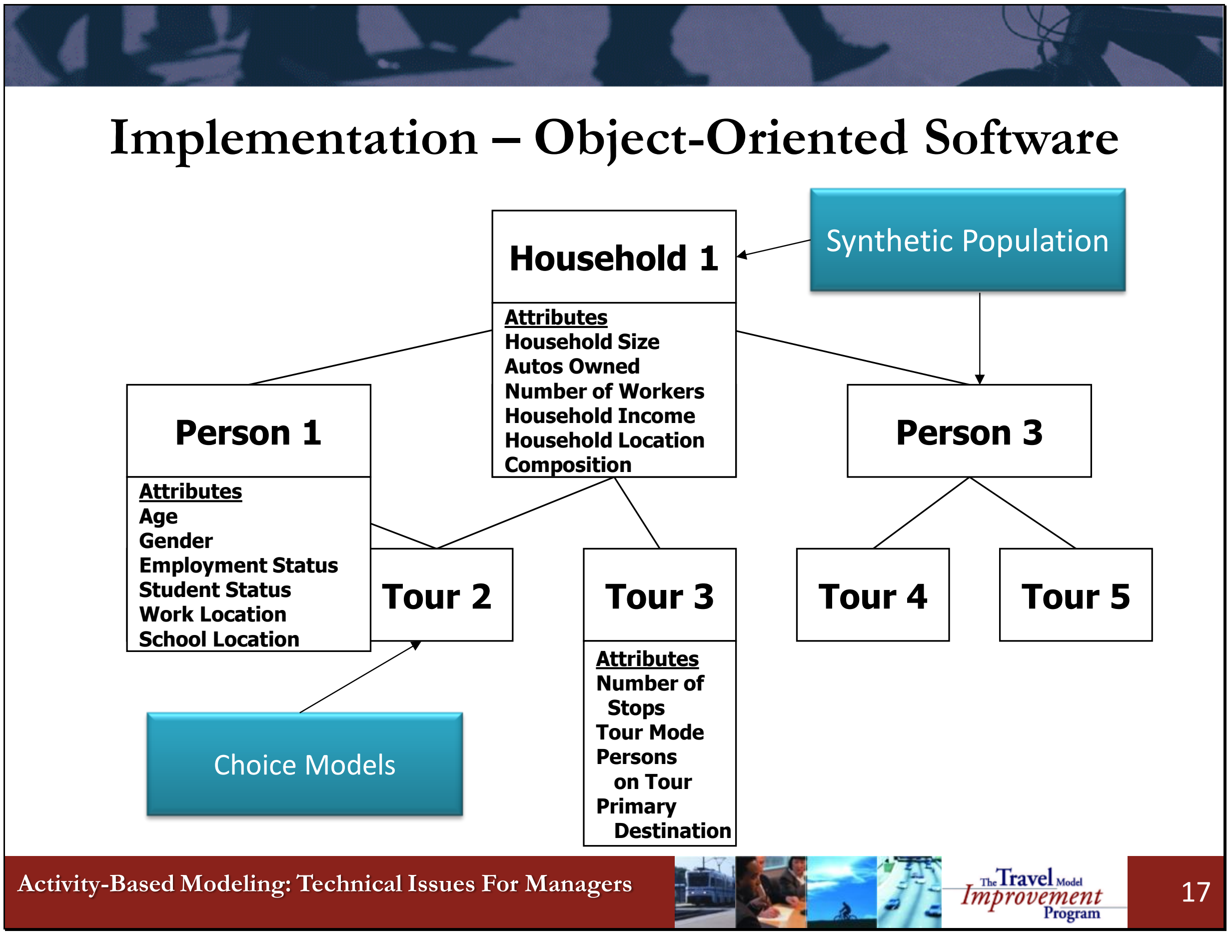


This graphs illustrates a portion of an activity-based model system in which tour-level decisions condition trip level choices. The model is applied from the top-down, with tour-level choices made first. In order to estimate the upstream models, however, it is necessary to first estimate the models downstream. This is particularly important when the models are structured as a series of nested choices, which most activity-based models do. The composite utilities (“log-sums”) of lower level choices are used as explanatory variables in the choices above them.

This might be the ideal process for model development. Let’s take a look at some other options.

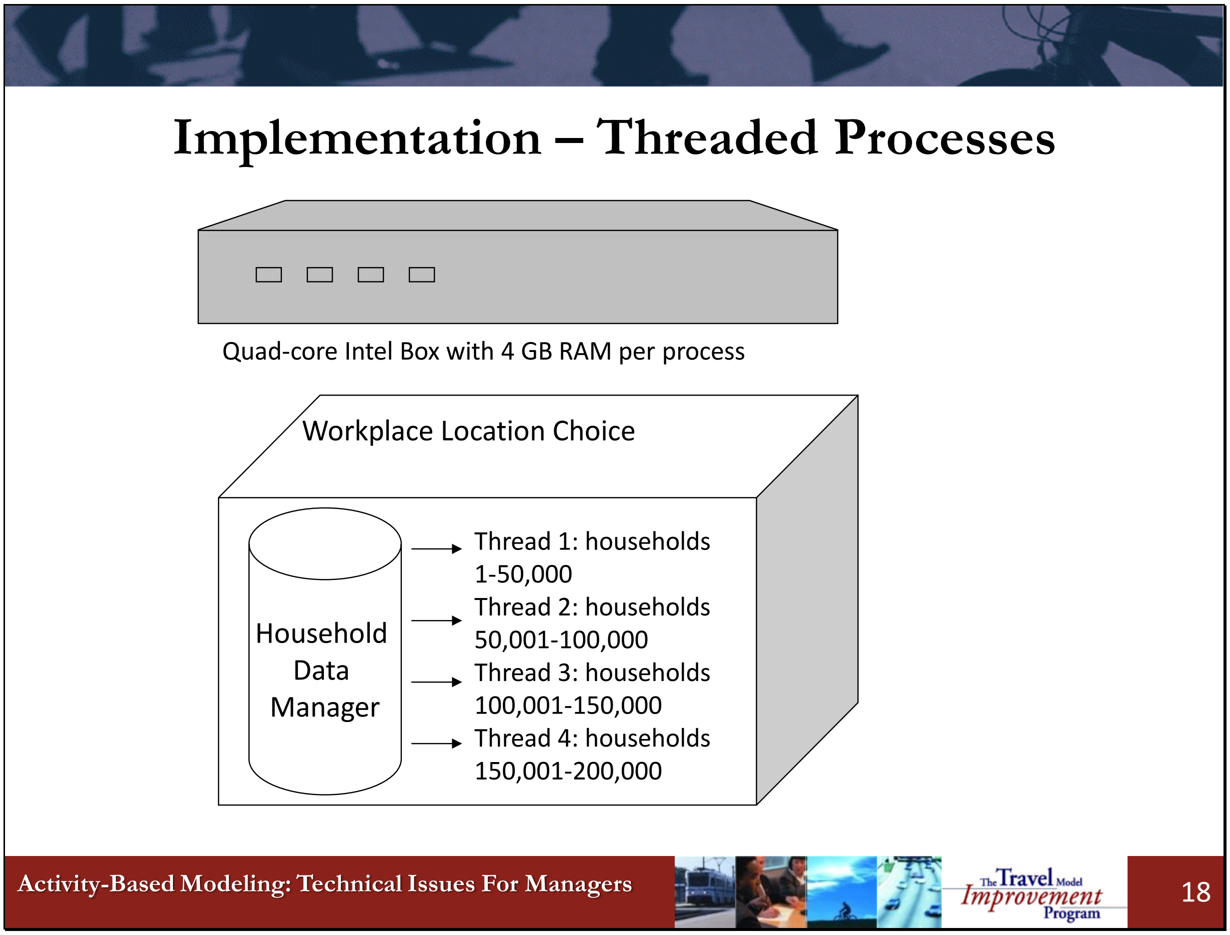


In the second webinar we talked about alternative paths to developing activity based modeling systems. One of these options was to transfer a model system from another region. That is, we would borrow model structure and initial parameter estimates. For certain model components, such as activity generation modules, there tends to be a high degree of transferability, because household and person types have the most influence on model parameters and people tend to be similar across regions. Other model components, however, such as tour or trip destination choices, are heavily driven by the unique spatial structure and transportation system service characteristics of each region, and really require estimation on locally derived data. This might be a refinement step that follows the initial model transfer. For all model components, calibration based on available local data is highly recommended.

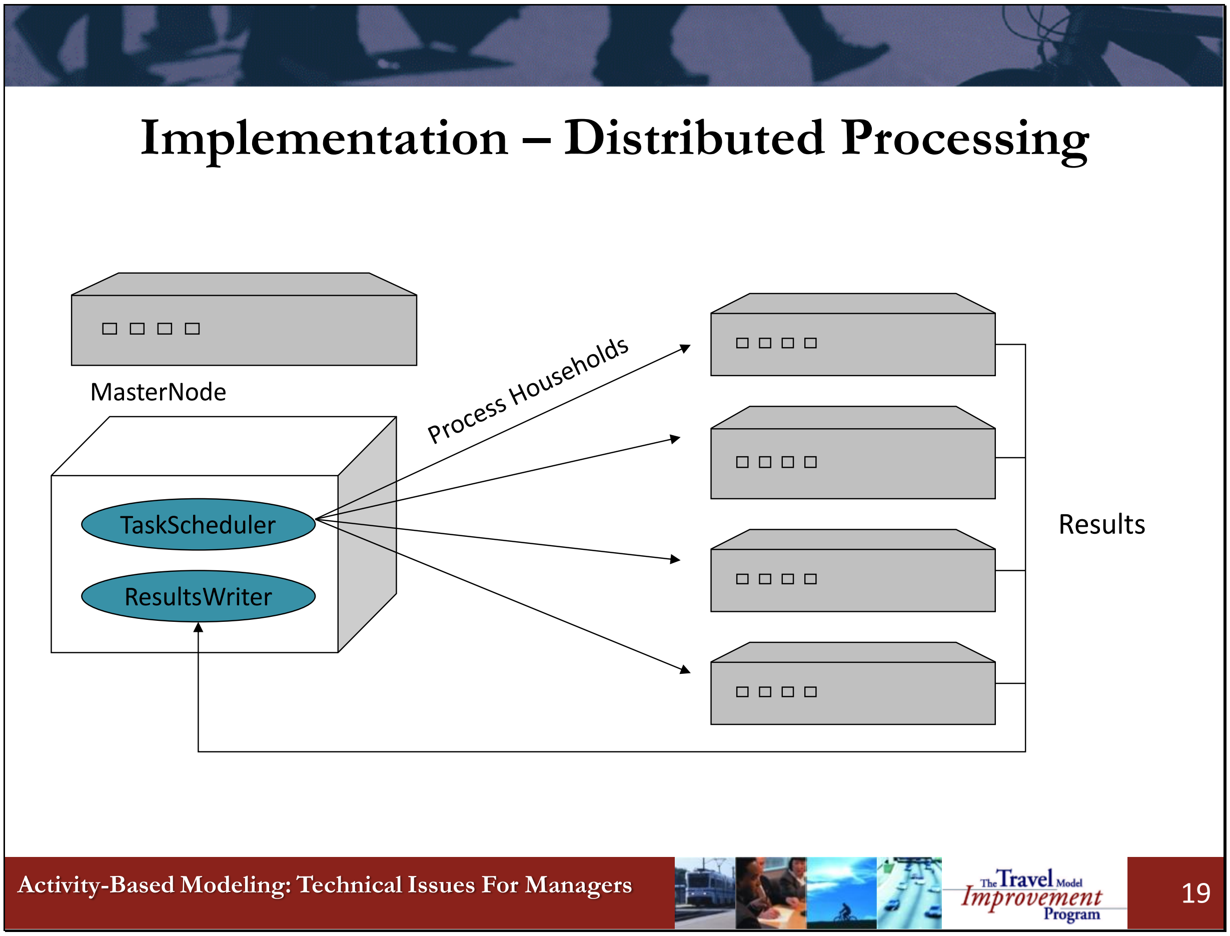


Before we can calibrate a model, we need to code it into an application program. As mentioned in the first webinar, activity-based models utilize some form of simulation to predict outcomes. Remember, we are simulating the movements and decisions of persons and households. The trip-based modeling structure of creating numerous matrix files, simply does not work in an activity-based modeling system in which you have to track individuals.

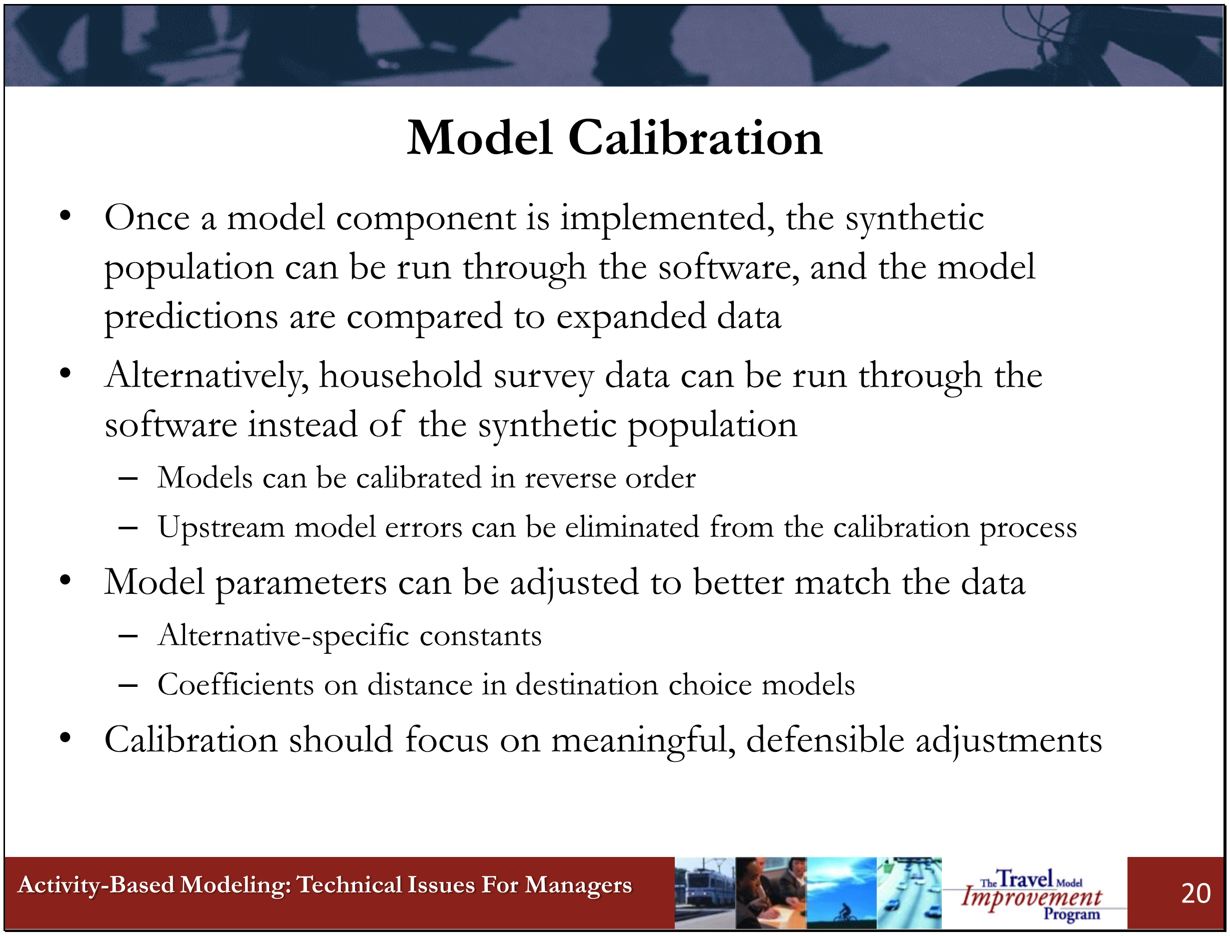
Most activity-based models rely upon object-oriented software for implementation. A paradigm referred to as composition in object oriented software are used to represent relationships between data items, such as households, which have attributes such as household size, autos owned, and possibly other attributes. Persons belong to households, and are also attributed with certain characteristics. Persons make tours, which consist of trips. Note that some of these attributes are read from the synthetic population, and some of these attributes are filled in by choice models.



The ability to process each household separately makes it easy to run activity-based models across multiple processors. Often, models are threaded to take advantage of multiple cores on a single computer. In this example, each of four cores are tasked with performing workplace location choice for a set of 50,000 households.



If even more computing power is needed (or desired), models can also be distributed across multiple computers to decrease run times (or allow for greater complexity, disaggregation, or behavioral fidelity in the model system). This graphic shows how a single computer sends bundles of households to a cluster of worker machines, which then process the choice models on those households and returns the results to the master computer which writes those results to disk.



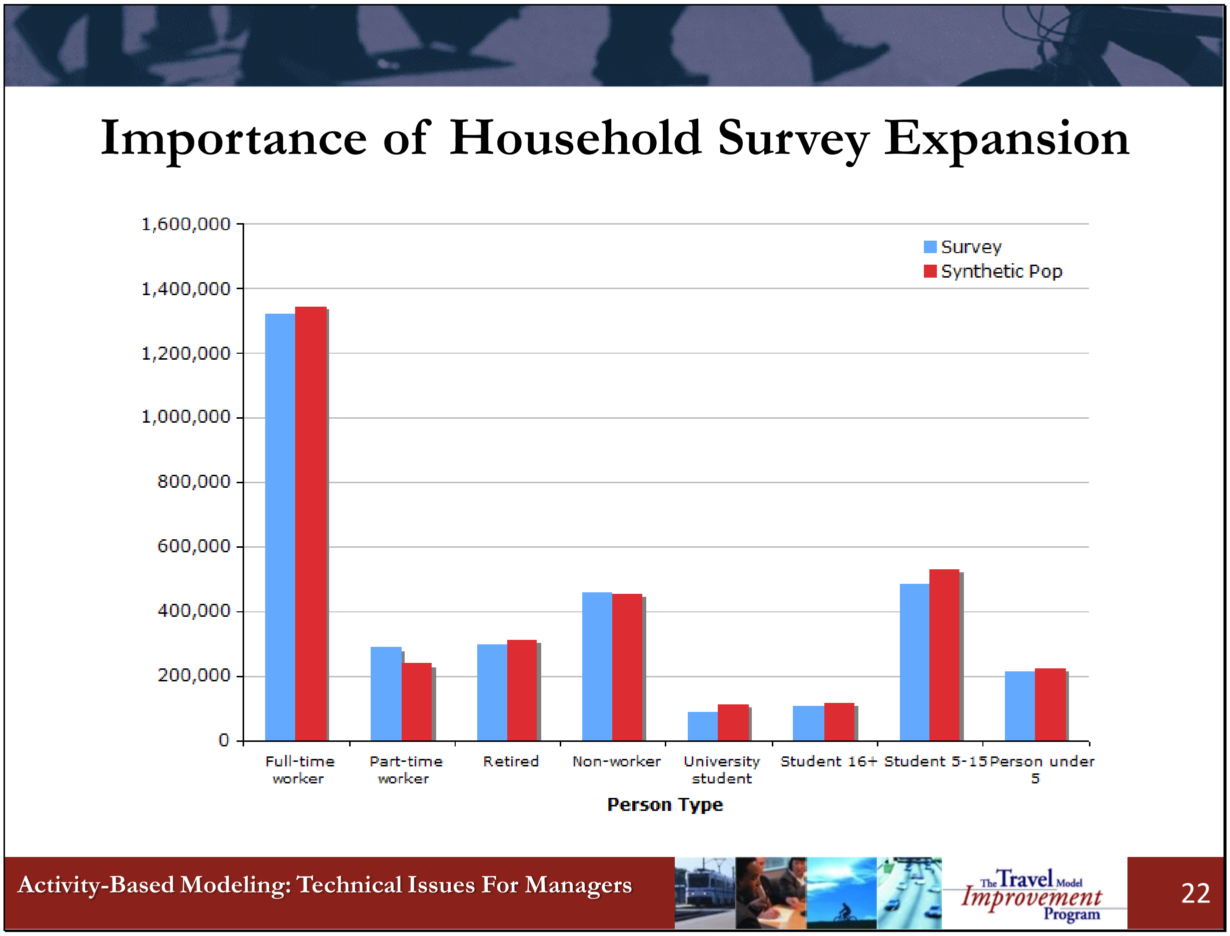
Model components are typically implemented in stand-alone software such as C++, C#, or Java. Once a model component is implemented, the synthetic population can be run through the software, and the model predictions are compared to expanded data. Model parameters (alternative-specific constants) can be adjusted so that predictions to better match observed data. Alternatively, household survey data can be run through the software instead of the synthetic population. This provides the ability to calibrate models from bottom-up, rather than top-down, because the survey has the observed choices for all of the upstream model components. In this way, upstream model errors can be eliminated from the calibration process. Model calibration should focus on meaningful, defensible adjustments to model parameters.

