

SACSIM/05

Activity-Based Travel Forecasting Model for SACOG

Featuring *DAYSIM*—the Person Day Activity and Travel Simulator

Technical Memo Number 1

Model System Design

July 31, 2006—Draft 2

Prepared for

Sacramento Area Council of Governments

Prepared by

John L. Bowman, Ph. D.

Transportation Systems and Decision Sciences

28 Beals Street, Brookline, MA 02446 USA

+1-617-232-8189 John_L_Bowman@alum.mit.edu <http://JBowman.net>

MARK BRADLEY

BRADLEY RESEARCH & CONSULTING

524 Arroyo Ave., Santa Barbara, CA 93109, USA.

+1-805-564-3908 mark_bradley@cox.net

Table of Contents

Introduction and Model System Overview.....	3
Model system	3
Model variables	4
Main differences between SACOG design and San Francisco model	6
Project scope.....	7
Population Synthesizer	7
Base year synthetic population.....	8
Forecast year synthetic population	9
Long Term Choice Simulator and DaySim	10
Long term choice models	12
Day level models	12
Tour level models.....	12
Trip/stop level models	14
Accessibility variables.....	15
Validation and Calibration	19
Stage 1: Compare activity based model results and reported results for HH survey sample (Development project—March 13 thru August 12)	19
Stage 2: Implement and validate Population Synthesizer (Development project—August 12 deadline)	20
Stage 3: Compare activity based model results for synthetic population to reported results for expanded household survey (Implementation project—deadline to be determined)	21
Stage 4: Calibrate long term components of the activity based model system with census and SACOG counts (Implementation project—deadline to be determined).....	21
Stage 5: Calibrate integrated model system with expanded on-board survey, census, screenlines and other counts (Implementation project—deadline to be determined)	22
Appendix 1—DaySim Model Features	24
Appendix 2—DaySim Variables	27

Introduction and Model System Overview

This is the first in a series of technical memos being produced according to a work program in which Mark A. Bradley and John L. Bowman are developing the activity-based demand model components of a new travel demand forecasting model system for the Sacramento Area Council of Governments (SACOG). The attributes of the model to be delivered were initially prescribed in “Final Land Use and Transport Modeling Design Report”, prepared for Sacramento Area Council of Governments (DKS Associates, Mark Bradley Research and Consulting, HBA Spectro Inc. and John L. Bowman, Ph.D., December 27, 2001), and subsequently enhanced and revised in “Addendum to the SACOG Model Design Report” (John Bowman and Mark Bradley, May 14, 2003), and the memo to Gordon Garry entitled “Overview of Structure and Procedures for New Travel Model Development” (Mark Bradley and John Bowman, January 15, 2005). This memo further refines and documents the design. It stands alone, so that reference to the prior documents is not necessary. The design is still subject to adjustment and refinement as it is reviewed and implemented.

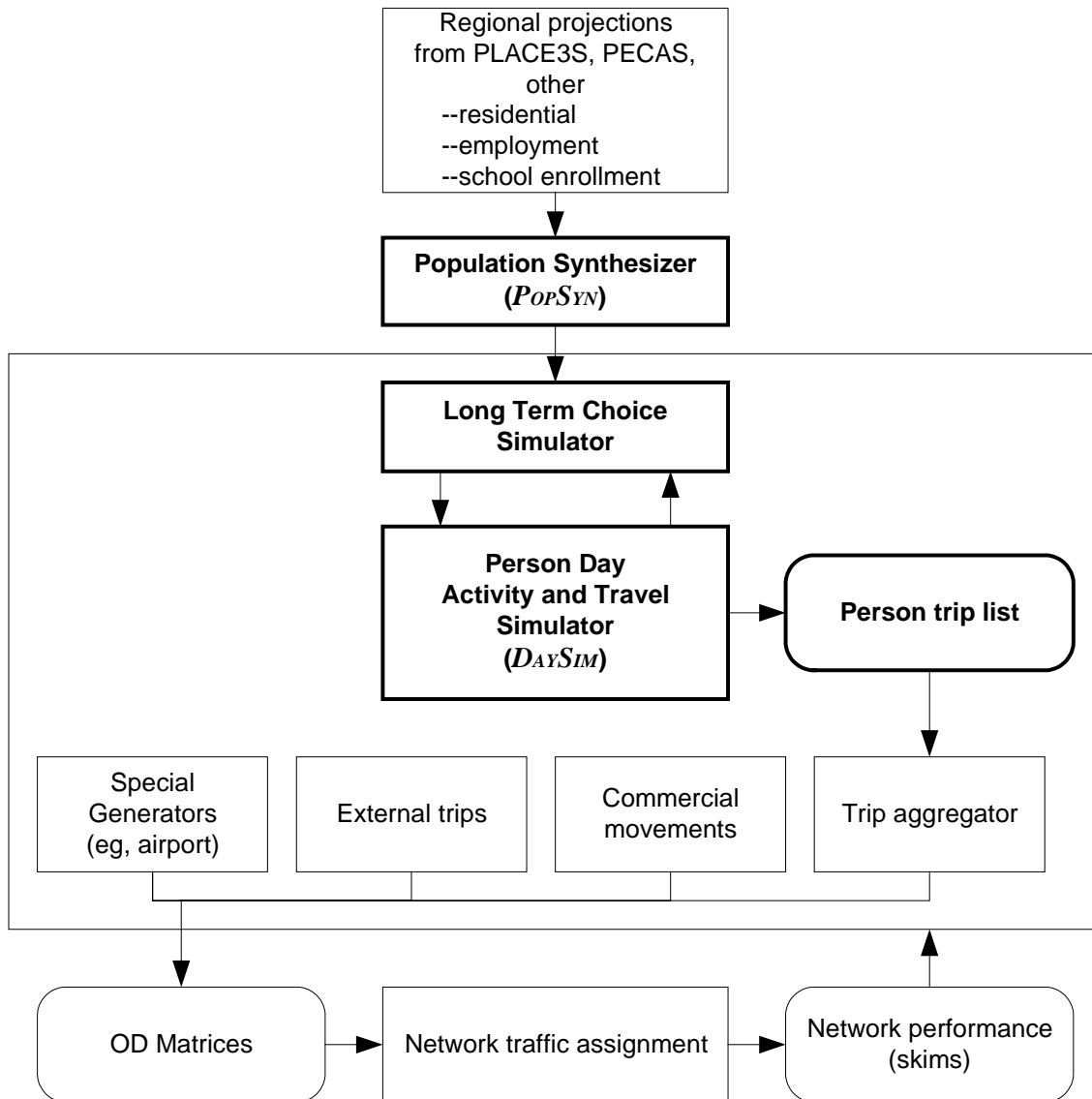
Model system

Figure 1 shows the major components of the new travel forecasting model system. The Population Synthesizer (PopSyn) creates a synthetic population, comprised of Census PUMS households, that is consistent with regional residential, employment and school enrollment forecasts. Longterm choices (work location, school location and auto ownership) are simulated for all members of the population. The Person Day Activity and Travel Simulator (DaySim) creates a one-day activity and travel schedule for each person in the population, including a list of their tours and the trips on each tour.

The trips predicted by DaySim are aggregated into trip matrices and combined with predicted trips for special generators, external trips and commercial traffic into time- and mode-specific trip matrices. The network traffic assignment models load the trips onto the network. Traffic assignment is iteratively equilibrated with the Long Term Choice Simulator, DaySim and the other demand models. As shown here, the regional forecasts are treated as exogenous. In subsequent implementations, it is anticipated that the travel forecasting model will be embedded in PECAS, the regional economic and land development model, so that the long range PECAS forecasts will depend on the activity-based travel forecast of DaySim.

20060719: As implemented, PopSyn, Long Term Choice Simulator and Person Day Activity and Travel Simulator are all included in the same executable program called DaySim.

Figure 1: New SACOG Regional Travel Forecasting Model System



Model variables

Table 1 lists the variables that will be produced by the models. The variables are at five different levels: household, person, person day, tour and trip. The table also lists the range of values that will be used for each output variable. **Table 1** contains only the most elemental variables. Other variables are computed from these, including characteristics of the household, such as household size and number of workers, that are aggregates of person characteristics, and characteristics of the day pattern or tour, such as number and purpose of tours and trips, that are aggregates of trip characteristics. Still more output variables can be computed in combination with the network and/or zonal data, such as the VMT traveled by a person.

