Transit Modeling Update

Trip Generation Review and Recommended Model Development Guidance

Contents

1 Introduction

1.1 Purpose and Background

The purpose of the Transit Modeling Update project is to specify, within FSUTMS and associated support systems, the changes necessary to improve the preparation of transit demand forecasts to a point consistent with federal expectations, and to incorporate state of the practice techniques and tools through a prototype model application. The Tallahassee Capital Region Transportation Planning Agency (CRTPA) model was chosen as the prototype FSUTMS model application.

The review and proposed recommendations are not limited to the mode choice model, because a sound foundation for the mode choice model starts with an appropriate definition of key travel markets throughout the entire model system. At the trip generation stage the key considerations include the stratification by trip purposes and household classes, as well as the identification of special resident and non-resident travel markets.

The FSUTMS trip generation stage identifies the following person trip markets:

- Internal, permanent resident trips
- Internal, non-permanent resident trips
- External-to-external trips
- Internal-to-external trips

Each of these markets has its own unique characteristics; in fact, they represent quite distinct trip-making units, and their travel patterns are sufficiently different that it is recommended that they be treated as completely separate markets, not just for trip generation but for all model steps. Locations such as hospitals, airports, stadiums, convention centers, and visitor attractions, among others, may constitute important transit markets as well, and may need to be singled-out for special attention on a region-by-region basis. In addition to the person trip markets, FSUTMS includes commercial vehicle trips and other ancillary models that help to estimate congestion levels well.

This guidance addresses exclusively the modeling of internal, permanent resident trips. This is the most important urban transit market in the majority of Florida regions.

Trips that start or end outside of the model area are rarely important from a transit modeling perspective, and for this reason the modeling of external-to-external and external-to-internal trips will not be discussed as part of this Task. It is however acknowledged that in some Florida regions long distance internal-external trips may constitute a significant market for inter-regional transit modes such as Amtrak. These trips are generally best modeled with a mega-region or statewide model.

The seasonal resident population is, of course, significant in some parts of the state. The techniques used to model seasonal resident trip generation are similar to those used to model permanent resident trip making. The key difference between the two populations is in estimating the size of the seasonal population, depending on the model area. Understanding trip pattern differences, whether in terms of the destinations visited, or likelihood of using certain modes, should inform the specification and validation of models specifically developed for these travelers.

Special market segments, which may comprise both permanent and seasonal residents in addition to visitors, are discussed in Section 7. These special markets include ground access trips for air travel, visitor (tourist) travel, and special event travel.

While the focus of this task is on transit modeling, it is recognized that the model will be used for multiple types of project and policy evaluations, and therefore the recommendations provided here aim at preserving the existing highway project model functionality.

1.2 Model Development Terminology

As indicated above the purpose of this guidance is to provide guidelines and recommendations for model specification. At its most comprehensive, model specification can be understood as the outcome of three related model development tasks--estimation, calibration and validation.

> *Model estimation* is the process of using household or trip records from a survey and econometric methods to compute model parameters and test alternative model forms and specifications. The classic example is mode choice estimation, where on-board data, household survey data and skims are used to estimate the in-vehicle time, cost, and other level of service attributes of a multinomial mode choice model.
>
> *Model calibration* is the process of verifying that the model reproduces current travel patterns well. One example of model calibration is the process of specifying alternative-specific constants in a mode choice model. The constants are specified so that the model matches mode shares, transit ridership, district-level patterns of trip flows, and other patterns of travel behavior.
>
> *Model validation* is the process of comparing the model output to independent data and verifying that it is able to replicate them well. Independent data are those not used in the estimation or calibration of the model. Often times the only independent data source are traffic counts; for this reason model validation has come to be understood as verifying that the forecasted traffic volumes match observed counts well. However *useful validation*, as defined by the Federal Transit Administration (FTA), seeks to verify that the model exhibits a consistent story about travel behavior, and that it responds in a reasonable manner to change. Thus, part of validating the model includes verifying that the values of time obtained via model estimation are plausible and reasonable, or that the relative differences among alternative-specific constants are commensurate with the expected differences in un-included attributes among modes and market segments. In addition, sensitivity tests, up to and including both forecasts and backcasts, can be considered useful validation tasks.

1.3 Trip Definition

Throughout this guidance a trip is simply defined as a person moving in time and space between two different locations. Because the National Household Travel Survey (NHTS) Florida Add-On dataset is the basis for the prototype model specification, in practical terms this guidance adopts the definition of trips used by NHTS during the trip diary retrieval process: *a trip is any time a person went from one address to another, including stops made for any reason, such as buying gas or taking someone somewhere, but not including stops made just to change the type of transportation.*[1] An extensive list of activity purposes was used to record the purpose of each trip. Among these purposes are serve passenger trips, which include transporting someone, picking someone up, dropping someone off, and taking someone and waiting. Some of these serve passenger trips were combined with other trips, as is customary for a trip-based model. The intent is to eliminate intermediate stops, and thereby reduce the number of non-home based trips. A detailed description of the NHTS data processing tasks will be included in the Tallahassee Prototype Model Memorandum.

2 Trip Market Segmentation

2.1 Trip Purposes

A standard set of trip purposes for transit modeling includes the following:

- Home-Based Work (HBW)
- Home-Based College/University (HBCU)
- Home-Based School (HBSC)
- Home-Based Shop (HBSH)
- Home-Based Social/Recreation (HBSR)
- Home-Based Other (HBO)
- Non-Home-Based Work (NHBW)
- Non-Home Based Other (NHBO)

Three new purposes are proposed, HBCU, HBSC and a sub-classification of non-home based trips into NHBW and NHBO. Transit ridership by college and university students can be an important transit travel market in areas that include large 2-year and 4-year colleges, particularly where transit improvements are planned to serve the college campus. While HBSC trips are typically not a focus of transit modeling, some transit systems do carry a relatively large number of middle and high school students. It is recommended that HBSC trips be treated separately given their unique trip distribution and mode choice patterns. Finally, the NHBW trips can be an important market segment for some types of transit projects, such as central city circulators and streetcars. This classification of NHB trips also allows for a higher probability of choosing a short walk trip when making a trip for lunch at work, and allows for capturing time-of-day differences between non-home-based trips made when at work versus those that are part of a non-work journey. It also lends itself towards the use of direct demand models to forecast transit ridership in downtown areas.

[1] NHTS Telephone (CATI) Questionnaire. Extended Interview.

Some advanced trip-based models further stratify the HBW trips into direct and complex or strategic trips. A complex HBW trip includes one or more incidental stops on the way from home to work (or from work to home), while a direct trip consists of a single journey between home and work. This classification stems from a desire to capture tour effects in a trip-based framework. The rationale is to avoid splitting a complex HBW tour into home and non-home based components, due to the loss of explanatory power for the non-home based trips, and the loss of consistency between the choices made for the two legs of the tour. Direct trips are thought to be more amenable to transit use, while transit ridership for complex HBW trips can be nearly negligible. In reality, significant pitfalls remain when trying to model HBW trips in this fashion, because some level of inconsistency in the definition of chosen and available modes for the complex trip is unavoidable. The integrity of the HBW trip and its underlying rationale can be preserved, to some extent, by judicious use of trip-linking techniques when preparing the household-interview survey data for model estimation. In regions where a desire to examine complex trip patterns (not just for HBW) is important, due to the types of projects being studied, a tour-based approach is highly recommended.

2.2 Household Classification

Best practice for trip-based production models are household cross-classification models. While at the trip generation stage the model can afford to be expansive in terms of the variables used to cross-classify the households, it is also desirable to maintain a parsimonious specification, to avoid spurious effects that stem from estimation based on small sample sizes. It is important to base the cross-classification models on the set of variables most likely to explain not just trip production in the region, but also relevant for trip distribution and mode choice. And it is equally important to use variables that can be forecasted with a level of effort consistent with agency resources and available data.

Household attributes that have been shown to correlate with transit ridership and that are typically used in trip-based models include household income, auto ownership, and/or car sufficiency. Low income travelers and people living in household with no cars are considered transit-captives and typically make up a large fraction of local bus riders. Depending on the socio-economic makeup of the population, the prevailing levels of highway congestion, and the quality of the transit service, a significant fraction of express bus and rail riders may consist of higher income travelers, as well as choice riders, i.e., people who use transit more out of convenience than out of necessity.

Household lifecycle attributes have been shown to be important as well, though they tend to be less frequently observed in trip-based models. Lifecycle attributes include age of the head of household and presence and age of children.

2.2.1 Household Classification Attributes

FSUTMS uses cross-classification models for home-based trip production, based on the following three variables and categories:

- Type of dwelling unit (single family, multi-family, hotel)
- Household size (1, 2, 3, 4 and 5+ persons per dwelling unit)
- Auto ownership (0, 1 and 2+ autos per dwelling unit)

The auto ownership stratification is carried forward to the trip distribution and mode choice models, and therefore defines the mode choice model socio-economic market segmentation.

FSUTMS uses 'occupied dwelling units' as the trip-making unit. It has become more common to conceptualize the basic trip-making unit as a 'household', following U.S. census nomenclature, as well travel behavior research on households consisting of related persons. Households reside in housing units; therefore an occupied housing unit is generally equivalent to an occupied dwelling unit. Note that group quarters are not counted as housing units; the travel associated with populations living in group quarters needs to be accounted for separately. Note as well that a household may consist of one or more families, unrelated adults living together, or a combination of both. This document uses the concept of households, implicitly understood as persons in occupied dwelling units. We propose to treat households, therefore as a single type, since the other market stratification variables we will use account for most, if not all the trip-making differences between single and multi-family households.