Model Calibration

- Similar calibration comparisons as in a trip-based model
  - Trip generation rates -> tours and stops by purpose
  - Trip distribution -> tour destination choice and intermediate stop location choice
  - Mode choice -> tour and trip mode choice
  - HB work trips -> usual workplace (Census JTW data)
- Different model components in an activity-based model
  - Joint travel and mobility models
- Person-level calibration versus aggregate calibration
- Importance of appropriate synthetic population and well-expanded observed data

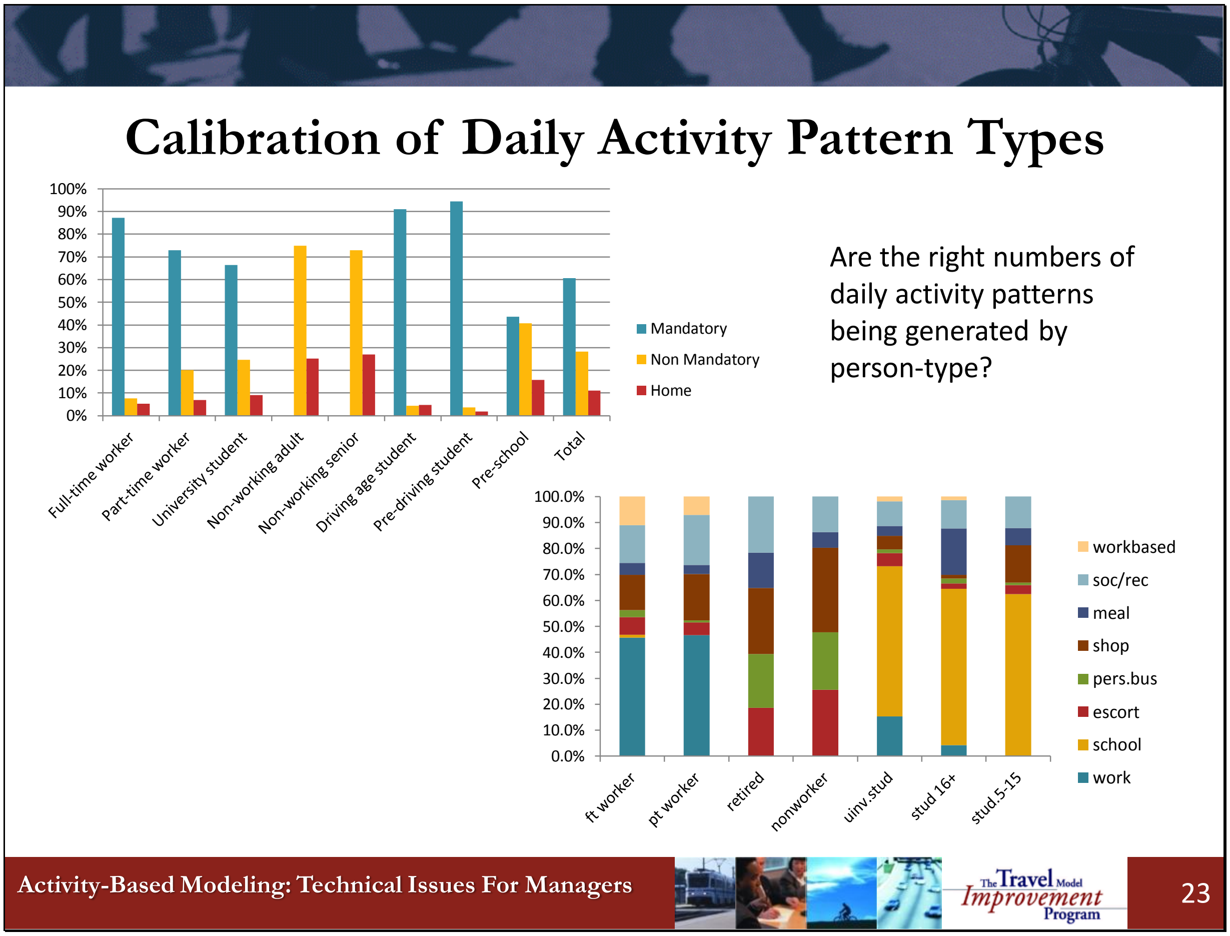
The calibration of a activity-based model is quite similar to that of a trip-based model. The goal of the calibration is to verify that the model reproduces certain patterns of travel behavior, and there are parallels between the types of summaries used to calibrate a trip-based model and the summaries used to calibrate an activity-based model. There are, however, some key differences:

- First, an activity-based model typically consists of a larger number of sub-models (for example tour and trip mode choice, instead of just trip mode choice), and in some cases models rarely included in trip-based models (joint travel and mobility models, for example).
- Second, it is often desirable to calibrate in the same order as model estimation, from the bottom-up. The reason for this is the same. So that composite utility variables from lower-level models represented their calibrated values in the upper-level models.
- Third, in an activity-based model, calibration targets specific person types, rather than aggregate market segments of households.

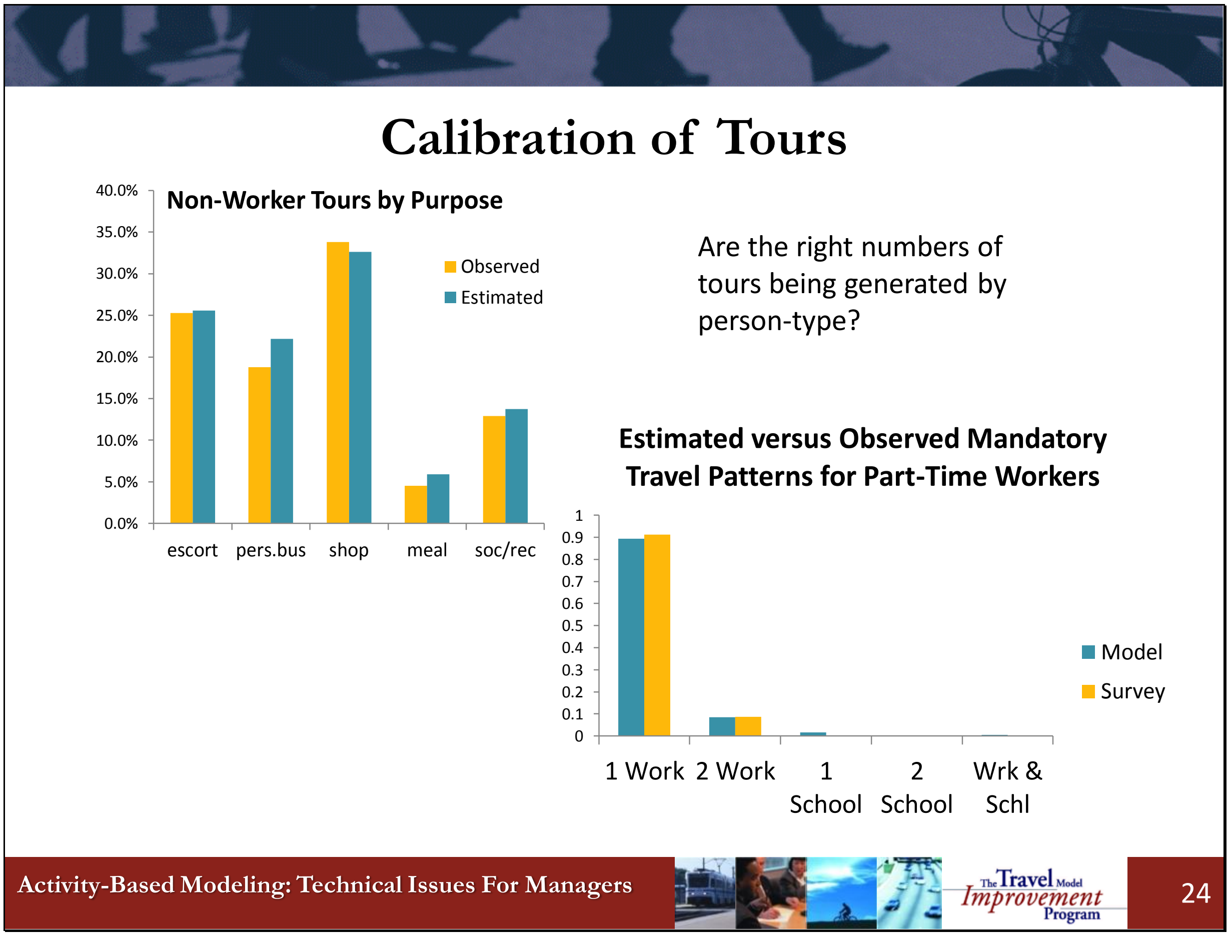


Activity-based models rely heavily upon person-level characteristics, such as work status, age, and household role, in explanatory variables. Often, activity-based models gauge model goodness-of-fit based upon these market segments. For example, does the model send the right number of full-time workers to work on an average weekday?

It is very important that the synthetic population and the calibration target values derived from expanded survey data agree in terms of attributes and attribute levels (e.g., age, income groups, auto sufficiency, etc.) before models are calibrated. Otherwise, the model calibration summaries will be difficult to interpret and compare to the household survey. This slide shows a summary that was put together for the Seattle activity-based model, comparing the number of persons by person-type in the synthetic population to the household survey. It shows a very good fit between the two data sets by person type. Additional summaries would include household attribute distributions (such as households by income, size, etc) and geographic distributions (such as households by district).



Model calibration of tour generation model components includes both the overall daily activity pattern and the numbers of tours by purpose. These calibration checks are often performed by person type. The chart at the top left of the screen shows a calibration summary for a daily activity pattern model whose alternatives are mandatory (at least one out-of-home mandatory tour), non-mandatory (at least one out-of-home mandatory tour) and stay-at-home. The summary shows the share of persons by type who engage in each of these activity patterns. A similar comparison would be created for the observed data, and constant terms can be adjusted to improve fit.



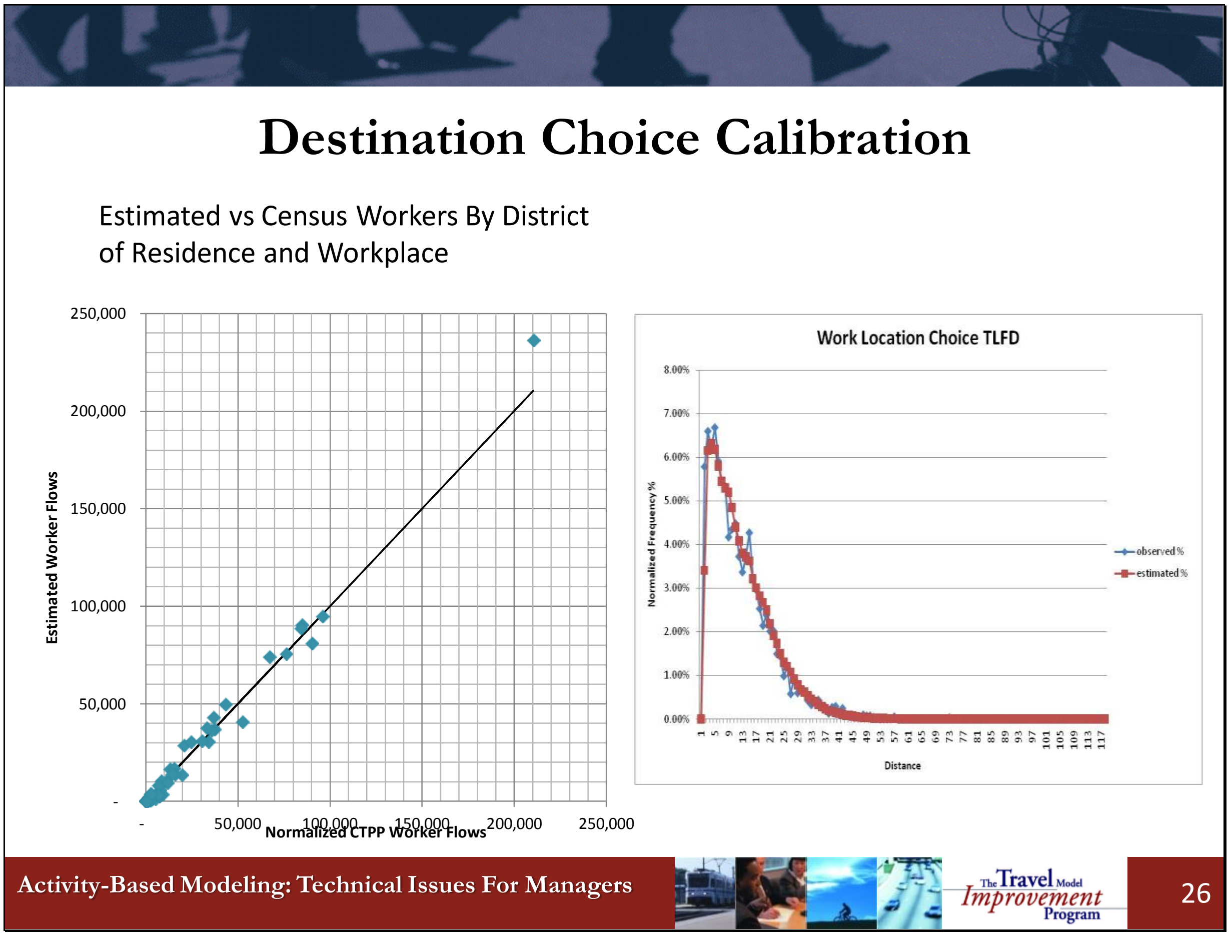
Once the daily activity pattern model is calibrated, the exact number of mandatory tours can be calibrated by person-type, in order to ensure that the right numbers of tours are generated.

Adjustments to Alternative-Specific Constants

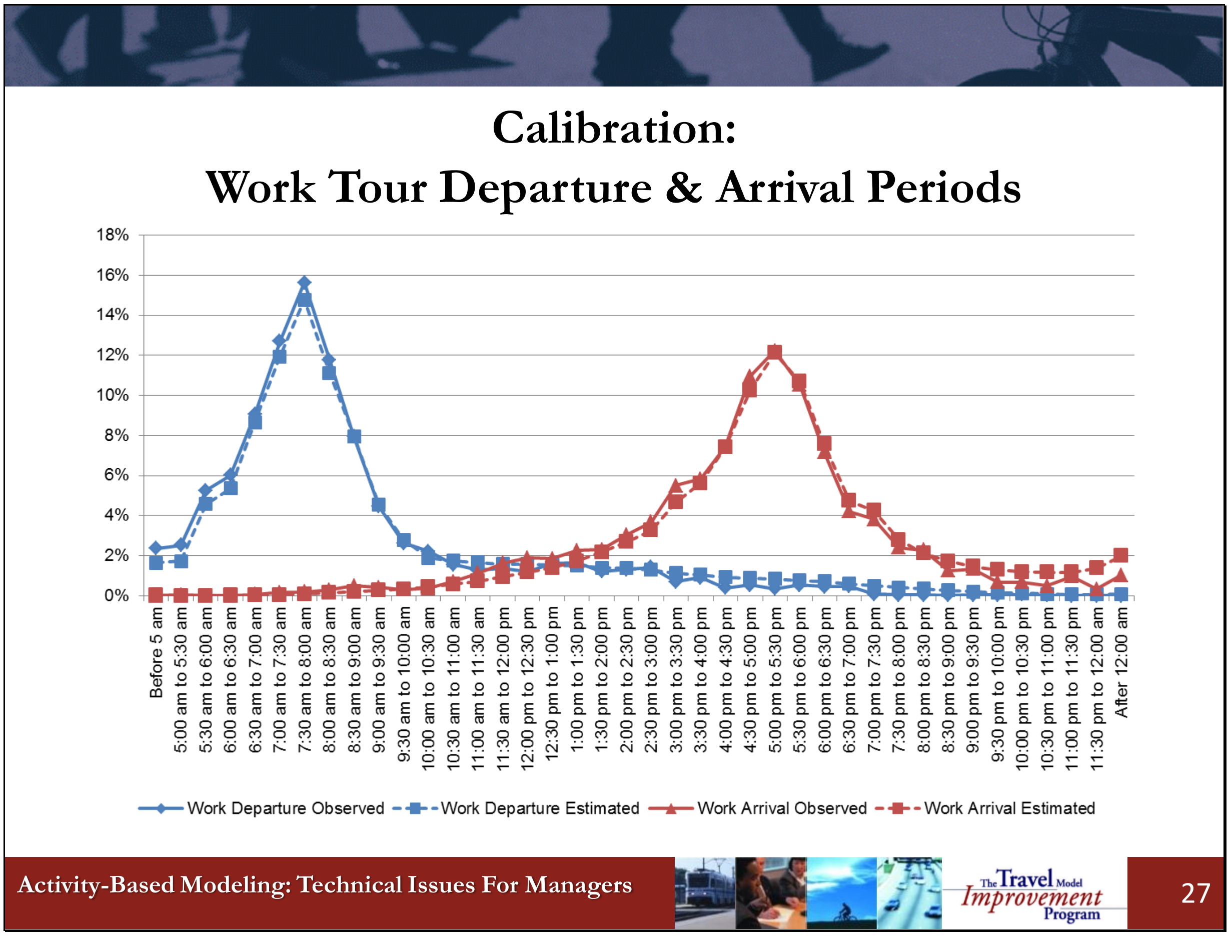
	Estimated	Calibrated
Stay At Home		
Full-time worker	0.0000	-0.3567
Part-time worker	-0.8884	-1.2451
Retired	-0.9572	-1.3139
Non-worker	-0.8739	-1.2306
University student	-1.0212	-1.3778
Student 16+	-0.3650	-0.7217
Student 5-15	-0.4576	-0.8143
Person under 5	-1.1964	-1.5531
Work Tour		
Alternative Specific Constant	1.0705	1.1701
Full-time worker	0.0000	-0.3000
Part-time worker	-1.3157	-1.1157
School Tour		
Alternative Specific Constant	-2.8699	-3.3350
Full-time worker	0.0000	-1.8000
University student	2.2136	0.2136
Student 16+	2.6766	4.6766
Student 5-15	2.5194	3.5194
Person under 5	0.0000	0.3500

