DaySim Standard Technical Documentation

Introduction

DaySim Standard[1] is a travel demand microsimulation software package that works in conjunction with any of a number of traffic and transit assignment packages to forecast a population's response to changes in transport infrastructure or policy. DaySim Standard uses an integrated system of discrete choice models to simulate long term choices for each household, and the activity and travel choices for a 24 hour period for each household member. It uses 48 half-hour time periods across the day as the basic units of temporal resolution, and uses either individual parcels of land or block-sized microzones as the basic units of spatial resolution.

This document provides technical information about how the DaySim Standard microsimulation operates and about its component models.

DaySim Standard Model System Overview

DaySim Standard follows the day activity schedule approach developed by Bowman and Ben-Akiva (2001). Its features include the following:

- The model uses a microsimulation structure, predicting outcomes for each household and person in order to produce activity/trip records comparable to those from a household survey (Bradley, et al, 1999).
- The model works at four integrated levels—longer term person and household choices, single day-long activity pattern choices, tour-level choices, and trip-level choices
- The upper level models of longer term decisions and activity/tour generation are sensitive to network accessibility and a variety of land use variables.
- The model allows the specific work tour destination for the day to differ from the person's usual work location.
- The model uses seven different activity purposes for both tours and intermediate stops (work, school, escort, shop, personal business, meal, social/recreation).
- The model predicts locations down to the individual parcel or microzone level.
- The model predicts the minute that each trip and activity starts and ends, using an internally consistent scheduling structure that is also sensitive to differences in travel times across the day (Vovsha and Bradley, 2004).
- The simulation achieves consistent realistic results by conditioning submodel choices upon previously simulated outcomes, including the use of rigorous available time window accounting.
- Logsums are used in the upper level models to account for expected utility differences associated with available modes, destinations and times of day.

Figure 1 is a flow diagram showing the simulation sequence and relationships among DaySim's component models, which are also listed in **Table 2**. The models themselves are numbered hierarchically

---

[1] DaySim software is available in two standard versions, Standard and Household. DaySim Household, which is not described in this documentation, explicitly models household outcomes within the day, such as household day pattern type, and various types of joint travel.

in the table; subsequently in this document, parenthetical numerical references to models refer to these numbers.

The DaySim model hierarchy and simulation sequence embody assumptions about the relationships among simultaneous real world outcomes. Outcomes from models higher in the hierarchy are simulated before those lower in the hierarchy and treated as known in the lower level models. Outcomes that are thought to be higher priority to the decision maker are placed higher in the hierarchy. The model structure also embodies priority assumptions that are hidden in the hierarchy, namely the relative priority of outcomes on a given level of the hierarchy. The most notable of these are the relative priority of tours in a pattern, and the relative priority of stops on a tour. These relative priorities determine the sequence of simulation, and outcomes simulated earlier are treated as known when subsequent outcomes are simulated.

At any given point in the simulation, known outcomes condition the model in two ways. First, they constrain the availability of alternatives. For example, the modes available for an intermediate stop are constrained by the chosen tour mode. Second, they can influence choice through the utility function. For example, as more intermediate stops are simulated for a tour, the likelihood of generating another stop declines. One of the ways that DaySim implements both types of conditioning is through "available time window" accounting. Throughout the simulation DaySim keeps track of the time periods required for already simulated outcomes. Then the available time windows are used to limit alternative availability and influence choices. The formal hierarchical structure and simulation sequence provide what has been referred to by Vovsha, Bradley and Bowman (2004) as downward vertical integrity.

Just as important as downward integrity is the upward vertical integrity that is achieved by the use of composite accessibility variables to explain upper level outcomes. Done properly, this makes the upper level models sensitive to important attributes that are not yet known because the lower level outcomes are not yet known, most notably travel times and costs. It also captures non-uniform cross-elasticities caused by shared unobserved attributes among groups of lower level alternatives sharing the same upper level outcome.