From a transit modeling perspective it is recommended that the FSUTMS stratification be revised to include household income, for the reasons stated below. In addition, from a best practice perspective, it is desirable to base the HBW production models on the number of workers per household.

While auto ownership may be essential to better understand and model mode choice behavior, this single variable is not sufficient to support the behavioral requirements of the HBW trip distribution and mode choice models. For the trip distribution model, worker earnings or household income are the primary descriptors that assist in associating worker and employer characteristics in a coherent manner--for example, connecting lower income blue collar workers with employment types that employ such workers. For mode choice, low income riders often make up a large share of the local bus ridership. Low income is also often used as a proxy for other variables that are not amenable to explicit treatment, such as immigrant status (recent immigrants are typically more transit-dependent), and ethnicity (which tends to correlate with income). However when income is used, alone, in the trip distribution model, the critically important contribution of auto ownership is absent from the mode choice model formulation. While there is some correlation between low income and zero-car household travelers, these two attributes are by no means inter-changeable. In urban, central city areas, for example, many households choose not to own cars as a lifestyle decision, rather than purely out of economic necessity. Conversely, many low income households choose to own a car, even when living in areas well-served by transit, given the availability of very low cost vehicles and the inconvenience of relying solely on transit for all travel. In suburban and rural areas a car is often times the only practical means of daily transportation.

An additional consideration is car sufficiency, instead of the simpler auto ownership measure. The likelihood of using transit may be better explained by the availability of autos ***relative*** to the size and/or number of workers in the household, than by auto ownership alone. For example, the likelihood of a transit trip in a two-worker, one auto household is higher than in a one-worker, one auto household. Car sufficiency is defined relative to the number of workers in the household for HBW trips, and to number of persons (or adults, if known) in the household for non-work trips, as follows:

- Zero car households
- Car-insufficient households: fewer vehicles than workers / persons
- Car-neutral households: equal number of vehicles and workers / persons
- Car-sufficient households: more vehicles than workers / persons

The most comprehensive stratification for transit modeling purposes, and the approach recommended by FTA, is a combination of household income and car sufficiency. Assuming four household income classes and four car sufficiency classes results in 16 market segments, which is clearly a prohibitively large number of market segments for a trip-based model. However, many of the joint categories are relatively infrequent and/or not critical from a modeling perspective. A possible HBW classification for mode choice, for example, would be the following:

i. Zero Car Households – All Income
ii. Cars less than Workers – Low Income
iii. Cars less than Workers - Medium Income
iv. Cars less than Workers – High Income
v. Cars equal to or greater than Workers - Low Income
vi. Cars equal to or greater than Workers - Medium Income
vii. Cars equal to or greater than Workers – High Income

An equivalent specification for non-work trips would use 'Persons' instead of 'Workers'.

Because car sufficiency is a derived attribute, the household classification needs to be based on auto ownership, household income and number of household workers, for HBW, and autos, income and persons, for non-work purposes. The trip rates however can be estimated using the more parsimonious classification, to avoid small sample sizes in select cells, as shown in Section 4 below.

2.2.2 Household Classification Methods

A category analysis or cross-classification approach to trip production models requires the use of a household submodel to prepare the household attribute forecasts in the proper format. Typically trip production rates are derived on a per-household basis, for various types of households. The application of such rates requires corresponding estimates of households by category for each Traffic Analysis Zone (TAZ).

In order to support the trip market stratification described above, the Tallahassee demonstration model considers stratifications by household size, workers per household, household income, and auto ownership. These stratifications are also known as the household distributions. The household size, workers and income distributions will be discussed in this section. The next section is devoted to the auto ownership model.

There are two stages of household distribution involved. The first stage involves independently distributing households by each of the classification variables. These independent one-way distributions are often referred to as marginal distributions. The second stage involves the multi-way distribution (cross-classification) of households which is referred to as the joint distribution. The joint distribution is controlled to the marginal distributions. For this reason the marginal distributions are sometimes also referred to as control totals.

Three techniques are typically used to forecast joint household classifications:

- Iterative proportional fitting
- Household classification lookup table
- Population synthesis

<u>Iterative Proportional Fitting (IPF) Models</u>. IPF, also known as fratar or Furness, is a matrix balancing procedure. An IPF household classification model starts with a regional or subarea-based joint distribution of households, and forecasts TAZ based distributions using TAZ-level marginals as control totals. The regional joint distribution can be developed with National Household Travel Survey (NHTS) data, Public Use Micro Sample (PUMS) data, or a more local household survey, if available. The TAZ-level marginal distributions can be obtained from census data, for a base year, and would need to be forecasted, for all future years. The IPF technique operates TAZ by TAZ, adjusting the regional distribution until it meets the control totals (marginals) for each TAZ. This is the simplest method to apply and relies on readily available data. It has two potential downsides. First, where the only available data source is PUMS, it must be assumed that the regional distribution applies locally as well. Second, the marginal distributions, which are expressed at the TAZ level, must be forecasted for each analysis year. The lookup models described below are an attempt at overcoming this forecasting burden.

<u>Household Classification Submodels</u>. A classification submodel combines the IPF technique above with a method for generating the TAZ-based marginal distributions. The method relies on the stability and strength of relationships between the distribution and its mean value. For example, the distribution of households with respect to household size within an area is strongly related to average household size; the distribution of households with respect to the number of workers in the household is strongly related to average workers per household, and so on. The household classification submodel is essentially a series of lookup curves; based on the average household size in a TAZ, the model forecasts the proportion of 1, 2, 3 and 4+ person households. The classification curves can be developed using Census or NHTS data, and must exhibit two properties:

i. For all average household size values, the sum of households by integer size category must equal 100%.
ii. Using the integer size percentage distribution, the calculated average persons per household must equal the observed average.

The model inputs are zonal averages, such as average household size, average workers per household, average auto ownership, and median income index. The latter is typically the ratio of the median TAZ income to the regional median income. These ratios can be assumed to remain constant for all analysis years, or they can be adjusted to reflect expected shifts in household composition. The other model input is total households by TAZ. The curves are used to construct marginal household distributions for each TAZ, which in turn are used as control totals in an IPF process, along with a seed regional or county-based joint distribution, to develop joint household distributions for each TAZ.

These household submodels have the advantage of being relatively easy to develop and maintain. They have a long history in travel demand models, and the relationship between the distribution and its mean has been shown to be stable.

Several variations of these models have been developed. For example, to account explicitly for the correlation between the number of workers in a household and the household income, the income classification model can be developed conditional on the worker classification. That is, separate income classification curves are developed for 0 worker households, 1 worker households , etc.

Population Synthesis. Unlike the two methods described above, population synthesis is a disaggregate forecasting technique. Instead of forecasting a joint distribution of households, it forecasts individual household records that, when aggregated, meet certain pre-specified control totals. While a disaggregate population is not required for a trip-based model, population synthesis is a natural extension of the household submodels described above, and often times comprises a first step towards an Activity-Based Model. In very general terms, population synthesis consists of two steps: first, multi-dimensional joint distributions of households are generated at the TAZ level, using methods similar to those described above; second, for each TAZ, household records are drawn at random from the PUMS dataset corresponding to the TAZ Public Use Micro Area (PUMA), so that the list of TAZ households replicates the joint distribution for the TAZ. Because the sampled household records contain the full set of attributes available in the survey data, population synthesis provides forecasting models with a richer set of household variables. It can also forecast person-level attributes along with household attributes, as well as use person-level control totals (such as population by age) and household composition of lifecycle variables explicitly in the population forecasting process.

The approach recommended for the Tallahassee demonstration is classification submodels. The joint household distribution of households by number of workers, income and household size was prepared with the Florida 5-Year American Community Survey (ACS) PUMS data, using these category values:

- 0, 1,2, and 3+ workers per household
- Less than $25,000, $25,000-$50,000, $50,000-$75,000, and $75,000 or more annual household income
- 1, 2, 3, and 4+ persons per household

The income ranges are expressed in $2009, and represent approximately household income quartiles--see Table 2.1. Another consideration in the selection of the household income ranges is the income categories available for select Census Transportation Planning Package (CTPP) tables derived from the 5-year ACS. Consistency with the income ranges used in the census tabulations is desirable to facilitate the use of these data for model calibration and validation. Income groups could also be defined so as to maximize the ability of the model to describe differences in travel behavior, such as trip rates, auto ownership, work trip distribution, and transit use.

For the Tallahassee prototype the joint household distribution was developed using data for the entire state. It may be possible to customize it to smaller geographic areas by indexing the distribution by PUMA region.

Table 2.1 Florida State Household Income (2009 NHTS)

Income ($2009)	Households		
	Count	Expanded	Cumulative Expanded (%)
<$5,000	269	177,481	2.7%
$5,000 - $9,999	623	378,853	8.6%
$10,000 - $14,999	827	411,881	15.0%
$15,000 - $19,999	1,002	450,691	21.9%
$20,000 - $24,999	794	350,062	27.3%
$25,000 - $29,999	1,134	525,966	35.4%
$30,000 - $34,999	630	324,050	40.4%
$35,000 - $39,999	1,028	453,050	47.4%
$40,000 - $44,999	492	231,504	51.0%
$45,000 - $49,999	914	376,502	56.8%
$50,000 - $54,999	414	217,348	60.2%
$55,000 - $59,999	791	336,470	65.4%
$60,000 - $64,999	300	155,791	67.8%
$65,000 - $69,999	655	252,782	71.7%
$70,000 - $74,999	312	125,445	73.6%
$75,000 - $79,999	640	257,905	77.6%
$80,000 - $99,999	1,167	501,497	85.4%
> = $100,000	2,410	948,006	100.0%
Total	14,402	6,475,284	

The household classification submodels were developed using the 5-year ACS dataset. The average household attributes (size, workers, etc.) and proportions (number of 1-person households, 2-person households, etc.) were computed for each census tract. The current ACS release does not include block group level tabulations. Even if the data were available at block group level, there may not be enough households within block groups to compute a representative household distribution, and therefore the group level proportions may exhibit excessive variation within each average category. The following ACS tabulations were used:

- Table B08202: Household Size by Number of Workers in Household
- Table B08301: Means of Transportation to Work (for total workers)
- Table B19001: Household Income
- Table B19013: Median Household Income

Since a tabulation of persons living in households is not yet available, an estimate was constructed from Table B08202 by assuming on average 4.7 persons per household in households with 4+ persons.