	Estimated	Calibrated
Escort Tour		
Alternative Specific Constant	-3.5706	-3.5321
Full-time worker	0.0000	-0.0400
Part-time worker	0.2330	0.3830
Personal Business Tour		
Alternative Specific Constant	-1.9002	-2.0880
Full-time worker	0.0000	-0.0400
Part-time worker	0.1670	0.3170
Shop Tour		
Alternative Specific Constant	-3.1667	-2.9325
Part-time worker	0.3400	0.4900
Meal Tour		
Alternative Specific Constant	-4.9213	-3.8873
Full-time worker	0.0000	-0.0400
Part-time worker	0.0000	0.2000
Social / Recreational Tours		
Alternative Specific Constant	-2.3216	-2.3718
Full-time worker	0.0000	-0.0400
Part-time worker	0.1241	0.2741
Interaction Effects		
Work + Shop	-1.0889	-0.5889
Work + Meal	-0.6221	-0.1221
Personal Business + Shop	-0.2805	-0.0805

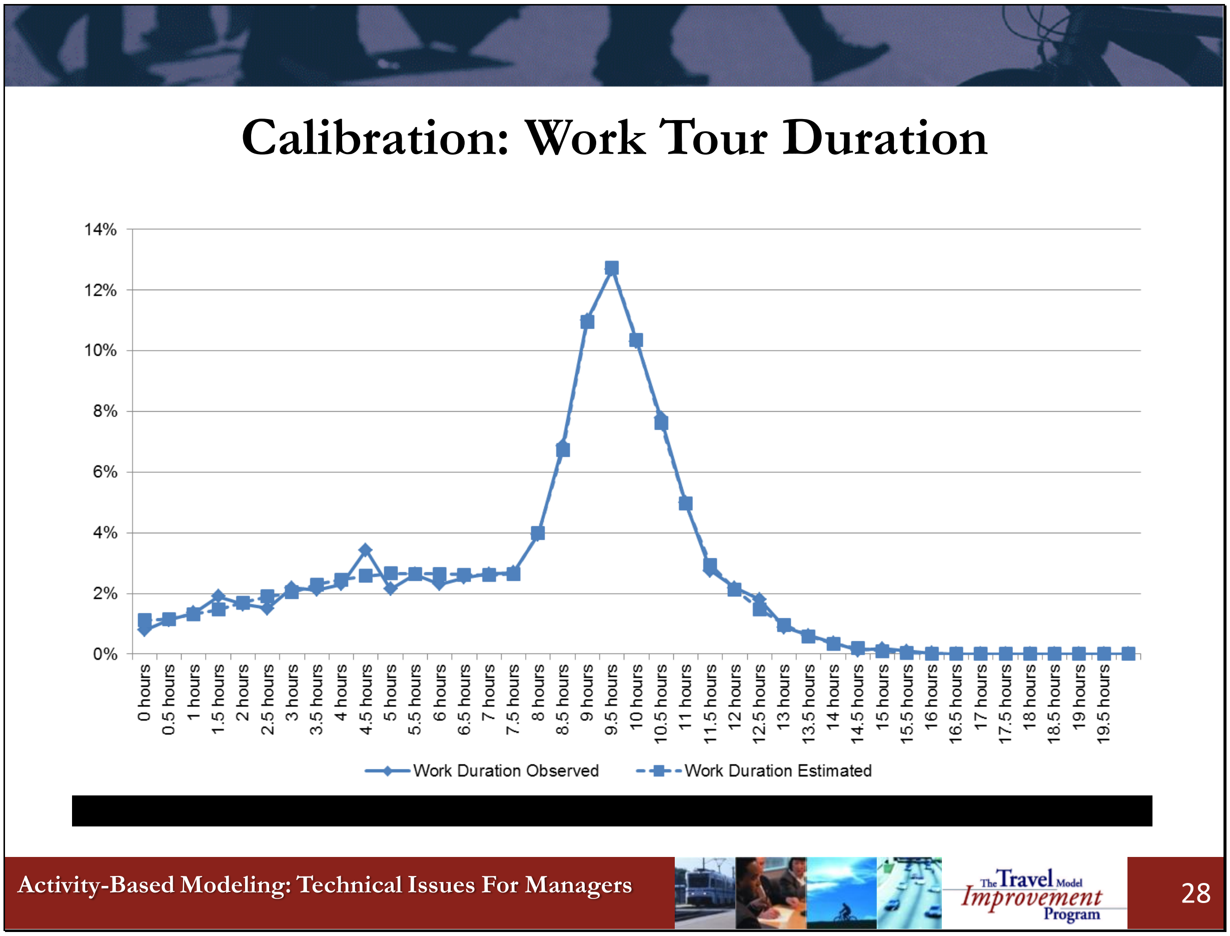
This slide is an example of alternative-specific constant adjustment for the DaySim daily activity pattern model developed for Seattle. There are a number of choice dimensions handled by this model, including number of tours by purpose, and number of stops by purpose. This table summarizes the adjustments made to the alternative-specific constants that represent the number of tours by purpose. There are constants for the core alternatives, such as work, school, escort tours, and other tour purposes, as well as constants by person-type and tour purpose.



Destination choice models need to be calibrated for both primary tour destination as well as intermediate stop locations. Shown on this slide are some model calibration summaries for the San Diego work location choice model. The results were compared to census worker flows at a district level (as shown on the left) as well as trip length frequency distributions (as shown on the right). Note that the model did not require any K-factor adjustments!



Time-of-day choice models should also be compared to survey data to ensure that tours begin and end at times consistent with observed travel patterns. Typically this calibration is done by tour purpose.



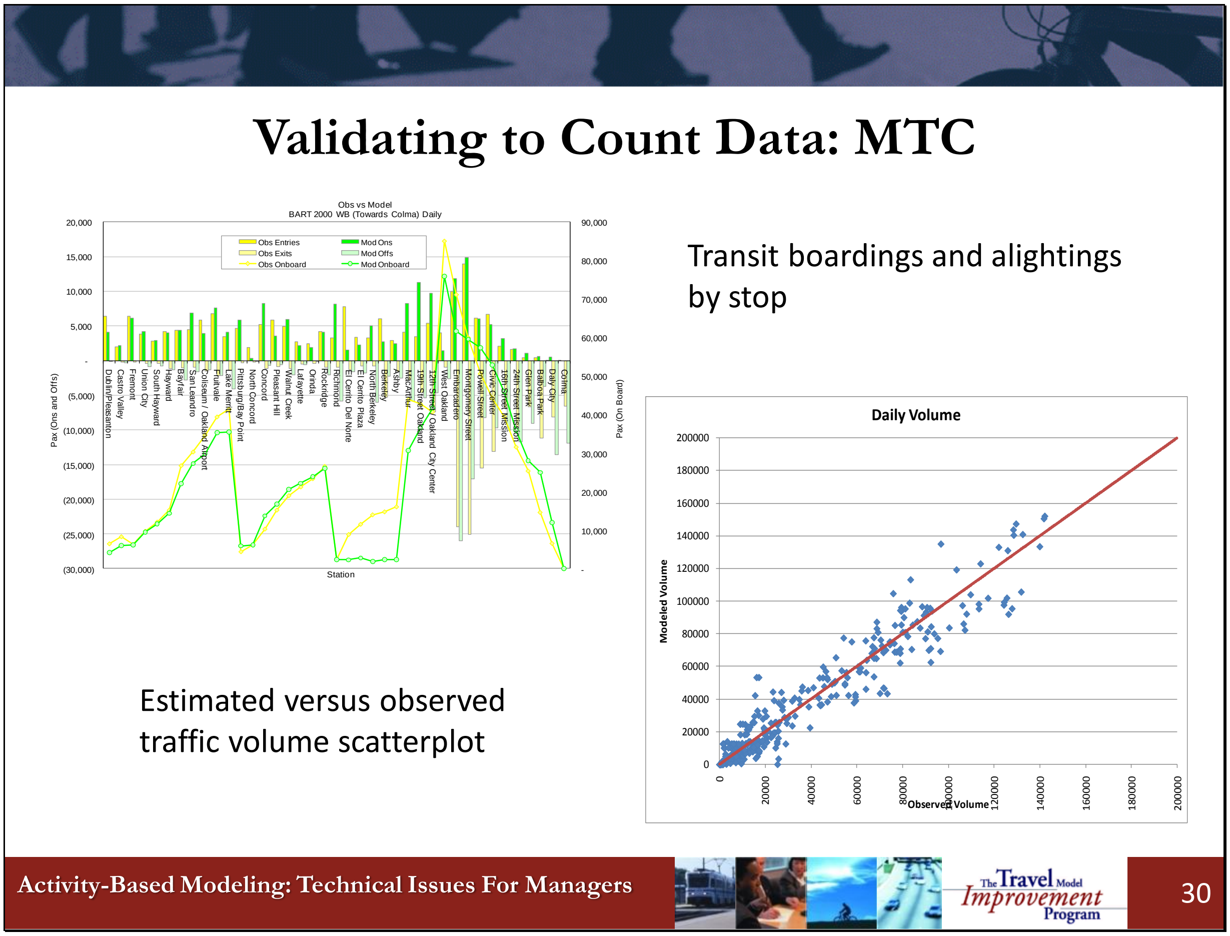
This shows a calibration summary on work tour duration. The estimated and observed distributions line up very well, thus it is difficult to distinguish between solid (observed) and dashed (estimated) lines.

Model Validation

- Check model performance against independent data sources
  - Typically traffic counts (by period) and transit boardings
- Sensitivity testing
  - Change network or land-use data
  - Does model respond appropriately?
- Look for systematic errors
  - Software bug?
  - Illogical coefficients?
  - Missing market segments?
  - Missing variables?

Once all of the model system components have been thoroughly calibrated, individually and as a system, it is time to validate the base-year model. The general approach is very similar to validating a trip-based model system, but considers more information.

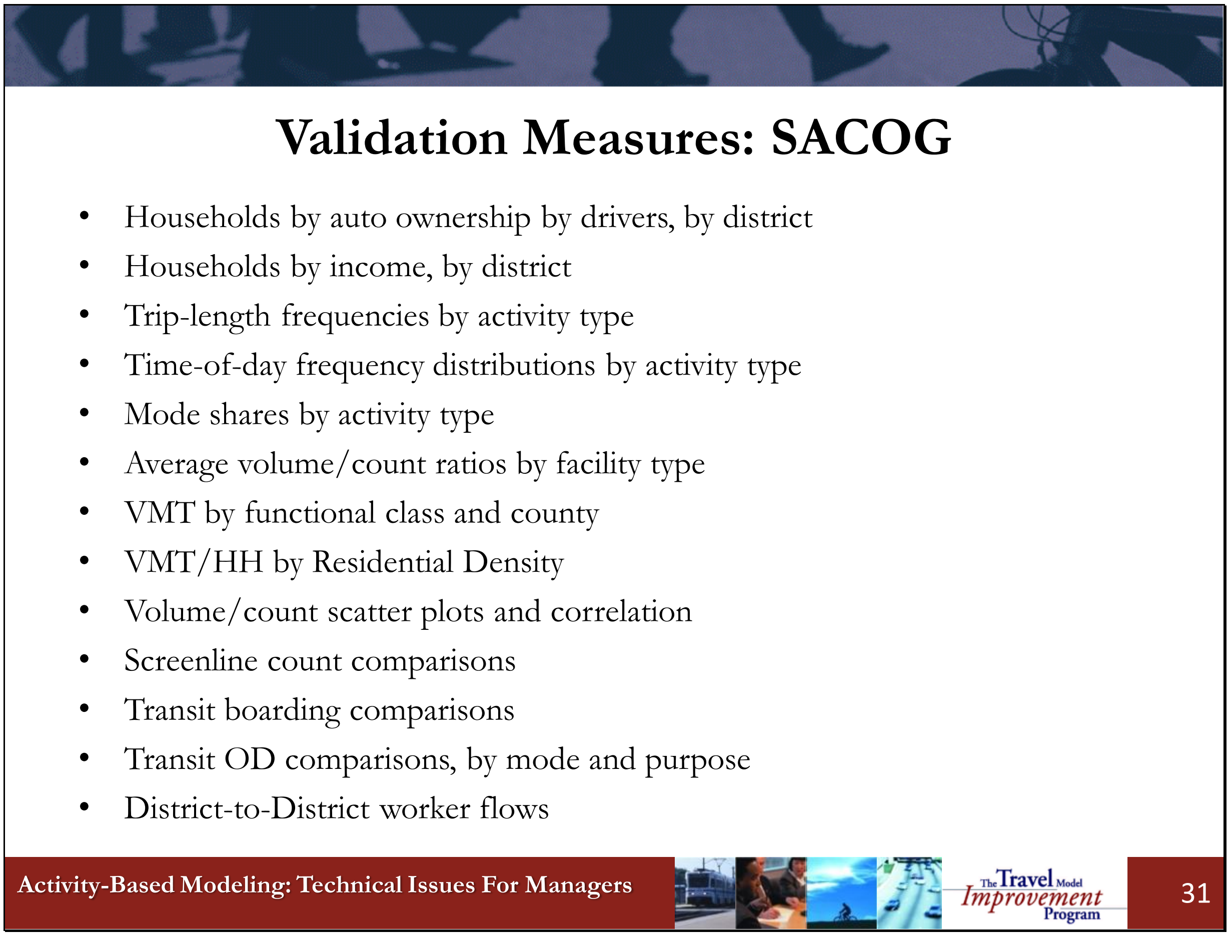
The process of model validation should be thorough. Traffic counts and transit boardings can be compared against model estimates for an independent assessment of model performance. Additionally, the model should be tested against various policies to better understand its suitability for policy\project assessment. The process of validation should inform a careful adjustment of parameters, variables, or model enhancements to better match observed travel patterns or ensure logical policy responses. Model validation does <u>not</u> mean adjusting network speeds to better match traffic volumes!



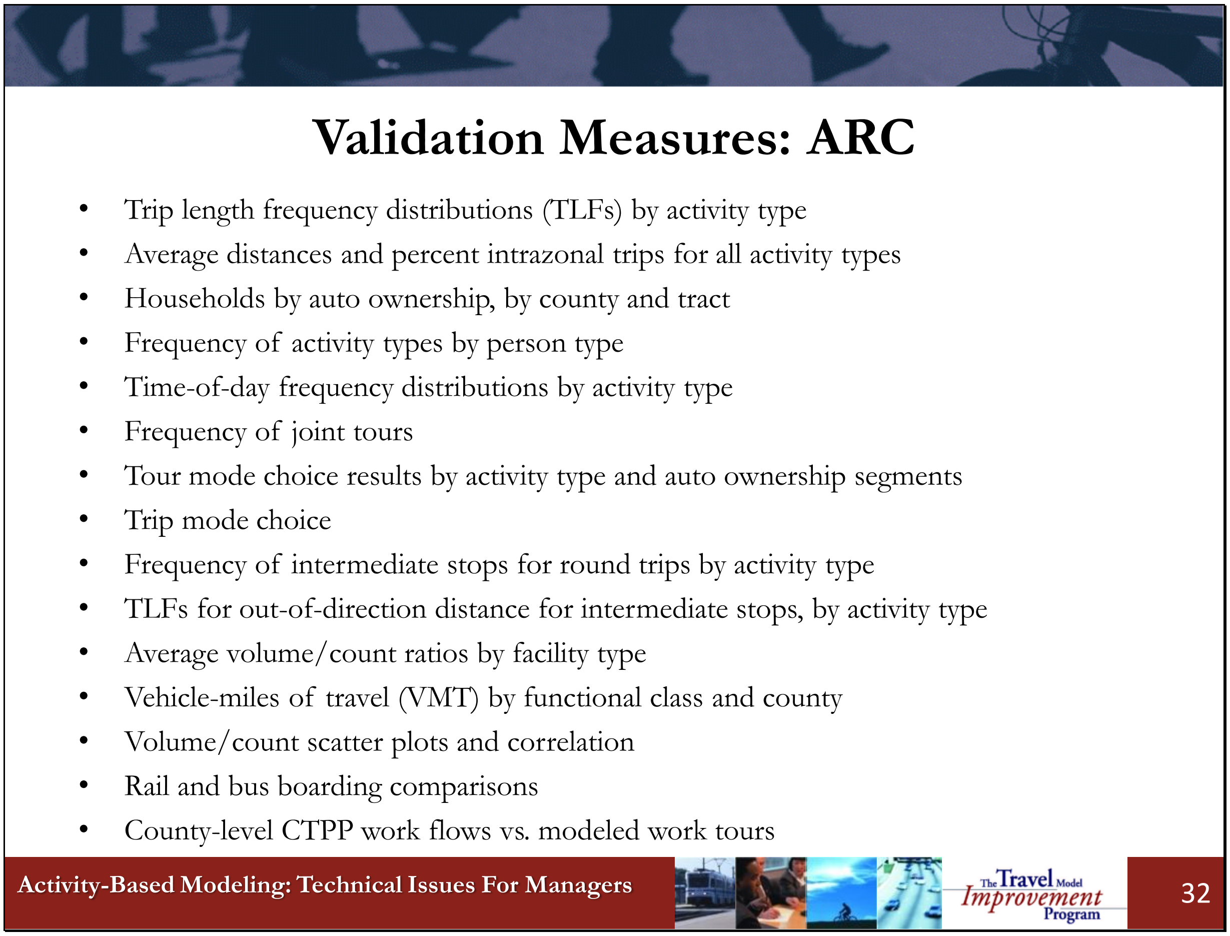
Activity-based models can be validated in the same way that trip-based models are validated. As with trip-based models, we always want to examine our observed data to make sure it is representative of actual conditions. For example, traffic counts should be from close to the same time period as each other and to the time of the household survey. Volume and speed profiles from ITS detector data should also reflect representative conditions. So, days when detectors were malfunctioning or results were affected be construction projects should not be used.