SACOG Activity-Based Travel Forecasting Model

Featuring *DaySim*—the Person Day Simulator

Technical Memo No. 1: **Model System Design**

Note: The Table 1 list of variables may be superceded by the lists in the DaySim users manual (Tech memo 10).

Table 1—Elemental variables produced by PopSyn and DaySim

Level VARIABLE ID	Variable Description	Range of Values
Household		
SAMPN	household ID number	
HHSIZE	# persons in HH	0-10
TOTVEH	# vehicles in HH	0-4+
INCOME	total household income	
FAMILY	HH family status	single, family, nonfamily
HHPARCEL	household residence parcel	
Person		
PERNO	person ID number	
GEND	Gender	
AGE	Age	0-98+
WORKER	employment status	employed, not employed
STUDENT	student status	University student, grade school student, nonstudent
HRSWORK	# hours worked per week	
WUPARCEL	usual work location parcel	
SUPARCEL	usual school location parcel	
Person Day		
Whhmm	Time availability for 30 minute period beginning hhmm (48 separate variables)	free, later part scheduled, early part scheduled, fully scheduled
Tour		
TOURNO	tour ID number (in simulation order)	
PDTYPE	primary destination purpose type	1-work 2-school 3-escort 4-per.bus 5-shopping 6-meal 7-social/rec 8-home
OPARCEL	Tour origin location parcel	Home parcel for home-based tours Work tour destination location for work-based tours
PDPARCEL	Primary destination loc. parcel	
MMODE	tour main mode (may be an aggregated set of the 9 modes)	1 -drive-transit-walk 2 -walk-transit-drive 3 -walk-transit-walk 4 -school bus 5 -car-shared ride 3+ 6 -car-shared ride 2 7 -car-drive alone 8-bike 9-walk
Trip		
TOURHALF	Trip tour half	1st, 2 nd
TRIPNO	Trip ID within tour half (outward from primary dest)	
SOTYPE	Trip origin purpose type	see tour primary destination purpose
SDTYPE	Trip destination purpose type	see tour primary destination purpose
SOPARCEL	Trip origin parcel	
SDPARCEL	Trip destination parcel	
SOTIME1	Trip origin arrival time	144 10-minute time periods
SOTIME2	Trip origin departure time	144 10-minute time periods
SDTIME1	Trip destination arrival time	144 10-minute time periods
SDTIME2	Trip destination departure time	144 10-minute time periods
SMODE	Trip mode	see tour main mode

Main differences between SACOG design and San Francisco model

The SACOG model retains the key structural advantages of the SFCTA model, including:

- 1 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.
- 2 The model works at 4 integrated levels—longer term person and household choices, single day-long activity pattern choices, tour-level choices, and trip-level choices.
- 3 The upper level models of longer terms decisions and activity/tour generation are sensitive to network accessibility and a variety of land use variables.

The SACOG design includes a number of key enhancements relative to the SFCTA model:

- 1 The model uses 7 purposes for both tours and intermediate stops (work, school, escort, shop, personal business, meal, social/recreation), while the SFCTA model uses only 3 for tours (work, school and other), and 1 (other) for all intermediate stops.
- 2 The model allows the specific work and/or school tour destination for the day to differ from the person's usual work and school location. That is not the case in the SFCTA model, where the specific day's work or school location was modeled as if it were the usual location (separate data on both were not available).
- 3 The model predicts locations down to the single parcel level, whereas the SFCTA model works completely at the TAZ level. This provides important advantages in modeling mode choice and in capturing land use and accessibility effects.
- 4 The model predicts the time that each trip and activity starts and ends to the nearest 30 minutes, using an internally consistent scheduling structure that is also sensitive to differences in travel times across the day. The SFCTA time-of-day model used a much simpler structure with only 5 time periods across the day and no sensitivity to travel times and congestion.
- 5 The accessibility variables used in the upper level models are approximations to a true expected utility structure, with single variables ("aggregate logsums") encapsulating differences across different modes and destinations. The accessibility variables used in SFCTA are mode-specific with rather arbitrary definitions—e.g. the number of retail jobs that are accessible within 30 minutes by transit.
- 6 Explicit treatment of time-space constraints in destination choice and mode choice to prevent the prediction of impossible combinations of choices.

Project scope

The current work program consists of the development of PopSyn, the Long Term Choice Simulator, and DaySim, the highlighted portions of **Figure 1**. While this work program is in progress, another work program will start in which the other model components are developed, and the entire model system is integrated and calibrated within a Citilabs CUBE operating framework.

The remainder of this memo describes the design features of PopSyn, the long term choices and DaySim. These include a description of each model component, definitions of the variables included in the simulated output, details about accessibility variables employed to help integrate the model system, and the sampling procedure used in the destination choice models. It also includes a validation procedure for PopSyn, the Long Term Choice Simulator and DaySim, as well as a calibration procedure for the entire model system.

Population Synthesizer

This model/procedure produces a list of household and person records from the PUMS microdata. Each household is defined in terms of income and household size, plus the age, gender, employment status and student status of all household members. Using CTPP and STF tables in the base year, appropriate numbers of each type of household are allocated to each TAZ. In forecast years, these numbers are adjusted according to demographic forecasts from the land use model and any additional sources. Parcel level inputs on residential land use are used to further allocate households to parcels.

The current plan is to implement and use the flexible validatable population synthesizer (PopSyn) that is under development for the Atlanta Regional Commission. If it is not satisfactorily debugged in time for the SACOG project, we will instead implement a simpler population synthesizer similar to those that were implemented for Portland and San Francisco County, coded with similar functionality to the ARC synthesizer, but with less built in flexibility and testing capabilities. **Figure 2** provides a schematic of PopSyn, showing key inputs and outputs for the base year and a forecast year, and the procedures are described in the next two subsections.

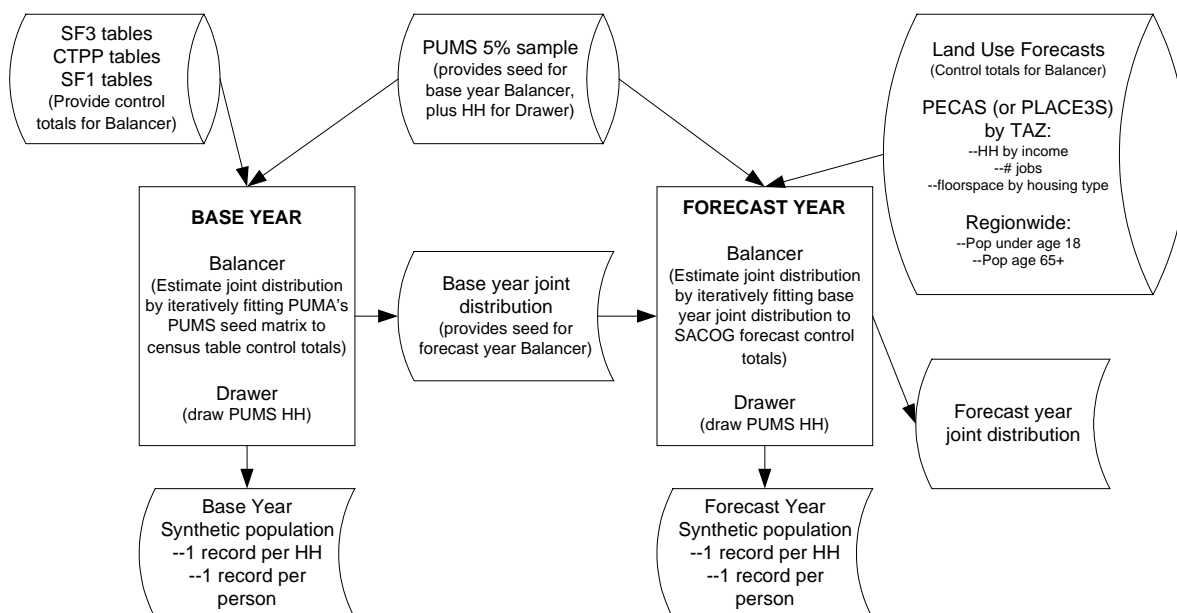
NOTE:: The ARC population synthesizer is not being used. A synthesizer within DaySim uses as input a three dimensional distribution of households (by income by size by number of workers), draws households from PUMS, and places them on parcels. The 3-D distribution comes from CTPP 1-75 for 2000, and from SACOG forecasts for other years. The following information about the ARC PopSyn does not apply to SACOG.