The process to develop the classification models will be illustrated using the household size submodel:

i. Compute average persons per household for each census tract.
ii. Compute, for each census tract, the percent of 1, 2, 3, and 4+ person households.

iii. Plot these data to visually check the underlying relationships and flag any outliers (**Error! Reference source not found.**).
iv. For each average household size value (rounded to the first decimal), aggregate the number of household in each classification category. Compute the fraction of household in each average household size category or bin.
v. Manually adjust the raw lookup table as needed. For example, fill-in missing values resulting from small sample size. This is the first-cut at the classification curves.
vi. Compute the estimated average household size for each bin--the sum product of the fraction of households in each size class and the number of persons per household in the class.
vii. Adjust the classification curves so that the estimated average household size per bin is equal to the observed average, while the sum of household fractions within each average bin is 100%.
viii. Interpolate the lookup table so it is indexed by 2-decimal average household size values.
ix. Calibrate the curves so that, when applied to the Tallahassee model area, the resulting marginals match census control totals.

The adjustments required to smooth out the curves are a combination of manual and automatic adjustments. The automatic adjustments can be calculated using this formula:

$$p_{i+1}^{k,m} = p_i^{k,m} \times \frac{e^{\delta^m \times b^k}}{\sum_k e^{\delta^m \times b^k}}$$

where:

- *i* is the adjustment iteration number,
- b^k is the number of persons per household in each class (1 person, 2 persons, etc.),
- *m* is an index for the bins used to compute average household size,
- $\delta^m$is the percent error in the estimated average household size, for each bin,
- *p* is the fraction of households in bin *m*, class *k*.

The estimated final household size classification model is shown in Figure 2.2. Table 2.2 shows the corresponding lookup values, prior to the linear interpolation step.

Alternatively, 3rd degree polynomials or other flexible functional forms can be fit to the data in Figure 2.1, while ensuring that the two constraints indicated above are met.

The worker classification submodel and the household income classification submodel developed with the ACS data for the Tallahassee prototype are shown in Figure 2.3 and Figure 2.4, respectively.

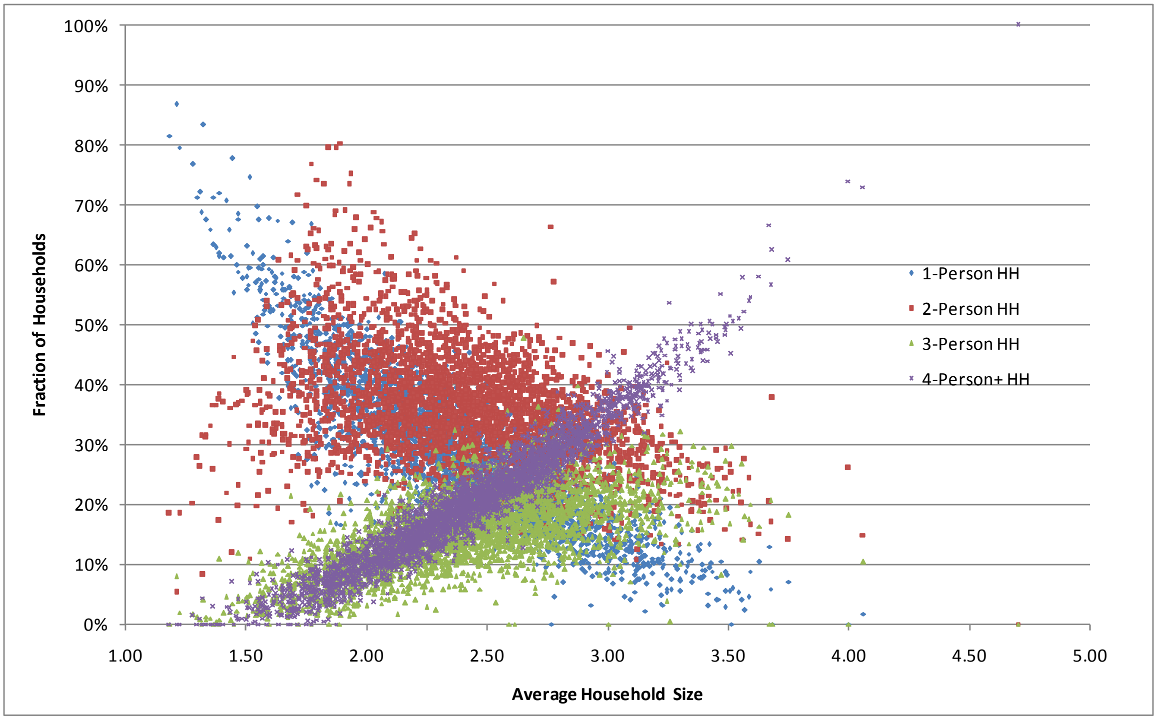

Figure 2.1 Household Proportions by Average Household Size, Florida Census Tracts

