SACSIM/05

Activity-Based Travel Forecasting Model for SACOG

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

Technical Memo Number 11

Impedance and Accessibility Effects

July 31, 2006 – Draft 1

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

Introduction

Upward Integration of Hierarchical Activity-based Models
or
Sensitivity to Impedance and Spatial Attributes in Activity-Based Models

A frequent critique of some trip-based models is that, for some aspects of travel choice that are sensitive to travel level of service (LOS), the travel demand models are NOT sensitive to travel level of service. This often includes trip generation and time-of-day aspects, and if the model is not properly equilibrated with the traffic assignment model, then it also includes trip distribution and mode choice. A second criticism, heard less often but still important, is that sometimes the techniques used to make the models sensitive to LOS are flawed, yielding inaccurate sensitivity to LOS. For example, if a destination choice or trip distribution model uses auto travel time and ignores transit time, then it is sensitive to auto LOS, but completely insensitive to transit LOS. It is even possible to construct composite accessibility measures that are sensitive in the wrong direction: improve transit service and transit mode share predictions fall. One argument in favor of properly constructed nested logit models has been that their logsum variable captures LOS effects at the upper level of a nested model in a way that takes into consideration all lower level alternatives and avoids counter-intuitive effects (Ben-Akiva and Lerman, 1985).

A similar critique can be made for models' sensitivity to spatial attributes, such as the distribution of employment, housing and other activity opportunities. Although proper sensitivity to these does not necessarily require equilibration with a traffic assignment model, since these attributes depend on land development processes, it still requires integration of the model components. For example, although a person or household's trip generation rates may depend significantly on the distribution of activity opportunities, model sensitivity to activity opportunities may be limited to the trip distribution or destination choice model. It is possible to use ad hoc measures, such as the density of activity opportunities of certain types within certain distances, to capture sensitivity of trip generation to activity opportunities. Here again, however, it is easy to make the models inappropriately sensitive in ways that bias predictions, and the use of logsum measures have been highly regarded as perhaps the best available means of capturing composite effects that cannot be measured directly in a model.

In recent years, activity-based models have been widely praised as being behaviorally more realistic than traditional trip-based models. They are supposed to achieve this by modeling aspects of choice that trip-based models ignore, by integrating the choices made by an individual over the course of a day, and in some cases, by integrating choices made by members of the same household. To the authors' knowledge, all current practical activity-based models with a behavioral foundation, including those under development, model the many aspects of choice by breaking the outcome into a conditional model hierarchy or a chain of models. Models lower in the hierarchy (or later in the chain) take as given the outcomes higher in the hierarchy. This achieves what has been referred to by Vovsha, Bradley and Bowman (2004) as downward vertical integrity.

Done properly, it assures that lower level models adhere to constraints imposed at higher levels, and makes the lower level models indirectly sensitive to all variables that directly affect the upper level outcomes.

Just as important as downward vertical integrity is upward vertical integrity (In this paper we ignore the likely problems caused by choosing an inferior hierarchy or forcing a network of interrelated choice aspects into a hierarchy or chain). Upward vertical integrity comes from making the upper level models appropriately sensitive to variables that affect the upper level outcome, but can't be measured directly because they differ among the undetermined lower level model outcomes. In formal nested logit hierarchies the upward integrity comes from the logsum, the composite measure of expected utility across the lower level alternatives.

One of the key contributions made by Bowman (1995) when he first developed a hierarchical model system representing a person's entire day, was his demonstration that the model of a person's choice of overarching day activity pattern can be made sensitive to transportation level of service, via logsum variables based on nested logit concepts. By doing that he also gave evidence that the choice of day activity pattern is indeed sensitive to transportation level of service. In other words, he demonstrated the need for upward vertical integrity and a way to achieve it.

Unfortunately, the strength of the logsum variable as a composite measure rests in a feature that makes it computationally expensive, and essentially infeasible with very large and detailed hierarchical model systems: it requires the calculation of utility for every single alternative in the hierarchy below the level being modeled. In order to model the highest level outcome, utilities of all alternatives in the entire hierarchy must be computed.

Therefore, none of the practical activity-based models implemented since Bowman's prototype (with the exception of the first models implemented at Portland METRO) have used logsums at the highest levels of the model system. Instead, they have resorted to the kinds of ad hoc measures that have been criticized in trip-based models, measures that tend to ignore or distort important indirect effects on upper level outcomes, most notably the effect of transportation level of service and land development attributes on the whole day activity pattern. For example, it is common to use a measure such as the number of jobs accessible within a certain amount of time by a certain mode, with separate variables calculated for auto and transit. However, the variables are correlated in the data, preventing the accurate identification of separate parameters; one of the parameters tends to dominate, and as a result the model gives too much weight to that mode, and ignores the effect of other modes.