Upward vertical integration is a very important aspect of model integration. Without it, the model system will not effectively capture sensitivity to travel conditions. However, when there are very many alternatives (millions in the case of the entire day activity schedule model), the most preferred measure of accessibility, the expected utility logsum, requires an infeasibly large amount of computation. So, in some cases DaySim Standard uses alternate approaches to capture the most important accessibility effects with a feasible amount of computation. One approach involves using logsums that approximate the expected utility logsum. They are calculated in the same basic way, by summing the exponentiated utilities of multiple alternatives. However, the amount of computation is reduced, either by ignoring some differences among decisionmakers, or by calculating utility for a carefully chosen subset or aggregation of the available alternatives. The approximate (also called aggregate) logsum is pre-calculated and used by several of the model components, and can be re-used for many persons. The approximate logsum is a tour mode-destination choice logsum, and it is used in situations where information is needed about accessibility to activity opportunities in all surrounding locations by all available transport modes at all times of day.
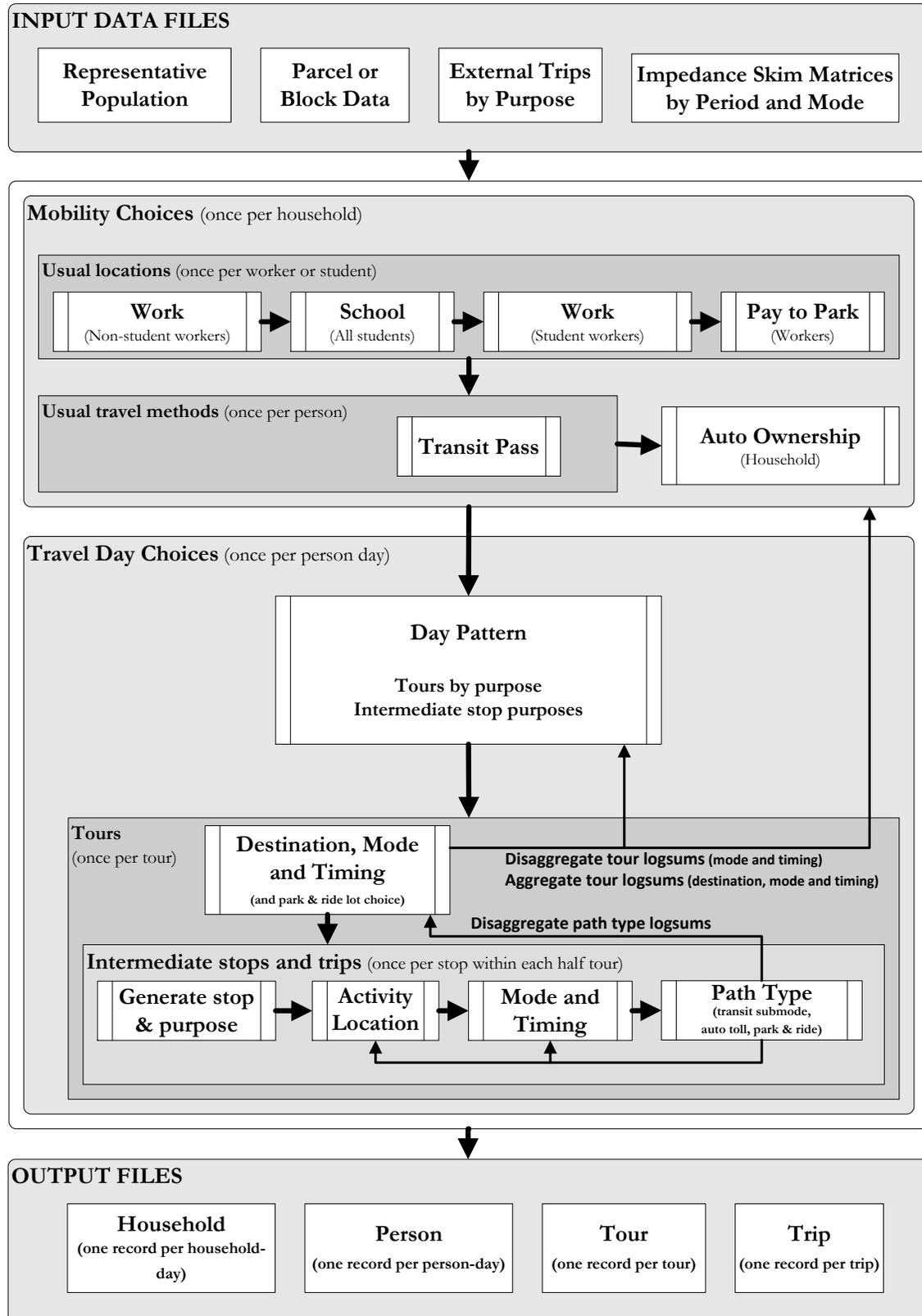
INPUT DATA FILES

| Representative Population | Parcel or Block Data | External Trips by Purpose | Impedance Skim Matrices by Period and Mode |

Mobility Choices (once per household)

Usual locations (once per worker or student)

| Work (Non-student workers) | School (All students) | Work (Student workers) | Pay to Park (Workers) |

Usual travel methods (once per person)

Transit Pass

Auto Ownership (Household)

Travel Day Choices (once per person day)

Day Pattern

Tours by purpose
Intermediate stop purposes

Tours (once per tour)

Destination, Mode and Timing (and park & ride lot choice)

Disaggregate tour logsums (mode and timing)
Aggregate tour logsums (destination, mode and timing)

Disaggregate path type logsums

Intermediate stops and trips (once per stop within each half tour)

| Generate stop & purpose | Activity Location | Mode and Timing | Path Type (transit submode, auto toll, park & ride) |

OUTPUT FILES

| Household (one record per household-day) | Person (one record per person-day) | Tour (one record per tour) | Trip (one record per trip) |

Figure 1: DaySim Standard Flow Diagram

Model #	Model Name	Level	What is predicted
1.1	Usual Workplace Location	Worker	Workplace location zone and parcel
1.2	Usual School Location	Student	School location zone and parcel
1.3	Pay to Park at Work	Worker	Whether (1) or not (0) worker must pay to park at or near the regular work location
1.4	Transit Pass Ownership	Person	Whether (1) or not (0) the person aged 16+ owns a transit pass
1.5	Auto Ownership	Household	Auto ownership
2.1	Day Activity Pattern	Person-day	0 or 1+ tours for 7 activity purposes. 0 or 1+ stops for 7 activity purposes
2.2	Exact Number of Tours	Person-day	For purposes with 1+ tours: 1, 2 or 3 tours.
3.1	Tour Primary Destination	(Sub)Tour	Primary destination zone and parcel (models are purpose-specific)
3.2	Work-Based Subtour Generation	Work Tour	Number and purpose of any subtours made during a work tour
3.3	Tour Main Mode	(Sub)Tour	Main tour mode (models are purpose-specific)
3.4	Tour Time of Day	(Sub)Tour	The time period arriving and the time period leaving primary destination (models are purpose-specific)
4.1	Intermediate Stop Generation	Half Tour	Number and activity purpose of any intermediate stops made on the half tour, conditional on day pattern
4.2	Intermediate Stop Location	Trip	Destination zone and parcel of each intermediate stop, conditional on tour origin, destination, and location of any previous stops
4.3	Trip Mode	Trip	Trip mode, conditional on main tour mode
4.4	Trip Departure Time	Trip	Departure time within 30 min. periods, conditional on time windows remaining from previous choices
5.1	Path Type	Tour or Trip	(need to describe this succinctly)

Table 1: Component Models of DaySim Standard

The other simplifying approach involves simulating a conditional outcome. For example, in the tour destination choice model, where time-of-day is not yet known, a mode choice logsum is calculated based on an assumed time of day, where the assumed time of day is determined by a probability-weighted Monte Carlo draw. In this way, the distribution of potential times of day is captured across the population rather than for each person, and the destination choice is sensitive to time-of-day changes in travel level of service.