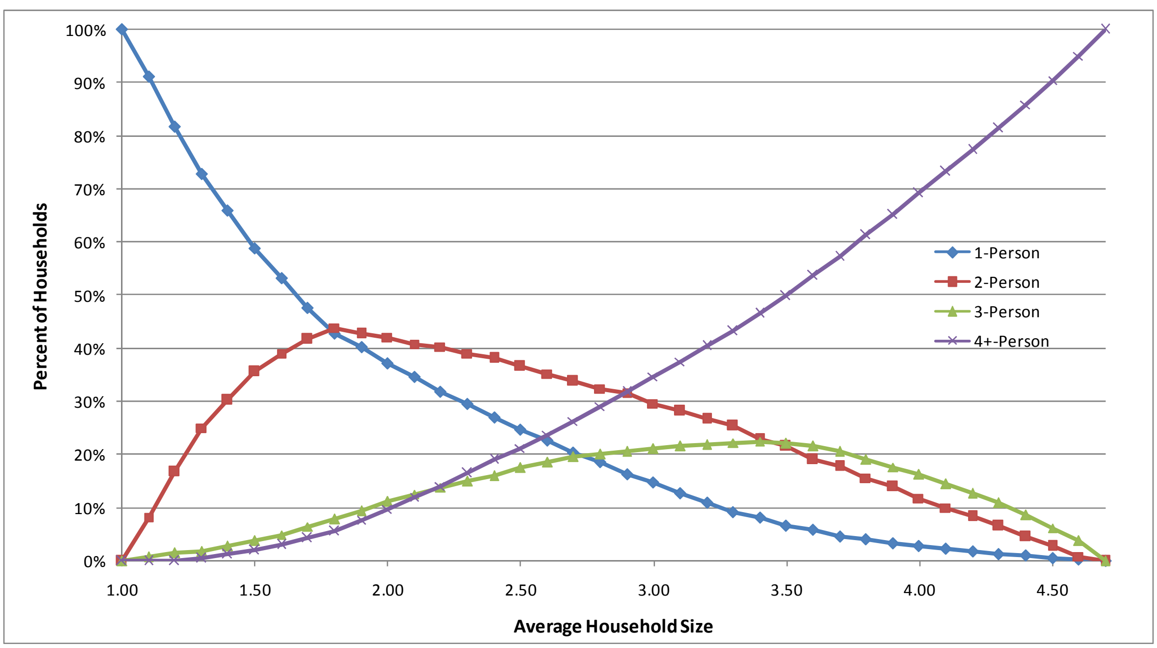

Figure 2.2 Household Size Classification Model

Table 2.2 Household Size Classification Model

Average Household Size	1-Person Households	2-Person Households	3-Person Households	4+ Person Households	Total Households	Computed Average
1.00	100.0%	0.0%	0.0%	0.0%	100.0%	1.00
1.10	91.1%	8.4%	0.5%	0.0%	100.0%	1.10
1.20	82.0%	15.8%	2.3%	0.0%	100.0%	1.20
1.30	72.5%	25.7%	1.2%	0.6%	100.0%	1.30
1.40	65.6%	31.0%	2.0%	1.4%	100.0%	1.40
1.50	58.7%	35.6%	3.8%	1.9%	100.0%	1.50
1.60	55.0%	35.8%	5.7%	3.5%	100.0%	1.60
1.70	48.5%	40.3%	6.8%	4.4%	100.0%	1.70
1.80	42.7%	43.9%	7.8%	5.6%	100.0%	1.80
1.90	40.2%	43.0%	8.8%	8.0%	100.0%	1.90
2.00	37.1%	42.1%	11.1%	9.7%	100.0%	2.00
2.10	34.7%	40.8%	12.5%	12.0%	100.0%	2.10
2.20	31.8%	40.3%	13.8%	14.1%	100.0%	2.20
2.30	29.9%	38.4%	15.0%	16.7%	100.0%	2.30
2.40	28.0%	36.8%	15.9%	19.3%	100.0%	2.40
2.50	24.5%	37.5%	16.7%	21.4%	100.0%	2.50
2.60	23.0%	35.1%	17.8%	24.1%	100.0%	2.60
2.70	20.0%	34.5%	19.1%	26.3%	100.0%	2.70
2.80	18.8%	32.3%	19.6%	29.3%	100.0%	2.80
2.90	16.1%	31.6%	20.4%	31.8%	100.0%	2.90
3.00	15.5%	29.2%	19.8%	35.5%	100.0%	3.00
3.10	12.6%	29.4%	20.0%	38.0%	100.0%	3.10
3.20	11.4%	26.5%	21.3%	40.8%	100.0%	3.20
3.30	10.3%	25.0%	20.1%	44.6%	100.0%	3.30
3.40	8.9%	21.7%	22.5%	46.9%	100.0%	3.40
3.50	6.2%	21.9%	22.1%	49.7%	100.0%	3.50
3.60	7.5%	18.4%	19.0%	55.1%	100.0%	3.60
3.70	6.0%	17.5%	17.7%	58.7%	100.0%	3.70
3.80	5.5%	15.0%	16.9%	62.5%	100.0%	3.80
3.90	4.6%	12.9%	16.5%	66.0%	100.0%	3.90
4.00	3.3%	11.6%	15.7%	69.4%	100.0%	4.00
4.10	2.0%	10.0%	14.9%	73.1%	100.0%	4.10
4.20	1.7%	8.4%	12.6%	77.3%	100.0%	4.20
4.30	1.1%	6.6%	10.9%	81.4%	100.0%	4.30
4.40	1.0%	4.5%	9.3%	85.3%	100.0%	4.40
4.50	0.6%	2.7%	7.7%	89.0%	100.0%	4.50
4.60	0.6%	0.8%	6.5%	92.2%	100.0%	4.60
4.70	0.0%	0.0%	0.0%	100.0%	100.0%	4.70

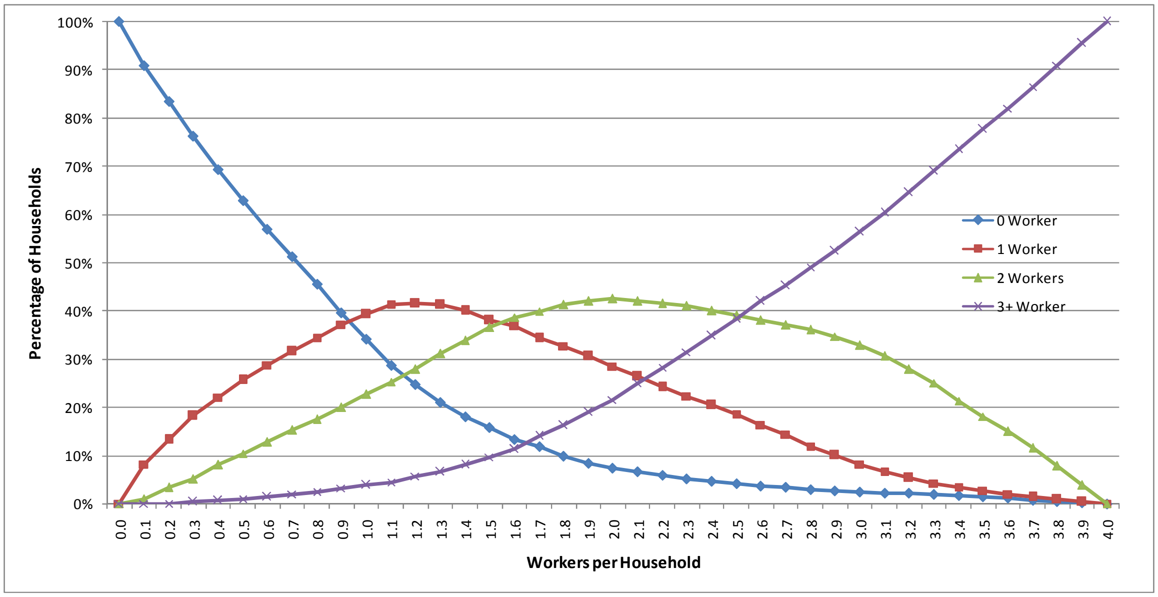

Figure 2.3 Household Workers Classification Submodel

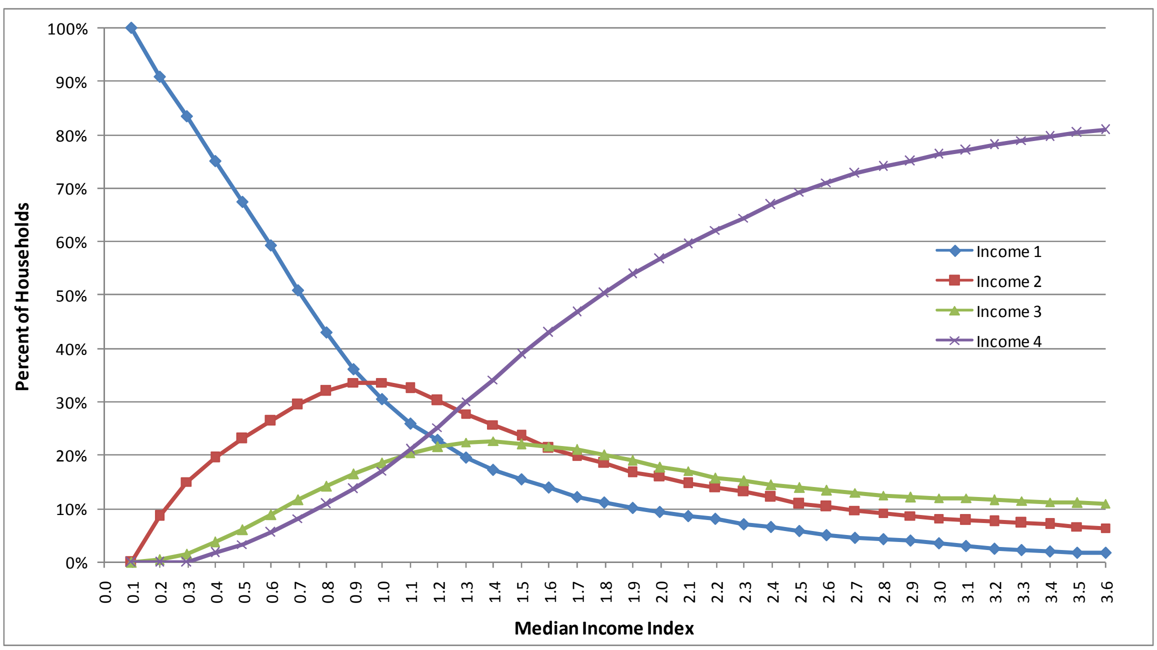

Figure 2.4 Household Income Classification Submodel

2.2.3 Group Quarters Population

Students living in group quarters are excluded from the household classifications above. To account for their travel in the model, each student living in a dormitory can be counted as a 1-person, 0-worker, low income, 0-car household for all trip purposes except HBCU. A unique HBCU trip rate should be assigned to the college student population. Counts of students in group quarters at the census tract level can be obtained from the 2010 census data.

3 Auto Ownership Model

Although any of the household classification methods described above can be used to forecast auto ownership, it has become common practice in the United States to estimate and apply disaggregate, logit-based auto ownership models. These discrete choice models have the advantage of allowing the use of transit and/or non-motorized accessibility terms and density to explain auto ownership, in addition to the use of household attributes such as income, size and type of housing unit. Auto ownership models are relatively simple to estimate and apply. In Florida they can be estimated using the NHTS data, along with model-based measures of transit and non-motorized accessibilities. While the models are estimated using disaggregate data (i.e., household records), they are applied using aggregate data (i.e., TAZ-based estimates of all explanatory variables). As long as the household classification models provide a joint distribution of households across all variables in the auto ownership model, no aggregation error is incurred when applying the model in this fashion.

Estimation of an auto ownership model for Florida would involve the following steps:

i. Prepare the base model estimation data. This consists of the sample of NHTS household records that have a valid home TAZ, and valid responses for all household attributes to be used in the model estimation process (number of household vehicles, number of workers, household income, household size, presence and/or age of children, for example). The estimation file consists of one record per household, with all the household attribute information attached.
ii. Develop measures of transit and/or non-motorized accessibility. Possible ways to measure accessibility include destination choice logsums calculated using only transit or non-motorized modes, total employment within a 30 minute or 60 minute transit trip, or within a 10 minute walk, among many others.
iii. Optionally, other explanatory variables may be included in the model. In California, for example, the recent legislation mandating targets for greenhouse gas emission reductions has prompted Metropolitan Planning Organizations (MPOs) to develop models sensitive to land use form. The model for the Southern California MPO (SCAG) includes a measure of mixed use density in its auto ownership form, such that increasing density results in a lower likelihood of multiple car households. As such, this land use form variable helps to locate the zero-car households in areas of high density, contributing to a better identification of the transit travel market. Another variable that has been found to be important to explain auto ownership in the SCAG region is type of housing unit (single family and multi-family), with households living in multi-family housing units exhibiting lower auto ownership, on average.

iv. Estimate the model, using maximum likelihood estimation software such as R, STATA or ALOGIT.
v. Develop the application script. Since the auto ownership model takes the form of a multinomial logit model, it can be easily applied using Cube/Voyager's XCHOICE function.
vi. Calibrate the model. Model calibration is the process of adjusting the constants so that the aggregate, estimated regional shares match observed shares developed from Census, NHTS or local household survey data.