Scant attention has been given to this problem since Bowman's prototype, and good solutions are still needed. Solutions far superior to the best current approaches should be possible. In private correspondence, Vovsha has suggested that in the context of a microsimulation that is iterated to achieve equilibration between demand and traffic assignment, it may be feasible to use actual utilities of the simulated chosen alternative from prior iterations instead of calculating the logsum every iteration, since the logsum represents the expected maximum utility, and the utility of the simulated outcome represents the simulated maximum utility. It may be appropriate to retain simulated outcome utlilities from all prior iterations and use a moving average to reduce random fluctuations and improve convergence. Another possibility that might work, again in the context of iterative microsimulation, is to use logsums but only re-simulate a fraction of the activity schedules, or only recalculate a fraction of the logsums, during each iteration.

In a model system called DaySim that has been developed for the Sacramento Area Council of Governments (SACOG), two other techniques are used in an effort to achieve better upward vertical integrity. The basic idea of the first technique used for SACOG is to avoid the use of a logsum when applying an upper level model by treating as given a conditional outcome that is not known, and would otherwise require the calculation of a logsum from all possible conditional outcomes. The assumed conditional outcome is selected by a Monte Carlo draw using approximate probabilities for the conditional outcome. Rather than making every simulated outcome sensitive to variability in the conditional outcome, sensitivity is achieved across the population through the variability of outcome in the Monte Carlo draws. This technique is used to include time-of-day sensitivity in the tour destination choice models, along with tour mode choice logsums. In this way, the destination choice models are sensitive to variations in transport level of service and spatial attributes across all possible combinations of time-of-day and mode, with the affects approximately weighted by the joint time-of-day and mode choice probabilities.

The basic idea of the second technique is to calculate an approximate, or aggregate, logsum. It 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. The approximate logsum is pre-calculated and used by several of the model components, and can be re-used for many persons. The categories of decisionmakers and the aggregation of alternatives are chosen so that in all choice cases an approximate logsum is available that closely approximates the true logsum. In essence, this is a sophisticated ad hoc measure that is intended to achieve most of the realism of the true logsum at a small fraction of the cost. Two kinds of approximate logsums are used, an approximate tour mode-destination choice logsum and an approximate intermediate stop location choice logsum.

The approximate tour mode-destination choice 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 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 for car availability. 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 uses a particular available time window size and time period combination. With these simplifications, it is possible to pre-calculate a relatively small number of logsums for each TAZ, and use them when needed at any point in the simulation of any person's day activity schedule.

The approximate intermediate stop location choice logsum is used in the activity pattern models, where accessibility for making intermediate stops affects whether the pattern will include intermediate stops on tours, and how many. Four logsums are calculated for each OD zone pair, distinguished by tour mode (transit or auto) and time of day (peak or offpeak). Each logsum is calculated across all possible intermediate stop zones, each stop's utility is a function of travel time and zonal attractiveness, and zonal attractiveness is a function of employment and school enrollment, taken from an estimated purpose-non-specific location choice model.

The following table lists the models in the SACOG model system, in numerical order from top to bottom of the conditional hierarchy. For each model, the table identifies how travel impedance and spatial attributes affect the model, including the use of direct measures, true mode choice logsums, simulated conditional outcomes, and aggregate logsums. The parameters estimated for these variables vary widely in their statistical significance and their level of impact on the model's predictions. At the time this white paper is being written, the model system is just beginning to be validated and sensitivity-tested, so the effectiveness of the techniques for achieving upward vertical integrity are not known. However, the estimation results provide some reason for optimism that the techniques provide an improvement over those that have been employed in other existing activity-based models.

References

Ben-Akiva, M., and Lerman, S. R. (1985). *Discrete choice analysis: Theory and Application to Travel Demand*, Cambridge, Massachusetts: MIT Press.

Bowman, John. L. (1995) *Activity based travel demand model system with daily activity schedules*, Master of Science Thesis in Transportation, Massachusetts Institute of Technology.

Vovsha P, Bradley M, Bowman J. (2004) Activity-Based Travel Forecasting Models in the United States: Progress since 1995 and Prospects for the Future. Paper presented at the *EIRASS Conference on Progress in Activity-Based Analysis*, Maastricht, The Netherlands.

