Travel Model Two Development: Calibration and Validation Data Summary Needs

Technical Paper

Metropolitan Transportation Commission with Parsons Brinckerhoff, Inc.

June 12, 2014

Table of Contents

List of Figures

List of Tables

1 Introduction

MTC is rebuilding the representation of demand in our travel model. This follows an overhaul of the representation of supply, which is still in progress[1]. The demand development work will first adapt existing model structures from peer agencies, then calibrate these structures to Bay Area conditions and assess their performance. This first step will result in a modeling system referred to as *Travel Model Two*. Next, we will use information gleaned from the development process to design the next version of the demand models.

This technical paper identifies the specific data summaries needed to calibrate and validate the *Travel Model Two* system. The document is organized as follows. In the next section, potential calibration and validation datasets are catalogued. Then, the proposed calibration summaries are presented. Validation summaries are then presented, followed by a discussions of how longitudinal and cross-data (source) analysis will be carried out. The GPS data collected as part of the home interview survey is then discussed. Commentary on the approach to modeling transit follows. A discussion of potential model enhancements concludes the document.

[1] Documentation on the progress is available here: http://analytics.mtc.ca.gov/foswiki/TravelModelTwo/Development.

2 Datasets

Travel Model Two will be calibrated to year 2010 conditions and validated against year 2000, 2005 and 2010 conditions. Additional calibration checks will be made against 2000 results if concerns regarding over-fitting emerge. The remainder of this section identifies datasets that will be used in this development effort.

Household travel survey data

The primary source of calibration data will be the California Household Travel Survey conducted by the California Department of Transportation (Caltrans)[2]. This dataset will be used to derive calibration targets for most of the model components including work and school location choice, daily activity pattern, tour frequency, destination choice and time-of-day. The survey field work was conducted from February 2012 to February 2013. The dataset includes about 9500 Bay Area households, of which 3500 were equipped with wearable GPS devices.

Census Data

Census data will be used to create observed distributions of automobile ownership levels and county-to-county worker flows that will be used in both model calibration and validation. This information will be derived from the American Community Survey for the 2010 calibration/validation and the 2005 validation, and from the Decennial Census for the 2000 validation.

Transit on-board survey data

Because transit usage is relatively rare in the Bay Area, the household travel survey will have a small number of records for each market segment of interest. Transit on-board surveys will be used to better inform the share of travelers using transit by access mode, tour purpose, and stop purpose. MTC has recently established a transit on-board surveying program, which aims to survey the vast majority of the Bay Area's transit riders. Table 1 summarizes the effort to date.

---

[2] Additional information is available here: http://www.dot.ca.gov/hq/tsip/otfa/tab/chts_travelsurvey.html.

For those operators who have yet to be surveyed as part of the MTC program, we will try and obtain the best available data to inform the year 2010 calibration. Two key "outside" (of the MTC program) sources include the 2006 San Francisco MUNI and 2008 BART surveys.

Table 1: MTC On-Board Survey Program Schedule and Data Availability

Agency	Technology*	Survey Date	Data Available
Alameda County (AC) Transit	Local, Express	Fall 2012	Yes
County Connection	Local, Express	Spring 2012	Yes
Golden Gate Transit	Local, Express, Ferry	Fall 2013	Yes
Livermore Amador Valley Transit Auth. (LAVTA or Wheels)	Local, Express	Fall 2013	Yes
Petaluma Transit	Local	Spring 2012	Yes
San Mateo County Transit (SamTrans)	Local	Spring 2013	Yes
City of Santa Rosa Transit (CityBus)	Local	Spring 2012	Yes
Sonoma County Transit	Local	Spring 2012	Yes
Union City Transit	Local	Spring 2013	Yes
Water Emergency Transportation Authority (WETA)	Ferry	Fall 2013	Yes
Napa County Transit (Vine)	Local	Spring 2014	No
Altamont Corridor Express (ACE)	Commuter	Spring 2014	No
Tri Delta Transit	Local	Spring 2014	No
Bay Area Rapid Transit (BART)	Heavy	Fall 2014	No
Caltrain	Commuter	Fall 2014	No
SF Muni	Local, Light	TBD	No
Santa Clara Valley Transportation Authority (VTA)	Local, Express, Light	TBD	No

* – Local is local bus service; Express is express bus service; Light, Heavy, and Commuter refer to types of rail service

Traffic count data

Traffic count data will be used to perform link level validation of assigned traffic volumes. The Caltrans Performance Evaluation and Monitoring System (PeMs) data, which monitors freeways, are available for the 2005 and 2010 model years. Other counts are provided by Caltrans for state-owned facilities from permanent station counts for each model year (2000, 2005, and 2010). Also, local jurisdictions provide counts as part of the federal highway performance monitoring system (HPMS). Many of the HPMS counts will need to be located on the model network before they can be used in validation.

Transit boarding data

Transit boarding data is available for all three validation years from most operators, often at the line level of detail. BART and Caltrain station-to-station flows are also available for each validation year. In addition to assessing the performance of transit model components, these will be used to verify the accuracy of the transit on-board survey expansion factors. Clipper card data can be used – by establishing a minimum – to verify the across-operator transfers revealed in the on-board surveys.

Parking data

The SF Park database can be used to verify estimated parking prices in San Francisco. BART should be able to provide parking demand estimates for the year 2010 validation effort.

Table 2 presents all the different data sources identified for use in calibration and validation for the three model years.

Table 2: Core Datasets for use in Calibration and Validation

Dataset Type	Year 2000	Year 2005	Year 2010
Household travel survey	Bay Area Travel Survey 2000	n/a	2012-2013 California Household Travel Survey
Transit on-board survey	n/a	n/a	2012-2014 MTC Transit Survey Program; 2006 Muni Survey; 2008 BART Survey
Census data	Decennial CTPP	ACS 2006-2010	ACS 2008-2012
Traffic count data	Caltrans	PeMS; Caltrans	PeMS; Caltrans; HPMS
Transit boarding data	Operator-provided	Operator-provided	Clipper; Operator-provided
Toll Bridge and HOT Origin	n/a	n/a	FasTrak
Parking	n/a	n/a	SFPark; BART Lots

3 Calibration Summaries

Travel Model Two is an implementation of the coordinated travel regional activity-based modeling platform (or CT-RAMP). The model structure and coefficients are largely borrowed from the regional model recently developed for the San Diego Association of Governments. Figure 1 graphically depicts the model structure. In general, the calibration process involves iteratively adjusting the alternative-specific constants in each model such that predictions match the expected outcomes. Importantly, this process in not done blindly. Rather, it is done with full consideration of retaining appropriate model sensitivity and behavioral realism. Further, calibration, validation, and sensitivity testing will be done iteratively, with refinements made throughout.

Table 3 presents the data comparisons expected for each model component.