Auto availability models developed originally for St. Louis and Southern California were transferred for use in the Tallahassee prototype. The SCAG model exhibits many desirable features, including use of logsums to measure transit accessibility, and the presence of several important household attributes (size, income, workers, type of housing unit). From the St. Louis model we borrow the application script, originally developed for Cube/Voyager. The model forecasts the likelihood of 0, 1, 2, 3 and 4+ autos per household as a function of household size, household income level, number of workers, population density, and transit accessibility. The model parameters are shown in **Error! Reference source not found.**. Note that all the household attribute variables are entered as indicator (dummy) variables, and that the reference category for each household attribute assumes a parameter value of 0.

Table 3.1 Auto Ownership Model Parameters

	1 Auto		2 Autos		3 Autos		4+ Autos	
	Coeff.	**t-stat**	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**
Household Income								
$0-$25,000	-2.8138	-6.8	-4.7168	-11.3	-5.5354	-12.9	-6.1442	-13.4
$25,000 - $50,000	-1.3453	-3.2	-2.6003	-6.2	-3.0210	-7.1	-3.6313	-8.3
$50,000 - $75,000	-0.3288	-0.7	-0.7827	-1.8	-0.9589	-2.1	-1.1941	-2.6
Household Size								
2 Person HH			+2.0175	+33.1	+1.9849	+18.5	+1.4450	+8.8
3 Person HH			+1.8057	+22.6	+2.4022	+19.7	+1.6413	+8.9
4+ Person HH			+2.0975	+27.8	+2.3327	+19.5	+2.3295	+13.6
Workers in HH								
1 Worker HH	+0.8839	+10.5	+1.0585	+10.9	+1.1472	+9.2	+1.3116	+7.1
2 Workers HH	+0.4965	+3.5	+1.5991	+10.9	+1.8583	+11.1	+2.1258	+9.8
3+ Workers HH			+0.7428	+4.1	+2.7108	+14.1	+3.7370	+15.6
Multi-Family Housing	-0.3262	-3.9	-1.0705	-11.6	-1.7900	-15.6	-2.1969	-13.2
Mixed Density	-0.0494	-2.2	-0.0731	-3.2	-0.1034	-4.3	-0.1181	-4.5
Walk Accessibility								
0 jobs								
1 to 1,000 jobs	-0.3500	-2.7	-0.3800	-2.8	-0.5800	-2.6	-0.5800	-2.6
1,000 to 5,000 jobs	-0.4003	-1.8	-0.4003	-2.5	-0.6417	-2.4	-0.6417	-2.4
5,000 to less than 10,000	-0.4570	-1.6	-0.4570	-1.6	-0.6622	-2.1	-0.6656	-2.6
more than 10,000 jobs	-0.4793		-0.4793		-0.7216		-0.7216	
Transit Accessibility								
Less than 12								
12 to less than 13	-0.0201	-2.1	-0.0201		-0.0300		-0.0300	
13 or higher	-0.0582	-7.0	-0.0897	-8.2	-0.0897	-9.1	-0.0897	
Constant	+4.3911	+9.9	+4.2830	+9.6	+3.3968	+7.4	+2.8727	+5.9

Mixed density is a combined indicator of population, employment and street intersection density. It is highest when all three densities are high, and as shown in Table 3.1, high mixed density results in lower auto ownership.

$$Mixed\ Density = \ln\left[\frac{Intersections \times (a \times Employment) + (b \times Households)}{Intersections + a \times Employment + b \times Households}\right]$$

In the formula above, *a* and *b* are normalization factors, calculated as the ratio of average intersections to average employment, and average intersections to average households, respectively. For each TAZ, the number of intersections, employment and households are measured over a 1/2 mile circle centered on the TAZ centroid, instead of the entire TAZ area. This is referred as a floating density, and has the advantage of capturing neighborhood density better than a TAZ average. The floating density captures density variations better than a standard TAZ density when employment, household and intersection totals are available at a smaller geographic unit, such as census block group level. When the TAZs are small relative to the 1/2 mile circle, either way yields approximately the same mixed density indicator.

Walk accessibility is measured as the number of jobs that are within a 10 minute walk of a TAZ. Transit accessibility measures the proximity of employment via transit and is calculated using a destination choice logsum type of measure, shown below.

$$Transit\ Logsum = Ln\big(-0.025 \times Transit\ travel\ time + \ln\,(Employment)\big)$$

The auto ownership model can be calibrated to each MPO model area by adjusting the constant term to match observed local auto ownership model shares. Auto ownership model calibration targets consist of the total number of households observed for each auto ownership level. The 5-year ACS data at the census tract level was used for the Tallahassee prototype (see Table 3.2).

Table 3.2 Auto Ownership Model Calibration Targets, Tallahassee Region

	Number of Occupied Households in Model Area (5-Year ACS)	
Auto Ownership per Household	Total Households	Percent
0	9,060	6.4%
1	50,874	36.2%
2	53,291	38.0%
3	19,434	13.8%
4+	7,660	5.4%
Total	140,319	100%

The model area includes all of Gadsden, Jefferson, Leon and Wakulla counties.

Optionally, as part of the calibration and validation effort other model parameters may be adjusted to match 2-way distributions of auto ownership and other household attributes. The 5-year ACS data provide estimates of household auto ownership by household size and by number of workers in the household, which may be used to validate the model base year forecast. Local household surveys may provide similar data, in addition to tabulations of auto ownership and household income, and auto ownership as a function of population density or transit accessibility. Care should be taken when using household survey data to ensure that the household totals match census totals.

The model parameters can be easily updated and customized with local data, following the estimation process described above. The Florida NHTS sample is large enough to estimate the household attribute variables for the entire state, and may even be sufficient to estimate local parameters for some of the larger sub-regions. However, because the statewide TAZs are relatively large, the derived measures of accessibility may not be fine enough for the urban areas, which potentially may result in biased and/or insignificant accessibility parameters. To overcome this problem, a two-step estimation process may be used: first the entire model is estimated using the entire state sample; then, the estimation file is rebuilt using urban MPO TAZs to locate the households and compute density and accessibility indicators, and the model is re-estimated holding the household attribute parameters constant. Another possible approach is to estimate the model from the start using MPO TAZ systems and derived measures of accessibility, thus pooling the data from several MPOs to maintain as large a sample size as possible. Yet another approach is to start with the statewide TAZ estimation and then adjust the model parameters as part of the model calibration process for each MPO.

4 Trip Production Model

For the Tallahassee demonstration a limited number of cross-classification models were explored. The models estimated include:

- HBW productions: number of workers and household income
- HBW productions: number of workers, household income and auto ownership
- HBW productions: household income, car sufficiency and type of dwelling unit
- Non-Work productions: household size and household income
- Non-Work productions: household size, income and auto ownership

The process of estimating these cross-classification models using the Florida NHTS data is as follows:

i. Prepare the household data. This step involves flagging household records with missing cross-classification attributes, such as missing values for household income, as well as flagging households with illogical or extreme attribute values.
ii. Prepare the trip data. This step involves performing several tasks to clean up and code the trip records, such as:
    a. Remove all weekend trips.
    b. Flag records with missing trip or household attribute values.
    c. Recode any attribute values that do not match the model’s specifics.
    d. Code trip purposes.
    e. Code external trip flags.
    f. Code origin and destination TAZs.
    g. Code production/attraction flags.
iii. Develop a weekday average expansion factor. The NHTS trip expansion factor is annualized, so it needs to be adjusted to represent average weekday trips. This can be accomplished by

removing the weekend day trips from the sample, and dividing WTTRDFIN by 240 (number of weekdays in a year, adjusted for holidays).

iv. The trip rates may also be estimated using un-expanded trip records. Where the expansion factors are based on the same variables used by the classification model, as is common in home interview surveys, expanded and un-expanded data yield similar trip rates. The NHTS expansion factor however is only partially based on household size, therefore use of expanded data is recommended.

v. Tabulate households by each of the desired cross-classifications. Develop both raw frequency and expanded households. The former are used to identify classification cells with few observations. In general, trip rates based on fewer than 30 household observations are unreliable. This table constitutes the denominator of the trip rate calculation. Note that all households are included in the tabulation, not just those that report weekday trips. The household cross-classifications are shown in Table 4.1 and Table 4.2 respectively for the work and non-work models.

vi. Tabulate trips by each of the desired cross-classifications and for the relevant trip purpose. This table constitutes the numerator of the trip rate calculation.

vii. Evaluate the estimated trip rates for reasonableness. It may be necessary to aggregate some categories to obtain reasonable, statistically valid trip rate estimates. A 2-tailed t statistic can be used to test whether two trip rates are statistically equivalent. When this method fails to produce a logical pattern, some rates may need to be asserted.

viii. Validate the estimated model by comparing aggregate trip productions to control totals developed from the NHTS survey data. The control totals will include observations that were omitted from the estimation data, due to missing income or trip purpose, for example.

The proposed trip rates for the Tallahassee prototype model are shown in Table 4.3 to Table 4.9. The 3-way classifications resulted in valid, reasonable trip rates for most household categories. For HBW the largest variation in trip rates is observed as a function of the number of workers in the household, as expected. Some variation is observed also across household income and auto ownership. Note that the 0-car household rates were asserted given the small number of 0-car households with workers in them. Similarly, for the non-work purposes the largest variation in trip rates is observed across household size, with smaller variation noted across household income and auto ownership.