Base year synthetic population

By far the best available detailed information about households comes from the US census. Therefore, the model system is set up to use a census year (2000) as the base year for model forecasts, and PopSyn is designed to extensively use census data to create the base year SynPop. Census SF1, SF3 and CTPP tables provide rich information about the distribution of various important household characteristics within each census block [SF1] or block group [SF3, CTPP]. Many of these tables are multidimensional; that is, the table provides information about the joint distribution of two or more important variables. PopSyn is set up so that it can synthesize a base year population that matches any number of desired multidimensional SF1, SF3 and CTPP distributions at the TAZ level of detail.

The distribution of households is synthesized through an iterative proportional fitting (IPF) procedure called ‘Balancer’ that is like a traditional Fratar procedure for balancing trip ends, except the ‘cells’ of the joint distribution are defined by household characteristics and the control values can apply to any designated subset of cells. For the base year, Balancer’s ‘seed’ distribution is the joint distribution observed in the census 5% Public Use Micro Sample (PUMS). The PUMS distribution is used because each PUMS household has enough data available to assign it precisely to one household demographic category (HHCat) defined jointly by several different variables. This allows us to define HHCats to take advantage of the SF1, SF3 and CTPP tables, and still have a reliable seed distribution. Since PUMS data is stripped of detailed geographic information, the seed distribution for each TAZ is the distribution of the PUMA to which it belongs.

Figure 2: Basic inputs, processes and outputs of population synthesizer (PopSyn)



Once Balancer determines the distribution of households by HHCat within TAZ, then the second major step in PopSyn—HHDrawer—creates the SynPop by drawing, for each TAZ, the correct number of households of each HHCat from the PUMS households with matching HHCat and PUMA. Then, parcel level inputs on residential land use are used to further allocate households to parcels. Since the number of households determined by Balancer is fractional, HHDrawer is preceded by a procedure that ‘integerizes’ the IPF results, preserving the distribution as much as possible. Also, since the number of households within a particular HHCat for a given PUMA may be small, Drawer is set up to draw from similar PUMAs if the same household would otherwise be drawn more than a prescribed number of times. PUMA similarity and the maximum number of times that a household may be drawn is specified in the control file.

In summary, PopSyn creates the base year SynPop in two steps called Balancer and HHDrawer. Balancer is an iterative proportional fitting procedure that estimates the base year distribution of households by household demographic category (HHCat) for each TAZ. HHDrawer is a sampling procedure that populates each TAZ by drawing the correct number of households of each HHCat from census PUMS data. For the base year, PopSyn matches exactly the targets determined by census SF1, SF3 and CTPP tables at the TAZ level, while preserving to the extent possible the full multi-dimensional distribution observed in PUMS at the PUMA level.

Forecast year synthetic population

PopSyn uses the same two steps, Balancer and HHDrawer, to synthesize the population for a forecast year, but it uses regional forecasts from (PLACE3S or PECAS) as input instead of census data. Balancer creates a forecast population distribution that matches the following SACOG forecasts: (a) households by income category in each TAZ, (b) number of jobs held by employed persons living in each TAZ, (c) floorspace by housing type in each TAZ, (d) number of persons aged 65 and older in the region, and (e) number of persons aged 0-17 in the region. Like the base year, PopSyn’s forecast inputs come from input parameters in its control file, so it would be possible, without software programming, to fairly quickly and inexpensively adjust PopSyn to match other regional forecasts.

Since the available forecast year information can be quite limited, and the distribution of household and personal characteristics change gradually over time, Balancer is set up to preserve the base year distribution as much as possible while matching the above-described forecast control totals. That is, Balancer uses the base year distribution created by PopSyn as its seed distribution for the forecast year. However, since the distribution at the TAZ level of geography may not be very stable over time, Balancer’s seed distribution for each TAZ is a blend of the TAZ, census tract and PUMA base year distributions. The exact blend for each TAZ depends on the sizes of the TAZ and its tract, and is determined by easily changed parameters in the control file; the bigger the TAZ, the more heavily it weighs in the blend.

Long Term Choice Simulator and DaySim

Operationally, the Long Term Choice Simulator and DaySim are integrated into the same software program. One reason for this is that, unlike PopSyn, the Long Term choice outcomes are influenced by accessibility arising from daily activity and travel choices. For simplicity, subsequent discussion will treat the long term choice models as a part of DaySim. **Figure 3** presents the DaySim model hierarchy, embedded within the program looping structure in which the models will run. Program loops are bounded by lines starting with ‘Begin’ and ‘End’, and indentation indicates embedded sub-loops. The models themselves are numbered. For each household, the long term choice models (1.2-1.4) run first. Then, a loop runs for each person, in which their day pattern (models 2.1-2.2) is simulated. Within that loop, each tour of the pattern is simulated in turn (models 3.1-3.4), and each stop is simulated within each tour (models 4.1-4.4). Work-based tours are modeled as tours, but at the same level of priority as stops on the way to and from work.

The next subsections describe each of the model types. Additional details about each model can be found in tabular form in **Appendix 1**, including the model type, output variables, and important variables that it uses. **Appendix 2** provides a detailed list of variables produced by the DaySim models, including for each a reference to the model that produces it.

NOTE: The tech memos for the various models may supercede the following descriptions.

Figure 3—DaySim models (numbered) within the program looping structure

Begin

```
{Read run controls, model coefficients, TAZ data, LOS matrices,
    population controls, and Parcel data into memory}
{Draw a synthetic household sample if specified}
{Pre-calculate destination sampling probabilities}
{Pre-calculate (or read in) TAZ aggregate accessibility arrays}
{Open other input and output files}
{Main loop on households}
  {Loop on persons in HH}
    {Apply model 1.1 Work Location for workers}
    {Apply model 1.2 School Location for students}
    {Apply model 1.1 Work Location for students}
  {End loop on persons in HH}
  {Apply model 1.3 Household Auto Availability }
  {Loop on all persons within HH}
    {Apply model 2.1 Activity Pattern (0/1+ tours and 0/1+ stops)
      and model 2.2 Exact Number of Tours for 7 purposes}
    {Count total home-based tours and assign purposes}
    {Initialize tour and stop counters and time window for the person-day before looping on tours}
    {If there are tours, loop on home-based tours within person in tour priority sequence,
      with tour priority determined by purpose and person type}
      {Increment number of home-based tours simulated for tour purpose (including current)}
      {Apply model 3.1 Tour destination}
      {If work tour, apply model 3.2 Number and purpose of work-based sub-tours}
      {Loop on predicted work-based sub-tours and insert then tour array after current tour}
      {Apply model 3.3 Tour mode}
      {Apply model 3.4 Tour primary destination arrival and departure times}
      {Loop on tour halves (before and after primary activity)}
        {Apply model 4.1 Half tour stop frequency and purpose}
        {Loop on trips within home-based half tour (in reverse temporal order for 1st tour half)}
          {Increment number of stops simulated for stop purpose (including current)}
          {Apply model 4.2 Intermediate stop location}
          {Apply model 4.3 Trip mode}
          {Apply model 4.4 Intermediate stop departure time}
          {Update the remaining time window}
        {End loop on trips within half tour}
      {End loop on tour halves}
    {End loop on tours within person}
    {Write output records for person-day and all tours and trips}
  {End loop on persons within household}
{End loop on Households}
{Close files}
{Create usual work location flow validation statistics}
```

End.

Long term choice models

Work location (1.2) and School location (1.3)

These are essentially destination choice models, but they determine the longer term choice of usual work and school locations (parcel within TAZ). These, along with residence location, tend to structure a person's spatial activity patterns. The choice is primarily a function of travel accessibility across all modes and land use characteristics in and surrounding each possible TAZ and parcel. Key segmentation variables include income for workers and age group for students. In the model sequence, work location conditions the school location for most workers, but for university and young driving age students, school location conditions work location.

Auto availability (1.4)

This model is applied at the household level, and determines the number of vehicles available to the household drivers. Key variables are the numbers of working adults, non-working adults, students of driving age, children below driving age, income, auto and non-auto accessibilities to work and school locations, and more general pedestrian, transit and auto accessibility to retail and service locations.

Day level models

Day activity pattern (2.1-2.2)

This 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.

Tour level models

Within each tour, three main models are used, to first simulate the tour's destination, then the beginning and ending period of the tour's primary activity, and finally the main mode used for the tour. For work tours, the number of work-based subtrips is also modeled, after destination choice, and before timing and travel mode.