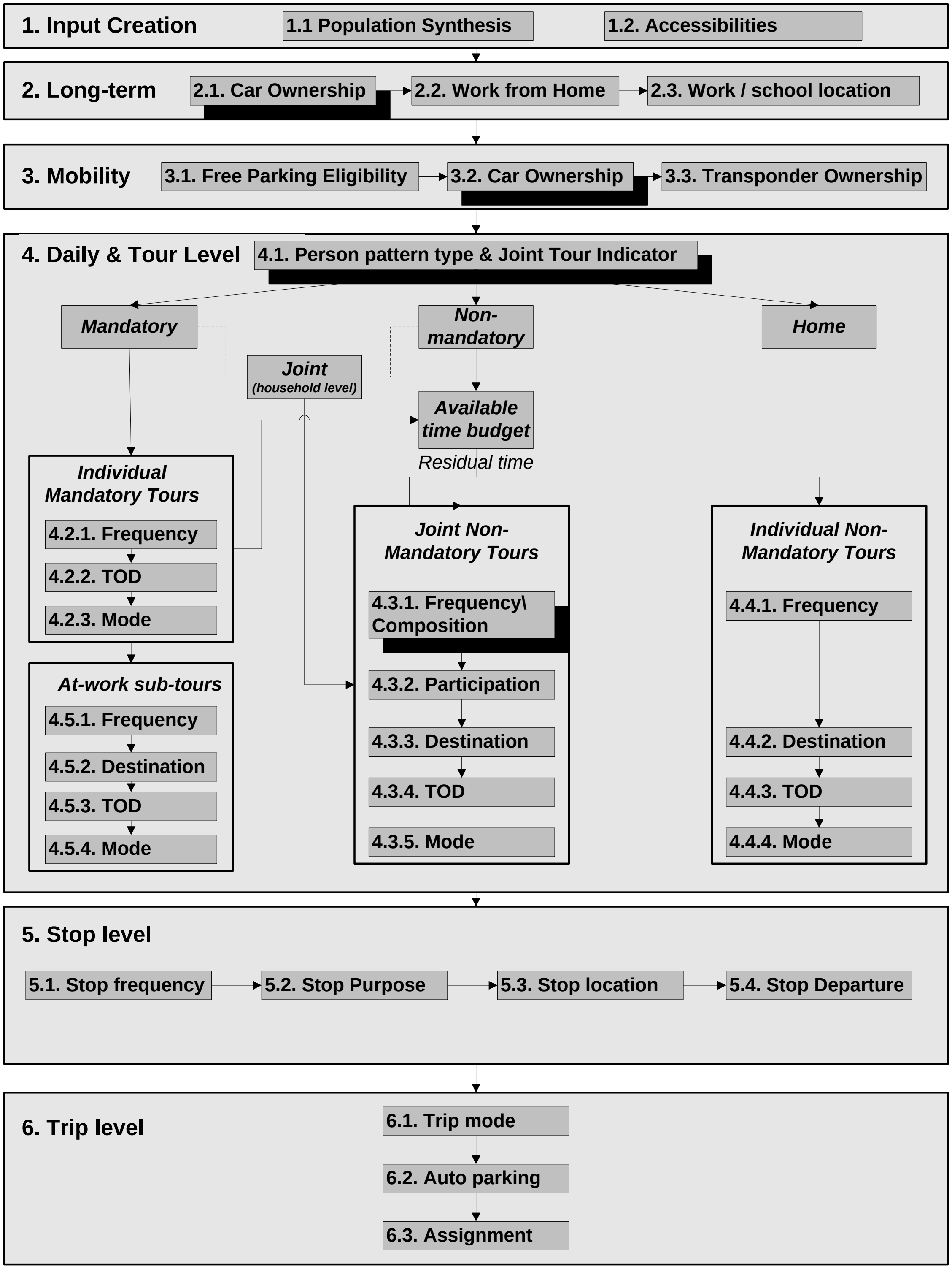

Figure 1: MTC Travel Model Two Design

Table 3: Data Summaries needed for Model Calibration

Model	Summary	Datasets
Work Location Choice	Average distance and frequency distributions for all workers	Household travel survey
	Average distance and frequency distributions for workers with a designated work location outside the home	Household travel survey, CTPP, ACS 5-year PUMS
	Frequency distributions for workers by county of residence and county of work	Household travel survey, CTPP, ACS5-year PUMS
	Scatter plot of residence PUMA by work PUMA	Household travel survey, CTPP, ACS5-year PUMS
	Distance frequency distributions by distance from regional CBD	Household travel survey, CTPP, ACS5-year PUMS
University, School Location Choice	Average distance and frequency distributions by grade category	Household travel survey
	Students by county of residence and county of school	Household travel survey
Automobile Ownership	Households by automobile ownership category	ACS 5-year PUMS
	Households by automobile ownership category and number of workers	ACS 5-year
	Households by automobile ownership category and household income	ACS 5-year
	Households by automobile ownership category and PUMA	ACS 5-year PUMS
	Zero automobile households by tract	ACS 5-year
Parking Reimbursement	Workers by level of parking reimbursement and PUMA	Household travel survey
Coordinated Daily Activity Pattern	Number and share of persons by person type	Household travel survey, ACS 5-year PUMS
	Share of persons by activity pattern and person type	Household travel survey
	Share of households by presence of fully joint tours and household size	Household travel survey
Mandatory Tour Generation	Share of mandatory tour generation alternatives by person type	Household travel survey
Fully Joint Tour Generation, Composition, and Participation	Share of joint tour generation alternatives (frequency and purpose combinations)	Household travel survey
	Tour frequency shares by household size	Household travel survey

Model	Summary	Datasets
	Share of tour composition alternatives by household size	Household travel survey
	Share of tour composition alternatives by number of persons participating	Household travel survey
Individual Non-Mandatory Tour Generation	Share of tours by purpose, frequency, and person type	Household travel survey
	Number of tours by person type	Household travel survey
Non-Mandatory Tour Location Choice	Average distance and distance frequency distribution	Household travel survey
	Tours by origin and destination county	Household travel survey
Tour Time-of-Day Choice	Share of tours by departure, arrival, and duration half-hour period and purpose	Household travel survey
Tour Mode Choice	Tours by purpose, mode, and automobile sufficiency	Household travel survey, Transit on-board survey
	Tours by purpose, mode, origin county, and destination county	Household travel survey, Transit on-board survey
Stop Frequency	Tours by number of outbound and inbound stops and tour purpose	Household travel survey
	Trips per tour and person type	Household travel survey
	Stops by tour mode	Household travel survey
Stop Location Choice	Stops by tour purpose and out-of-direction distance	Household travel survey
	Stops by distance to tour origin and primary destination	Household travel survey
Stop Purpose and Stop Duration	Fixed, observed distributions	Household travel survey
Trip Mode Choice	Trips by tour purpose, tour mode, and trip mode	Household travel survey, Transit on-board survey
	Transit trips by access mode and trip distance	Household travel survey, Transit on-board survey
	Transit trips by origin county, destination county, access mode, and line-haul mode	Household travel survey, Transit on-board survey

4 Validation Summaries

Table 4 presents the data summaries to be used in validation.

Table 4: Data Summaries needed for Model Validation

Model	Summary	Datasets
Roadway Assignment	Vehicles by facility type and area type	PeMS, Caltrans, HPMS
	Observed versus estimated scatter plots by time of day	PeMS
	Percent route mean square error by facility type and area type	PeMS, Caltrans, HPMS
	Key location summaries (county lines, other screen-lines, bridges)	PeMS, Caltrans, HPMS
Transit Assignment	Transit boardings by route and operator	Transit operator boardings
	BART and Caltrain station-to-station movements	Transit operator boardings
	Transfer rate check (2010 only)	Transit on-board surveys, Clipper
	Parking utilization at rail stations	BART Parking data
Additional Checks	Toll payment by Zip Code of residence (2010 only)	FasTrak
	Number of stops per tour by tour purpose	Household travel survey
	Average number of tours per person by person type	Household travel survey

5 Longitudinal analysis

The availability of household travel surveys for multiple years provides the opportunity to perform cross-survey comparisons in an attempt to understand how travel behavior has changed over time. The 2000 Bay Area travel survey has already been coded in the appropriate formats for performing calibration summaries, though certain differences between the *Travel Model One* and *Travel Model Two* model structures will require some additional survey processing. The 2012/13 California household travel survey will need to be transformed into appropriate formats, and tables can be created comparing various household travel summaries. Multiple years of Census data can be utilized in a similar fashion to spot meaningful trends in automobile ownership levels and/or worker flows.