In many other cases within the model system, true expected utility logsums are used. For example, tour mode choice logsums are used in the tour destination models.

Component Models of DaySim Standard

Most of the default models in DaySim Standard were estimated using data from the 1999 Sacramento Area Household Travel Survey, fielded by NuStats. The survey was a fairly standard place-based one-day travel diary survey, very similar to most other regional household travel surveys carried out in the US prior to the advent of GPS-based surveys. It is a straightforward exercise to re-estimate the model coefficients using data from another region, using the built-in feature of DaySim that allows it to generate ALOGIT estimation data and control files when it is run in 'estimation mode' with survey data, instead of simulating choices when it is run in 'simulation mode' with a synthetic population.

The following sections describe key aspects of the various models in DaySim Standard. Similar models are grouped together, for ease of presentation. The following attached embedded DaySim Standard Model Specifications Excel workbook includes a worksheet for each model, providing detailed values, descriptions and notes of the default coefficients estimated from the Sacramento data.

Daysim Standard
Default Model Spec

PathType Model (5.1)

As shown in **Figure 1** and **Table 1**, the PathType model is at the lowest level of the choice hierarchy. It models choices that are conditioned by a known combination of tour or trip origin, destination, mode and time of day; it uses Level of Service skim data accordingly; and it provides logsums representing generalized times to the destination, mode and timing choice models, as well as to the other upper level models via the logsums. More specifically, the PathType model performs the following functions in a consistent way:

- For a given mode/origin parcel/destination parcel/time of day, it determines if a valid path is available via one or more path types for that mode. ("Valid" meaning that there is a network path, and that the total travel time is less than a user-defined maximum.) The path can be one-way (for trip-level models) or round trip for two different times of day (for tour-level models)
- For each possible path type, a utility is determined, using the tour-specific time and cost coefficients (VOT) as well as additional time weights provided by the user.
- If one or more path types are available, a logsum across those path types is calculated and passed back for use in higher-level models such as mode choice or time of day choice.
- The travel time, cost, and distance via a chosen path type are also generated. For most uses, they are deterministic, via the path type with the best utility, although at the trip-level where the path type is predicted for the final simulated trips, a stochastic choice can be simulated instead.

A good deal of technical detail on the path type models is provided in the embedded DaySim Standard Model Specifications Excel workbook, including PathType parameter values that were used in the

estimation of the other DaySim models. Some highlights for specific modes include:

- For bicycle, the user can define additive weights for distance on specific types of links, to calibrate the usage of different facility types.
- For auto, the user can define different VOT ranges for the skim matrices, and also specify the size of a constant term to be used for toll routes to calibrate/reflect resistance to using tolled facilities.
- For walk, bicycle, and auto, parcel-based or microzone-based adjustments are made to get a more accurate estimate of distance and travel time for short trips (particularly intra-zonal trips for which the network skims provide little useful information)
- For transit, the user can define additive in-vehicle time weights, as well as PathType-specific constants, in order to calibrate the usage of different types of transit services (as well as vary their attractiveness in higher level models such as mode choice)
- For transit, access and egress walk distance are determined based on parcel-specific or microzone-specific walk distances to the nearest stops, and the user can change parameters related to the maximum walk distance and the characteristics of walking to direct paths versus paths that involve transfers.
- For park and ride, the model is similar to the transit model, but substituting drive access time for walk access at the home end. (Park and ride is always evaluated round trip, assuming the same lot on both halves of a tour.)
- For park and ride, DaySim will search across all park and ride lots and find the one that provides the best utility for the given O/D/mode/path type/times of day. Alternatively, the user can find the best park and ride lot node with other software outside of DaySim and provide a matrix of the best park and ride lot for each O/D pair.
- Even if the user does not define different path types for a mode, the path type model will be used for the single, default path type in order to calculate the generalize time utility for that alternative. This ensures that the calculations are done consistently whether or not there are multiple path types available. For example, only one PathType is available for the walk and bike modes. However, DaySim uses the PathType model for those modes and it could be used to evaluate multiple path types for them—an example would be to use completely separate bike skims for path types with and without Class 1 or 2 bike lines.

Activity Generation Models

Day Activity Pattern (2.1)

The Day Activity Pattern Model is a variation on the Bowman and Ben-Akiva approach, jointly predicting the number of home-based tours a person undertakes during a day for seven purposes, and the occurrence of additional stops during the day for the same seven purposes. The seven purposes are work, school, escort, personal business, shopping, meal and social/recreational. The pattern choice is a function of many types of household and person characteristics, as well as land use and accessibility at the residence and, if relevant, the usual work location. The main pattern model (2.1) predicts the occurrence of tours (0 or 1+) and extra stops (0 or 1+) for each purpose, and a simpler conditional model (2.2) predicts the exact number of tours for each purpose. The "base alternative" in the model is the "stay at home" alternative where all 14 dependent variables are 0 (no tours or stops are made).