Traffic volume estimates at a link level should be compared to traffic counts. Since activity-based models typically assign to more time periods, counts by time-of-day are helpful. Transit boardings can also be compared to observed ridership at a route or route-group level. This slide shows comparisons between the Metropolitan Transportation Commission’s activity-based model traffic and transit counts to observed data for the San Francisco-Oakland Bay area. The plot on the upper left shows transit boardings and alightings for BART by station as bars, and total passengers along the BART route as a line plot.

The scatter plot compares estimated versus observed traffic volumes across all count locations on a daily basis. Similar plots, not shown, were created by time of day. Discrepancies in the estimated versus observed comparisons suggested further analysis (particularly of network coding) were necessary. In one case, shortages of volumes lead to the development of a special generator model for San Francisco airport.



This slide is a list of key calibration and validation measures that have been used by the Sacramento Council of Governments in evaluating the quality of its activity-based modeling system. These measures were actually developed when the agency was comparing its new activity-based modeling system to its trip-based modeling system. As a result, these measures a quite similar to what the agency would use when validating its trip based model.



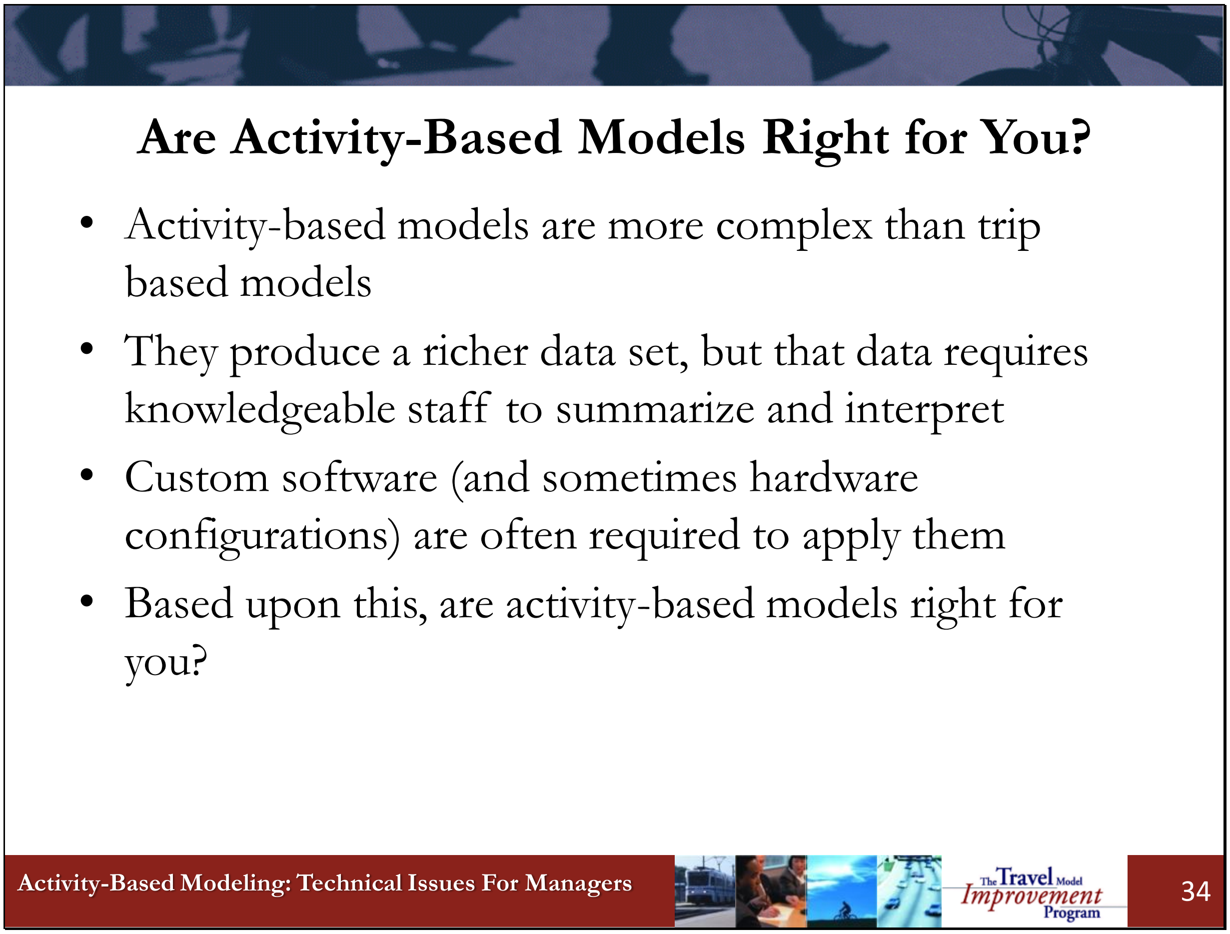
Here is another list of validation measures, developed by the Atlanta Regional Commission for its activity-based modeling system. As you can see, this list goes well beyond the usual trip-based measures to consider measures related to activities and tours.

It should be noted that for most non-work tour-based measures, the only available observed data is usually household survey results. In order to get a truly independent estimate of these tour-based measures (a source not used in estimation and calibration), it is necessary to either use a holdout sample of the current survey (which is rarely available), or to apply the model to an older household survey, combined with a back-casting exercise. If independently gathered GPS data were available, this might also be used to collect tour-based information.

Knowing that validating an activity-based model will entail a bit more detail, agencies should be prepared to budget for any extra data collection. Much of the data, however, is from sources already familiar to agencies from their trip-based model work. The extra effort is mostly in using more information from existing sources.

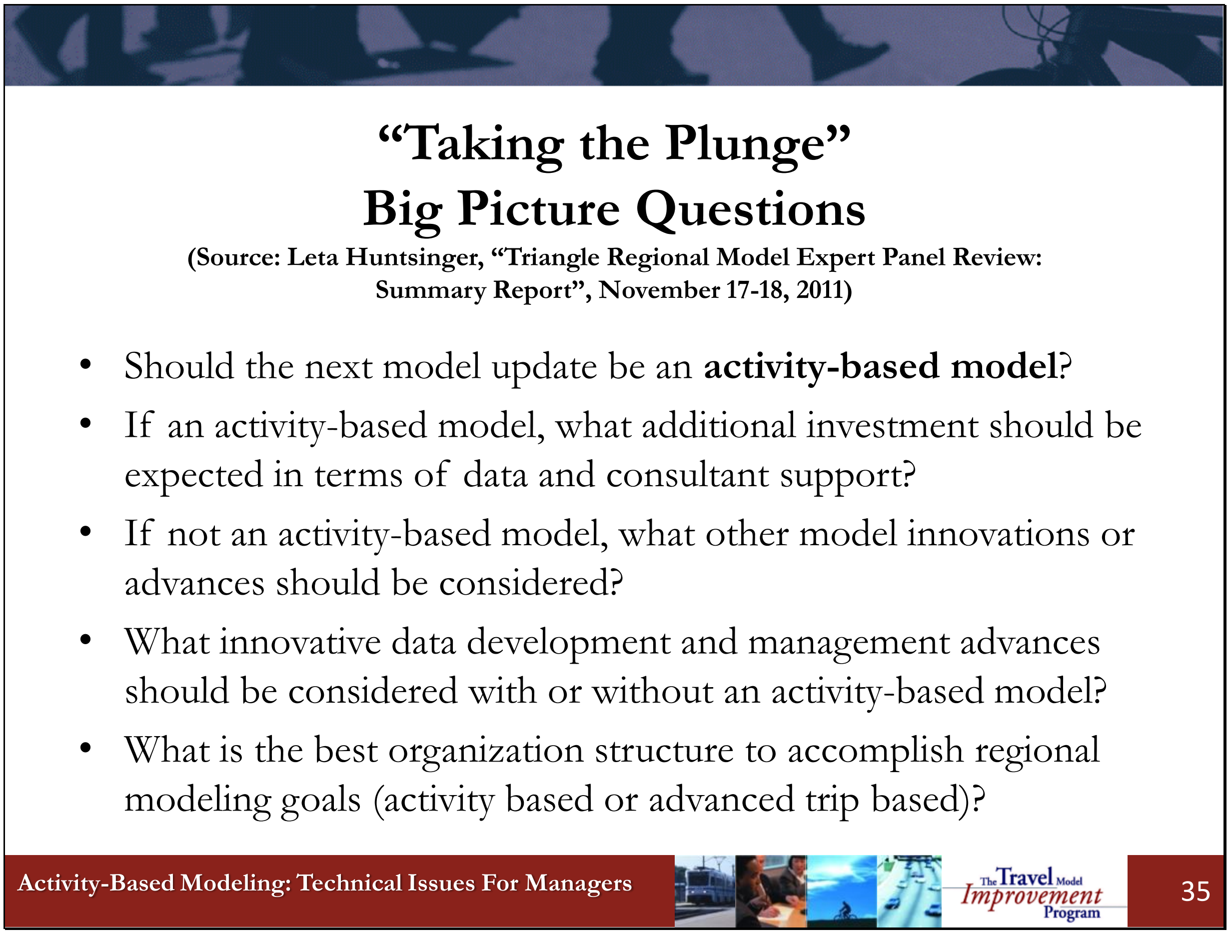
Questions and Answers

The Travel Model Improvement Program



In the second half of today's webinar, we're going to focus on the decision of whether or not an agency should develop an activity-based model and, if so, what questions do they need to answer.

Activity-based models can take advantage of a much wider range of data than can trip-based models, and are more complex as a result. They produce a richer dataset for analysis, but that data requires knowledgeable staff to summarize and interpret it effectively. Custom software, and sometimes hardware solutions that include distributed computing, can be required to apply activity-based models. Taking all of this into consideration, you may be wondering whether an activity-based model is right for your agency.



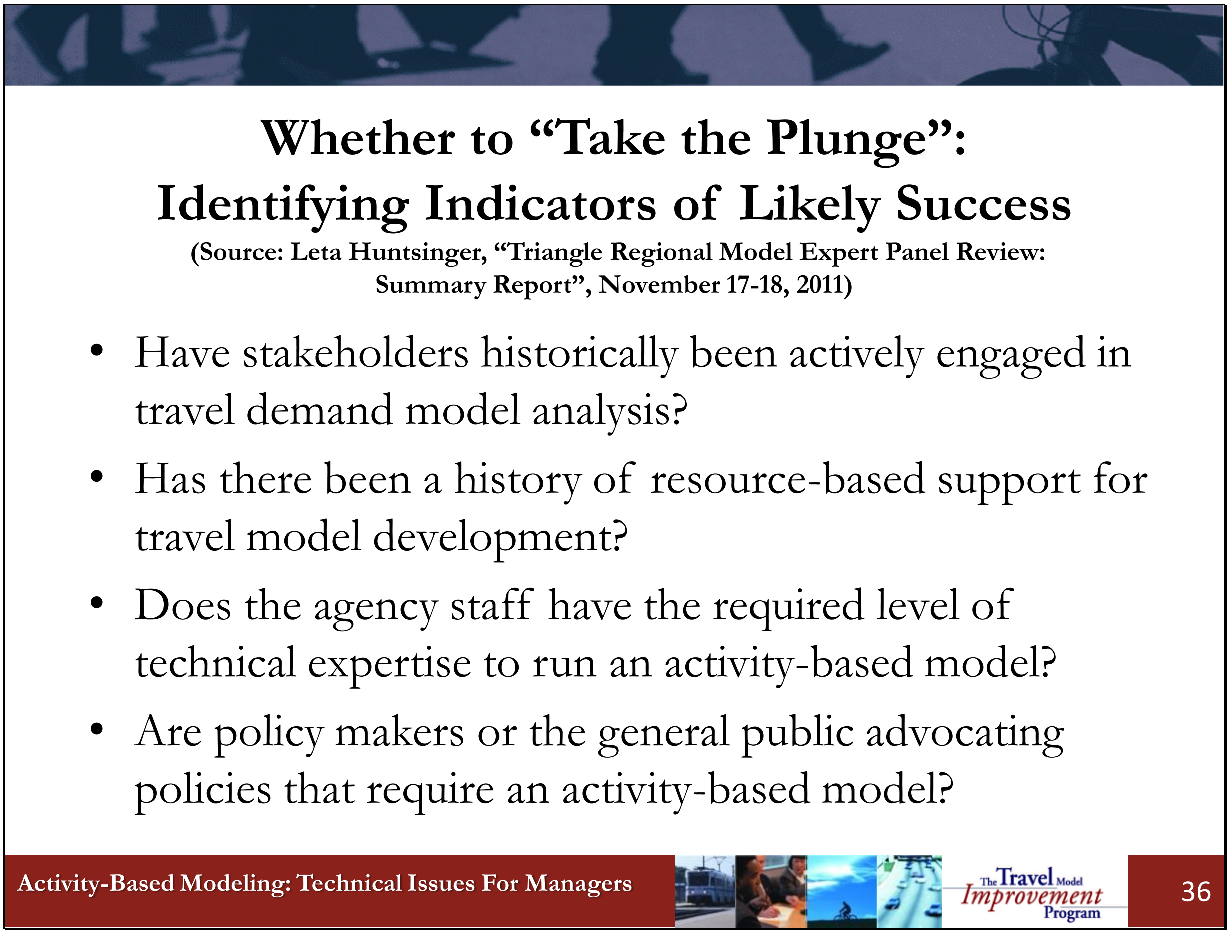
In this series, we talk a lot about the details of developing and using activity-based modeling systems. Our hope is that by sharing this information with you, that it will lead to more informed, better decisions about the development of an activity-based modeling system. It's relatively easy, however, to get caught up (if not lost) in these details. Taking a big-picture perspective, there are probably some fundamental questions that an agency should by asking when considering "taking the plunge" towards an activity-based model system.

The following list of general questions was developed by an Expert Panel for the Triangle Regional Model (North Carolina) in November 2011:

- Should the next regional model update be an activity-based model?
- If the region moves towards an activity-based model, what additional investment should be expected in terms of data and consultant support?
- If the region does not move forward with an activity-based model, what other model innovations or advances should be considered?

- What innovative data development and management advances should be considered with or without an activity-based model?
- What is the best organization structure to accomplish regional modeling goals (whether activity based or advanced trip based)?

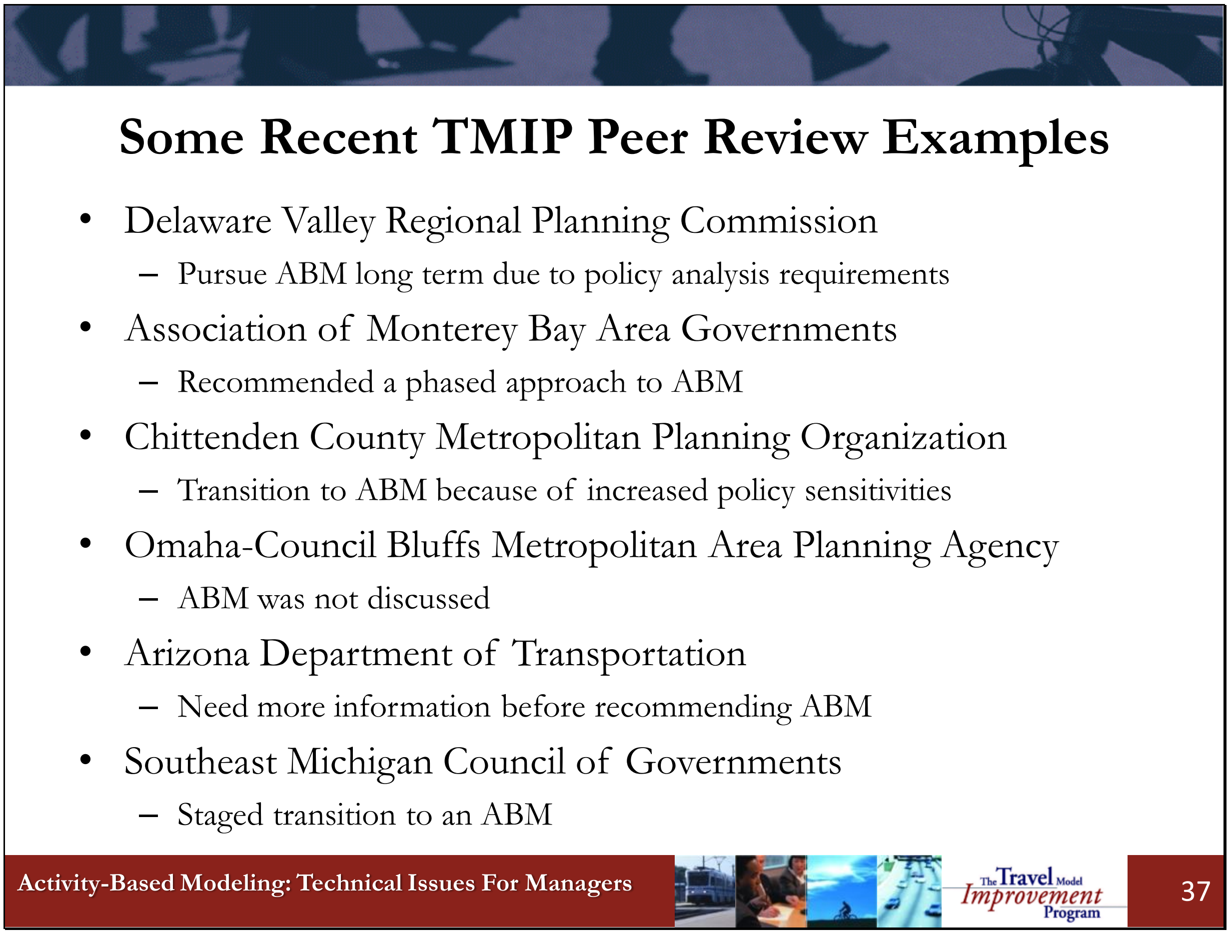
In this webinar, we can provide information that will help to answer most of these questions. Organizational structure is a tough one, though, and something extremely specific to any agency engaged in travel modeling, with its unique form of governance and relationships to stakeholders. For example, in the Triangle region (Raleigh-Durham-Chapel Hill), modeling is provided by a Service Bureau, located at North Carolina State University, and serving two MPOs, with support from NCDOT. Quite different arrangements exist in other states and regions. Organizational structure is often developed to address many other issues beyond modeling. Consequently, we will not attempt to provide answers to questions of the best organizational structure, other than to note that it is an important consideration.



We can, however, attempt to answer a number of other questions that might help an agency assess the likelihood of success in developing an activity-based modeling system. This list of questions was also derived from recommendations made by the Expert Panel in the Triangle Regional Model review.

- Have stakeholders historically been actively engaged in travel demand model analysis?
- Has there been a history of resource-based support for travel model development?
- Does the agency staff have the required level of technical expertise to run an activity-based model?
- Are policy makers or the general public advocating policies that require an activity-based model?

The answers to these first two questions should be relatively obvious to the agency asking them, based on first-hand knowledge of the relationships between an agency and its stakeholders and the history of funding for modeling programs. The answers to these last two questions might be less obvious, because they require familiarity with activity-based modeling systems, which most persons in the agency won’t have had.