Table 5 and Table 6 identify the different data summaries that will be performed to understand longitudinal trends in travel behavior using the survey and the Census datasets respectively.

Table 5: Data summaries for analyzing longitudinal trends in travel behavior using household travel survey

Dimension	Summary
Work and School Location Choice	Share of workers who work from home
	Average work distance and distance category frequency distributions
	Distance frequency distributions by distance from regional CBD
	Average school distance and distance category frequency distributions
Daily Activity Pattern	Share of persons by activity pattern and person type
Tour Frequency	Share of mandatory tour generation model alternatives by person type
	Share of non-mandatory tours by purpose, frequency, and person type
	Total number of individual non-mandatory tours by person type
	Total tours by person type and average number of tours per person
Tour Destination Choice	Average distance and distance category frequency distributions
	Tours by residence county and destination county
Tour Time of Day	Share of tours by departure, arrival, and duration half-hour period and purpose
Tour Mode Choice	Tours by tour purpose, mode and auto sufficiency
	Tours by tour purpose, mode and origin/destination county
Intermediate Stop Frequency	Share of tours by number of outbound and inbound stops and tour purpose
	Stops per tour by tour purpose
	Trips per tour and person type
	Stops by tour mode
Stop Location	Stops by tour purpose and out-of-direction distance
	Intermediate stops by distance to tour origin and destination
Trip Mode Choice	Trips by tour purpose, tour mode and trip mode

Table 6: Data summaries for analyzing longitudinal trends in travel behavior using Census data

Dimension	Summary
Work Location Choice	Workers by county of residence and county of work
Automobile Ownership	Households by automobile ownership category
	Households by automobile ownership category and number of workers
	Households by automobile ownership category and household income
	Households by automobile ownership category and PUMA

6 Cross-data comparisons

The different datasets identified should be compared across sources for similar years to ensure that data sets are consistent and to judge the reliability of calibration summaries when there are differences.

The following comparisons will be undertaken to ensure that there is no inherent bias in any of the calibration targets or model inputs:

1. County-to-county worker flows from the ACS data and 2012/13 household travel survey;
2. Employment by PUMA from socio-economic data from ABAG and the ACS data;
3. Share of workers who work from home by county form ACS data and the 2012/13 household travel survey; and,
4. Transit trips from the on-board surveys and 2012/13 household travel surveys.

In addition to the summaries described above, it is also useful to compare data summaries from MTC to other regions to ensure that calibration targets are within reasonable ranges.

7 GPS Day Pattern, Tour and Stop Comparisons

The availability of a GPS dataset for a sub-sample of 2012/13 household travel survey households provides the opportunity to determine whether there is systematic under-reporting of tours and/or stops, and develop corrections for such under-reporting. The analysis involves linking the reported (via the web or phone) place data and the GPS place data thus creating a unified tour file. Missing tours or stops can then be easily counted based on inconsistencies between the two data sources. For ease of creation of the unified tour file it is important that the GPS sub-sample and the CATI-retrieved dataset have similar data structures and a common place ID which can be used to link the two datasets.

8 Modeling Transit Path Choice

There are two important differences between how transit path choice is modeled in *Travel Model Two* relative to *Travel Model One*: application software and the path choice procedures. Each is discussed below, followed by a detailed description of the *Travel Model Two* approach.

Application Software

The *Travel Model One* and *Travel Model Two* modeling systems rely on custom-written CT-RAMP application software and commercial Citilabs (Cube) application software. In *Travel Model One*, transit paths are created completely in Cube, selected in CT-RAMP (as part of the mode choice models), and assigned in Cube. In *Travel Model Two,* the paths will be created in Cube and CT-RAMP, selected in CT-RAMP, and assigned (to transit routes) in Cube. Cube will handle the TAP-to-TAP path building and assignment; CT-RAMP will handle the rest.

In *Travel Model Two*, each transit path has three components:

- Access movement from origin MAZ to first-boarding TAP;
- Ride movement from first-boarding TAP to last-alighting TAP;
- Egress movement from last-alighting TAP to destination MAZ.

Path Choice Procedures

In *Travel Model One*, up to ten transit paths are presented to each traveler as part of the mode choice models. These paths are differentiated by access/egress mode (walk only has five paths, drive to or from transit has the other five) and technology preference (five technology preferences under each of the two access/egress categories). The technology preference applies different weights to different paths in an effort to reveal a plausible way through the transit network in which certain transit technologies are utilized. For example, the first path excludes all transit service except local bus service. This approach is likely to reveal a path through the network that includes only local bus. Next, light rail and ferry service is added to the network, and these technologies are revealed to the traveler as moving a bit faster than they actually do, thus increasing the likelihood that a plausible path involving light rail and/or ferry service is identified. Next, express bus is added to the network, artificially made to look faster, and a third way through the network is found. The fourth and fifth paths repeat this process with heavy rail and commuter rail, respectively.

The primary benefits of this approach are as follows: (a) duplicate paths can be quickly identified and removed (e.g., if the in-vehicle time for express bus is zero for the express bus path, it is eliminated); and, (b) path weights can be tuned, as needed, to improve validation results (e.g., if no ferry paths are revealed to travelers, weights and/or ferry's place in the hierarchy can be adjusted).

The drawback of this approach is that it explicitly assumes that competition across transit technologies improves a traveler's perception of transit but competition within transit technologies does not.

As described in more detail in the next section, *Travel Model Two* will attempt to build, select, and assign transit paths without relying solely on technology labels to identify different ways through the transit network. We propose to implement an approach in which multiple transit paths are enumerated for each origin-destination micro-zone pair. The initial approach will utilize revised software which eliminates technology labels, while still potentially utilizing skimming techniques that rely upon a modal hierarchy. This approach is described in more detail below. Ultimately, the modal hierarchy will be replaced with a software technique that ranks all potential paths according to their utility and exposes the transit choice (and higher-level choices) to a given set of ranked paths.

Travel Model Two Approach

The detailed implementation approach for *Travel Model Two* is as follows.

Step One: Identify the origin and destination TAP sets

For the origin MAZ, a set of potential origin TAPs are identified using a distance threshold. TAPs that provide inferior access to a route accessible by another route are eliminated (e.g., travelers who can walk one block to access route A are not permitted to walk three blocks to access route A). Figure 2 presents an example in which a single origin/destination pair has a choice of three boarding TAPs (numbered 1, 2, 3) and three alighting TAPs (numbered 4, 5, 6).

Step Two: Identify TAP-to-TAP paths

For each possible boarding-TAP-to-alighting-TAP pair, Cube is used to build N potential paths through the network. The goal in designing the path weights are to capture heterogeneity in path preferences revealed in on-board surveys and aid model validation by revealing specific-types of paths to travelers. Traditional technology-based weights may be introduced as needed, but will not be relied on exclusively. Rather, we will attempt to create plausible paths that reveal different preferences for transferring, waiting, and using more reliable services. In Figure 2, three paths are shown for each TAP pair and identified with separate colors. Note that the path weights will be used identify the best way through the network for each path category; the path costs will then be computed with a set of coefficients that are consistent across path categories.

Step Three: Remove duplicate TAP-to-TAP paths

Moving away from technology-based path weights makes the task of identifying and removing duplicate paths a bit more difficult. It is possible that the path weights we specify may result in two or more identical TAP-to-TAP paths. If the level-of-service matrices are identical for two paths, we will consider one redundant and remove it from the set of potential paths. This step will be done in CT-RAMP or Cube.