Some aggregation of the less frequent household classes was necessary to avoid unreasonable rates due to small sample sizes. For example, the HBW 0-car trip rates were first estimated across all income and worker levels, and then modified to build in a trend over number of workers. This process was necessary because the vast majority of 0-car households has no workers in them, and therefore do not contribute information to the HBW trip rate calculation. In fact, for all the models the 0-car household rates were estimated by collapsing the income classification. Other classes where some of the trip rates were asserted includes the 1-person, 3+ auto households, and the 3-person, 1-auto, med/high income households. The former are relatively infrequent in the population so it is expected that the survey will contain few of these households. The latter however should be more frequent, so it is somewhat surprising that the NHTS data exhibits few of these observations. A comparison of the NHTS household cross-classification with 2010 PUMS data would help to understand whether some household classes are under-represented in the NHTS data.

Table 4.1 Household Sample Size for Home-Based Work Cross-Classification Models

Household Income	Household Autos	Number of Workers in Household			
		Wrks0	Wrks1	Wrks2	Wrks3+
$0-$24999	Veh0	412	57	5	0
$25000-$49999		62	18	8	2
$50000-$74999		12	6	1	0
$75000+		13	6	2	1
$0-$24999	Veh1	1,551	410	24	2
$25000-$49999		1,208	557	52	3
$50000-$74999		356	246	23	2
$75000+		264	179	30	2
$0-$24999	Veh2	451	277	61	7
$25000-$49999		716	657	261	18
$50000-$74999		357	466	351	9
$75000+		493	822	904	19
$0-$24999	Veh3+	113	94	38	13
$25000-$49999		160	252	170	54
$50000-$74999		123	214	243	63
$75000+		196	468	623	195

Table 4.2 Household Sample Size for Home-Based Non-Work Cross-Classification Models

Household Income	Household Autos	Household Size			
		Size1	Size2	Size3	Size4+
$0-$24999	Veh0	352	91	11	20
$25000-$49999		52	22	7	9
$50000-$74999		16	1	1	1
$75000+		9	8	4	1
$0-$24999	Veh1	1,235	601	88	63
$25000-$49999		930	769	76	45
$50000-$74999		299	284	24	20
$75000+		204	238	22	11
$0-$24999	Veh2	118	495	102	81
$25000-$49999		125	1,129	205	193
$50000-$74999		56	777	157	193
$75000+		55	1,420	292	471
$0-$24999	Veh3+	27	109	61	61
$25000-$49999		31	293	168	144
$50000-$74999		26	257	172	188
$75000+		26	612	378	466

Table 4.3 HBW Trip Production Rates

Household Income	Household Autos	Number of Workers in Household			
		Wrks0	Wrks1	Wrks2	Wrks3+
$0-$24999	Veh0	0.000	1.000	1.783	3.500
$25000-$49999		0.000	1.000	1.783	3.500
$50000-$74999		0.000	1.000	1.783	3.500
$75000+		0.000	1.000	1.783	3.500
$0-$24999	Veh1	0.000	1.085	2.272	4.389
$25000-$49999		0.000	1.332	3.145	4.389
$50000-$74999		0.000	1.145	2.172	4.389
$75000+		0.000	0.954	2.094	4.389
$0-$24999	Veh2	0.000	1.581	2.974	4.102
$25000-$49999		0.000	1.604	3.171	4.102
$50000-$74999		0.000	1.572	2.863	4.102
$75000+		0.000	1.330	3.111	4.102
$0-$24999	Veh3+	0.000	1.595	3.130	5.075
$25000-$49999		0.000	2.297	3.109	5.075
$50000-$74999		0.000	1.967	3.247	5.799
$75000+		0.000	1.551	3.119	5.799

Table 4.4 HB Shop Trip Production Rates

Household Income	Household Autos	Household Size			
		Size1	Size2	Size3	Size4+
$0-$24999	Veh0	0.575	1.510	1.617	1.951
$25000-$49999		0.575	1.510	1.617	1.951
$50000-$74999		0.575	1.510	1.617	1.951
$75000+		0.575	1.510	1.617	1.951
$0-$24999	Veh1	0.840	1.467	1.336	1.891
$25000-$49999		0.679	1.431	1.831	2.629
$50000-$74999		0.619	1.308	1.900	3.279
$75000+		0.769	1.183	2.016	2.076
$0-$24999	Veh2	0.842	1.491	1.460	1.700
$25000-$49999		0.654	1.473	1.772	1.786
$50000-$74999		0.798	1.309	1.865	1.687
$75000+		0.800	1.262	1.734	2.099
$0-$24999	Veh3+	1.169	1.259	2.693	2.225
$25000-$49999		1.169	1.785	2.186	1.766
$50000-$74999		1.169	1.288	1.625	2.643
$75000+		1.169	1.202	1.790	2.740

Table 4.5 HB Social/Recreation Trip Production Rates

Household Income	Household Autos	Household Size			
		Size1	Size2	Size3	Size4+
$0-$24999	Veh0	0.330	0.700	0.900	1.200
$25000-$49999		0.330	0.700	0.900	1.200
$50000-$74999		0.330	0.700	0.900	1.200
$75000+		0.330	0.700	0.900	1.200
$0-$24999	Veh1	0.428	0.801	1.061	1.450
$25000-$49999		0.423	0.912	1.061	1.450
$50000-$74999		0.330	1.114	1.061	1.450
$75000+		0.524	1.596	1.061	1.450
$0-$24999	Veh2	0.694	0.802	0.855	2.101
$25000-$49999		0.600	0.895	1.143	2.164
$50000-$74999		0.602	0.952	1.409	2.056
$75000+		1.053	1.045	1.660	2.434
$0-$24999	Veh3+	0.717	0.611	1.880	1.753
$25000-$49999		0.925	0.867	1.759	2.390
$50000-$74999		0.615	0.776	1.800	3.725
$75000+		0.836	0.888	1.803	2.801

Table 4.6 HB Other Trip Production Rates

Household Income	Household Autos	Household Size			
		Size1	Size2	Size3	Size4+
$0-$24999	Veh0	0.535	1.057	1.310	3.719
$25000-$49999		0.535	1.057	1.310	3.719
$50000-$74999		0.535	1.057	1.310	3.719
$75000+		0.535	1.057	1.310	3.719
$0-$24999	Veh1	0.611	1.451	2.946	3.155
$25000-$49999		0.577	1.581	2.542	3.155
$50000-$74999		0.696	1.877	2.438	3.155
$75000+		0.936	2.011	4.455	3.155
$0-$24999	Veh2	0.630	1.181	2.823	4.000
$25000-$49999		0.717	1.368	2.168	4.151
$50000-$74999		0.768	1.330	2.160	4.839
$75000+		1.277	1.605	2.772	4.989
$0-$24999	Veh3+	1.175	0.973	1.832	4.216
$25000-$49999		1.175	1.310	2.375	4.563
$50000-$74999		1.175	1.443	2.347	4.285
$75000+		1.175	1.319	2.039	4.813

Table 4.7 NHB Work Trip Production Rates

Household Income	Household Autos	Number of Workers in Household			
		Wrks0	Wrks1	Wrks2	Wrks3+
$0-$24999	Veh0	0.000	0.377	0.558	1.000
$25000-$49999		0.000	0.377	0.558	1.000
$50000-$74999		0.000	0.377	0.558	1.000
$75000+		0.000	0.377	0.558	1.000
$0-$24999	Veh1	0.000	0.562	1.047	3.592
$25000-$49999		0.000	0.816	1.795	3.592
$50000-$74999		0.000	0.827	1.600	3.592
$75000+		0.000	1.299	1.657	3.592
$0-$24999	Veh2	0.000	0.758	1.180	3.592
$25000-$49999		0.000	0.720	1.761	3.592
$50000-$74999		0.000	0.893	1.559	3.592
$75000+		0.000	1.058	2.500	3.592
$0-$24999	Veh3+	0.000	0.428	1.101	3.592
$25000-$49999		0.000	1.027	1.708	3.592
$50000-$74999		0.000	0.818	1.834	3.592
$75000+		0.000	1.141	2.341	3.592

Table 4.8 NHB Other Trip Production Rates

Household Income	Household Autos	Household Size			
		Size1	Size2	Size3	Size4+
$0-$24999	Veh0	0.335	1.300	1.700	2.000
$25000-$49999		0.608	1.300	1.700	2.000
$50000-$74999		0.608	1.300	1.700	2.000
$75000+		0.608	1.300	1.700	2.000
$0-$24999	Veh1	0.872	1.337	1.294	2.174
$25000-$49999		0.676	1.435	2.832	3.626
$50000-$74999		0.800	1.554	2.700	3.638
$75000+		0.835	1.992	2.644	3.609
$0-$24999	Veh2	1.001	1.175	1.781	2.997
$25000-$49999		0.544	1.310	1.454	2.579
$50000-$74999		0.704	1.617	1.815	2.687
$75000+		2.245	1.382	1.995	3.347
$0-$24999	Veh3+	1.864	0.905	2.572	3.735
$25000-$49999		1.406	1.586	1.595	3.011
$50000-$74999		1.400	1.366	2.362	2.844
$75000+		1.400	1.236	1.587	3.458

Table 4.9 HB College/University and HB School Trip Production Rates

Trip Purpose	Household Autos /Income	Household Size			
		Size1	Size2	Size3	Size4+
HB College/University	All	0.016	0.036	0.266	0.363
HB School		0.000	0.042	0.491	1.623

The proposed school and college trip production rates are expressed as a function of household size only. The incidence of college trips per household is too small to allow for a more disaggregate classification. Similarly, the incidence of k-12 school trips per household is negligible for 1 and 2-person households. For the 3 and 4+ person households the school trip rates are not expected to vary much as a function of income or auto ownership, given the lack of discretion associated with these trips.