Many household and person variables were found to have significant effects on the likelihood of participating in different types of activities in the day, and on whether those activities tend to be made on separate tours or as stops on complex tours. The significant variables include employment status, student status, age group, income group, car availability, work at home dummy, gender, presence of children in different age groups, presence of other adults in the household, and family/non-family status. Accessibility effects are noted in the detailed estimation results.

Person Exact Number of Tours (2.2)

A simpler model predicts the exact number of tours (1, 2 or 3) for any given purpose, conditional on making 1+ tours for that purpose. An important result is that, for most purposes, the accessibility variables have influence. This result indicates that the small percentage of people who make multiple tours for any given purpose during a day tend to be those people who live in areas that best accommodate those tours. Other people will be more likely to participate in fewer activities and/or chain their activities into fewer home-based tours.

DaySim Standard does not include explicit models of intra-household interactions for within-day choices. Although explanatory variables are used throughout the model system to take account of the characteristics of other household members, it does not explicitly link the activity patterns across individuals so that they travel together. Instead, the focus is placed on using fine level spatial detail (parcels or microzones) and temporal detail (30 minute periods), as well as on achieving upward integrity through consistent use of accessibility logsums at all levels of the model system.

Generation Model for Work-based Subtours (3.2)

The work-based subtour generation model determines, for a given work tour with known destination, the number and purpose of work-based subtours that occur in the midst of the workday. It is structured as a multinomial choice model with one alternative for each purpose, and one alternative for the 'no more subtours' alternative. It runs repeatedly ntil the 'no more subtours' alternative is chosen.

Generation Model for Intermediate Stops on Half-Tours (4.1)

For each tour, once its destination, timing and mode have been determined, the exact number of stops and their purposes is modeled for the half-tours leading to and from the tour destination. For each potential stop, the model predicts whether it occurs or not and, if so, its activity purpose. This repeats as long as another stop is predicted. The outcomes of this model are strongly conditioned by (a) the outcome of the day activity pattern model, and (b) the outcomes of this model for higher priority tours. For the last modeled tour, this model is constrained to accomplish all intermediate stop activity purposes prescribed by the activity pattern model that have not yet been accomplished on other tours.

The estimation results for this model indicate that accessibility measures are important in determining which stops are made on which tours, as well as the exact number of stops. An important feature of this model system is that we do not predict the number and allocation of stops completely at the upper pattern level, as is done in the Portland and SFCTA models, or completely at the tour level, as is done in other models such as those in Columbus and New York. Rather, the upper level pattern model predicts the likelihood that ANY stops will be made during the day for a given purpose, at a level where the substitution between extra stops versus extra tours can be modeled directly. Then, once the exact destinations, modes and times of day of tours are known, the exact allocation and number of stops is predicted using this additional tour-level information. We think that this approach provides a good balance between person-day-level and tour-level sensitivities.

Location Choice Models

Usual Work Location (1.1), Usual School Location (1.2) and Tour Primary Destination (Work 3.1a; Other 3.1b)

The dependent variable in the usual location and tour destination models is the parcel address or microzone where the activity takes place. Because of the large number of parcels or microzones in a region, it is necessary to both estimate and apply the location choice models using a sample of alternatives. The sampling of alternatives is done using two-stage importance sampling with replacement; first a Transport Analysis Zone (TAZ) is drawn according to a probability determined by its size and impedance, and then a parcel or microzone is drawn within the TAZ, with a size-based probability.

The usual work location model was estimated using all survey person records of employed persons, with the reported usual work location as the dependent variable. Similarly, the school location model uses all survey person records of students, with the reported usual school location as dependent variable. Some persons are both employed and student, so they provide observed outcomes for both models.

Some differences among the models come from the assumed model hierarchy in **Table 1**. For the usual work and school location models, the other long term choices, including auto ownership, the need to pay to park and work and the availability of a transit pass, are assumed to be unknown, based on the assumption that these are conditioned by work and school locations of household members, rather than the other way around. For the tour destinations, these long term choices are treated as given, and affect location choice. For university and high school students who also work, the usual school location is known when usual work location is modeled; for other workers who also go to school, the work location is known when usual school location is modeled. For the tour destination models, all usual locations are known.

There are additional structural differences among these models. For the two usual location models (work and school), the home location is treated as a special location, because it occurs with greater frequency than any given non-home location, and size and impedance are not meaningful attributes. As a result, both of these models take the nested logit form, with all non-home locations nested together under the conditioning choice between home and non-home. In the estimation data, all workers have a usual work location and all students have a usual school location, so the model does not have an alternative called "no usual location".

Because a large majority of work tours go to the usual work location, the work tour destination model has this as a special alternative. Therefore, the model is nested, with all locations other than the usual location nested together under the conditioning binary choice between usual and non-usual. Nearly all observed school tours go to the usual school location. Therefore, there is no school tour destination choice model.