Destination choice (3.1)

Similar to the work and school location models, these models determine the primary destination TAZ and parcel for home-based tours and work-based subtrips. For the primary tour destination, the logsum from the mode choice model across all modes is used as the main level of service variable.

The universal choice set of destinations is very large, including all parcels within the metropolitan area. In any given situation, some of the parcels will be infeasible, either because the location cannot be reached in the available time, or because the desired activity cannot be accomplished there. Also, for the sake of computational feasibility, the huge size of the choice set makes it necessary to sample alternatives when applying the destination choice models. A sampling procedure has been designed to deal with both of these issues. The available alternatives are sampled in a way that allows the probability of being drawn into the sample to be calculated for each drawn alternative. Statistical procedures are then used during model estimation and application to allow the sample to represent the entire set of available alternatives without biasing the results.

The chosen sampling procedure is called two-stage importance sampling with replacement. In the first stage, a TAZ is drawn with a known probability approximately equal to its chance of containing the chosen destination. Then, a parcel is drawn within that TAZ with a known probability approximately equal to its chance of being the chosen parcel within the TAZ. The two main criteria used in the design of the procedure are statistical soundness and computational efficiency. A later technical memo on the location choice models will document these procedures in detail.

Number and purpose of work-based tours (3.2)

For this model, the work tour destination is known, so variables measuring the number and accessibility of activity opportunities near the work site are expected to influence the number of work-based tours.

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), with the alternatives being drive to transit, walk to transit, school bus, car shared ride 3+, car shared ride 2, car drive alone, bike and walk.

Primary activity periods (3.4)

The dependent variables of this choice model are a pair of 30 minute time periods representing the times that the person arrives at and departs from the tour primary activity location. It therefore provides an approximation of both time-of-day and activity duration. Since entire tours, including stop outcomes are modeled one at a time, first for work and school tours and then for other tours, the periods away from home for each tour become unavailable for subsequently modeled tours. The time period of a work-based subtour is constrained to be within the time period of its parent tour.

Trip/stop level models

Although the presence of extra (intermediate) stops in the day pattern is determined in the pattern model, the exact number of stops for each purpose is a result of the stop level models. Within each tour, the stops are modeled one-by-one, first for stops before the tour destination, and then for stops after the tour destination. This is an iterative model structure, very similar to the one used in Model 3.2 for the number and purpose of work-based subtours.

Stops before the tour destination are modeled in reverse temporal sequence. First the possible participation in a stop is modeled simultaneously with the stop's purpose (4.1). If the stop occurs, then its location (4.2), and then its trip mode (4.3), and finally the 10-minute time period of the arrival at the tour destination (4.4) are modeled. These results also determine the time period in which the trip from the stop location begins, since the trip mode and travel level of service are known. If a stop occurs, then the possible participation and purpose of a prior stop are modeled, along with details of location, trip mode and timing. This continues, constructing the trip chain from the tour primary destination to the tour origin in reverse chronological sequence until the model predicts no more stops (at which point, the "final" trip between the "last" stop and the tour origin is modeled). The reason for modeling in reverse chronological sequence for the first half tour is the hypothesis that people aim to arrive at the primary destination at a particular time, and adjust their tour departure time so as to enable completion of the desired intermediate stops. After the trip chain for the first half-tour is modeled, the trip chain for the second half-tour back to the tour origin is similarly modeled, but this time in regular chronological order.

Number and purpose of intermediate stops (4.1)

Throughout the construction of the trip chains, the making of intermediate stops by purpose is accounted for, so that as stop purposes called for by the pattern model are accomplished, the likelihood of additional stops decreases.

Intermediate stop location (4.2)

For intermediate stop locations, the main mode used for the tour is already known, so the choice is primarily a tradeoff between the additional deviation and impedance of making another stop by that mode versus the accessibility to additional land use opportunities in alternative zones and parcels.

As with tour destinations, a sampling procedure is required for the stop location models, and a procedure has been designed that employs importance sampling with replacement. The exact procedure is different, however, because the sampling problem is more complex. For intermediate stops, the travel impedance affecting choice is a function of three locations instead of two: the intermediate stop location, as well as locations before it and after it in the half tour. This expands the number of relevant impedances geometrically. Therefore, a 3-stage importance sampling procedure has been designed. For each parcel to be drawn, first a stratum is drawn,

then a TAZ within the stratum, and finally a parcel within the TAZ. A later technical memo on the location choice models will document these procedures in detail.

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 trip mode alternatives are more precisely defined than the tour mode alternatives. In particular, for transit modes, the access and egress modes are both explicitly identified, resulting in the following set of alternatives: drive-LRT-walk, walk-LRT-drive, walk-LRT-walk, drive-BUS-walk, walk-BUS-drive, walk-BUS-walk, schoolbus, car-shared ride (possibly split between 2 and 3+ passengers), car-drive alone, bike and walk. The tour and trip level mode choice models are estimated simultaneously to ensure the most significant and consistent values for key travel time and cost coefficients.

Trip timing (4.4)

For intermediate stop locations, this model predicts either the departure time (for stops on the 2nd half tour) or the arrival time (for stops on the 1st half tour). The use of travel time variables in this model and model 3.4 allows us to capture peak spreading effects for car tours and trips.

Accessibility variables

NOTE: The tech memo on accessibility variables supercedes the following description.

Accessibility variables are discussed separately in this memo for two reasons. First, they are very important because they capture the sensitivity of activity and travel decisions to the utility of opportunities associated with conditional (and hence undetermined) model outcomes. For example, in a destination choice model, a logsum variable can capture the expected utility of the available travel mode alternatives. This is a very important aspect of model integration. Without it, the model system will not effectively capture sensitivity to travel conditions. Second, 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, this section describes a carefully designed approach for capturing the most important accessibility effects with a feasible amount of computation. The approach involves using two basic techniques to substitute for a pure logsum in cases where the logsum computation is very costly and a substitute can provide much of the benefit. First, in some cases, an approximate logsum is used. This is a variable that is calculated in the same basic way as a true logsum, by calculating the utility of multiple alternatives, and then taking expectation across the alternatives by calculating the log of the sum of the exponentiated utilities. 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. Second, in some cases where the attractiveness of a location alternative may depend on the accessibility near the location for pursuing secondary activities, directly measurable attributes of

SACOG Activity-Based Travel Forecasting Model

Featuring *DaySim*—the Person Day Simulator

Technical Memo No. 1: **Model System Design**

the location are used instead of logsums for the potential secondary stops themselves. Such attributes include indicators of pedestrian friendliness and density of activity opportunities in the neighborhood.

The remainder of this section will discuss the accessibility variables used for each component of DaySim. However, one approximate logsum, which is pre-calculated and used by several of the model components, is explained first, and then referred to as needed in the subsequent discussion.

The approximate mode-destination choice logsum

This logsum 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 the day. Because of the large amount of computation required for calculating a true logsum for all feasible combinations in these three dimensions, an approximate logsum is used with several simplifications. First, it ignores socio-demographic characteristics, except sometimes it distinguishes between situations where a car is available and those where it isn't. Second, it uses aggregate distance bands for transit walk access. Third, sometimes it uses a logsum for a composite or most likely purpose instead of calculating it across a full set of specific purposes. Finally, instead of basing the logsum on the exact available time window of the choice situation, and calculating it across all of the available time period combinations within the window, it either uses a particular available time window size and time period combination, or a weighted average of attributes for several time period combinations. With these simplifications, it is possible to pre-calculate 39 approximate logsums for each TAZ, and use them when needed at any point in the simulation of any person's day activity schedule. **Table 2** lists the models in which this approximate logsum is used, along with the variations required in each of the four dimensions discussed above.