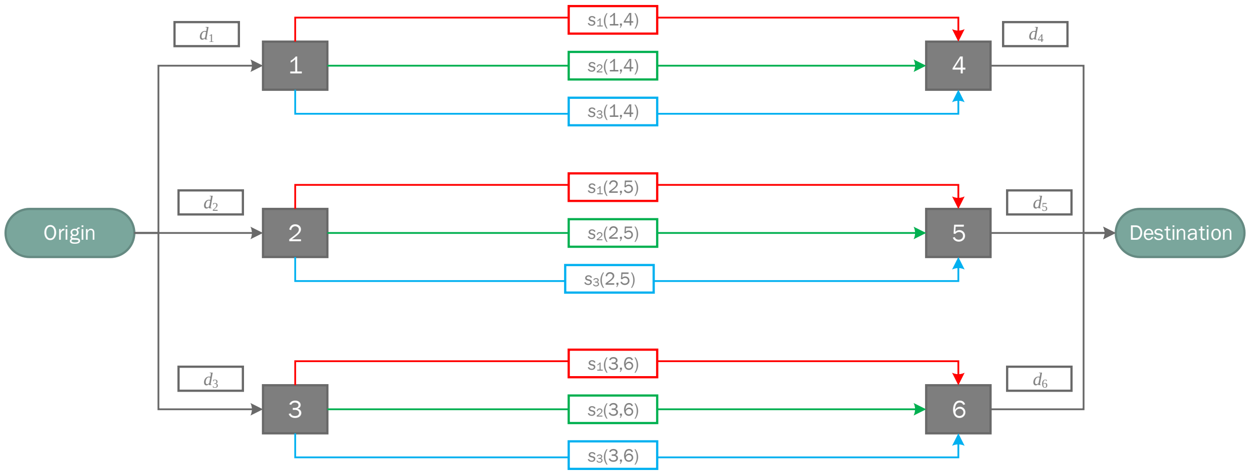

Figure 2: Transit Path Choice Example

Step Four: Identify the N-best origin to destination paths

In CT-RAMP, the best origin MAZ to first-boarding TAP to last-alighting TAP to destination MAZ will be selected for each of three skim sets. This will be done by specifying one or more sets of "cost" parameters and identifying the unique best path across the set that includes all potential TAP-to-TAP pairs. The selected three paths, if available, will then populate the path choice level of the mode choice models under each access/egress nest (e.g., walk, bike, park, or dropped off/picked up).

To illustrate this step, Figure 3 populates Figure 2 with costs, noting separate costs for each MAZ-to-TAP, TAP-to-TAP, and TAP-to-MAZ movement. For TAP pair (1,4), the three different paths are unique; for TAP pairs (2,5) and (3,6), two of the three paths are duplicates and are removed. The result is shown in Figure 4. It is important to note that the selection of which path to drop is arbitrary – e.g., in the (2,5) set, either the red or the green path in Figure 3 can be eliminated.

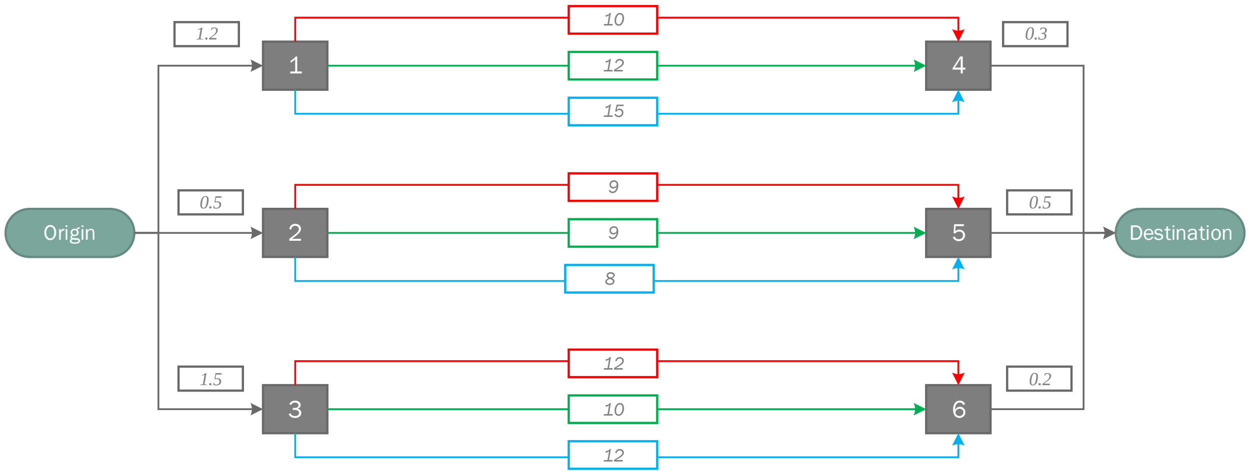

Figure 3: Three path sets between each TAP pair

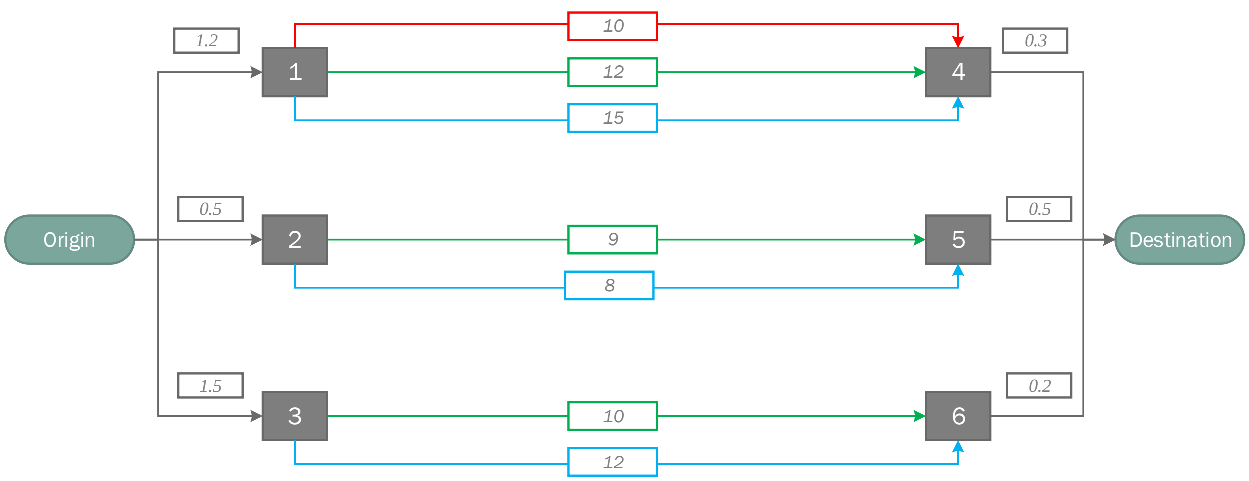

Figure 4: Unique path sets between each TAP pair

The final step involves evaluating the remaining path costs and retaining the best values for each skim set. Ultimately, we would like to implement a procedure in CT-RAMP in which the best N paths are retained, regardless of the skims that the paths were found in. This would require significant software changes, but would provide the full range of transit options and most closely replicate transit traveler information systems such as Google transit paths.

Verifying network accuracy

The availability of on-board survey data for multiple operators in the Bay Area provides an opportunity to verify the accuracy of the transit network representation. This will first require that on-board survey data is properly geocoded and expanded; individual on-board survey trips will then be assigned to the transit network using the *Travel Model Two* software. In this network verification phase, it will be verified if the TAP that is chosen by the user is actually being made available in the choice set generated by the software. This will help identify issues in TAP coding where multiple stops gets absorbed into a single TAP. Specific records that cannot be assigned to their surveyed route would be investigated further to determine the cause of the error.