TMIP conducts peer reviews around the country for planning agencies who are considering improvements to their travel demand forecasting models. These peer reviews include a series of recommendations from the peer review panel to the agency. In four of the last six peer reviews by TMIP, there was a recommendation to develop activity-based models, based primarily on the need for increased policy sensitivities that are found in activity-based models. In Arizona, the panel suggested that more information was needed before a recommendation on activity-based models could be made (perhaps this webinar series will help to fill that request!). In the Omaha region, activity-based models were not discussed or recommended. Of the four who recommended activity-based models, the size ranges from very small (Burlington and Monterey) to quite large (Philadelphia and Detroit). Philadelphia is proceeding with the development of their activity-based model and has initiated a consultant contract for this development in 2012, which will last three years. AMBAG has selected a consultant for their model update, which will remain as a 4-step model. Burlington and Detroit are more recent and have yet to proceed with their recommendations.

These examples show the range of recommendations and responses to initiating an activity-based model. Requesting a TMIP peer review can be a measured and thorough means to determine if you are ready to take the plunge. I am also going to present a list of considerations for whether an activity-based model is right for your agency.

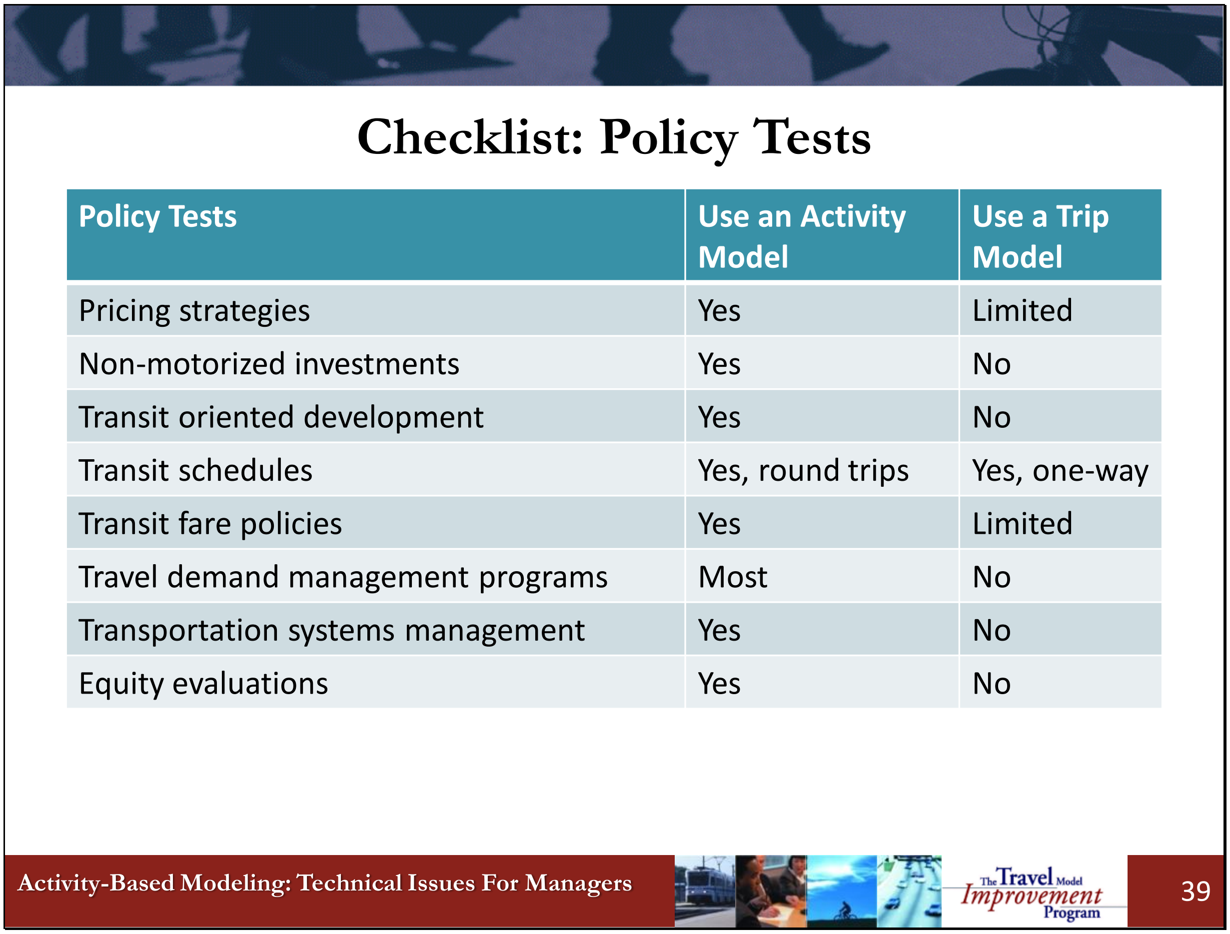
Checklist: Staff Expertise

Skill Sets	Need for Activity Model	Need for Trip-Based Model
Facility with simulation programming and outputs	Essential	Not necessary
Knowledge of discrete choice model structures, behavior, utility theory	Essential	Desirable
GIS manipulation of land use parcel data	Essential/Desirable	Desirable
SQL scripting and database manipulation	Essential/Desirable	Not necessary

We have developed a series of checklists so that you can do a self assessment. In this slide, we've listed skill sets that are particularly useful for an agency engaged in activity-based model development:

- Facility with simulation programming and outputs
- Knowledge of discrete choice model structures, behavior, utility theory
- GIS and land use coverage – this may be less important in a purely zone-based system, but essential for parcels or micro-zones.
- SQL scripting and database manipulation – this is essential if the application software is linked to an SQL database and highly desirable, even if it is not, so as to data mine model outputs.

These are skill sets that an agency may already possess, but which are much more important for activity-based modeling. While consultants may be able to supplement staff expertise, most agencies find it much more cost-effective to have as much of this expertise as possible in-house.



Checklist: Policy Tests

Policy Tests	Use an Activity Model	Use a Trip Model
Pricing strategies	Yes	Limited
Non-motorized investments	Yes	No
Transit oriented development	Yes	No
Transit schedules	Yes, round trips	Yes, one-way
Transit fare policies	Yes	Limited
Travel demand management programs	Most	No
Transportation systems management	Yes	No
Equity evaluations	Yes	No

The other question we want to answer is whether policy makers or the public are advocating policies that require an activity-based modeling approach. In this first session in this series, the Executive Perspective, we provided several examples of policies that were best addressed through an activity-based approach. In the second session, Institutional Issues for Managers, we described how the additional information provided by activity-based models differs from that of trip-based models and makes activity-based models more appropriate for answering certain types of questions. This list shown here summarizes these policies contexts:

- Pricing strategies
- Non-motorized investments
- Transit oriented development
- Transit schedules
- Transit fare policies
- Travel demand management programs
- Transportation systems management

- Equity evaluations

Checklist: Impacts of Interest

Impacts of Interest	Use an Activity Model	Use a Trip Model
Induced Demand	Yes	No
Emissions for a Household	Yes	No
Peak spreading	Yes	No
Start/stop emissions by time of day	Yes	No
Vehicle type and choice	Yes	No
Equity by income group	Yes	No

To drill down a bit further, we can imagine a number of performance measures that modelers would like to be able to predict. In the checklist shown here, we've listed several measures that activity-based models are designed to predict and for which trip-based models are ill-suited. These are somewhat independent of specific policy initiatives and are relevant to common agency work programs, such as long range planning, conformity analysis, and environmental justice.

You've Decided to Take the Plunge—Now What? Assembling Resources

Resource Considerations	Activity Model	Trip Model
Staff resources	Training needed	Trained
Executive support	Preferred	Assumed
Quality data	Higher expectations	Reasonable
Cost	Set approach to resources available	Updates required
Approach	Transfer or build	Update
Innovation	Lots of possibilities	More difficult
Development	Consultant support or in-house or hybrid	
Funding	Work program budget or external grants	Work program

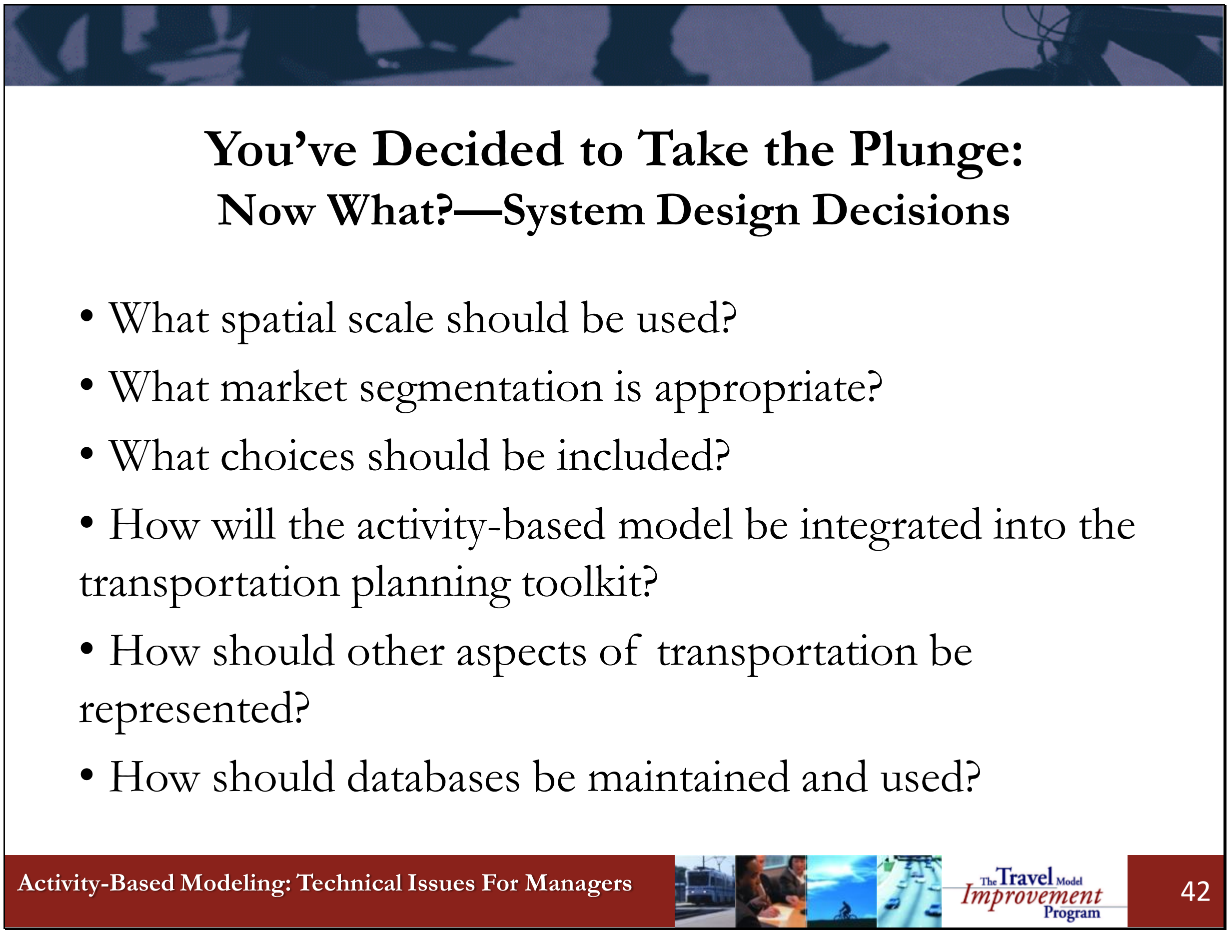
Activity-Based Modeling: Technical Issues For Managers

The Travel Model Improvement Program

41

In the second session, Institutional Issues for Managers, we discussed resource issues relevant to the development of an activity-based travel demand model. Here, we have listed some of the resources considerations that an agency should consider when undertaking an activity-based model development project. These include:

- Staff resources
- Executive support
- Quality data
- Cost
- Approach
- Innovation
- Development
- Funding



If, and when, you may decide to proceed with developing an activity-based model, there are some important system design decisions to consider. These include details about the activity-based models, such as what spatial scale should be used, what type of market segmentation you want to include, what travel behavior choices you want to represent.

These also include details about how the activity-based model will fit into the broader travel modeling system, such as how will it be integrated with other transportation planning software, how will other aspects of transportation (freight, long distance travel, visitor travel) be represented, and how data used in the model will be maintained over time.

Page 43

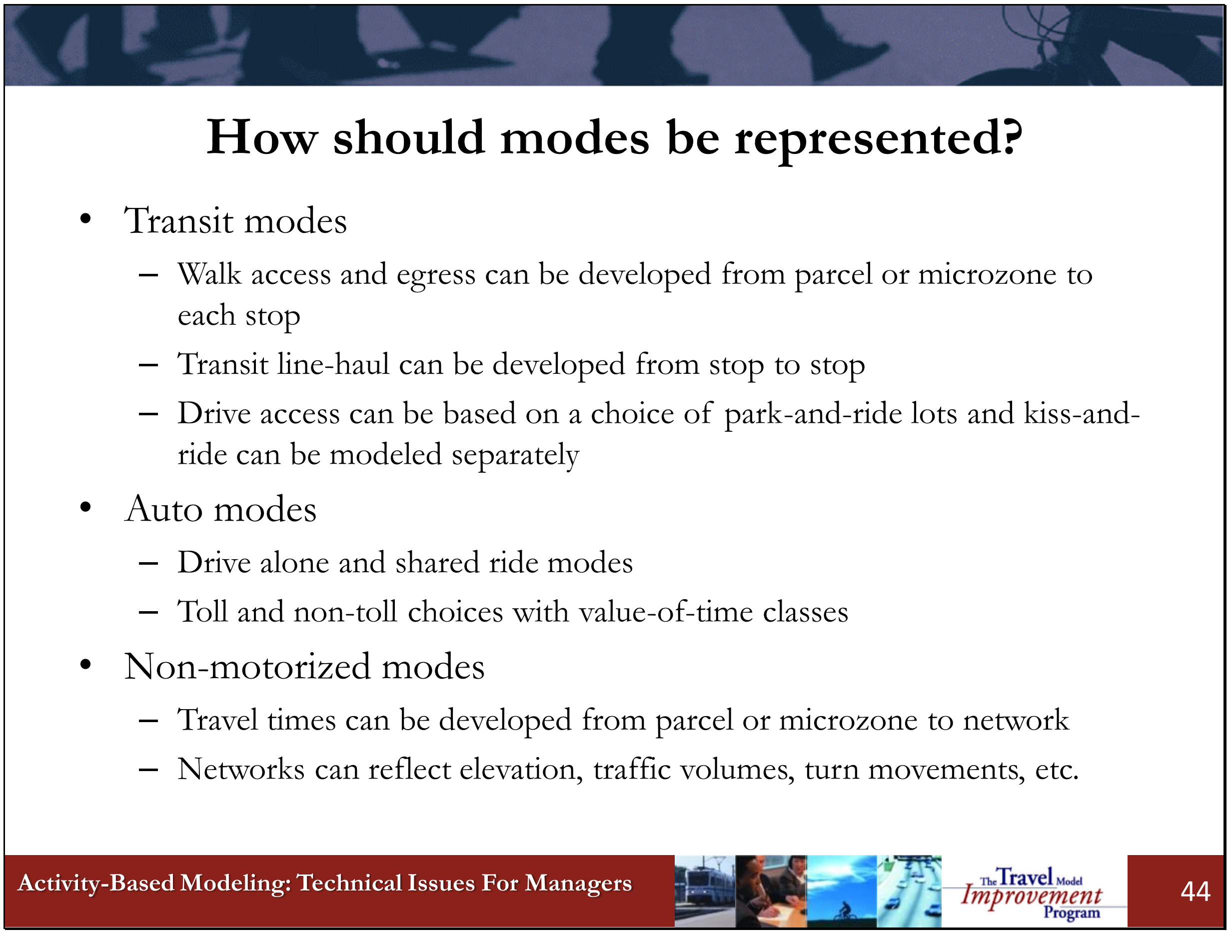
What spatial scale should be used?

	Zones	Microzones	Parcels
Spatial Detail	Heterogeneous Land Uses	Primarily Homogenous Land Uses	Individual Land Use
Data Sources	Existing travel demand model	Census data for households; LEHD or local sources for employment	Tax assessor data combined with Census and local data
Travel times and costs	Aggregate	Use microzones for transit access and non-motorized modes	Use parcels for transit access and non-motorized modes
Measures of Attractiveness	Aggregate	Partially disaggregate	Disaggregate
Level of Effort	Least	Can be automated from available data	Most

Activity-Based Modeling: Technical Issues For Managers — The Travel Model Improvement Program — 43

Determining the best spatial scale for an activity-based model is a critical design decision, which affects most aspects of the modeling system. The spatial scale is not a decision unique to travel modeling either, since land use modeling, economic modeling, performance measurement, air quality, operations are all affected by the spatial scale chosen and may indeed have different spatial scales. One of the bigger benefits of activity-based models is the additional spatial detail and if traditional traffic analysis zones are used, then these benefits are not achieved. To date, there are three spatial scales that could be used – zones, micro-zones, and parcels. Micro-zones are approximately the size of census blocks and may be about 10:1 scale with zones. Parcels are the smallest scale possible and may be several hundred or a thousand to 1 scale.

Data sources will vary depending on which spatial scale you choose, but each level has national or state data sources that can be used. The spatial scale selected can vary depending on what function in the model you are evaluating. For example, travel times and costs may be developed at a zonal scale even though micro-zones or parcels are used for land use data. Some activity-based models have modified the network skimming process to take advantage of micro-zone or parcel level detail for walk to transit, walk or bike modes at the origin and destination end of the trips and use stop to stop impedances for the line-haul portion of the transit trip. This represents a hybrid of spatial scales to take advantage of additional detail without bogging down the process too much. The level of effort is dependent on the spatial scale chosen and the additional effort involved in developing impedance measures if you use a hybrid approach.