Table 2: Models using approximate logsums, and their approximating categories

Model	Car availability	Walk dist to transit	Purposes	Time window	Time period combo
Work location School location Tourdestination	Available	Short (<¼ mi), Medium (¼-½ mi), Long (½+ mi)	Composite nonwork	½ hr	Weighted avg of attributes across five 1-period time combos (early-early, am- am, MD-MD, PM-PM, Night-Night)
Stop location	Available	Short, Medium, Long	Composite nonwork	½ hr	Midday-Midday (mdmd) PM peak-PM peak (pmpm) Night-Night (ntnt)
Auto availability	Available Not available	Short, Medium, Long	Composite nonwork	2+ hrs	Weighted avg of attributes across the 15 time period combos
Activity pattern	Available	Short, Medium, Long	7 specific purposes	2+ hrs	Purpose-specific weighted avg of attributes across the 15 time period combos

Calculation of the approximate logsums and estimation of the parameters for the calculations

The set of 39 approximate logsums is calculated for each TAZ as follows:

Calculate aggregate size variables for a composite non-work purpose for three subgroups of parcels in each TAZ defined by three ranges of walk distance to transit (less than ¼ mile, ¼ to ½ mile and more than ½ mile). Estimate a non-work tour mode-destination choice model without

socioeconomic explanatory variables, using destination zones defined by the distance-range categories, and using survey tours with short available time windows. Calculate 12 logsums from this model for different assumed combinations of walk distance to transit at origin (short, medium, long) and time period combo (weighted avg, MD-MD, PM-PM, and Night-Night). For the weighted average, use mode-dest choice attributes composed as a weighted average, where the weights are the relative frequencies of each 1-period time combination.

Similarly, calculate size variables and estimate a mode-destination choice model using all non-work survey tours, regardless of available time window size. Calculate 6 logsums from this model for different assumed combinations of auto availability (available, not available) and walk distance to transit at origin (short, medium, long), using a weighted average of attributes across 15 time period combinations.

Similarly, calculate size variables and estimate a mode-destination choice model with purpose-specific parameters, using all tours except those to usual work or school location. Calculate 21 logsums from this model for different assumed combinations of purpose (7 purposes) and walk distance to transit at origin (short, medium, long), using a purpose-specific weighted average of attributes across 15 time period combinations.

NOTE: The exact number and definition of these aggregate logsums may be adjusted upon further analysis of the properties of the logsum variables.

Accessibility variables used in DaySim model components

Work location, school location, and auto availability

The work location model uses the individual's tour mode choice logsum from the home parcel to each of the sampled parcel locations, using the urban and transportation system attributes for a work tour, and a sample-based time-of-day weighted average of mode choice utility attributes across the 15 feasible time period combinations. This approach is far less time consuming to calculate than a two-stage logsum that calculates time-of-day and mode utility for all possible combinations. The school location uses an analogous logsum for school tours.

In order to capture the effect of accessibility for work-based sub-tours and intermediate stops near the work location, these models also use the approximate mode-dest logsum for nonwork tours originating at the work (or school) location, as well as attributes of the work (or school location) indicating pedestrian friendliness and density of nearby activity opportunities.

The auto availability model uses logsum accessibility to the chosen work location and school locations of workers and students, calculated with and without an auto available. These logsums can be derived from the logsums calculated for the work location and school location models, as long as the with and without auto components of the logsums are kept separate. The auto availability model also uses the approximate logsums for nonwork tours originating at the home location, and directly measured attributes of the home location and the usual work and school locations.

Activity pattern

The activity pattern model uses the logsums previously calculated for the chosen work and/or school location, to capture the effect of accessibility on the probability of going to work or school on any given day. For other purposes (and for work or school when there is no usual work or school location) the pattern model uses the purpose-specific approximate mode-dest logsums. The model also uses attributes of the residential, usual work and usual school locations to capture the accessibility for short tours or intermediate stops near those key locations.

Tours

In the tour model hierarchy, destination choice conditions time-of-day choice, which conditions mode. The destination and time-of-day models incorporate time-of-day variations in mode accessibility. For the time-of-day choice, 15 mode choice logsums are calculated, one for each of 15 begin and end time period combinations, using five aggregate time categories: before AM peak, AM peak, MIDDAY, PM peak, and evening. By assuming accessibility equivalence of the before AM and evening periods, the number of logsums would drop from 15 to 13. For the destination choice, the logsums use a sample-based weighted average of LOS attributes across all time period combinations.

Other attributes of the sampled destinations (such as distance to bus & LRT stops, and sidewalk density) are also used, to help capture the accessibility for short tours or intermediate stops in the neighborhood.

Stops

The measurement of accessibility at the stop level is fairly simple because it is at the lowest level of the model's conditional hierarchy. At this stage, the main destination, approximate time-of-day and mode of the tour are known, and the stop models determine the stop location, trip mode, and timing for each trip segment on both half-tours, from the tour destination back to the tour origin. So, the stop location model can use a direct measurement of travel times and costs for the tour's main mode to the sampled stop location. In addition it uses the approximate mode-dest logsum and other attributes to measure attractiveness of trips in the vicinity of the sampled stop. The trip mode model uses direct measures of times and costs, and the timing model is essentially a stop duration model, not dependent on accessibility.

Validation and Calibration

20060719: This section should be superseded by a document describing the actual validation and calibration procedures and results.

This section updates and expands the outline of model validation and calibration procedures described in our January 15 memo. It covers work in the current development project as well as in the subsequent implementation project. The validation and calibration procedures described here will assure that the model system properly replicates the best available estimates of the characteristics of the population and its activity and travel behavior on a typical weekday in the base year. (They do not validate the sensitivities of aggregate forecasts to changes in various factors affecting activity and travel., although that will need to be a part of any follow-on implementation project.)

Stage 1: Compare activity based model results and reported results for HH survey sample

Because activity-based microsimulation models operate on a record-by-record basis and put out the same variables for each person that are found in a household travel survey, it is possible to internally validate the results of the model implementation directly against the survey records that are used to estimate the models. The procedure is as follows:

- Before any models are estimated (by March 27), program the structure of the entire model system, including all input and output variables, but not yet including specific model utility functions or coefficients. Run parameters will allow the model system to be run for all or a specified subsample of either the survey sample or the synthetic population.

The demand model's application program with full input and output functionality (and a fairly realistic preliminary output data set) will be available to SACOG, Citilabs and any other consultants who can begin setting up any remaining input data (freight, externals, etc.), as well as the procedures for network assignment, equilibration, external calibration, validation, and visualization in parallel with our model estimation. Then, when they and we have completed those parallel activities, we can together proceed immediately into the actual external calibration and validation work.

- In the first pass, use the household survey sample records as the input to the program, with the survey providing both input and output variables. Using appropriate expansion weights on these records, this will provide an initial "naïve" forecast, as well as providing the basis against which estimated models are compared.
- As each model is estimated, program the model logic and utility functions to provide forecasts. The observed aggregate distribution across alternatives can then be compared to the predicted distribution, and the two will match virtually perfectly if the model has been estimated and implemented properly. By testing each model this way, both one at a time and

SACOG Activity-Based Travel Forecasting Model

Featuring *DAYSIM*—the Person Day Simulator

Technical Memo No. 1: **Model System Design**

in combination with other models, we can be confident that the models are performing correctly and avoid problems that might otherwise crop up later when they would be more difficult to isolate.

The observed and predicted distributions of the activity and travel diaries will be compared at all levels of the model system. **Table 3** provides a tentative list of the aggregate measures that will be compared at each level.

**Table 3: Possible Stage 1 validation checks
(observed versus predicted for the HH survey data)**

Level	Validation checks
Work Location	# workers by work end RAD by HH income # workers [regionally] by travel time to work mean travel time to work [by home end RAD] mean travel time to work [by work end RAD] OD matrix [RAD] by person type, HH type
School Location	# students by school end RAD by primary & other #students [regional] by travel time to school by person type mean travel time to school [by home end RAD] mean travel time to school [by school end RAD] OD matrix [RAD] by person type
Auto Availability	Compare by RAD and regionally: #HH by #vehicles avail, HH size & HH income #HH by #vehicles avail, #workers & HH income #HH by #vehicles avail and #age16+
Activity Pattern	#patterns by pattern type, person type, HH type and RAD #patterns by amounts of travel time, activity time, travel+activity time, person type
Tour	#tours by purpose, #stops, mode, time period, person type, HH type, RAD quartiles of tour distance by #stops, purp, mode, time period, person type, HH type, RAD tour OD matrix (RAD) by purpose, #stops, mode, time period, person type, HH type time arriving at work by work end RAD tour mode by purpose and subregion (Davis vs all other RADs) mode to work by HH income by work end RAD mode to work by HH income by home end RAD
Trip/ Stop	#stops by tour purp, stop purp, mode, time, person type, HH type, RAD quartiles of trip distance by tour purp, stop purp, mode, time, person type, HH type, RAD trip OD matrix [RAD] by tour purp, stop purp, mode, time, person type, HH type, RAD transit trip origins by RAD, transit mode, purpose transit trip destinations by RAD, transit mode, purpose daily transit trips, perhaps by submode

Stage 2: Implement and validate Population Synthesizer

Since the population synthesizer provides the population on which the activity based model acts, it must be implemented and validated first. Current plans call for the implementation and use of the flexible population synthesizer (PopSyn) that is under development for the Atlanta Regional Commission. (If the ARC PopSyn implementation is delayed or encounters difficulties, then we will implement a simpler, less flexible population synthesizer like those that were developed for Portland and San Francisco.) PopSyn includes a validation module (Validator) that calculates statistics, and presents them in graphical form, comparing the synthetic population (SynPop) to census table values for various household and person characteristics at various levels of geographic and demographic aggregation.