Determining the path building parameters

The unique aspect of the proposed approach is that between any given TAP pair three sets of paths are computed based on different path building parameters. This introduces heterogeneity/taste variation in the path choice. It is important that these be representative of the population that is being modeled. Hence, these path building parameters need to be recovered in a systematic way from the observed data. We propose estimating a path choice model to this end.

The primary source of data for this exercise would come from the transit on-board survey. For each trip record, all possible sets of path alternatives need to be generated using the *Travel Model Two* software. Segmentation by user class can be handled in three different ways. One approach would assume that we are modeling pre-defined user classes. Using a path choice model we can estimate the best set of parameters and then assert additional weights on certain components based on user class characteristics. A second approach would involve using latent class clustering after performing a factor analysis/principal components analysis to classify the on-board survey records into user classes – the user classes so determined can be labeled based on their characteristics. Then the path choice model specification can be segmented by user class. The third approach involves specifying a structural equation model of class membership and path choice. This has the advantage of being more consistent in the sense that that both membership and path building parameters are determined simultaneously.

9 Population Synthesizer: PopSynIII

The PB-developed `PopSynIII` software will be used to create the synthetic population for *Travel Model Two*; *Travel Model One* used the original `PopSyn` software. `PopSynIII` accommodates both household- and person-level controls. Three key innovations of `PopSynIII` include:

- **Controls at multiple levels of geography**. An algorithm called "meta-balancing" reconciles inconsistencies across controls at up to three levels of geography.
- **Controls can be weighted**. Users may have different levels of confidence in different controls. The software accommodates weights via an entropy-maximizing formulation.
- **Discretization**. Each household in the source sample is weighted and the optimal weights are always non-integers and very often close to zero (e.g., 0.01). The software intelligently discretizes these fractional weights, resulting in a ready-to-use synthetic population.

The proposed controls for the population synthesis is shown in Table 7.

Table 7: Population Synthesis Controls

Control	Categories	Geography	Data Source
Number of HHs	N/A	MAZ	Census/ Socio-economic forecast
Number of HH by income (year 2000 dollars)	0-25k, 25k-60k, 60k-120k, 120k+	MAZ/Tract	ACS Distributions/ Socio-economic forecast
Number of HH by HH size	1,2,3,4,5,6+	MAZ/Tract	Census/ Socio-economic forecast
Number of HH by workers	0,1,2,3+	MAZ/Tract	Census/ Socio-economic forecast
Number of persons by occupation	Management, Professional, Services, Retail, Manual, Military	PUMA	PUMS Distributions
Number of persons by age	0-14, 15-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75-84, 85+	PUMA/County	Socio-economic forecast/ PUMS distributions

The seed data for the synthesis is to be sourced from 2007-2011 5 year ACS PUMS data (which is geo-coded to the 2000 PUMA geographies) for California:

- Population file: http://www2.census.gov/acs2011_5yr/pums/csv_pca.zip

- Household file: http://www2.census.gov/acs2011_5yr/pums/csv_hca.zip

10 Enhancements to MTC Travel Model

This section describes some of the recent advancements in the CT-RAMP activity based modeling framework that can be incorporated into the next version of MTC travel model, i.e., *Travel Model Three*.

Transit Enhancement features

Recently, the transit model component of the Chicago Metropolitan Agency for Planning (CMAP) ABM was overhauled to make the model system sensitive to a wider array of policy and service variables than traditional transit demand models. These variables include so-called “modern” variables such as reliability, safety, cleanliness, crowding level, potential for productivity, ease of boarding, etc. Here we describe some of these additional features. The MTC and CMAP approaches to modeling transit is, to a certain extent, similar. Particularly, the movement away from traditional labeled alternatives to non-labeled alternatives in the mode choice model and the representation of space using MAZ-TAP systems are at the core of the two CT-RAMP implementations[3]. Given this similarity, we believe that some of the more advanced features in CMAP can be gainfully extended into the MTC model.

Table 8 lists out the transit enhancement components for CMAP. The various components were implemented in two phases as shown.

[3] Access modes represents the distinct options (Walk, PNR, and KNR). The alternatives in the lower level nest within these access modes are determined by individual path-building rules as discussed earlier.

Table 8: Transit enhancements for CMAP ABM

Component	Phase 1	Phase 2
Advanced “non-labeled” mode-choice	❶	❷
Transit access / spatial resolution		❷
Station characteristics	❶	❷
In-vehicle characteristics	❶	❷
Capacity constraints		❷
Crowding effects		❷
Service reliability		❷
Transit frequency / wait time	❶	❷
Fare / cost structures	❶	❷
Individualized transit path choice		❷
Mobility attributes and modality		❷

To understand where the CMAP implementation deviates from the MTC approach we have to take a close look at how the virtual transit path is calculated. In the current implementation, MTC uses a fixed distance threshold to determine accessible TAPs. For CMAP, the access and egress legs (MAZ-TAP legs) of the VTP are individualized by using an individual propensity to walk. This is determined by randomly drawing from a propensity curves of traveler’s age (see Figure 5). The propensity gets translated to walk speed, walk weight and walk threshold based on an (asserted) lookup table. In this fashion, each individual sees a different set of TAPs and perceives the walk experience to TAPs differently.

Figure 5: Propensity to Walk by Person Age

For the TAP to TAP component of the VTP, for MTC it is proposed that three TAP to TAP paths be built using parameters that are estimated for broad user classes but the best alternative among all classes gets exposed to an individual in the mode choice model (see Modeling Transit section for details). The TAP to TAP path building for CMAP is also undertaken for three user classes but all these paths are retained (some additional differences are discussed later in this section). Before the mode choice model, the traveler is classified using a class membership model into one of the three user classes and all the paths in that class gets presented as alternatives to the individual. We will now discuss some of the other features in Table 8.

Table 9: Characteristics used in the Virtual Path Building

Characteristic	Orig TAP (access)	TAP-to-TAP	Dest TAP (egress)
Station type	●	●	●
Real-time information	●	●	●
Formal parking capacity	●		●
Informal parking capacity	●		●
Parking cost	●		●
Parking lot walk time	●		●
KNR convenience category	●		●
Buffered area crime rate	●		●
Buffered retail density	●		●
First boarding fare	●		
Boarding (traversal) time		●	
Ease of boarding		●	
Station cleanliness		●	

Table 9 shows the full set of characteristics that are considered in the TVPB procedure. The parameters used in TAP to TAP path building are primary attributes of station type. Each transit station is grouped into one of the five categories: pole, shelter, plaza, station and major station. Based on this, the station gets assigned a wait convenience factor (waiting versus in-vehicle time), real time information factor (makes waiting less onerous), boarding/transfer time (time required to traverse the station) and station cleanliness factor. These gets further multiplied by the individual perception factor based on which user class the person belongs to.

In addition to the parameters just discussed the boarding time and ease of boarding are determined based on the station type. The base boarding time for each station gets multiplied by an individual perception factor that is also dependent on cleanliness. There is a further penalty applied based on the line specific ease of boarding as shown in Figure 6.