Details of Implementation

Aggregate accessibility logsums are used for several upper level models in the system, as shown in the next to last column in Table 1. The form is that of mode-destination choice logsums to indicate the accessibility of various zones for non-mandatory activity purposes. To make it feasible to use such measures, they are pre-calculated for a limited number of segments. Those segments are each combination of:

Non-mandatory tour purpose:

(1) Home-based personal business

(2) Home-based shopping

(3) Home-based meal

(4) Home-based social/recreation

(5) Home-based escort

(6) All home-based purposes combined

(7) Work-based

Car availability segment:

(1) Child age under 16

(2) Adult in HH with no cars

(3) Adult in HH with cars, but fewer cars than drivers

(4) Adult in HH with 1+ cars per driver

Transit accessibility:

(1) Origin is within ¼ mile of transit stop

(2) Origin is more than ¼ mile from transit stop, but walk to transit is available

(3) Walk to transit not available

In total, this makes 7 * 4 * 3 = 84 combinations for each origin zone.

So, the simplified mode and destination choice models include only those variables that are defined by those segments. Other simplifications include:

- Only TAZ-based information is used, and no parcel-based land use information.
- Drive to transit, school bus and bike are all omitted, and shared ride is a single mode. This leaves 4 modes: WT – Walk to Transit, SR – Shared Ride 2+, DA – Drive Alone, and WK – Walk.

The resulting estimates for mode choice are shown in Table 8 of Tech Memo 4.

The application of these models has been programmed, and incorporated into a routine that calculates mode/destination choice logsums from every possible origin zone for each of the 84 segment combinations. This application code for precalculating the accessibility logsums essentially applies two steps of a 4-step zonal aggregate travel demand model system:

- Loop on origin zones
    - Loop on 84 tour purpose/car availability/transit accessibility segments
        - Loop on destinations zones and calculate mode choice utilities, mode choice logsums, destination choice utilities and accessibility logsum

A second routine also calculates intermediate stop logsums for car tours. As shown in the last column in Table 2.1, these measures are used in the pattern models (2.1 and 4.1) to make intermediate stops more likely between zone pairs where useful stop locations can be conveniently reached. This routine takes longer to run than the first one described above, because it uses 3 nested zone loops:

- Loop on time periods (peak, off-peak)
    - Loop on origin zones
        - Loop on destination zones
            - Loop on all intermediate stop zones and calculate intermediate stop location choice logsum using the formula below

Logsum is the log of the sum over all zones of : Size * exp (- 2 * extra time / 6.0 minutes)

Where Size is a weighted function of various attraction variables (the size variable function estimated for the composite non-mandatory tour purpose in the aggregate destination choice model), and Extra time is the auto travel time from the origin zone to the stop zone plus the auto travel time from the stop zone to the destination zone, minus the direct auto travel time from the origin zone to the destination zone (i.e. the detour time required to make the stop on the way from the origin to the destination).

Table 1: Impedance and spatial attribute effects in SACOG's DaySim activity-based model hierarchy

	Model	**Direct measures of travel impedance**	**Direct measures of spatial attributes**	**Tour mode choice logsum**	**Simulated conditional outcomes**	**Aggregate tour mode-destination choice logsum**	**Aggregate intermediate stop location choice logsum**
1.2	Usual Work Location	Distance. Distance from school.	Employment, enrollment, households. Parking & employment mix. Grid connectivity.	Yes.		At destination.	
1.3	School Location	Distance.	Employment, enrollment, households.	Yes.		At destination.	
1.4	HH Auto Availability	Distance to transit stop.	Parking price near home. Commercial employment near home.	To work. To school.		At home.	
2.1	Day Activity Pattern		Mixed use density near home. Intersection density near home.	For work & school.		At home.	Yes.
2.2	Number of Tours (by purpose)			For work and school tours.		At home.	
3.1	Tour Destination	Distance. Distance from work. Distance from school.	Employment, enrollment, households. Parking & employment mix Grid connectivity.	Yes.	Primary activity periods	At destination.	
3.2	Number & purpose of work-based tours		Commercial employment near work. School enrollment near work.				
3.3	Tour Primary Activity Timing (begin and end time periods)		Mixed use density	Yes.			
3.4	Tour Mode	All LOS variables	Parking costs. Transit accessibility. Mixed use density.				
4.1	Number & Purpose of Intermediate Stops		Grid connectivity X commercial employment at tour dest.				For auto-based tour modes.
4.2	Stop Location	Generalized time. Distance. Distance from tour origin. Distance from tour destin.	Employment, enrollment, households. Parking & employment mix.				
4.3	Trip Mode	All LOS variables	Parking costs. Transit accessibiity.				
4.4	Trip Departure Time	Travel times					