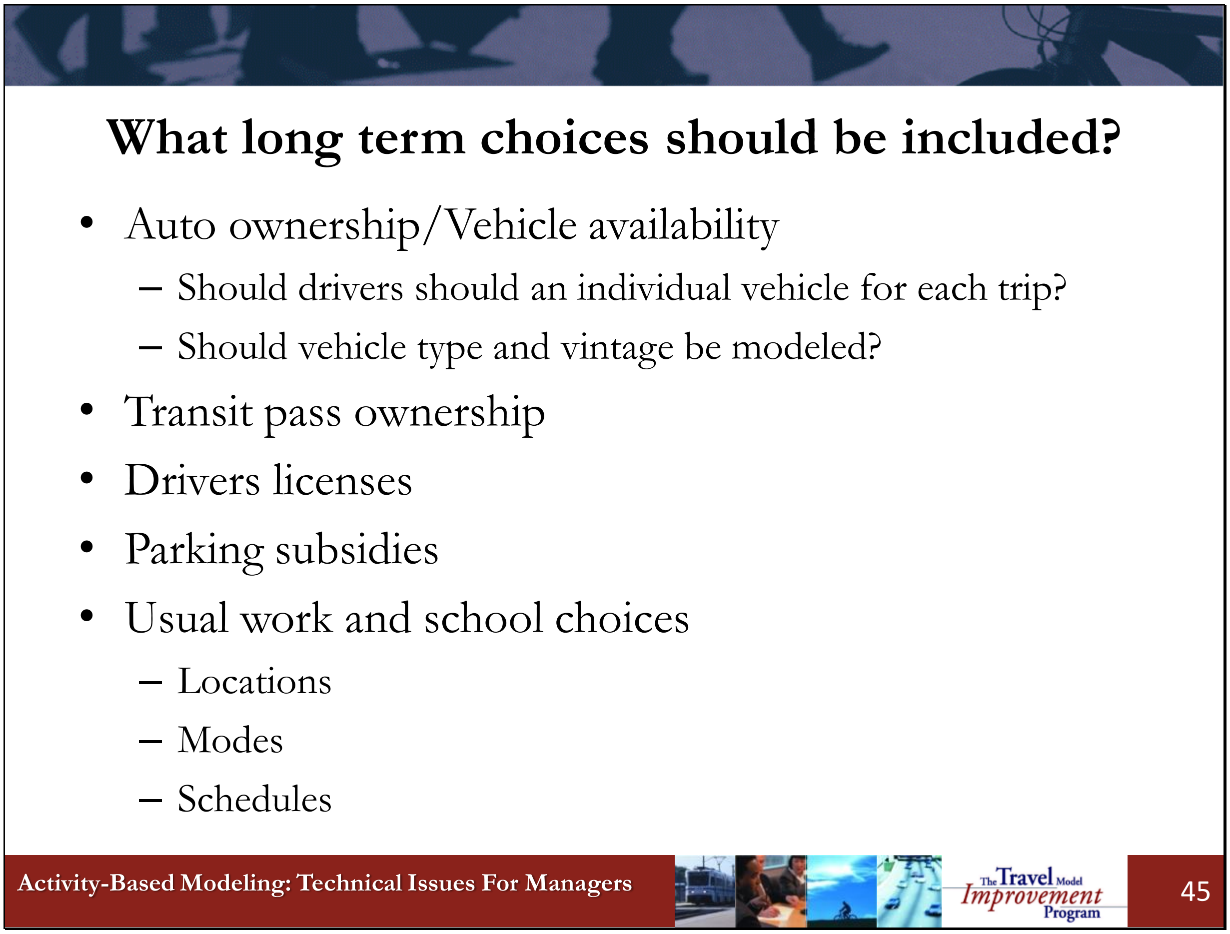


Another design question for activity-based models is which modes are represented, and at what spatial scale. The spatial scale is particularly important for walk, bike and walk to transit modes, since these are short trips and significantly affected by small inaccuracies in travel time or distance caused by zonal centroids or networks that are missing local streets.

It is preferable to develop transit access and egress from parcels or micro-zones to the transit stops, but impractical to develop parcel to parcel or micro-zone to micro-zone time matrices, so hybrid approach to develop transit times as a combination of stop to stop and parcel or micro-zone to stop works well. Drive access to transit can be separated according to park-and-ride and kiss-and-ride trips in order to account for parking capacity issues at the parking lots.

Auto modes are commonly separated by occupancy to represent drive alone and shared ride modes. Additionally, many activity-based models that are used have further segmented auto trips into toll and non-toll choices and different value of time classes to account for characteristics of travel that may lead someone to pay (or not pay) a toll for any particular trip.

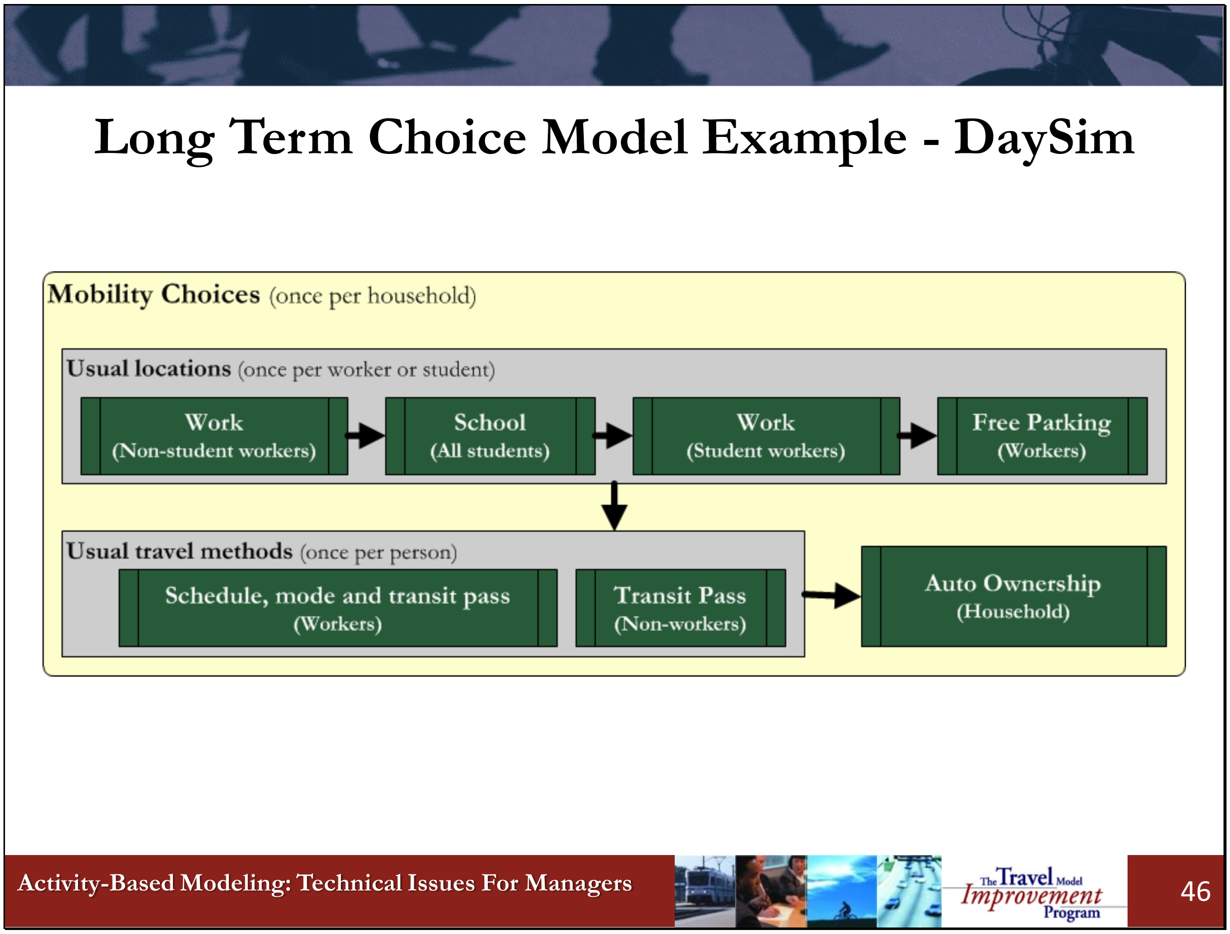
Non-motorized modes can benefit from the additional spatial detail of parcels or micro-zones. In many cases, highway networks have been supplemented with pedestrian or bike facilities. In addition, non-motorized networks have added characteristics, such as elevation, traffic volumes, and turn movements that affect pedestrian and bicycle movements.



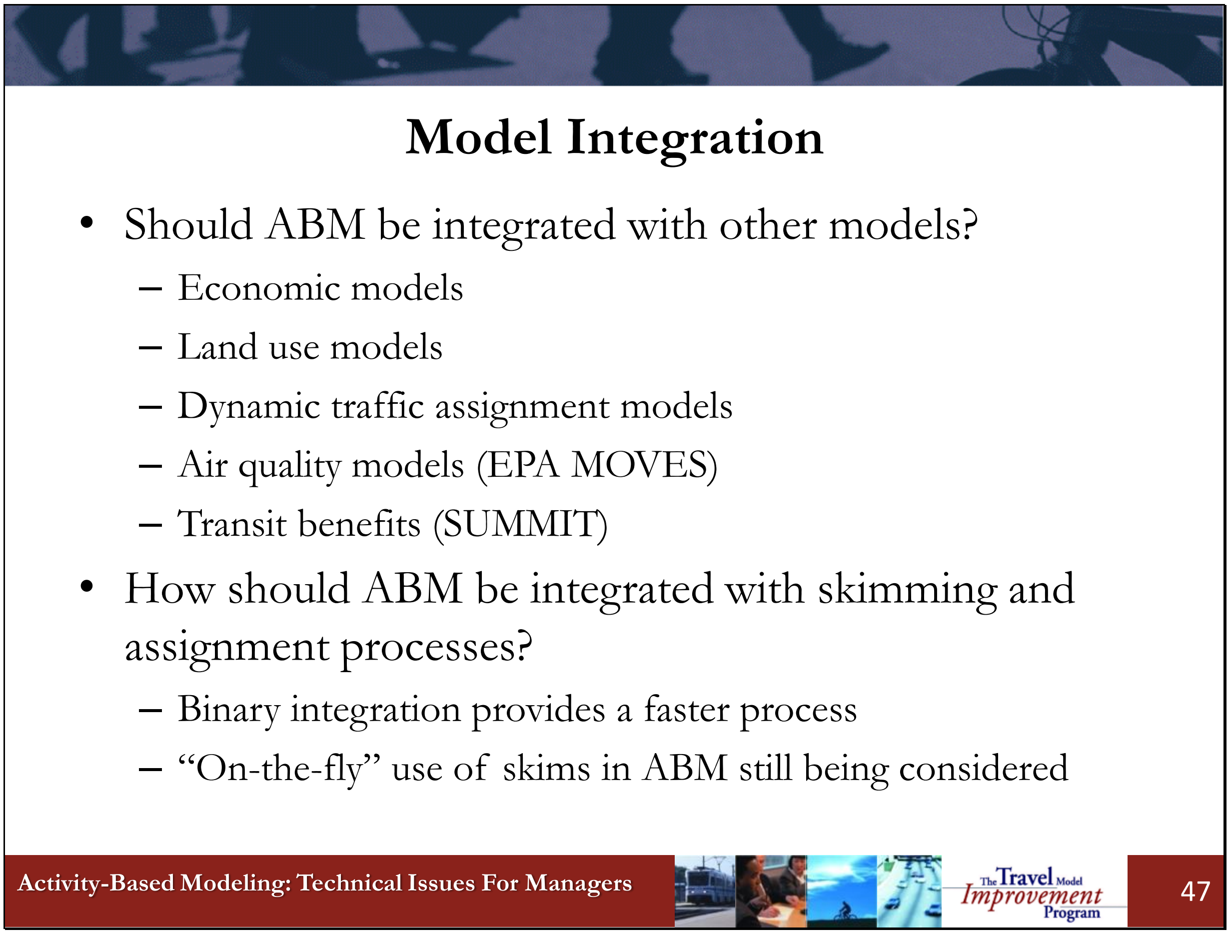
The earliest activity-based models had auto ownership and work location models included in them, but more recent models have benefited from additional long term choices that can be used in subsequent short term travel choices. Auto ownership models are also expanding to include the choice of a specific vehicle for each trip and to identify the vehicle type and year for subsequent use in air quality modeling. Travelers may choose the more energy efficient vehicle for longer distance trips or the larger vehicle for trips with more passengers and air quality model results will be affected by these choices.

Transit passes and drivers licenses are strong indicators for someone to choose transit or auto modes and so are useful choices to include in the mode choice model. In addition, if a person owns a transit pass, then the price of subsequent transit trips is zero. Parking subsidies also affects the cost of travel and the likelihood that someone will choose to drive. Parking cost is already included in most activity-based models and recognizing who bears the cost is important to include as well.

Work and school location choice models are typically included in activity-based models. The may include the choice of usual work and school choices first, and then whether a traveler will go to the usual work or school location on the travel day. Some people will stay home or may go to an outside meeting. The usual work and school location can affect mode choices and schedules for daily routines, but on some days the person may go to a different location.

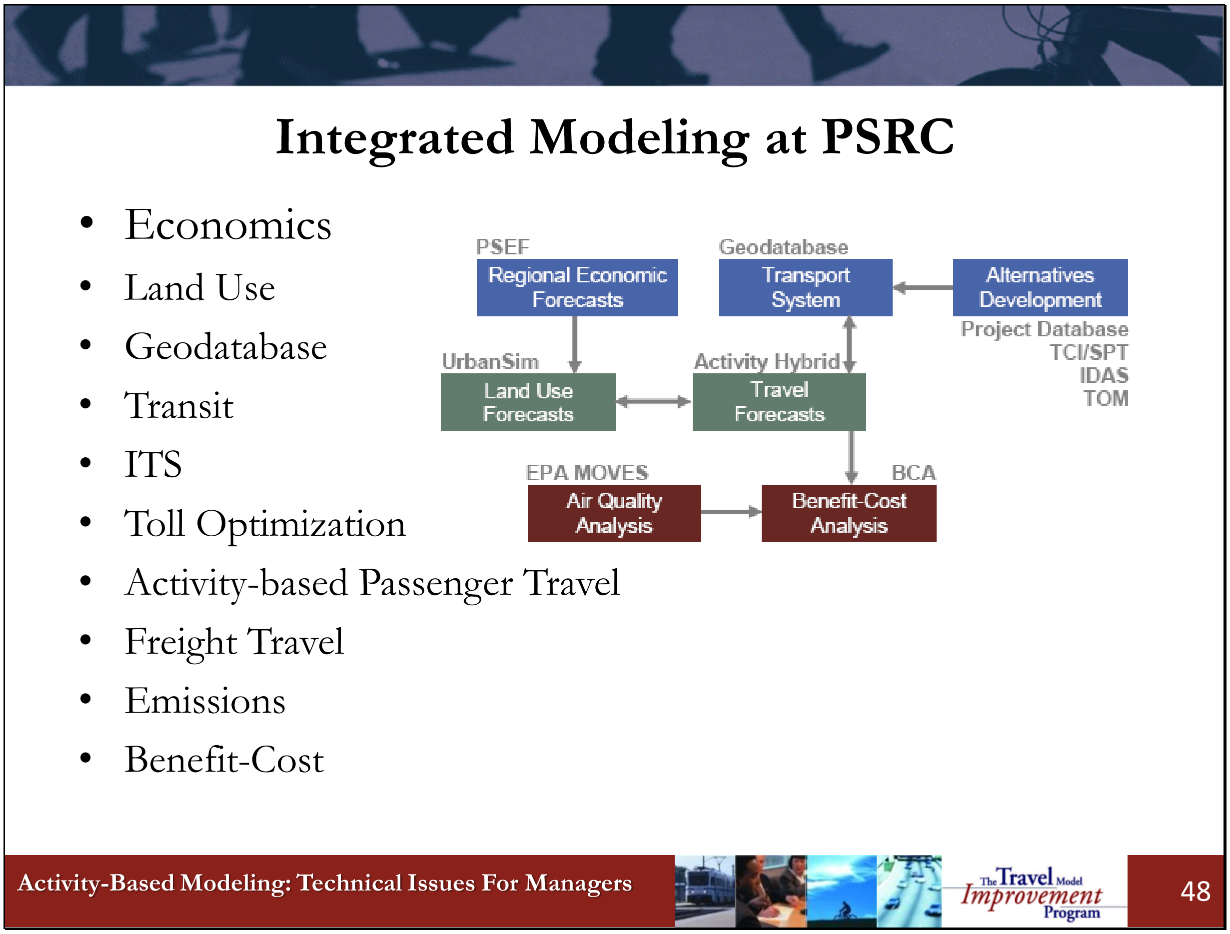


This is an example of a model (DaySim) that includes all of these additional long term choices. In this example, usual work and school locations are estimated first, then work locations for student workers, then free parking is identified for all workers. Once these initial choices are made, usual schedule, mode and transit passes are determined for workers, and transit passes are determined for non-workers. Finally, auto ownership for each household is estimated. All of these choices are retained as characteristics of the household and persons for further use in downstream models.

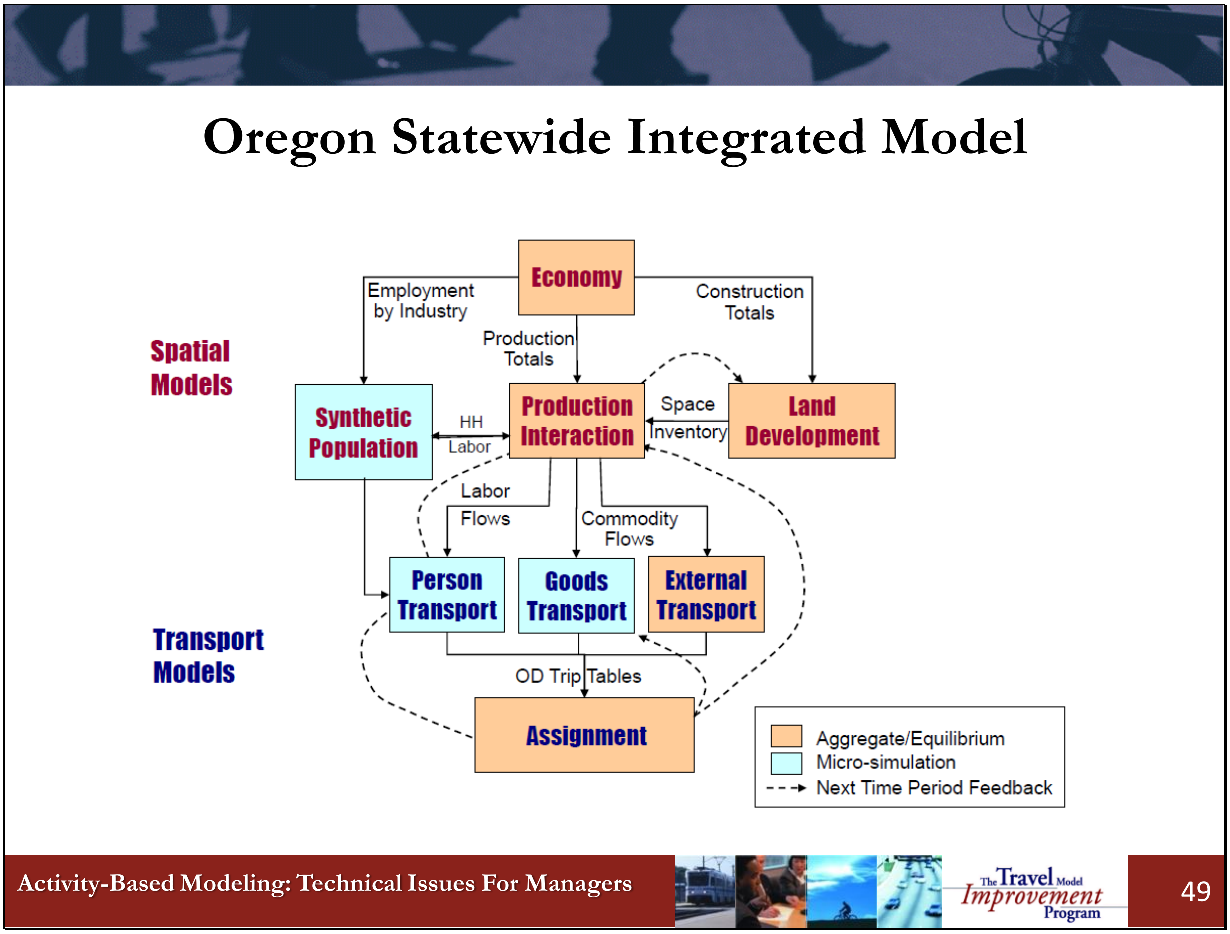


Activity-based travel demand models do not operate in a vacuum, nor do they answer all the transportation questions on their own, so integration of activity-based models is expected. The design question is which type of integration is the most useful. Currently, activity-based models have been integrated with land use models, economic models, dynamic traffic assignment models, air quality models like MOVES, and transit benefit models like SUMMIT. These are the most common models to integrate with an activity-based model, but there may be other types as well.