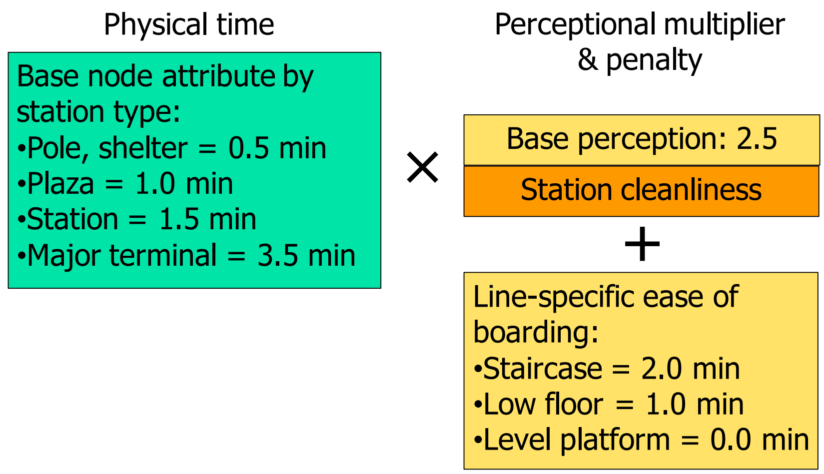

Figure 6: Perceived Boarding Time and Ease

For wait time, the base physical wait is calculated as a product of the line headway, a fraction based on the passenger arrival profile at the transit stop/station, and effective headway multiplier that is used to constrain line boarding capacity. Average extra wait associated with transit unreliability (currently applied for buses only) is added to the physical time. This gets multiplied by the wait convenience factor discussed before and a factor based on cleanliness.

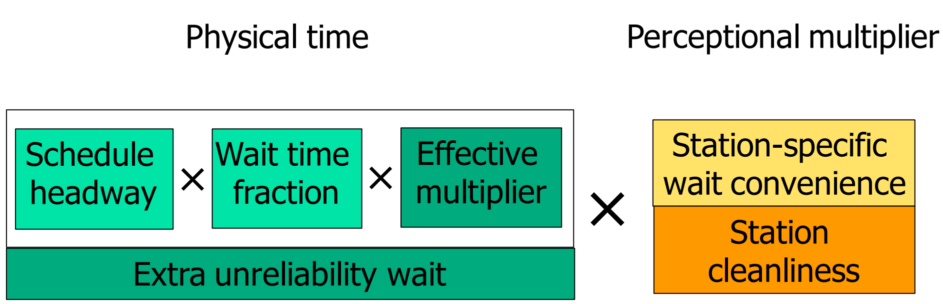

Figure 7: Wait Time Calculation in Transit Path Building

A rich set of factors come into play when calculating the in-vehicle time including seating comfort, productivity, cleanliness, on-board amenities as well as socio-economic compatibility between riders. The EMME-based transit skimming procedure is used to determine the base physical time and these gets multiplied by a perceptional multiplier and subsequently by an additional multiplier that reflects specific conditions for a particular trip.

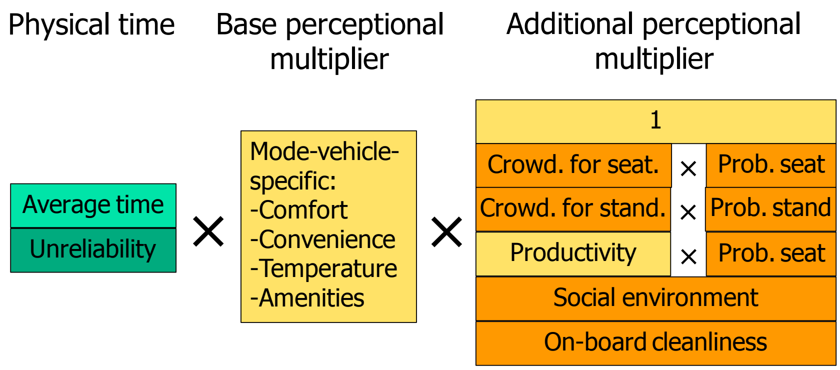

Figure 8: Calculation of In-Vehicle Time for Transit Path Building

Importantly, most of the perceptional multipliers used in this implementation are asserted. Data collection and estimation work could improve the robustness and credibility of these factors.

The individualization of the path building parameters captures the heterogeneity in choice to a great extent. However, the associated computational burden is very high and it could lead to prohibitive runtimes. Hence, we propose a model that borrows some of the features discussed here and ties in with the current transit model.

- The MAZ to TAP and TAP to MAZ can be completely individualized using the propensity to walk curves. This would entail computing the TAP choice set based on the walk distance threshold as opposed to current approach of using a fixed set of TAPs for each MAZ.
- A class membership model that segments the population into three distinct user classes as a precursor to the mode choice model.

- The three set of paths between each TAPs are all retained and the person gets to choose only from the path sets within her/his user class.

Escorting Kids to School

Escorting children to school is a common travel arrangement in a household with school children. It is important to model this aspect of travel behavior as it constrains the travel patterns of the adult household members in terms of mode and time-of-day in order to accommodate dropping-off or picking-up children to/from the school. Figure 9 shows the modified framework being deployed for Maricopa Associations of Government (MAG) that accounts for the school escort decisions.

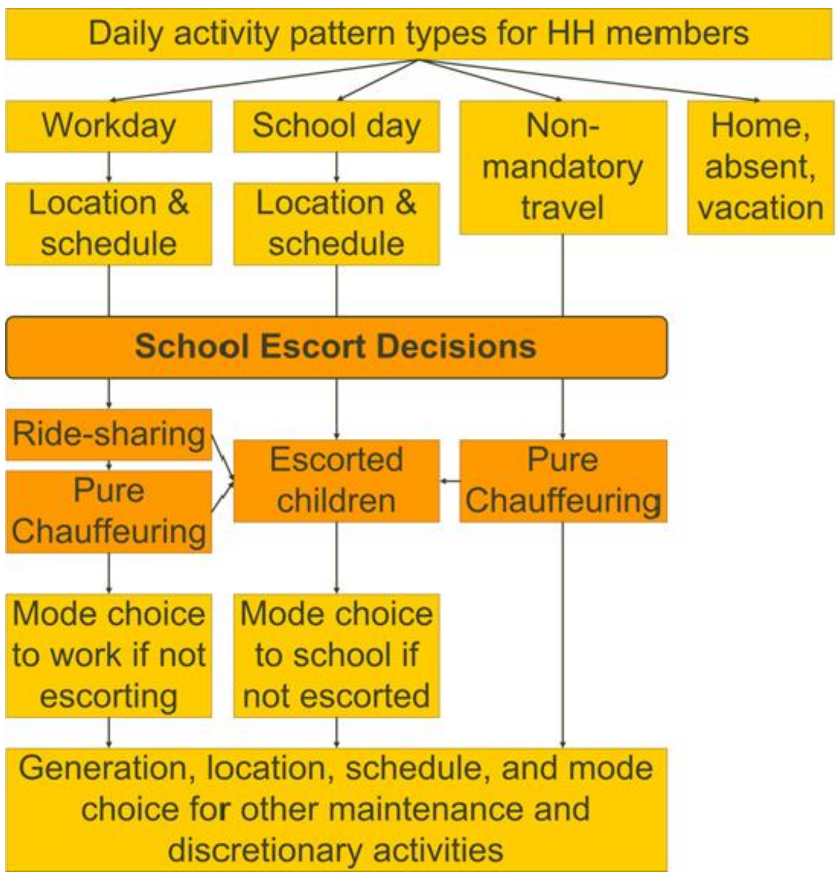

Figure 9: Modified ABM framework for accommodating School Escort Decisions

The model is placed after the long-term choice models (such as work and school location), medium-term choice models (such as auto-ownership) and preferred time of day choice for mandatory tours. In order to limit the proliferation of number of alternatives the chauffeuring decisions is modeled in an iterative fashion. First the choice of the outbound escorting arrangement is independently determined. Next the inbound chauffeuring decision is modeled conditional on the outbound leg followed by determining the outbound arrangement conditional on the inbound leg and so on till an equilibrium is reached. This approach is more formally known as Blocked Gibbs Sampling.