Since there are no modeled usual locations for activities other than work and school, the destination choice model of all remaining purposes is simply a multinomial logit model. **Figure 2** depicts the structures of the usual location and tour destination models.
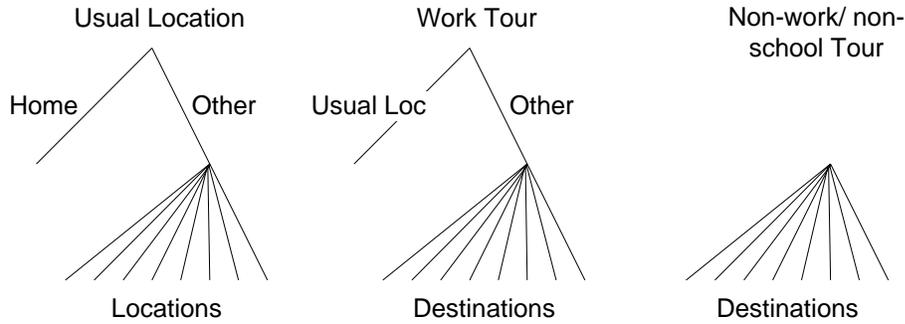
Figure 2: Structure of the usual location and tour destination models

The utility function of each regular location alternative includes a regular utility component and a size function component. Equation 1 shows the form of the utility function, with size function included:

$$V_{in} = \sum_{k=1}^{K^v} \beta_k x_{ink} z_{nk} + \mu' \ln \sum_{k=K^v+1}^{K^v+K^s} \exp(\beta_k) x_{ink} z_{nk} \qquad (1)$$

where:

V_{in} is the systematic utility of parcel alternative *i* for tour *n*,

K^v is the number of utility parameters,

K^s is the number of size parameters,

β_k, $k = 1, 2, ..., K^v + K^s$ are the utility and size parameters,

x_{ink} is an attribute of parcel alternative *i* for tour *n*,

z_{nk} is a characteristic of tour *n*,

μ' is a scale parameter measuring correlation among elemental activity opportunities within parcels or microzones (1—no correlation, 0+--high correlation)

Table 2 provides an overview of the variables (alternative attributes and person/tour characteristics) used in the utility and size functions to explain choice in the models. The left-hand column lists the alternative attributes for the binary choice (special vs. regular alternative) as well as for the conditional MNL choice among regular parcel alternatives. To the right is a column for each of the four models, and in each model's column are the characteristics associated with each of the applicable attributes.

Two important variables in all of these models are the disaggregate mode choice logsum and network distance. The logsum represents the expected maximum utility from the tour mode choice, and captures the effect of transportation system level of service on the location choice. Distance effects, independent of the level of service, are also present to varying degrees depending on the type of tour being modeled. In nearly all cases, sensitivity to distance declines as distance increases; in some cases this is captured through a logarithmic form of distance. In other cases, where there is plenty of data to support a larger number of estimated parameters, a piecewise linear form is used to more accurately capture this nonlinear effect.

In most cases the models include an aggregate mode-destination logsum variable at the destination. A positive effect is interpreted as the location's attractiveness for making subtours and intermediate stops on tours to this location. A mix of parking and employment, at both the TAZ and parcel or microzone level,

as well as street connectivity in the neighborhood, attract workers and tours for non-work purposes. Also, parcel-based or microzone-based size variables and TAZ-level density variables affect location choice.

Attributes	Usual work location	Work tour destination	Usual school location	Non-work tour destination
Binary choice	**Home vs other**	**Usual vs other**	**Home vs other**	**not applicable**
Constants	by person type*	Byperson type* tour type	Byperson type* HH size	
Disaggregte logsum among regular locations	Yes	Yes	yes	
Conditional MNL choice among regular locations				
Disaggregate mode choice logsum to destination	Yes	Yes	Yes	Yes
Piecewise linear driving distance function	For fulltime workers		For children under age 16	ByPurpose Priority Pattern type
Natural log of driving distance	For other than fulltime workers by person type* income	Byperson type* tour type	For persons age 16+ by person type*	Bytour type income person type* time available
Distance from usual work location		Yes	for not student aged	
Distance from usual school location	for student aged	for student aged		Yes
Aggregate mode-dest logsum at destination	By person type	By person type	By person type	By purpose
Parking and employment mix	For daily parking in parcel and in TAZ	for daily parking in parcel and TAZ		For hourly parking in parcel and TAZ by car availability
Ratio of neighborhood nodes with 3 or 4 entering links	Yes	By car availability		By car availability
employment, enrollment and households by category:	byperson type income	Byperson type Income	by person type	by purpose (and by 'kids in household' for escort tours)
--Neighborhood density	--yes	--yes	--yes	--yes
--Parcel size	--yes	--yes	--yes	--yes
Person type categories in the models	full-time worker part-time worker not full- or part-time	full-time worker part-time worker not full- or part-time	child under 5 child 5 to 15 child 16+ university student not student aged	full-time worker part-time worker retired adult other adult university student child 16+ child 5 to 15 child under 5

Table 2: Utility function variables in the location choice models

Locations of Intermediate Stops (4.2)

At the time that a particular intermediate stop's location is modeled, information about the tour (origin, destination, time period arriving and departing the primary destination, and tour mode are known, and can be used to explain the location choice. Details about any stops nearer to the primary destination are also known, including the location, trip mode, and the time of departure toward the tour destination (or arrival from the tour destination on the second half-tour). Finally, details of previously simulated tours are known, including the time during which those tours occur.

However, at the time a stop's destination is modeled, several things are NOT known. These include the trip mode for the trip between this stop and the stop nearer to the tour destination, and the departure and arrival times of that trip, which will be modeled immediately after this stop's location. The arrival time

from the stop nearer to the tour origin (or departure time to that stop on second half-tour) is also not known because it will be modeled along with stop location and trip mode for the next stop further from the tour origin.

As a result of this modeling approach, two known locations serve as anchor points for calculating travel impedance. These are the stop location immediately toward the tour destination, which is called the stop origin (it is the tour destination itself for the first stop in a half-tour), and the tour origin. The choice of location involves comparing, among competing locations, (a) the impedance of making a detour to get there along the route between stop origin and tour origin, and (b) the location's attractiveness for the given activity purpose.