Activity-based models are typically also integrated with transportation planning software (like Cube, TransCAD, EMME, or VISUM) for developing skims and running assignments. This process involves running the path-building or assignment programs and then reading the matrices of travel times and costs for use in the activity-based model. The activity-based model, in turn, will produce new estimates of travel, which are aggregated to zones (or micro-zones) and assigned to highway and transit networks for assignment. Integration that bypasses input and output of these files and reads files directly in their native format are faster. In addition, reading skims “on-the-fly” when you need them in the modeling process, rather than reading and storing the full matrix, are still being considered to improve the integration of these systems.



One example of an integrated system is at the Puget Sound Regional Council in Seattle. PSRC has integrated a macro-economic model, a land use model, and an activity-based model. In addition, they have developed a geo-database to store and deliver all the data need for these models and a series of tools to build alternatives, including a transit service planning tool, an ITS evaluation tool and a toll optimization model. Once all the models are run, they have linked the results with a benefit-cost analysis tool and the EPA MOVES model to produce performance measures for all aspects of the system.

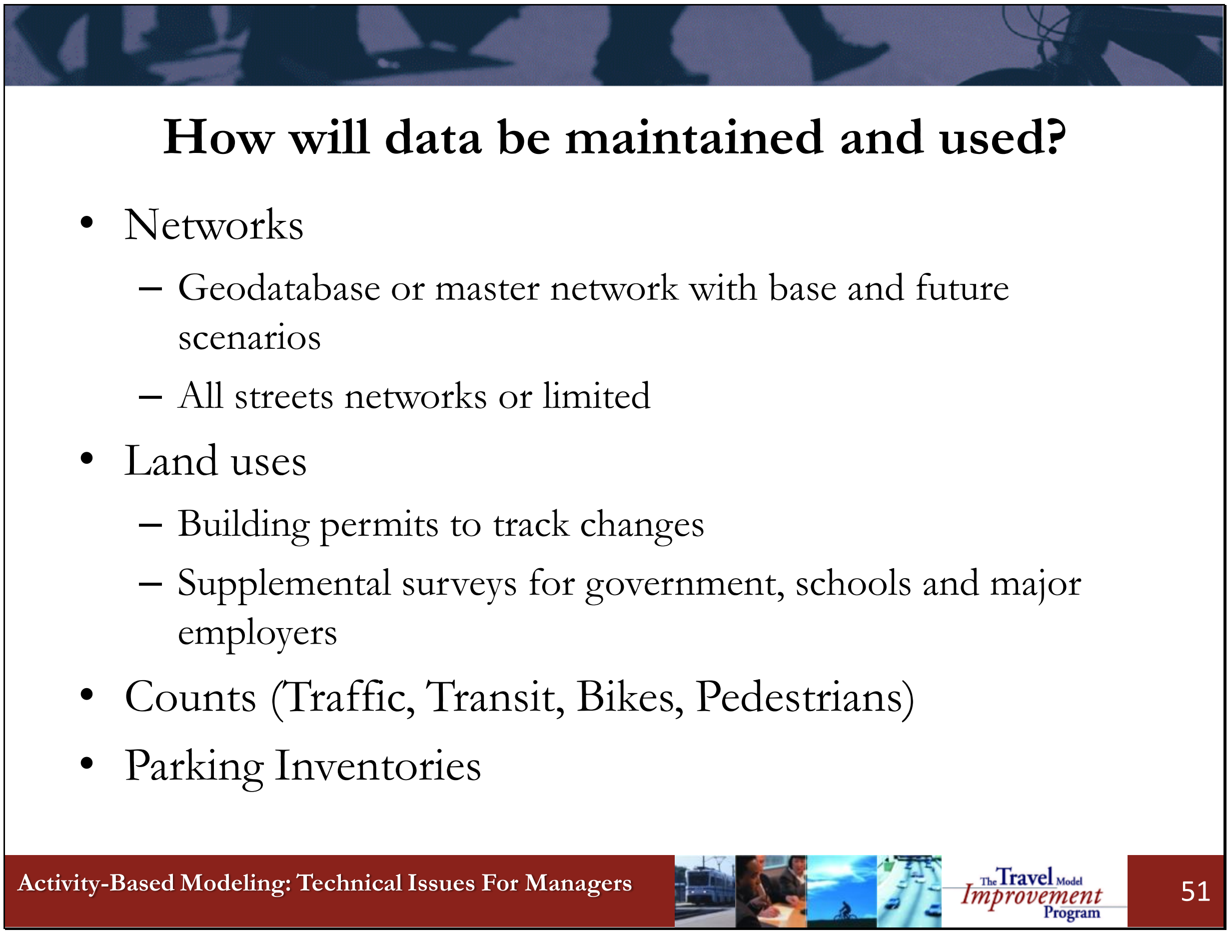


Another example of an integrated modeling system is the Oregon Department of Transportation. This model integrates the macroeconomic model with several land use models for production and consumption of goods and services, with the travel demand models for persons, goods, and external travel. Some of these models operate at a micro-simulation level and some operate at an aggregate level as shown in this diagram.

How will other travel be estimated?

Auxiliary Travel	Type	Method
Long distance	Airport	Airport models, long distance models or special generators
	External	Long distance models or externals
Non-resident	Visitor	Visitor models, long distance models or special generators
Commercial vehicles	Trucks, freight and goods	Commodity flow and vehicle touring models
	Taxis, shuttles, rentals	Vehicle touring or simple 4-step
	Service	Vehicle touring models or simple 4-step

Activity-based models typically produce resident passenger travel within a region of interest. Total travel in the region would need to include long distance travel, non-resident travel and commercial vehicles for the travel demand to be comprehensive. There are many different methods that can be employed for these types of trips, and they vary widely based on level of effort, accuracy needed, and the types of policies that may affect these trips. Many activity-based models are integrated with existing methods to develop these auxiliary trips, but future consideration can be given to improving the methods adopted for these trips.

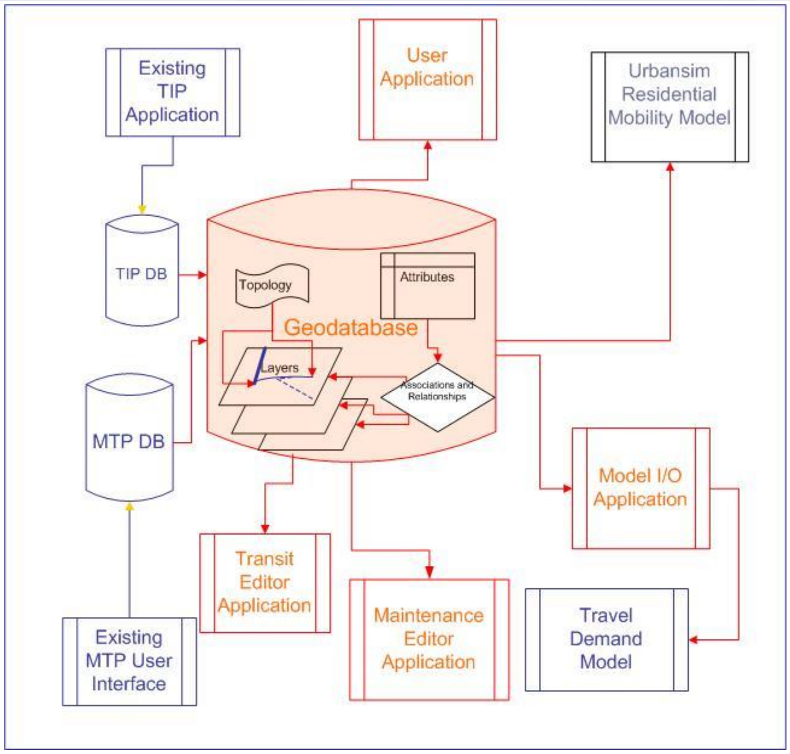


Data maintenance and usage, as any modeler knows, can be a time consuming process. Recent years have seen many improvements in the methods and databases used to maintain and update data for use in travel demand forecasting processes. The maintenance of networks involves storing information about possible future projects and combining the lists of projects into alternatives that can be used to generate a future network. Geo-databases and master networks have been developed to achieve consistency and accuracy in developing and maintaining networks. These geographic networks often have more detail than was used in travel demand networks, but can be brought into the process as more detail is added on the land use side.

Land use data maintenance systems involve tracking building permits to see what development has occurred against the forecasts. These can be maintained by local governments in charge of approving building permits. In addition, it is often necessary to conduct supplemental surveys of government and education employers, as well as major employers with multiple locations, to ensure accuracy of the employment data.

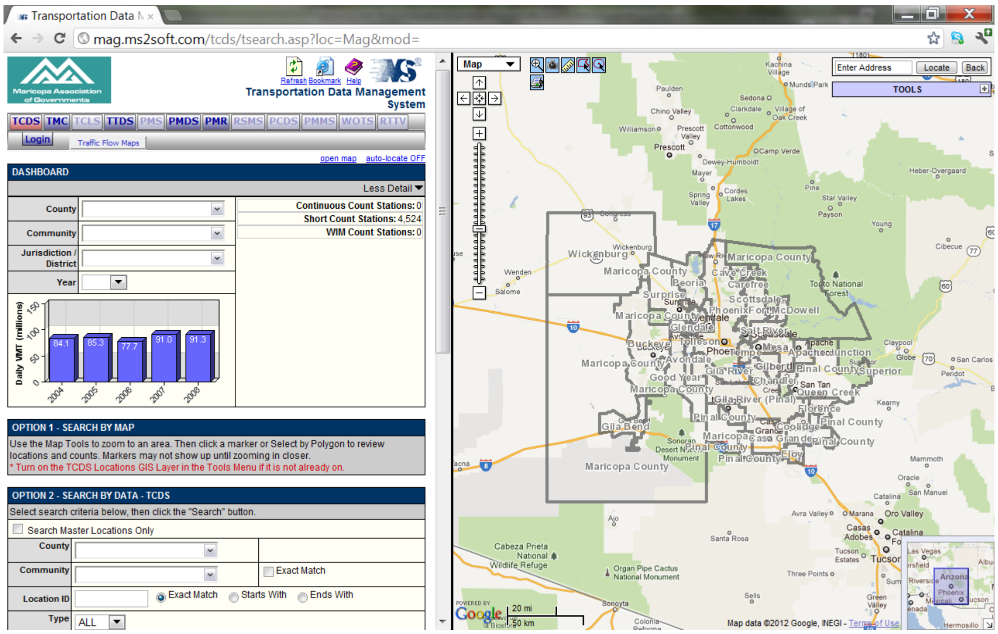
Other data that needs to be maintained for use in activity-based modeling are traffic and transit counts, as well as bike and pedestrian counts if these are collected. Some regions also conduct parking inventories, which are quite useful for parking choices in the activity-based model.

Geo-database at Puget Sound Regional Council

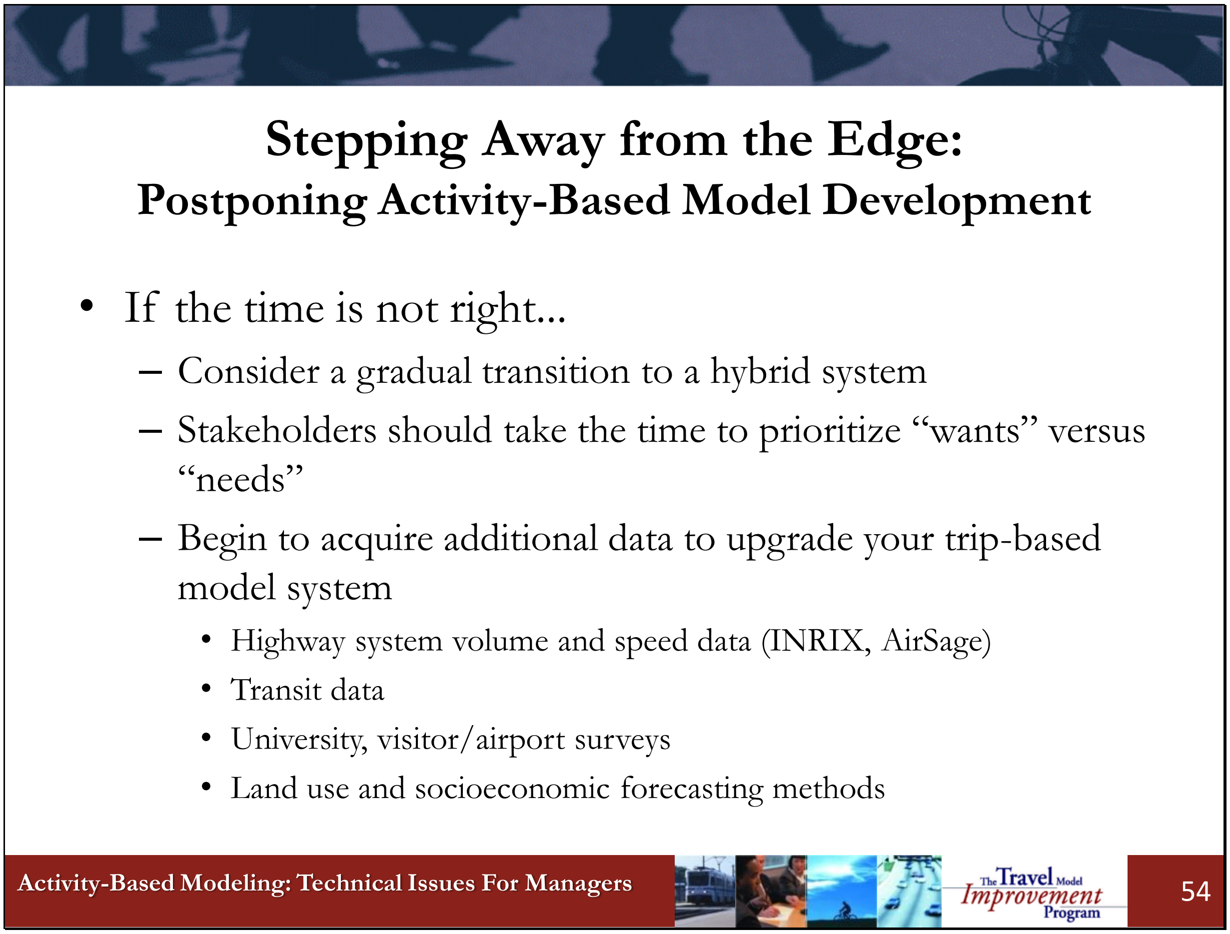


One example of a geo-database is one that was developed by PSRC for integration of model networks with the Metropolitan Transportation Plan database and the Transportation Improvement Plan database. This was built and integrated with editing tools and the ability to output networks to the travel demand forecasting model. The system was designed so that planners and non-modelers could code and maintain projects that could then be used directly in the modeling of alternatives. The geo-database was initially developed to maintain networks and then was expanded to include land use data.

Transportation Data Management System at Maricopa Association of Governments

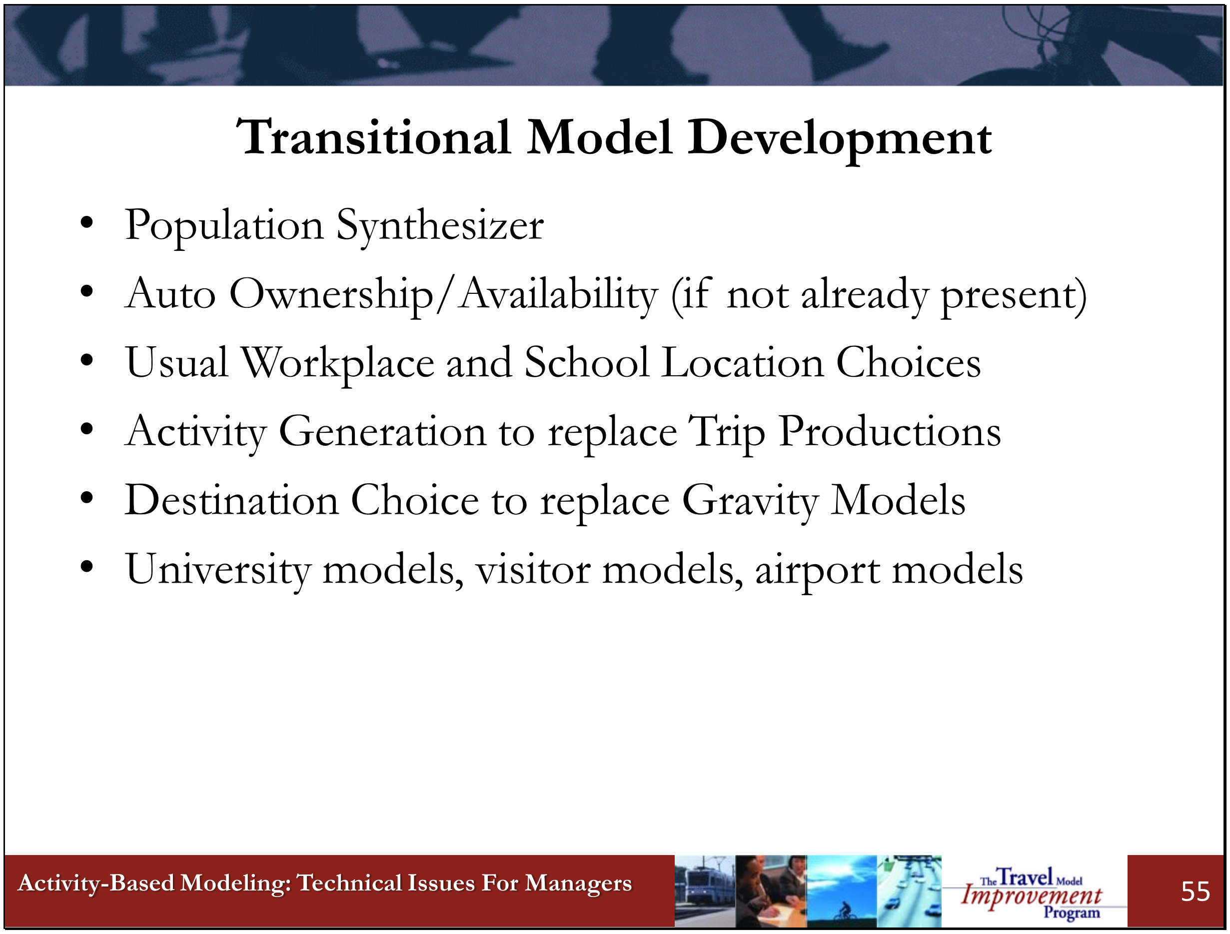


Another example of a maintenance database is one developed by Maricopa Association of Governments for transportation data. This system stores and maps traffic counts and turn movement counts, travel times, and other performance measures for use by member agencies. Data can be mapped or reported based on a series of filters and queries.