In the base year validation, PopSyn input will be the 2000 Census tables and the 2000 PUMS. Validation statistics will compare the resulting synthetic population to additional 2000 Census

tables, even along dimension that are not controlled for explicitly—e.g, more detailed age breakdowns or gender breakdowns, or household size distribution for large households. We can produce graphs comparing the synthetic population to the census data at various levels of aggregation:

- census tract (300 tracts in region)
- RAD (71 RADs, of quite different sizes)
- PUMA (xx PUMAs in region, all of approximately equal size)
- county (6 counties, of quite different sizes)

A more detailed description of this validation procedure will be documented in the forthcoming technical memo on the Population Synthesizer Design.

Stage 3: Compare activity based model results for synthetic population to reported results for expanded household survey

When the activity based models have been validated with the survey sample, and the synthetic population has been validated, we can substitute the synthetic population for the survey sample, and see how closely the new predictions match the expanded survey sample results at each step down the model chain. The aggregate statistics used for the comparison will be the same as in stage 1 validation. New discrepancies may appear if the synthetic population differs substantially from the expanded survey. Although it will be impossible to know for sure whether discrepancies are caused by weaknesses in the models or in the expanded HH survey distribution, this validation test may indicate the need to re-estimate some of the models or, possibly, to introduce calibration parameters to compensate for weaknesses in the survey sample.

Stage 4: Calibrate long term components of the activity based model system with census and SACOG counts

Stage 4 begins the formal calibration process. The base scenario predictions of the long term model components (work location choice, school location choice and auto availability) are compared to census data (work location, auto availability) and to SACOG inventories (school location). If significant discrepancies exist, then it may be necessary to include procedures that partially constrain the location choice models so that the model results match census totals. In application, the census constraints would be replaced by employment and school enrollment forecasts of the land use model. Alternatively, it may be possible to implement a few simple calibration factors. In the best case, the results would match closely enough without calibration.

Table 4 shows the tentative list of validation checks that will be used, and the corresponding census tables that will provide the control values.

**Table 4: Possible Stage 4 validation checks
(Long term outcomes—model predictions vs census and counts)**

Level	Validation checks
Work Location	# workers [by work end RAD] by HH income (ctpp2-30) # workers [regionally] by travel time to work (ctpp1-19) mean travel time to work [by home end RAD] (ctpp1-19) mean travel time to work [by work end RAD] (ctpp2-19)
School Location	# students by school end RAD by primary & other
Auto Availability	Compare by RAD and regionally: #HH by #vehicles avail, HH size & HH income (ctpp1-76) #HH by #vehicles avail, #workers & HH income (ctpp1-79) #HH by #vehicles avail and #age16+(ctpp1-68)

Stage 5: Calibrate integrated model system with expanded on-board survey, census, screenlines and other counts

After validation and calibration stages 1 through 4, the population synthesizer and activity based model will be ready for validation and calibration of the integrated model system. By this time, the CUBE-based model system application should be in place, integrating the population synthesizer, activity based model, commercial flows, special and external generators, and assignment models. A validation run will consist of running a base year 2000 scenario of the entire model system to an equilibrated state, and comparing aggregate results to the best available external information about the actual base year characteristics on a typical weekday.

We expect some substantial discrepancies to surface at this time. In particular, it is not until this stage that we will compare modeled results to screenline counts, the primary data that can help us compensate for underreporting of trips that probably occurred in the household travel survey. We will also have transit counts or expanded onboard survey results that may identify mode share discrepancies resulting from limited volumes of transit in the estimation data set. If estimates of bicycle OD flows or screenline counts are available, especially in Davis and at least one lower traffic area, then we will be able to calibrate these as well. Other discrepancies may also surface, such as time-of-day results between the counts and the model outputs.

Table 5 lists the types of calibration data that will be used, the corresponding validation checks that will be used, and the calibration action that might be employed to resolve discrepancies.

SACOG Activity-Based Travel Forecasting Model

Featuring *DAYSIM*—the Person Day Simulator

Technical Memo No. 1: **Model System Design**

**Table 5: Possible Stage 5 validation checks
(Travel outcomes—model predictions vs counts)**

Data Source	Validation checks	Calibration action
Screen line counts	vehicles crossing screenline by time period	Aggregate underreporting of trips across screenlines will require the use of calibration constants in the pattern model and/or stop frequency model to increase the incidence of trips in patterns. Uneven spatial or temporal distribution of discrepancies may require re-estimation of destination or time-of-day models to capture subregional idiosyncrasies
Transit on-board survey	transit trip Os by RAD, submode & purp transit trip Ds by RAD, submode & purp transit OD matrix by purp & time of day	Aggregate under- or over-reporting of transit trips relative to auto trips will require investigation of transit LOS calculations in data, and may require the use of mode-specific calibration constants. Submode-specific discrepancies may require re-estimation of mode choice models with submode-specific parameters. Subregional discrepancies will require investigation of transit LOS calculations subregionally. It may require re-estimation of mode choice models with subregion-specific variables, or the use of subregion-specific calibration constants.
Transit boarding counts	transit boardings by station or TAZ or RAD	See notes above for transit on-board survey.
Bike flow counts or OD survey	O-D flows for specific high bicycle traffic facilities, and OD pairs (if data is available)	
Pedestrian flow counts	O-D flows for specific high pedestrian traffic facilities and OD pairs (if data is available)	

SACOG Activity-Based Travel Forecasting Model

Featuring *DaySim*—the Person Day Simulator

Technical Memo No. 1: **Model System Design**

Appendix 1—DaySim Model Features

20060719: This table is probably out-of-date. It should be audited and updated if it is desired to include it somewhere in the final system documentation.

This appendix consists of a table listing all the basic models in DaySim. For each model it gives the model type, the dependent variables predicted by the model, and important other variables used in the model, especially spatial, temporal, and accessibility variables.

	Model	Model Type	Dependent Variables	Additional Variables Used in Model
1.1	Population synthesizer	Iterative Proportional Fitting (IPF)	Household variables Household ID Household size # in HH by Person type Number in HH employed Number in HH students Family/nonfamily code HH annual income Res location Person variables household ID Person ID Age Sex Person type Employment status Usual hrs worked per week Student status	TAZ-level marginals Parcel "sizes" based on dwelling types
1.2	Work location	Multinomial logit (MNL)	Work location (Parcel, TAZ, RAD,X,Y)	Sample of valid parcel locations Worker's work mode choice logsums from home parcel to sampled parcel locations, using weighted avg attributes across all time periods Non-work tour mode/dest. approximate logsum for sampled location (pre-calculated, generic nonwork purpose, for 1 of 2 transit accessibility levels within O parcel's TAZ, 1/2 hr time window, using weighted avg attributes across 5 one-period time windows) Size variables and other sampled parcel attributes, including..... for student over 15 or college student, school location conditions work location
1.3	School location	MNL	School location (Parcel, TAZ, RAD,X,Y)	Analog of 1.2 For other than student over 15 or college student, work location conditions school location