To illustrate the choice generation process, consider a household with 3 children going to school. Figure 10 enumerates the different need and bundling options that can be associated with the household. In the outbound direction, there are 15 combinations of bundling options and escorting needs which can create 0, 1, 2 or 3 school escorting tasks. For each escorting task there are up to 4 possible chauffeur assignments:

- 1st chauffeur, escorting on the way to/from work,
- 1st chauffeur, pure escorting as a separate home-based tour,
- 2nd chauffeur, escorting on the way to/from work,
- 2nd chauffeur, pure escorting as a separate home-based tour.

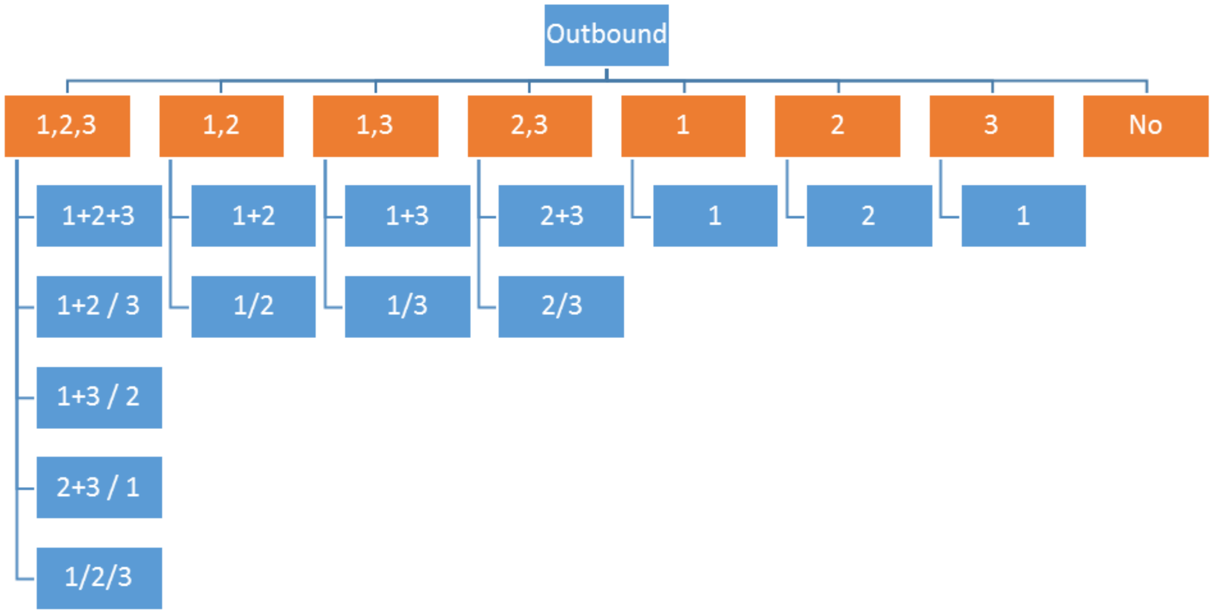

Figure 10: Need and bundling options in escort arrangements

The utility function is composed of several components including utility to the escortee (level of service characteristics of the different modal alternatives, person characteristics and household characteristics), chauffeur (dis)-utility (in terms of out of direction distance, time, time savings due to bundling and person characteristics) and symmetry component (to address the fact that escorting need and chauffeur allocation is typically similar in the outbound and inbound direction). Also, a traveling salesman algorithm is used to sequence stops when multiple children are bundled.

Incorporating this model into MTC's framework would involve substantial reworking of the software implementation as the sequencing of the different model component is very different from the existing approach. However, this would add great value in terms of capturing travel behavior accurately.

Stop Duration Model

A stop duration model was recently implemented for ARC to allocate the predicted tour duration to stops on the tour. In the overall model framework, the stop purposes and locations are generated before this model is applied. Then for every tour, the stop timings on tour is predicted in two stages. The first stage predicts the durations of the main activity, as well as the outbound and inbound stops in aggregate. The second stage predicts the durations of the individual stops within the outbound and inbound legs. The breakdown of the tour duration into stop durations are shown in Figure 11. Both stages use a multi-nomial logit model to apportion the tour/leg duration – the alternatives being every possible combination of durations for the different components that make up the tour/leg. Variables in the model included the tour purpose, free-flow travel time to stop, worker status, stop purpose and number of stops on tours etc.

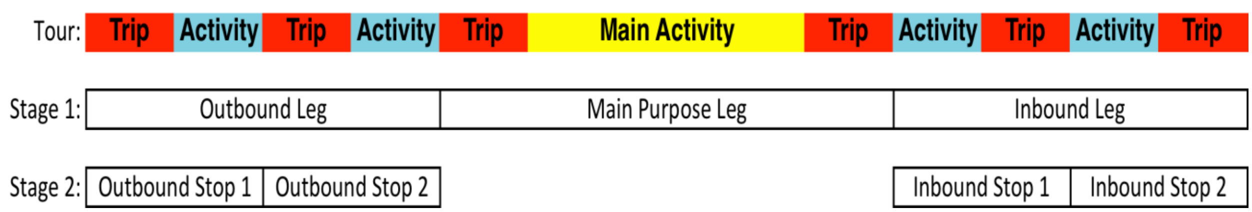

Figure 11: Tour timing breakdown

For the MAG ABM a different approach is being pursued using a Multiple Discrete Continuous Extreme Value (MDCEV) model. The model uses total tour duration as the budget constraint to predict the stop frequency and duration simultaneously.

Currently, the MTC model uses observed duration distributions to model stop duration. It would be useful to replace those with a model to predict the stop duration as it integrates better with the modeling system and would respond to changes in other model components.

Special market models – University and Visitor models

Special markets such as the university and the visitor segments have very unique characteristics that govern their travel choices and it is important to explicitly recognize this and model it within an ABM framework. For instance, major universities induce a substantial clustering of student population in rented apartments and dormitory settings (non-institutional GQ population) – a general population synthesis procedure cannot be expected to locate students nearby universities unless it is provided with the correct controls for the university segment. The travel patterns carried out by visitors in a region are also very unique as their tours tend to be primarily of either work or recreational type depending on the nature of the visit and they are usually dependent on taxi, rental car or transit. Some of the recent modeling efforts have incorporated these special markets into the ABM model system and are described briefly below.

Figure 12 shows the university model framework that was implemented for University of Oregon in the Eugene-Springfield region and Oregon State University in the Corvallis-Albany-Lebanon region in Oregon. The model system uses a tour based modeling approach where the travel induced by students in the population is modeled as independent tours with no constraint on an overall daily schedule. The travel behavior data is drawn from surveys targeting the university students while student estimates and land-use inventory is compiled from university databases and other online sources. A residential location choice model that uses distance from the university as the impedance measure and households by housing type as size term is used to locate the students within zones. This data is fed into the population synthesis as an additional control to generate a disaggregate set of student records with all person characteristics. Following this the model follows typical tour based modeling steps as shown in the figure. The model components shown in green are estimated, those in blue are distributions based on observed survey data and the models shown in yellow are asserted and calibrated.