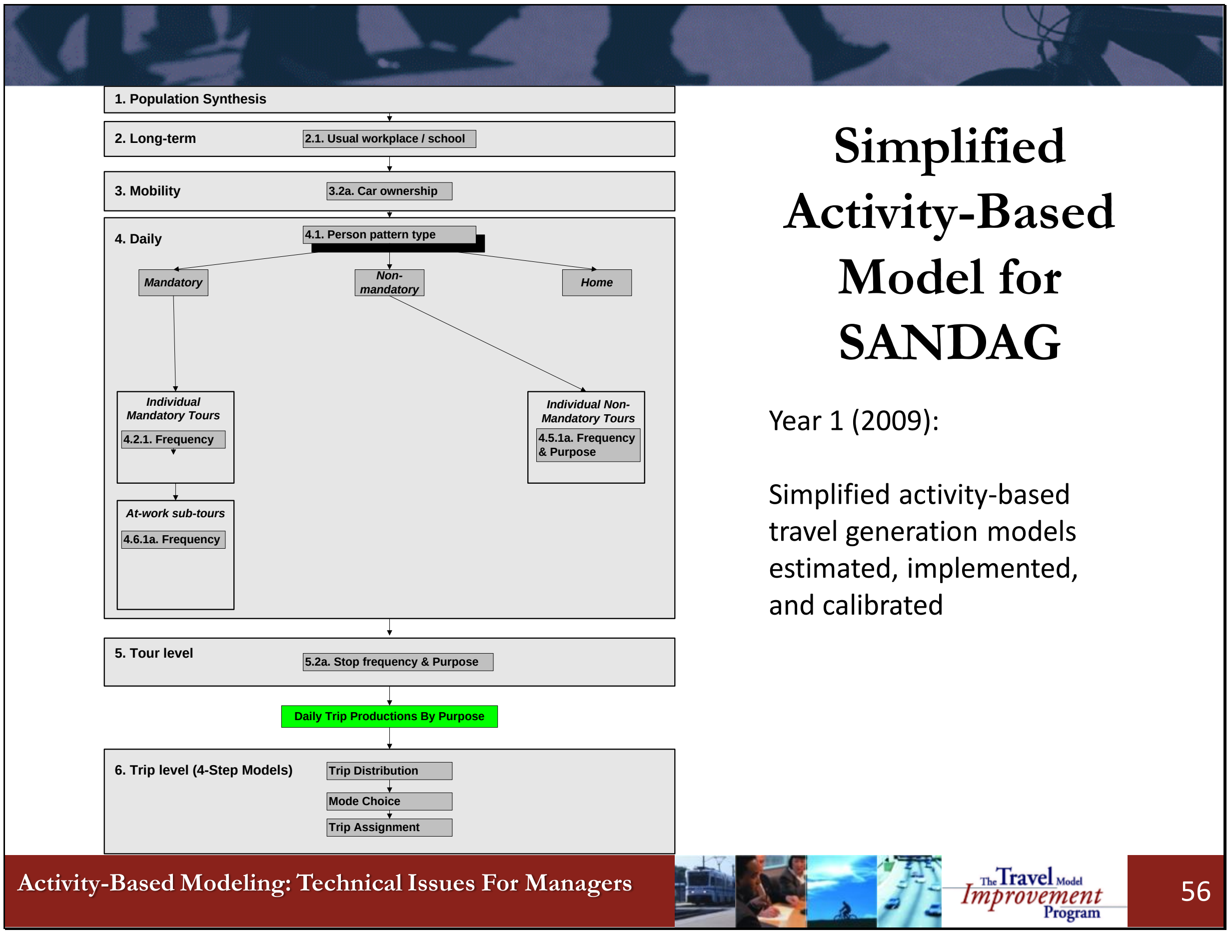
You may decide that the answers to the questions we have been asking in this second part of the session lead to the conclusion that an activity-based modeling system might be nice to have, but perhaps not completely necessary. Nevertheless, your agency might be interested in enhancing its modeling capabilities. In doing so, there are things that you can do that may better position your agency to in the future. For example:

- Consider a gradual transition to a hybrid system.
- Encourage stakeholders to prioritize "wants" versus "needs"
- Begin to acquire additional data to upgrade your trip-based model system

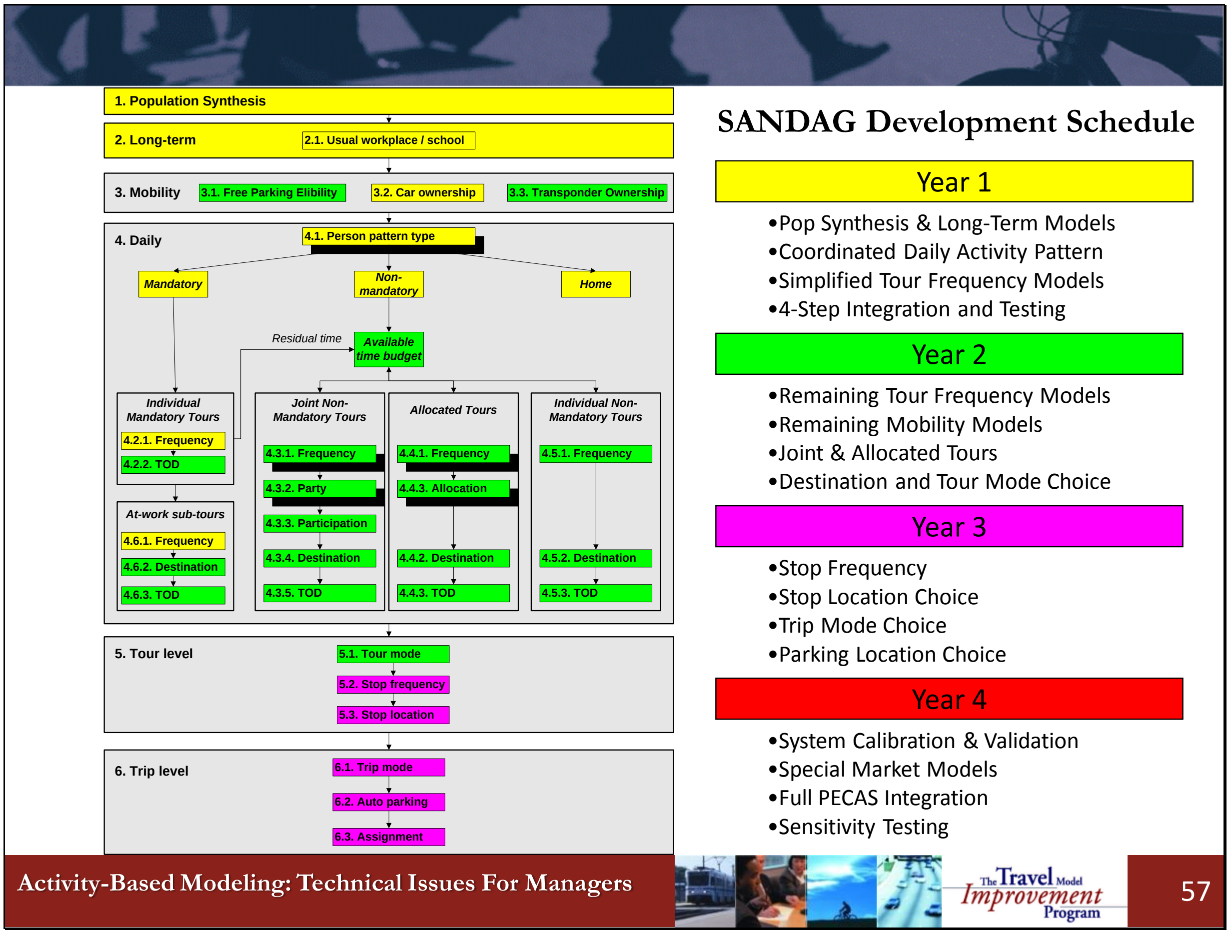


In addition, there are a number of model components that can be developed for a hybrid model system. Activity-based model components that can take the place of trip-based model components include:

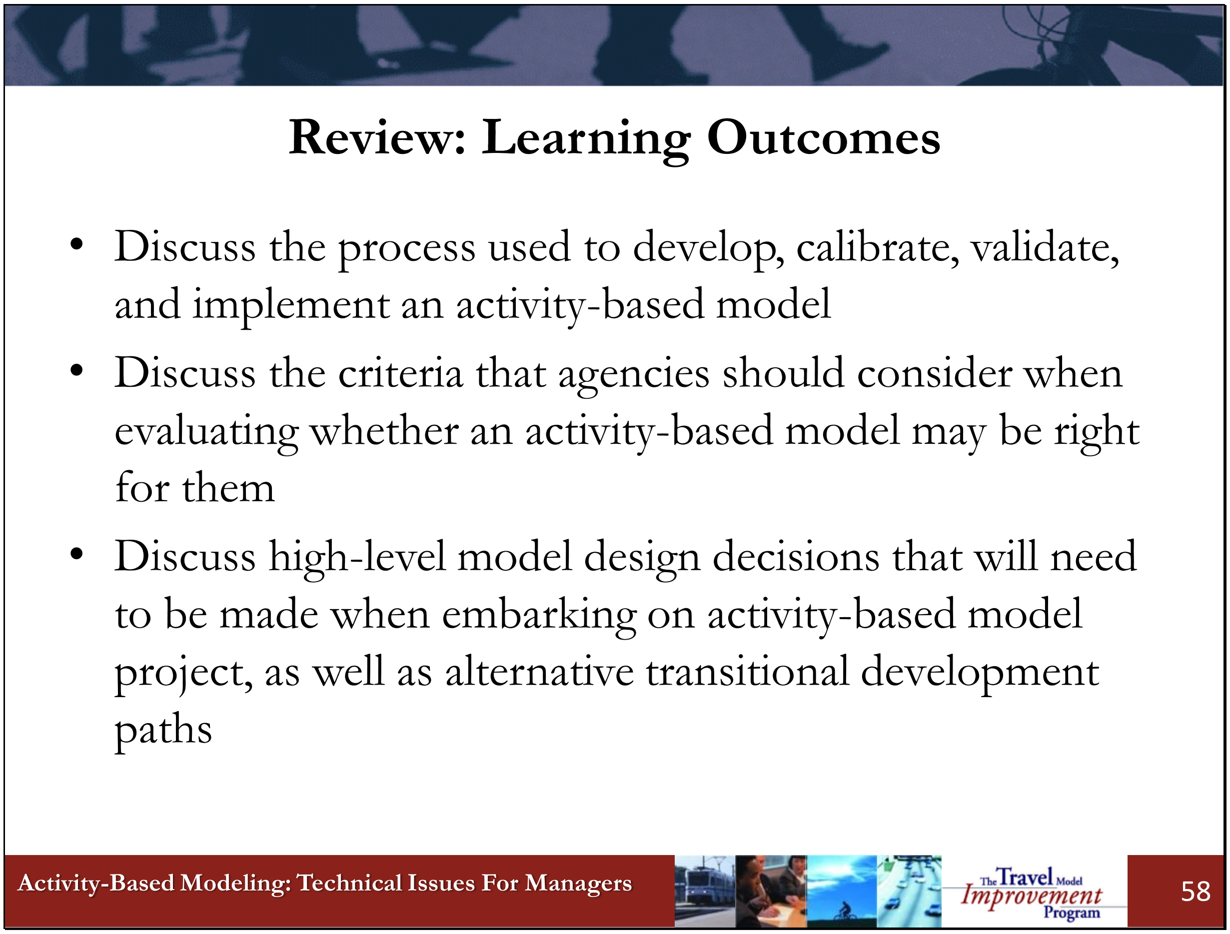
- Population Synthesizer
- Auto Ownership/Availability (if not already present)
- Usual Workplace and School Location Choices
- Activity Generation to replace Trip Productions
- Destination Choice to replace Gravity Models
- University models, visitor models, airport models



One example of a transitional model development effort is in San Diego, who developed a very simple activity-based model in the first year of their development. These models were fully estimated, implemented and calibrated for the base year.



In subsequent years, the remaining models were completed to supplement this initial development. This allows for an initial use of the models in the short term (one year) and subsequent completion of the full activity-based modeling system.



We'd like to conclude the main part of the presentation by reviewing the learning outcomes of this session. To recap, we hope that you are now better able to:

- Discuss the processes used to develop, calibrate, validate, and implement an activity-based model;
- Discuss the criteria that agencies should consider when evaluating whether an activity-based model may be right for them; and
- Discuss high-level model design decisions that will need to be made when embarking on activity-based model project, as well as alternative transitional development paths.

Questions and Answers

The Travel Model Improvement Program

2012 Activity-Based Modeling Webinar Series

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 Basics	**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 3 Questions and Answers

What is different between the data used for an ABM and the data used for a trip-based model?

Joel: Typically we use data from the same sources. Household survey data is very important, but in a trip-based model we collect a lot of data that we don't actually use. For example, data collected at a person level: age, gender, occupation, work status, usual work and school location, even if somebody doesn't go to work or school. In a trip-based model, we don't use that data as explanatory variables in the model system because we're working with aggregate trips. In an activity based model, we are generating travel for individual persons, so we can use that data as explanatory variables in the model system. But most of the data sources we rely upon to estimate, calibrate, and validate the model are similar between the two systems.

How would an activity-based model use on-board survey data differently from a trip-based model?

Joel: In any area where transit is at all an issue, on-board survey data is very important to collect well. A trip-based model works in origin-destination or production-attraction format. On-board surveys are usually collected in the same format. In an activity-based model, we are interested more in the context of the trip. We are interested in the entire tour within which the trip was taken. For example, if someone leaves home, goes to work, then to the gym, and then back to home, all on transit. If we perform an on-board survey we have an equal likelihood of intercepting that passenger on the way to work, on the way to the gym, or on the way home, but we are really interested not only in the one particular trip to one particular destination, but also the primary destination on the tour. Depending on the way the questions are asked, if we survey the person on the way to the gym, we may not get information about whether the person has been to work, and we may not get information about the home location of that traveler. Typically what we need to do to make an on-board survey more useful for an activity based model is to ask clarifying questions, such as whether the person had been to work during the trip away from home to allocate the trip to the correct tour purpose. The data is used in a similar way for model calibration as it is used in a trip based model to ensure the mode choice model is representing transit markets correctly.

Volume counts are being used for both calibration and validation. How should the data be broken out for each phase?

Maren: Volume counts refer to traffic counts and transit boarding counts. Some of this is semantic. We use these data both for validation and calibration. Initially, we compare them to the modeled volumes for highway and transit assignments, and then we use that comparison to go back and make adjustments to individual model constants in order to better match the highway and transit assignments. For example, if the AM peak highway counts are higher than the modeled volume, we might make adjustments to the full-time worker daily activity pattern or time of day models. These adjustments are highly interpretive because the models are not directly linked to the traffic volumes, but the highway counts offer a second independent source to inform the model calibration in addition to the household survey data. Calibration refers to the adjustment of the model coefficients, and validation refers to the comparison of the modeled and observed outcomes. So we don't split up the data. We are using the same data source in different ways.

When you talk about adjusting alternative-specific constants, how do you justify or prove the correctness of these constants?

Joel: Typically, the reason you need to adjust the constants is that the survey data used to estimate the models is not perfect. The expanded survey data and the modeled population usually end up with differences. There is some debate about whether the model might actually be more correct than the survey used to estimate the models. For example, in Atlanta, the survey did not include group quarters residents living near or on college campuses. In that case, we did not adjust the modeled trip length frequency distributions. But in other cases, we know that the model system is not reflecting the true behavior as well because the accessibility estimates are not perfect compared to the true accessibilities in the region and the survey. This discussion underscores why we avoid making large, unjustified adjustments to constants, because to do so would override the estimates of behavior taken from the survey data.

Do you ever compare the activity based model to results from a trip-based model to see if the activity based model is more accurate or policy-sensitive than the trip-based model?

Maren: Yes, I am aware of three agencies that have performed such comparisons. MORPC in Columbus did a very thorough comparison using grant money to look at the exact same networks and inputs using different scenarios in the future, evaluating their sensitivity, and also comparing the validations in the base year. In the case of validation, both models were very well-validated. There weren't significant differences. For the sensitivity to policies, there were differences. San Francisco and Sacramento are other regions that have tested differences in policy sensitivity. They looked at tolling, land use, transit oriented design. Stark differences were found. When looking at forecasts, some judgment is required in determining what is reasonable, but they were more able to explain the results from the activity based model, which were more intuitive. The trip-based model was interpreted to over-predict and under-predict certain changes relative to what was expected.

It seems multiple days of travel activity would provide a lot more valuable information for development of an activity based model. Is this happening--are we using multiple-day diaries to develop activity-based models?

Joel: One of the first activity-based travel surveys was done in Portland in the mid-nineties, and was a two-day survey. Both days were used, but the days were treated independently. Survey fatigue is an issue, and trip rates often decrease during the second day. GPS is an option for getting multiple days of data without fatigue. With car-based GPS, the unit can be left in the car for up to a week. With this method, we get all activities for an entire week, and information about habitual activities: for example work locations that are visited every day. We can infer activity locations from the non-habitual activities and get a richer data set, and this is used for calibration and validation. Research into behavior over multiple days is being performed in academia. We're not too far out from having models that build behavior in across multiple days. The advantage would be that certain policies might delay travel but not totally change behavior, such as license plate allocation or pricing policies.

For calibration and validation, is it more important to match traffic volumes or origin-destination patterns?

Joel: When we look at assignment results, we consider if we're getting the right amount of VMT for a particular period. Sometimes, a mismatch might be because of issues in the survey data expansion. Other times the mismatch might be due to a lack or inappropriate assumptions in the commercial vehicle travel. Screenlines from the traffic volumes can be used to get at origin-destination patterns. We can go back to destination choice models and make sure that the district-level summaries are consistent with the survey. Perhaps the survey sample was thin for a certain district movement, but the screenline reveals a more accurate picture, and adjustments might need to be made to better match the screenline. Then we might go into more detail, for example looking at traffic by facility type. These are similar procedures as one would follow for validating a trip-based model, but the focus is on whether there are systematic errors or issues in the model system that need to be addressed.