SACOG Activity-Based Travel Forecasting Model

Featuring *DAYSIM*—the Person Day Simulator

Technical Memo No. 1: **Model System Design**

	Model	Model Type	Dependent Variables	Additional Variables Used in Model
1.4	Household vehicles available	(MNL: 0, 1, 2, 3 or 4+ vehicles)	Number of vehicles available (0-4+)	<p>Workers' (all ftw & ptw) tour mode choice logsum to usual workplace with, and without, auto-based alts available, using weighted avg attributes across time periods</p> <p>Worker's (all ftw & ptw) parking price and walk accessibility [f(nodes w 4+ links,service+retail employment)]at usual workplace</p> <p>Student's (all uni & das) tour mode choice logsum to usual schoolplace with, and without, auto-based alts available, using weighted avg attributes across time periods</p> <p>Worker's (all uni and das) parking price and walk accessibility [f(nodes w 4+ links,service+retail employment)]at usual school loc</p> <p>Non-work tour mode/dest. approximate logsum for home (pre-calculated, generic nonwork purpose, for 1 of 2 transit accessibility levels within O parcel's TAZ, 2+ hr time window, with and without auto-based alts available, use weighted avg attributes across time periods)</p> <p>Residence parking price and walk accessibility [f(nodes w 4+ links,service+retail employment)]</p>
2.1	Activity pattern	Nested Logit (NL)	0/1+ tours and 0/1+ stops by 7 purposes	<p>Residence parcel attributes (walkability, densities, etc.)</p> <p>Usual work and/or school location parcel attributes (walkability, densities, etc.)</p> <p>Worker's work mode choice logsum from home parcel to usual work location, using weighted avg attributes across all time periods</p> <p>Student's school mode choice logsum from home parcel to usual school location, using weighted avg attributes across time periods</p> <p>Purpose-specific mode/dest. approximate logsums for home location (pre-calculated, 7 purposes, for 1 of 2 transit accessibility levels within O parcel's TAZ, 2+ hour time window, purpose-specific weighted avg attributes across time periods)</p>
2.2	Exact number of tours for 7 tour purposes	MNL or NL	Number of tours for 7 tour purposes	<p>Minimum # tours assigned to pattern so far, by purpose</p> <p>Residence parcel attributes (walkability, densities, etc.)</p> <p>Usual work and/or school location parcel attributes (walkability, densities, etc.)</p> <p>Worker's work mode choice logsum from home parcel to usual work location, using weighted avg attributes across all time periods</p> <p>Student's school mode choice logsum from home parcel to usual school location, using weighted avg attributes across time periods</p> <p>Purpose-specific mode/dest. approximate logsums for home location (pre-calculated, 7 purposes, for 1 of 2 transit accessibility levels within O parcel's TAZ, 2+ hour time window, purpose-specific weighted avg attributes across time periods)</p>
3.1	Tour destination	MNL	Tour destination (Parcel, TAZ, RAD,X,Y)	<p>Longest available contiguous time window.</p> <p>Sample of valid parcel locations</p> <p>Non-work tour mode/dest. approximate logsum for sampled location (pre-calculated, generic nonwork purpose, for 1 of 2 transit accessibility levels within O parcel's TAZ, 1/2 hr time window, using weighted avg attributes across 5 one-period time windows)</p> <p>Size variables of sampled parcels and corresponding size of 1/2 mile buffers</p> <p>Sampled parcel attributes: distance to bus & LRT stop, ped environment variables</p> <p>Purpose-specific mode choice logsums from home to sampled locations, using purpose-specific weighted avg attributes across time periods</p>

SACOG Activity-Based Travel Forecasting Model

Featuring *DAYSIM*—the Person Day Simulator

Technical Memo No. 1: **Model System Design**

	Model	Model Type	Dependent Variables	Additional Variables Used in Model
3.2	Number and purpose of work-based subtours	Recursive MNL [no more subtours or one more subtour with purpose p]	Tour participation (yes or no). Tour purpose (if yes). Tour ID (if yes)	Longest available contiguous time window for parent tour Non-work tour mode/dest. approximate logsum for tour destination (pre-calculated, generic nonwork purpose, for 1 of 2 transit accessibility levels within D parcel's TAZ, 1/2 hr time window using weighted avg attributes across 5 one-period time windows) Size variables of D parcel and corresponding size of 1/2 mile buffers D parcel attributes
3.3	Tour main mode	MNL or NL	Tour mode	For work tours: Number of work-based subtours TAZ-to-TAZ car, transit and walk/bike skim variables O and D parcel transit accessibility, pedestrian measures, etc. D parcel parking costs For WB tours, tour mode of parent tour
3.4	Primary activity start and end periods	MNL or NL	Arrival period at primary activity location (to within 30 minutes) Departure period from primary activity location (to within 30 minutes)	For work tours: number of work-based subtours Availability array of all 30 minute time periods across the day Travel time to and from tour dest by mode / time of day For WB tours, arrival and departure hour at tour destination
4.1	Stop participation and purpose	Recursive MNL [no more stops or one more stop with purpose p]	Stop participation (yes or no). If yes, then stop's purpose. Stop ID	Allocation deficit/surplus of stop purposes on pattern (conditioned by prior modeled tours and stops) Number of stops already assigned to pattern Stops by purpose already assigned to each segment of tour (before dest, work subtour, after dest) Accessibility around tour origin and adjacent prior modeled stop Tour/stop purpose interactions Tour mode Tour destination begin and end time periods (60 min) Stop departure time (10 min) from adjacent prior modeled location [if any] Available time window for tour segment between this stop and tour origin
4.2	Stop location	MNL	Stop location (Parcel, TAZ, RAD,X,Y) [if not last modeled trip in half-tour]	Sample of valid parcel locations Tour mode-specific LOS between adjacent prior modeled stop and tour origin (in direction of movement and through sampled location) Non-work tour mode/dest. approximate logsum for sampled locations (pre-calculated, generic nonwork purpose, for 1 of 2 transit accessibility levels within O parcel's TAZ, 1/2 hr time window in the time period determined by actual available time window, with auto-based alts available) Size variables of sampled parcels and corresponding size of 1/2 mile buffers Sampled parcel attributes: distance to bus & LRT stop, ped env vars Trip mode of adjacent prior modeled trip
4.3	Trip mode	MNL or NL	Trip mode	TAZ-to-TAZ car, transit and walk/bike skim variables O and D parcel transit accessibility, pedestrian measures, etc. D parcel parking costs Tour mode Mode used for trip from adjacent prior modeled location. Trip modes used for already modeled stops on tour (12 0/1 flags)
4.4	Intermediate stop arrival time	MNL or NL	Trip departure time (arrival time for 1st half-tour) Trip travel time and arrival (departure) time [determined by trip mode and departure time]	Availability array of all 30 minute time periods across the day Tour destination begin and end time periods (60 minute) Arrival time of adjacent prior modeled trip [if any] (departure time for 1st half-tour) Trip mode-specific travel times for different time periods (peak spreading model)

SACOG Activity-Based Travel Forecasting Model

Featuring *DaySim*—the Person Day Simulator

Technical Memo No. 1: **Model System Design**

Appendix 2—DaySim Variables

This appendix contains a table with the current complete list of DaySim variables, with the names as included in the estimation data files from the household survey. It includes the elemental variables and derived variables that will be output from DaySim application runs.

Variable	Description	Model (output) correspondence
SAMPN	household ID number	ID field
PERNO	person ID number	ID field
EXFAC	household expansion factor	1.1. Population synthesizer
PERSTYPE	person type	1.1. Population synthesizer
GEND	Gender	1.1. Population synthesizer
AGE	Age	1.1. Population synthesizer
WORKER	employment status	1.1. Population synthesizer
STUDENT	student status	1.1. Population synthesizer
HRSWORK	# hours worked per week	1.1. Population synthesizer
HHSIZE	# people in HH	1.1. Population synthesizer
TOTVEH	# vehicles in HH	1.4. Auto availability
INCOME	total household income	1.1. Population synthesizer
FAMILY	HH family status	1.1. Population synthesizer
HHWORK	# workers in HH	1.1. Population synthesizer
HHSTUD	# students in HH	1.1. Population synthesizer
HHFTW	# full time workers in HH	1.1. Population synthesizer
HHPTW	# part time workers in HH	1.1. Population synthesizer
HHRET	# retired adults in HH	1.1. Population synthesizer
HHNWA	# other adults in HH	1.1. Population synthesizer
HHUNI	# university students in HH	1.1. Population synthesizer
HHDAS	# driving age non-univ students in HH	1.1. Population synthesizer
HH515	# children age 5 to 15 in HH	1.1. Population synthesizer
HHCU5	# children under age 5 in HH	1.1. Population synthesizer
HHLOCID	household location id	ID field
HHATYPE	household urbanization type	1.1. Population synthesizer
HHRAD	household RAD	1.1. Population synthesizer
HHTAZ	household TAZ	1.1. Population synthesizer
HHXCORD	household X coord.	1.1. Population synthesizer
HHYCORD	household Y coord.	1.1. Population synthesizer
WULOCID	usual work location id	ID field
WUATYPE	usual work urbanization type	1.2. Workplace location
WURAD	usual work RAD	1.2. Workplace location
WUTAZ	usual work TAZ	1.2. Workplace location
WUXCORD	usual work X coord.	1.2. Workplace location
WUYCORD	usual work Y coord.	1.2. Workplace location
SULOCID	usual school location id	ID field
SUATYPE	usual school urbanization type	1.3. School location
SURAD	usual school RAD	1.3. School location
SUTAZ	usual school TAZ	1.3. School location
SUXCORD	usual school X coord.	1.3. School location
SUYCORD	usual school Y coord.	1.3. School location
HBTOURS	# HB tour records	2.1,2.2. Activity pattern
WBTOURS	# WB tour records	3.2. Work-based tour generation
HBWORKT	# home-based work tours	2.1,2.2. Activity pattern