Moreover, the production of school trips, whether K-12 or college, is more directly correlated with the number of school age persons in the household than household size. More appropriate household classifications for school trip modeling include presence and age of children, number of school age children, and/or number of college age person. Such HBCU and HBSC production models bring additional input data requirements, in the form of additional variables needed to support the model, for both base and forecast years.

In instances where the transit system carries a large number of students, this additional model maintenance burden is justified by the gains obtained for transit modeling purposes. In such cases care should be taken to collect on-board survey information from middle and high school students, rather than omit them from the sampling plan as is often the case. And the discounted fare that students pay should be reflected in the mode choice model.

HBCU trips, on the other hand, are challenging to model well. The college/university student population consists of many sub-populations with distinct travel patterns: students living on campus, students living in group quarters outside campus, unrelated adults sharing private residences, and part and full-time students living on their own or with their families, either as heads of households or as dependents. The likelihood of using transit among these various populations can be significantly different, depending not just on their living arrangements but obviously also on how well transit serves the college campus. An additional challenge in modeling HBCU transit trips may the lack of relevant observed transit travel patterns. In some cities the HBCU population tends to be concentrated in a few corridors, requiring a dedicated on-board sampling plan to capture sufficient observations for modeling purposes. In other cities, such as Miami, the college student population is dispersed throughout the model area, so care should be taken to obtain sufficient responses to build a matrix of observed HBCU transit trips.

One approach used in New Starts projects has been to directly develop a person trip matrix from local data, in lieu of the trip generation and distribution models. To build this matrix, local residence address information for each of the major junior colleges and universities is obtained, and then an average trip rate per student is used to compute trips.

5 Trip Attraction Model

The state of the practice for trip attraction models are linear regression models, similar to those used by FSUTMS. The models are estimated as regressions through the origin (i.e., with no constant term), using observed NHTS attractions for the relevant trip purpose and market segment as the dependent variable, and employment by type, households, and/or school enrollment as the explanatory variables, summarized at the TAZ level. These models are rarely stratified by household income or any other household attribute, which limits their ability to link job categories to the right mix of household workers.

Two alternative approaches to the linear regression models will be discussed.[2] The first approach uses CTPP Part 2 data to predict the share of workers in various household income and auto ownership classes. The second approach relies on disaggregate worker at-place-of-work data, obtained from a workplace survey, to develop an attraction cross-classification model. It is also worth mentioning that, when destination choice models are used for trip distribution, the attraction models are no longer used, since they are replaced by the size terms of the destination utility. It is however not uncommon to use an attraction model or its forecasts as the size term itself.

The CTPP Part 2 provides various tabulations of workers at place of work by TAZ, including workers by auto ownership and household income. These data can be used to compute the share of workers in each household class, and then used to estimate a logit share model that forecasts the share of workers as a function of the share of jobs by employment type and other explanatory variables, such as employment density and accessibility. The advantage of this model over a series of liner regressions is that, by estimating all model coefficients simultaneously, their values can be investigated to ascertain whether they move in a logical direction as a function of increasing income or auto ownership.

A workplace survey data provides disaggregate information that can be used to develop a cross-classification model of trip attractions, similar to the production classification models. Trip attractions are classified by employment type and household attributes such as auto ownership and household income. Then trip attractions for each household class are computed using trip rates stratified by employment type.

Since the CTPP tabulations are not yet ready, and in the absence of a workplace survey, standard linear regression models were developed for the Tallahassee prototype model. All attraction models were re-estimated using the most recent available data. The trip attractions were calculated from the NHTS data and summarized by statewide TAZ. The state InfoUSA employment data was used to compute employment by TAZ for the three broad categories used in FSUTMS: industrial employment, commercial employment and service employment.

Separate HBW attraction models were estimated for each income group, to improve the ability of the model to link work productions with the correct work attractions, by income level. As discussed above,

[2] Freedman, J. et al. Comparing Stratified Cross-Classification and Logit-Based Trip Attraction Models. Proceedings of the Seventh TRB Conference on the Application of Transportation Planning Methods. March 7-11, 1999, Boston, MA.

income stratification is important for the HBW trip distribution model. For all other purposes, a single attraction model was estimated for each purpose.

The process of estimating the attraction models is as follows:

i. Prepare the trip attraction data. In addition to the data clean-up and coding steps performed for the trip production models, a statewide TAZ needs to be assigned to each trip location. The NHTS data includes (x,y) coordinates in decimal degrees for most trip locations, so that the TAZ can be assigned by creating a point shape file and overlaying it with the statewide TAZ shape file. Trip locations outside of the state should be flagged and excluded from the attraction model estimation, so that only attractions that are part of trips internal to the model region are used in the estimation. Similarly, trip locations with missing coordinate information need to be flagged and excluded.
ii. Prepare the employment data. The InfoUSA employment data was reviewed and summarized into the three employment categories used by FSUTMS[3]. In addition, the employment records were allocated to statewide TAZs by overlaying (x,y) coordinates on the statewide TAZ shape file.
iii. The estimates of total households by TAZ developed for the household classification models are combined with the employment and attractions data to form the attractions estimation file. This file takes the form of a TAZ-indexed table.
iv. Estimate the regression models. This can be accomplished with virtually any statistical software package.
v. Validate the trip attractions model by comparing aggregate total attractions to control totals developed from the NHTS data. The control totals will include observations excluded from the estimation dataset, for example due to missing trip purpose or geocoded to a TAZ with zero employment.

The estimated trip attraction rates for Florida are shown in Table 5.1. As indicated above, the HBW attractions were stratified by household income and separate coefficients were estimated for each income level. The resulting coefficients show decreasing importance of commercial employment with increasing income. This is expected given that commercial employment includes retail employment, which comprises a higher proportion of lower income jobs than most other industries. This trend can be more clearly appreciated when the HBW attraction rates are expressed relative to the service employment rate, as shown in Table 5.2. When aggregated across the income levels, the HBW trip attraction rates add up to 0.9 to 1.1 attractions per employee. These attraction rates are on the low side of the typical range. This may be because some trip observations were deleted from the NHTS data, such as those with missing income. It is expected that the trip attraction rates will be adjusted upwards when the model is validated to Tallahassee attractions.

Households and Commercial employment tend to attract the most non-work trips, followed by service employment. Industrial employment exhibited negative trip attraction rates for HBO and NHBO trips,

[3] Processing InfoGroup Data -- Summary Report. Technical Memorandum. Parsons Brinckerhoff, Nov. 5, 2010.

and a positive but insignificant rate for HBSR. Therefore industrial employment will be excluded from all non-work models except NHBW.

The HBCU and HBSC attraction rates were transferred from the SCAG regional model, as school enrollment data was not readily available. These trip rates can be modified at a later date.

Table 5.1 Trip Attraction Rates

Trip Purpose	Industrial Employment		Service Employment		Commercial Employment		Households		School Enroll.	R^2
	Coeff.	t-Stat	Coeff.	t-Stat	Coeff.	t-Stat	Coeff.	t-Stat		
HBW Inc1	0.0663	1.3	0.0910	7.4	0.2218	7.1				0.07
HBW Inc2	0.4764	6.9	0.1943	11.5	0.3642	8.5				0.17
HBW Inc3	0.1457	2.6	0.1932	14.3	0.2560	7.4				0.15
HBW Inc4	0.5202	7.9	0.4225	26.5	0.3132	7.7				0.35
HBSh					3.2805	56.0				0.37
HBSR			0.2359	8.0	0.6814	8.6	0.7368	35.5		0.42
HBO			0.4910	12.1	1.3213	12.9	1.1578	39.8		0.51
HBSc									1.326	
HBU									0.549	
NHBW	0.6346	7.6	0.4534	22.1	0.7118	13.0	0.1773	12.3		0.45
NHBO			0.2443	7.5	1.947	23.6	0.7514	32.1		0.51

Table 5.2 HBW Trip Attraction Rate Multipliers

Trip Purpose	Common Trip Rate	Trip Rate Employment Multiplier: Industrial Employment	Service Employment	Commercial Employment
HBW Inc1	0.0910	0.728	1.000	2.437
HBW Inc2	0.1943	2.452	1.000	1.874
HBW Inc3	0.1932	0.754	1.000	1.325
HBW Inc4	0.4225	1.231	1.000	0.741

6 Trip Balancing

Standard practice in trip-based models is to balance trip attractions to productions on a regional basis. Where this method does not result in an adequate distribution of trip attractions, the underlying reasons should be explored, and a solution implemented accordingly. Subarea balancing may prove to be a feasible solution for non-work trips. Other possible solutions include stratified trip rates by area type or other indicator of density and land use form, a more disaggregate classification of employment types, and the use of special trip attraction generators.

Subarea balancing is not recommended for HBW trips, because it can mask real and important imbalances between jobs and housing availability.

School trips are an exception to the rule of balancing to productions. When there is high confidence in the school enrollment estimates, it is recommended that the HBCU and HBSC productions be balanced to attractions. This assumes that the HBCU trips produced by students living in group quarters are in included in the trip production totals.

The recommended guidance for the Tallahassee prototype is to balance all trip purposes to productions at the regional level, except for HBSC and HBCU which should be balanced to attractions.

7 Special Trip Markets

Trip-making populations other than permanent residents may constitute important trip markets for transit modeling in Florida. The extent to which these populations use the transit system today is one factor to consider when deciding whether and how to develop special market models for transit purposes. Another consideration is the extent to which future transit services may be planned to serve one or more of these populations.