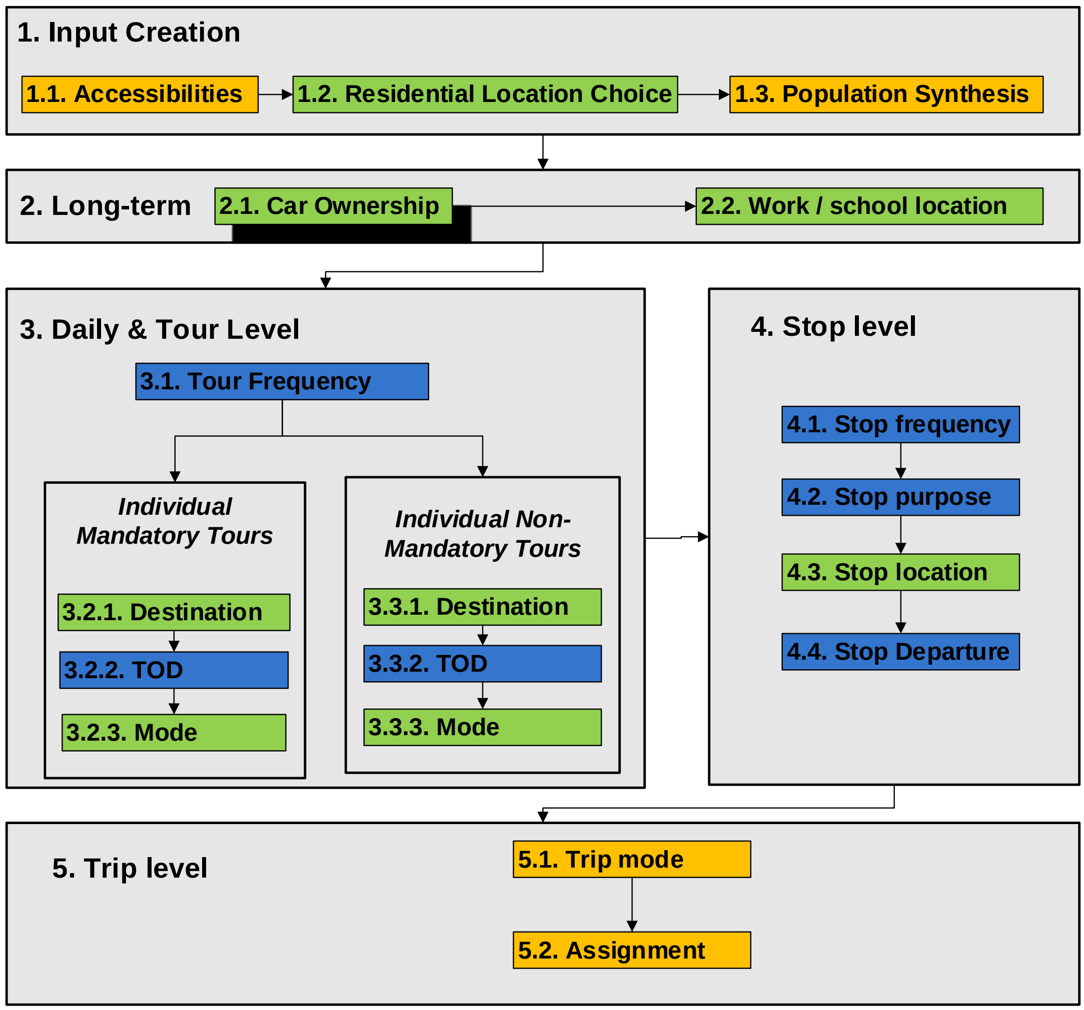

Figure 12: University Model Framework

The MTC modeling region has multiple major universities that could benefit from such a model including Stanford University and University of California, Berkeley. Such models can be used to test scenarios involving major changes in enrolment and also evaluate the impact of campus development strategies on travel demand. At a minimum, the population synthesizer can be modified to generate university students of major universities by applying a residential location choice model for each school. Then university tours for those students are modeled using a university-specific size terms such that students of those schools are sent to their respective MAZs for university travel.

Visitor models were implemented for both SANDAG and the Oahu MPO. These models were developed based on visitor surveys conducted in the region. Figure 13 shows the visitor model framework used for SANDAG. With the exception of destination choice and mode choice, all the models were based upon observed distribution from the survey. This model stratifies visitors into either a Business or Personal market segment and generates visitor parties by segment by applying separate occupancy rates to hotels and households. Each visitor party is attributed with an income level based on observed distributions and tours are generated. Each tour is then attributed with automobile availability and party size. The rest of the model follows a typical tour based modeling approach as shown. Again, this would be a good enhancement to the current MTC model.

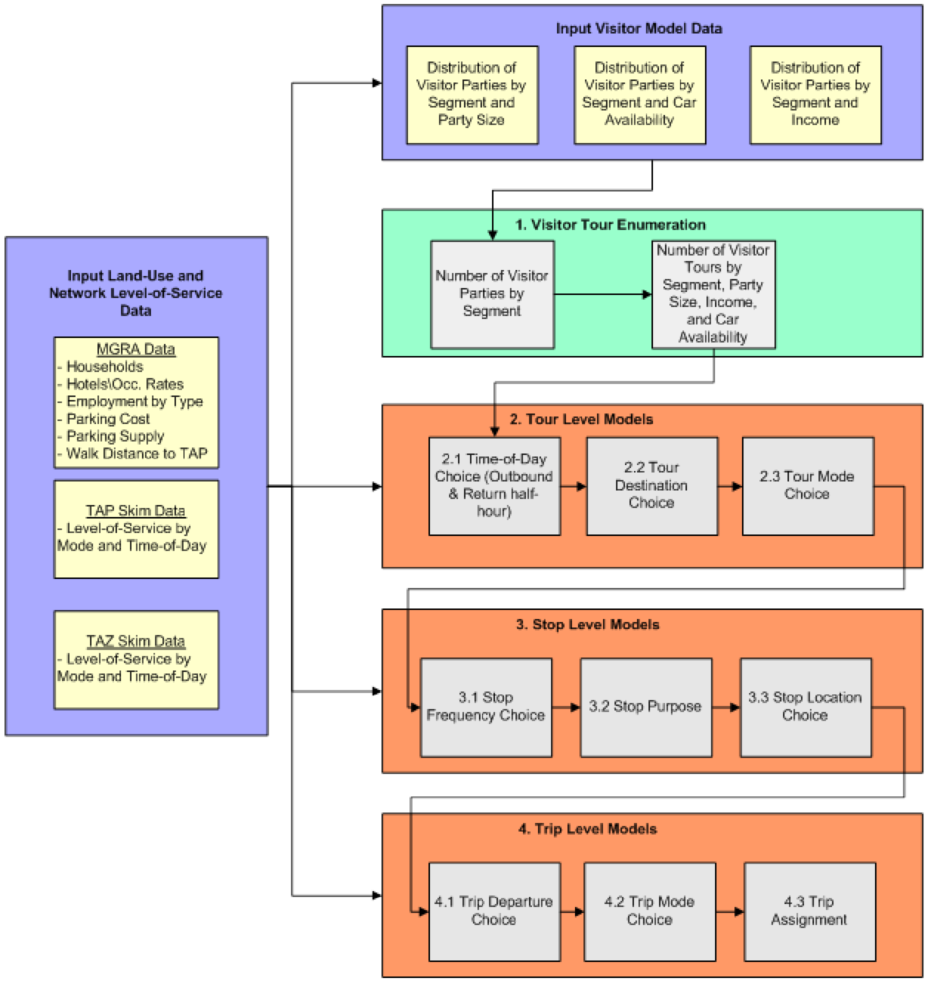

Figure 13: Visitor Model Framework