SACOG Activity-Based Travel Forecasting Model

Featuring *DAYSIM*—the Person Day Simulator

Technical Memo No. 1: **Model System Design**

Variable	Description	Model (output) correspondence
HBEDUCT	# home-based education tours	2.1,2.2. Activity pattern
HBESCOT	# home-based escort tours	2.1,2.2. Activity pattern
HBPERBT	# home-based pers.business tours	2.1,2.2. Activity pattern
HBSHOPT	# home-based shopping tours	2.1,2.2. Activity pattern
HBMEALT	# home-based meal tours	2.1,2.2. Activity pattern
HBSOVRT	# home-based socialrecre tours	2.1,2.2. Activity pattern
INWORKS	# intermediate work stops	2.1. Activity pattern, 3.2 WB tour generation, 4.1 Stop generation
INEDUCS	# intermediate education stops	2.1. Activity pattern, 3.2 WB tour generation, 4.1 Stop generation
INESCOS	# intermediate escort stops	2.1. Activity pattern, 3.2 WB tour generation, 4.1 Stop generation
INPERBS	# intermediate pers.business stops	2.1. Activity pattern, 3.2 WB tour generation, 4.1 Stop generation
INSHOPS	# intermediate shopping stops	2.1. Activity pattern, 3.2 WB tour generation, 4.1 Stop generation
INMEALS	# intermediate meal stops	2.1. Activity pattern, 3.2 WB tour generation, 4.1 Stop generation
INSOVRTS	# intermediate socialrecre stops	2.1. Activity pattern, 3.2 WB tour generation, 4.1 Stop generation
TOURNO	tour ID number (in simulation order)	ID field
TOURSEQ	tour chronological sequence number (within HB or WB)	ID field
STOURNO	subtour ID within tour (in simulation order, 0=HB)	ID field
PRNTTOUR	subtour parent tour ID (0=HB)	ID field
SUBTOURS	# of subtours (of parent tour for WB)	3.2. Work-based tour generation
PDTYPE	primary destination purpose type	2.1,2.2. Activity pattern
ODEPART	tour origin departure time	4.4 Stop before timing
PDARRIVE	primary destination arrival time	3.4 Primary activity periods, 4.4 Stop before timing
PDDEPART	primary destination departure time	3.4 Primary activity periods, 4.4 Stop before timing
PDDURA	primary destination activity duration	3.4. Primary activity periods
OARRIVE	tour origin arrival time	4.4 Stop after timing
OLOCID	tour origin location id	ID field
OATYPE	tour origin urbanization type	1.1. Pop. syn. or 3.1. Work tour destination
ORAD	tour origin RAD	1.1. Pop. syn. or 3.1. Work tour destination
OTAZ	tour origin TAZ	1.1. Pop. syn. or 3.1. Work tour destination
OXCORD	tour origin X coord.	1.1. Pop. syn. or 3.1. Work tour destination
OYCORD	tour origin Y coord.	1.1. Pop. syn. or 3.1. Work tour destination
PDLOCID	primary dest. location id	ID field
PDATEYPE	primary dest. urbanization type	3.1 Tour destination
PDRAD	primary dest. RAD	3.1 Tour destination
PDTAZ	primary dest. TAZ	3.1 Tour destination
PDXCORD	primary dest. X coord.	3.1 Tour destination
PDYCORD	primary dest. Y coord.	3.1 Tour destination
MMODETOW	Work parent tour main mode	3.4. Tour main mode
MMODE	tour main mode	3.4 Tour main mode
TRIPSH1	# trips in first tour half	4.1. Stops before generation
TRIPSH2	# trips in second tour half	4.1 Stops after generation
TOURHALF	Trip tour half	ID field
TRIPNO	Trip ID within tour half (outward from PD)	ID field
SOTYPE	Trip origin* purpose type	2.1. Activity pattern, 4.1 Stop generation/purpose
SDTYPE	Trip destination* purpose type	4.1 Stop generation/purpose
SOLOCID	Trip origin* location ID	ID field
SOATYPE	Trip origin* area type	3.1 Tour destination, 4.2 Stop location
SORAD	Trip origin* RAD	3.1 Tour destination, 4.2 Stop location
SOTAZ	Trip origin* TAZ	3.1 Tour destination, 4.2 Stop location
SOXCORD	Trip origin* X coord.	3.1 Tour destination, 4.2 Stop location
SOYCORD	Trip origin* Y coord.	3.1 Tour destination, 4.2 Stop location
SDLOCID	Trip destination* location* ID	ID field
SDATYPE	Trip destination* area type	4.2 Stop location
SDRAD	Trip destination* RAD	4.2 Stop location
SDTAZ	Trip destination* TAZ	4.2 Stop location

SACOG Activity-Based Travel Forecasting Model

Featuring *DAYSIM*—the Person Day Simulator

Technical Memo No. 1: **Model System Design**

Variable	Description	Model (output) correspondence
SDXCORD	Trip destination* X coord.	4.2 Stop location
SDYCORD	Trip destination* Y coord.	4.2 Stop location
SOTIME1	Trip origin* 1-departure or 2-arrival time	4.4 Stop timing
SOTIME2	Trip origin* 1-arrival or 2-departure time	4.4 Stop timing
SDTIME1	Trip destination* 1-departure or 2-arrival time	4.4 Stop timing
SDTIME2	Trip destination* 1-arrival or 2-departure time	4.4 Stop timing
SMODE	Trip mode	4.3 Stop mode
HBWORKTM	# home-based work tours made (incl.current)	Incremented before model 3.1 for tour
HBEDUCTM	# home-based education tours made (incl.current)	Incremented before model 3.1 for tour
HBESCOTM	# home-based escort tours made (incl.current)	Incremented before model 3.1 for tour
HBPERBTM	# home-based pers.business tours made (incl.current)	Incremented before model 3.1 for tour
HBSHOPTM	# home-based shopping tours made (incl.current)	Incremented before model 3.1 for tour
HBMEALTM	# home-based meal tours made (incl.current)	Incremented before model 3.1 for tour
HBSOVRTM	# home-based socialrecr tours made (incl.current)	Incremented before model 3.1 for tour
INWORKSM	# intermediate work stops made (incl.current)	Incremented after model 3.2 for WB tour or 4.1 for int. stop
INEDUCSM	# intermediate education stops made (incl.current)	Incremented after model 3.2 for WB tour or 4.1 for int. stop
INESCOSM	# intermediate escort stops made (incl.current)	Incremented after model 3.2 for WB tour or 4.1 for int. stop
INPERBSM	# intermediate pers.business stops made (incl.current)	Incremented after model 3.2 for WB tour or 4.1 for int. stop
INSHOPSM	# intermediate shopping stops made (incl.current)	Incremented after model 3.2 for WB tour or 4.1 for int. stop
INMEALSM	# intermediate meal stops made (incl.current)	Incremented after model 3.2 for WB tour or 4.1 for int. stop
INSOVRTSM	# intermediate socialrecr stops made (incl.current)	Incremented after model 3.2 for WB tour or 4.1 for int. stop
W0300	Time window availability 0300-0329	Reset after 4.4 Stop timing
W0330	Time window availability 0330-0359	Reset after 4.4 Stop timing
W0400etc....etc....
.....etc for 48 30 minute periods of the dayetc....
W2600etc....etc....
W2630.	Time window availability 2630-2659	Reset after 4.4 Stop timing

Notes:

For the first half tour, the trip "origin" is actually the destination in the usual sense, and vice versa.

Also, for the first half tour, the 4 times (SOTIME1,SOTIME2,SDTIME1,SDTIME2) are in reverse chronological order.

The urbanization type variables are left in at the moment as placeholders for the parcel ID# when that is available.