Special trip market populations include the following:

- Seasonal residents
- Visitors
- Air passengers
- Special events

Seasonal Residents. Florida seasonal residents are characterized by older, retired persons. The trip behavior of seasonal residents may be gleaned from the seasonal resident subsample of the Florida NHTS Add-On survey, possibly augmented with data from the resident population of similar characteristics to seasonal residents, in terms of age profile, employment status, household size and household income. The more challenging effort may be to quantify the seasonal resident population and to locate their households at the TAZ level.

There is very limited experience modeling seasonal urban travel in the United States. In Lake Tahoe, California, a seasonal resident travel model was developed using household survey data. The model follows a structure similar to the permanent resident model, albeit simplified where sample sizes were insufficient to support the more complex models. Counts of residences occupied by seasonal residents were obtained from the Lake Tahoe MPO, separately for winter and summer months.

Visitors. Tourists and other visitors are of course an important travel market in Florida. Currently FSUTMS accounts for hotel-based visitor travel. A visitor model is best informed by a focused survey effort, such as those conducted in Honolulu, Las Vegas, and Lake Tahoe. Important attributes of the visitor population include trip purpose (business, vacation, visiting family), length of stay, type of accommodation (hotel/motel, with friends/family, time shares, etc.), party size and party composition. Visitor travel patterns can be quite different depending on the type of travel and even the type of visited attractions. An understanding of the relevance of visitors as potential patrons of transit systems and of their travel patterns should be formed prior to attempting a visitor travel model for transit planning purposes.

The visitor travel model developed for Honolulu is shown in Figure 7.1. The nested logit model predicts trip frequency, destination, and mode choice. The most attractive feature of the model is the feedback or influence on upper level choices of lower level logsum values. In other words, improvements in accessibility to a particular destination, for example due to a new rail system, will improve the attractiveness of the destination and the likelihood that visitors will travel to the destination.

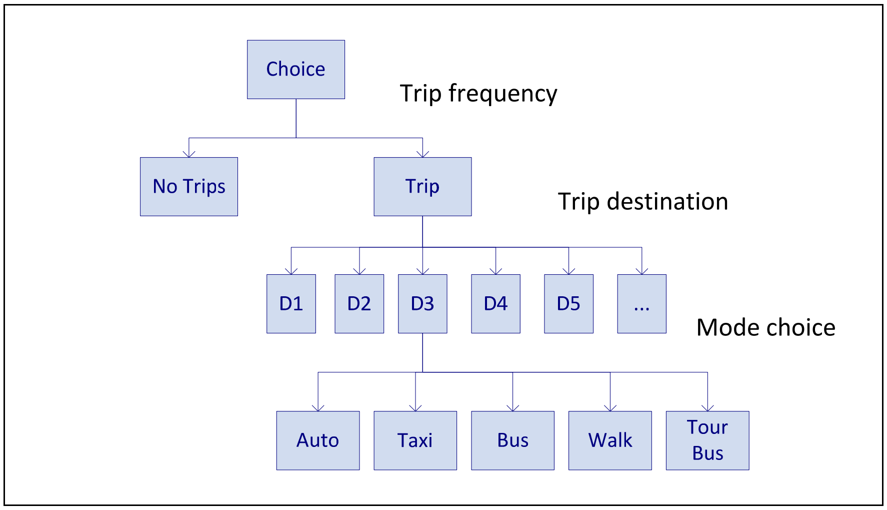

Figure 7.1 Honolulu Visitor Travel Model

Another example is the Portland Central City Visitor model.[4] It is a traditional four step model, with hotel-based trip generation, destination choice, mode choice, time-of-day factoring, and assignment. It uses mode choice logsums for destination choice, which is essential given the high frequency of walk and transit trips observed in downtown Portland. All model components were estimated with visitor data, collected with a survey that targeted visitors staying at downtown hotels. The survey included revealed preference questions, aimed at collecting a full day's worth of travel and personal and household information, and stated preference questions used to develop the mode choice model and compare and contrast preferences for transit relative to Portland residents' preferences.

Air Passengers. Where rail systems provide ground access to airports, the air passenger population becomes an important travel market. Several examples of airport ground access trip generation, distribution and mode choice models are available in the literature, sometimes combined with airport choice models. The trip generation component of these models is usually based on passenger enplanements and stratified by type of travel (business versus leisure) and residence (permanent versus visitor). As such, the trips are first 'generated' at the airport and then distributed to appropriate trip origins depending on the type of trip. It is critically important to accurately represent transit levels of service when forecasting these trips, along with all the various modal options not typically included in a regional model (shuttles, hotel courtesy vans, and taxicabs, for example). Level of service characteristics such as frequency by time of day, walk time to the airport terminal, accessibility to regional employment centers and visitor attractions, and regional transit connections can all make a substantial difference in rail-to-airport ridership.

[4] Parsons Brinckerhoff. Portland Central City Visitor Model. Model Design Report. June 2010.

One example of an airport ground access model is the Airport Passenger Travel Demand Model, developed for the Port of Portland in 2009.[5] This model forecasts the non-airport trip end, and ground access trip time of day and mode split as a function of passenger enplanements and readily available socio-economic data. The model was designed to be consistent with the data used by the Portland Metro model, but implemented as a stand-alone module. Trips are stratified by type of traveler (resident or visitor), trip purpose (business or leisure) and ground access trip end (internal or external to the Portland metropolitan area). For each these markets, the model forecasts trip origin/destination, time of day and mode choice--see. The model parameters were estimated with data from surveys administered to terminal users at Portland International Airport in June and September of 2008. The availability of disaggregate trip data permitted the estimation of destination choice and mode choice models. The model uses mode choice logsums in the destination choice models, an important attribute given that the airport is conveniently served by the region's light rail system.

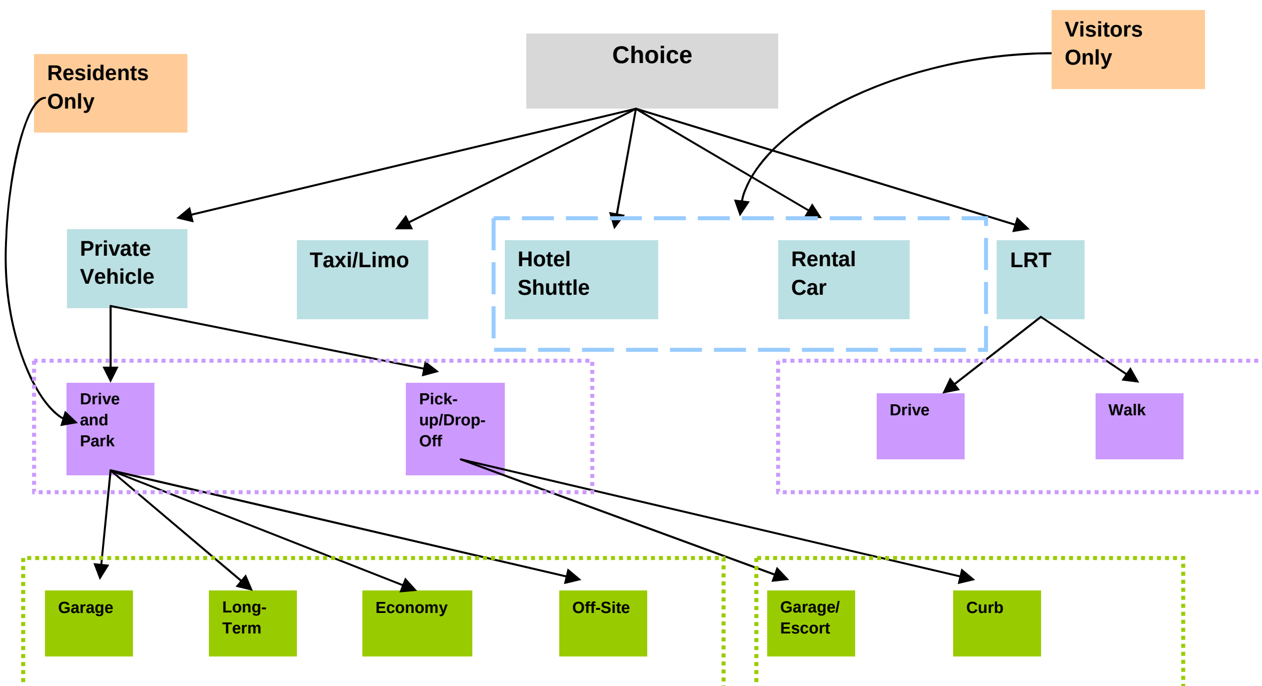

Figure 7.2 Portland Air Passenger Ground Access Mode Choice Model

Special Events. Travel by residents and visitors to events such as baseball games, festivals, convention centers and other similar venues falls under the umbrella of special event travel. Models to forecast transit ridership and user benefits for special events have been developed and used to support applications for FTA New Starts funding. One such example is the Phoenix EVENT model [6]. Starting with total attendance at each event, the model distributes the event trips to origins throughout the model area. A mode split model is then applied to forecast the share of transit trips to the event. Like the

[5] Parsons Brinckerhoff. Airport Passenger Travel Demand Model User's Guide. Portland, OR, 2009.
[6] Parsons, Brinckerhoff, Quade and Douglas. EVENT Model Development and Verification Report. Memo #6, Phoenix Model Enhancement for LRT Forecasting. December 1999.

Honolulu visitor model, the Event model uses mode choice accessibilities to forecast trip distribution. This model was developed in recognition of the importance of light rail to serve various event venues in the proposed corridor. The trip generation, distribution and mode choice models were estimated using data from a stated-preference survey administered to event patrons. The EVENT model has been transferred to other metropolitan areas, including Portland, OR, and Los Angeles, CA.