The dependent variable in the intermediate stop location model is the parcel or microzone where the activity takes place. As with the usual location and tour models, it is necessary to both estimate and apply the model using a sample of alternatives. The alternative sampling is done using a similar two-stage importance sampling with replacement. However, because there are two anchor points for the intermediate stop, and such stops are more likely to occur near one or the other of the anchor points, the sample consists of two subsamples, one surrounding each anchor point.

The model is a multinomial logit (MNL), with a utility function as described above for the usual location and tour models. Trip characteristics used in the model include stop purpose, tour purpose, tour mode, tour structure, stop placement in tour, person type, and household characteristics. The most important characteristics are the tour mode and the stop purpose. The tour mode restricts the modes available for the stop, and this affects the availability and impedance of stop locations. The availability and attractiveness of stop locations depend heavily on the stop purpose. Tour characteristics also affect willingness to travel for the stop, and the tendency to stop near the stop or tour origin. These trip and tour characteristics tend to overshadow the effect of personal and household characteristics in this model.

The main impedance variable is generalized time, as well as its quadratic and cubic forms, to allow for nonlinear effects. It combines all travel cost and time components according to assumptions about their relative values. Generalized time is used, instead of various separately estimated time and cost coefficients, because the intermediate stop data is not robust enough to support good estimates of the relative values. Generalized time is measured as the (generalized) time required to travel from stop origin to stop location and on to tour origin, minus the time required to travel directly from stop origin to tour origin. It is further modified by discounting it according to the distance between the stop origin and the tour origin. The discounting is based on the hypothesis that people are more willing to make longer detours for intermediate stops on long tours than they are on short tours.

Additional impedance variables used in the model include travel time as a fraction of the available time window, which captures the tendency to choose nearby activity locations if there are tight time constraints on the stop, and proximity variables (inverse distance), which capture the tendency to stop near either the stop origin or the tour origin.

Mode Choice Models

Tour Main Mode (3.3)

The tour mode choice model determines the main mode for each tour (a small percentage of tours are multi-modal). There are eight modes, although some of them are only available for specific purposes. They are listed below along with the availability rules, in the same priority order as used to determine the main mode of a multi-mode tour:

1. Drive to Transit: Available only in the Home-based Work model, for tours with a valid drive to transit path in both the outbound and return observed tour
2. Walk to Transit: Available in all models except for Home-based Escort, for tours with a valid walk to transit path in both the outbound and return observed tour periods.
3. School Bus: Available only in the Home-based School model, for all tours.
4. Shared Ride 3+: Available in all models, for all tours.
5. Shared Ride 2: Available in all models, for all tours.
6. Drive Alone: Available in all models, except for Home-based Escort, for tours made by persons age 16+ in car-owning households.
7. Bike: Available in all models, except for Home-based Escort, for all tours with round trip road distance of 30 miles or less.
8. Walk: Available in all models, for all tours with round trip road distance of 10 miles or less.

The following tour main mode models have been specified and estimated separately:
    WorkTourModeModel (3.3a)
    SchoolTourModeModel (3.3b)
    EscortTourModeModel (3.3c)
    OtherHomeBasedTourModeModel (3.3d)
    WorkBasedSubtourModeModel (3.3e)

Two land use variables came out as significant in many of the models, increasing the probability of walk, bike and transit:

Mixed use density: This is defined as the geometric average of retail and service employment (RS) and households (HH) within a half mile of the origin or destination parcel (= RS * HH / (RS + HH)). This value is highest when jobs and households are both high and balanced. High values near the tour origin tend to encourage walking and biking, while high values near the tour destination more often encourage transit use.

Intersection density: This is defined as the number of 4-way intersections plus one half the number of 3-way intersections minus the number of 1-way "intersections" (dead ends and cul de sacs) within a half mile of the origin or destination parcel. Higher values tend to encourage walking for School and Escort tours, where safety for children is an issue, and also to encourage walking, biking and transit for Home-Based Other tours.

A number of different nesting structures were tested. In the nesting structure that was selected there are three combined nests:
    (1)   Drive to Transit with Walk to Transit
    (2)   Shared Ride 2 with Shared Ride 3+
    (3)   Bike with Walk

These all gave logsum coefficients less than 1.0 but not significantly different from each other, so a single estimated nesting parameter applies to all 3 nests (as well as to the 2 additional "nests" that only have one alternative each: Drive Alone, and School Bus). No nesting was used for the Escort model, as it contains only 3 alternatives and is a very simple model.

Trip Mode (4.3)

The trip-level mode is conditional on the predicted tour mode, but now uses a specific OD pair and a time anchor, and also the trip mode for the adjacent, previously modeled trip in the chain. The majority of tours use a single mode for all trips, so this model only explains the small percentage of trips that are

made by modes other than the main mode. The most common occurrence of this is a Drive Alone trip that is made as part of a Shared Ride tour after the passenger has been picked up or dropped off. These cases are most common on Escort tours, where predicting the trip(s) that is Drive Alone is mainly a function of the half tour (away from home or towards home) and the time of day.

Table 3 shows the distribution in the household survey data of valid trip-level mode choice records versus the mode for the tour[2]. According to the hierarchy used to assign the main mode, only a subset of modes may be parts of certain tours. At the extreme end, all 1127 trips that are part of walk tours are walk trips.

Tour mode / Trip mode	drive-transit-walk	walk-transit-walk	school bus	car-shared ride 3+	car-shared ride 2	car-drive alone	bike	walk	Total
drive-transit-walk	*47*								47
walk-transit-drive	*23*								23
walk-transit-walk	9	*282*							291
school bus	2	3	*293*						298
car-shared ride 3+	7	17	62	*4,024*					4,110
car-shared ride 2	20	37	62	1,085	*5,895*				7,099
car-drive alone	15	13	9	719	1,852	*11,478*			14,086
bike		8		7	12	8	*437*		472
walk	11	58	23	116	113	64	22	*1,127*	1,534
Total	134	418	449	5,951	7,872	11,550	459	1,127	27,960

Table 3: Tour Mode (columns) vs. Trip Mode (rows)

As shown in **Table 4**, almost 85% of all trips have the same tour mode and trip mode. The purpose of the trip level mode choice model is to explain the other 15%. The table also shows that the large majority of those remaining trips are either shared ride 2 trips in shared ride 3+ tours, or else drive alone trips in shared ride tours. So, it is up to the trip level mode choice model to explain which trips in shared ride tours are shared ride and which are drive alone. There are also 205 trips (0.7%) that are shared ride trips in transit or (mainly) school bus tours. A further 407 trips are walk trips on non-walk tours (1.5%). This leaves only 0.3% of trips in all other combinations.

	Frequency	Percent
Tour mode and trip mode are the same	*23,606*	84.4
Tour and trip modes are different, but both by car	3,656	13.1
Shared ride trip on a transit/school bus tour	205	0.7
Walk trip on a non-walk tour	407	1.5
All other combinations	86	0.3
Total	27,960	100

Table 4: Summary of tour/trip mode combination types

For model estimation, trips were excluded for a number of reasons:

---

[2] Tables 3 and 4 come from the original estimation of the SACOG models in 2005. The DaySim Basic coefficients in the embedded workbook come from re-estimation in 2012 with the same basic data set, but including a few more observed outcomes because of enhanced data preparation procedures used at that time.

- Walk trips on walk tours were excluded because only one alternative is available
- Trips were excluded when they were the last chronological trip in the tour, and no previous trip in the tour had used the main tour mode. In those cases, by definition, the last trip must be the main tour mode. (This was 423 cases, or about 1.5%)
- Various other trips were dropped when the chosen mode was not available in the networks.

The embedded workbook shows the estimated coefficients for the model. The model includes three new types of variables that are specific to the trip-level data. A large proportion of the mixed drive alone (DA)/shared ride (SR) tours are tours that contain at least one escort (pick up/drop off) stop. The first set of variables (par # 161-170) attempt to explain which trips are shared ride by looking at both the origin purpose and destination purpose, as well as the time of day. Not surprisingly, trips to work in the morning after dropping someone off and from work in the afternoon before picking someone up are rarely shared ride trips. The opposite side of these are the positive SR coefficients for trips from home to drop off in the AM and trips from pick up to home in the PM. Trips from an escort stop back home in the AM, and to an escort stop in the midday tend not to be shared ride.

Parameters 100-107 are mainly positive coefficients for the likelihood that the trip mode is the same as the tour mode, regardless of mode. This is particularly true when the half tour only has one trip (no intermediate stops). In cases where the half tour has 2+ trips, the mode for the first outbound trip and last return trip are the least likely to be the same as the tour mode, although these effects are not strong.

Finally, parameters 115-131 apply to specific trip mode/tour mode combinations, all relative to the "base" trip mode of walk. The general pattern in these coefficients is:
- Relative to the walk mode, school bus (SB), shared ride (S3,S2) and drive alone (DA) are not likely as part of transit tours (DT,WT)
- Relative to the walk mode, shared ride (S2,S3) are more likely as part of school bus (SB) tours.
- The strongest positive switching is for S2 on S3 tours.
- Relative to the walk mode, drive alone (DA) is more likely on shared ride (S2, S3) tours.
- Relative to the walk mode, bike (BI) is not likely to be a part of any tours that are not bike-only tours.

Time of Day/Activity Scheduling Models

DaySim employs a method of modeling time of day developed by Vovsha and Bradley (2004). The time of day models explicitly model the 30 minute time periods of arrival and departure at all activity locations, and hence for all trips between those locations. DaySim Standard uses 48 half-hour periods in the day—3:00-3:29 AM, 3:30-3:59 AM, …, 2:30 AM-2:59 AM. It includes two types of time-of-day models, tour primary destination arrival and departure time, and intermediate stop arrival or departure time.

Tour primary destination arrival and departure time (3.4). For each home-based or work-based tour, the model predicts the time that the person arrives at the tour primary destination, and the time that the person leaves that destination to begin the return half-tour. Given the way that the activity diary data was collected, no tour begins before 3:00 AM or ends after 2:59 constant AM. The tour model includes as alternatives every possible combination of the 48 alternatives, or 48 x 49 / 2 = 1,716 possible alternatives. The model is applied after the tour primary destination and main mode have already been predicted. There are four separately estimated tour models in the embedded workbook of model specifications:

WorkTourTimeModel (3.4a)
SchoolTourTimeModel (3.4b)
OtherHomeBasedTourTimeModel (3.4c)
WorkBasedSubtourTimeModel (3.4d)

Intermediate stop arrival or departure time (4.4). For each intermediate stop made on any tour, this model predicts either the time that the person arrives at the stop location (on the first half tour), or else the time that the person departs from the stop location (on the second half tour). On the second (return) half tour, we know the time that the person departs from the tour primary destination, and, because the model is applied after the stop location and trip mode have been predicted, we also know the travel time from the primary destination to the first intermediate stop. As a result, we know the arrival time at the first intermediate stop, so the model only needs to predict the departure time from among a maximum of 48 alternatives (the same 30 minute periods that are used in the tour models). This procedure is repeated for each intermediate stop on the half tour. On the first (outbound) half tour, the stops are simulated in reverse order from the primary destination back to the tour origin, so we know the departure time from each stop and only need to predict the arrival time.

Once an arrival or departure time period has been simulated, DaySim randomly assigns a specific arrival or departure minute within the time period. Since entire tours, including stop outcomes, are modeled one at a time, first for work and school tours and then for other tours, the times away from home for each tour become unavailable for subsequently modeled tours.

A key concept in DaySim is the "time window". A time window is a set of contiguous minutes that are available for scheduling tours and stops. When a tour or stop is scheduled within an available time window, the portion of the window that it does not fill is left as two separate and smaller time windows.

Another key aspect in the time of day models is the use of shift variables. These are dummy variables interacted with the arrival time and the duration of the alternative. If the arrival shift coefficient is negative, it means that activities tend to be made earlier than the base case (because the shift coefficient causes later arrival time alternatives to have lower utility), and if it is positive, it means that activities tend to be made later. If the duration shift coefficient is negative, it means that activities tend to be shorter (because the shift coefficient causes longer duration time alternatives to have lower utility), and if it is positive, activities tend to be longer. No departure shift coefficient is estimated because the departure shift is simply the sum of the arrival shift and the duration shift (e.g. if the arrival shift is an hour earlier and the duration shift is an hour longer, the departure shift is 0). In the models, shift variables interact extensively with other characteristics of the person, day activity pattern and tour, as well as time-dependent attributes of the network, such as travel times and measures of congestion, to effectively represent their influence on time-of-day choice.

The time of day models also use a variety of variables to represent scheduling pressure, conditional on what other activities have already been scheduled or remain to be scheduled for the day. The overall scheduling pressure is given by the number of tours remaining to be scheduled divided by the total empty window that would remain if an alternative is chosen. The negative effect indicates that people are less likely to choose schedule alternatives that would leave them with much to schedule and little time to schedule it in. A similar variable is the number of tours remaining divided by the maximum consecutive time window. This is also negative, meaning that people with more tours to schedule will tend to try to leave a large consecutive block of time rather than two or more smaller blocks.

Relative travel times across the day also influence time of day choice. The travel time for each period is based on the network travel times for the various periods of the day. The variable is applied for both the

outbound half tour (tour origin to tour destination) and the return half (tour destination to tour origin). For auto tours, the time is just the in-vehicle time, while transit time is in-vehicle time plus first wait time, transfer time, and drive access time. Walk access/egress time is not included, as that does not vary by time period. These variables are not applied for walk, bike or school bus tours. Significant travel time effects were found for Work and Other tours and for Intermediate Stops, but not for School or Work-based Tours.

For Work tours, in both the AM and PM, the estimation results show a tendency to move the work activity earlier as the time in very congested conditions increases. For School tours and Work-based subtours, no significant congestion effects were estimated. For Other tours, times in the PM peak were found to shift both earlier and later with high congestion.

Long Term Mobility Choices

Pay to Park at Work (1.3)

This binary choice model predicts whether or not each person age 16+ owns a transit pass, as a function of person type, age, employment status, student status, and accessibility by transit from their home, workplace and/or school location. This model was estimated using data from the Seattle (PSRC) region, since this variable was not available in the SACOG 2000 survey data.

Transit Pass Ownership (1.4)

For each worker, this model predicts whether or not the person has to pay to park at/near their workplace—i.e. that they do not receive free or totally subsidized parking. It is a binary model, mainly a function of income, employment status, and the land use and parking supply around the workplace. If the model predicts that a worker does have to pay, then the parking cost at their workplace is determined by the average daily price for paid off-street parking in the (smaller) buffer around the work parcel. Otherwise, the parking cost is set at 0 (free).

Auto Availability (1.5)

This model is applied at the household level, and determines the number of vehicles available to the household drivers. It is structured as a multinomial logit (MNL) with five available alternatives: 0, 1, 2, 3, and 4+. Key variables are the numbers of working adults, non-working adults, students of driving age, children below driving age and income. Statistically significant policy variables affecting car ownership include mode choice logsums measuring accessibility to the workers' and students' usual work and school locations, a mode-destination choice logsum measuring accessibility from home to non-work activities, distance from home to the nearest transit stop, parking prices in the home neighborhood, and commercial employment in the home neighbourhood.

References

Bowman, J. L., and M. Ben-Akiva (2001) Activity-based disaggregate travel demand model system with activity schedules, Transportation Research Part A, 35 1-28.

Bradley, M., J.L. Bowman, and T. K. Lawton (1999) A Comparison of Sample Enumeration and Stochastic Microsimulation for Application of Tour-Based and Activity-Based Travel Demand Models,, Presented at European Transport Conference, Cambridge, UK.

Vovsha P and M. Bradley (2004) A Hybrid Discrete Choice Departure Time and Duration Model for Scheduling Travel Tours. Paper presented at the 83rd Annual Meeting of the Transportation Research Board, Washington, D.C.

Vovsha, P., M. Bradley, and J.L. Bowman. (2004) Activity-based travel forecasting models in the United States: Progress since 1995 and Prospects for the Future, Presented at the EIRASS Conference on Progress in Activity-Based Analysis, Maastricht.

Needed documentation enhancements

Here is a list of further work that should be done to the documentation:

1. Audit and improve the variable descriptions in the embedded workbook for OtherTourDestination and IntermediateStopLocation
2. Audit and correct Table 2 summarizing location choice variables
3. Move variable descriptions from SACOG Tech Memo 5 into the text and/or worksheet of the IntermediateStopLocation model, making sure that the text correctly describes the variables in the current version of the model.
4. Provide descriptions in the worksheets for the long term models, mode choice models and TOD models
5. Provide more extensive text for some of the models.
6. Include a detailed description of the aggregate logsum models and a sheet of default coefficients.