RTC Washoe Travel Demand Model

Development Documentation

Prepared by: Caliper Corporation

January 14, 2025

# Resident Productions

Productions are predicted at the person level using decision trees for the machine learning stack. Compared to traditional cross-classification matrices, they use the survey samples more efficiently and can consider more variables. While state of the art decision trees methods including bagging and boosting achieve slightly higher prediction accuracy, they are opaque. The models below are simple trees that are easy to understand and perform nearly as well.

Trip purposes in the model are first stratified by the tour type they occur on (work vs non-work), and are laid out below:

* Trips on work tours
  + Home-based work (W\_HBW)
  + Home-based other (W\_HBO)
* Trips on non-work tours
  + Home-based social/recreational (N\_HBSR)
  + Home-based other (N\_HBO)
  + Home-based school (N\_HBSCH)
  + Home-based shop (N\_HBSHP)

For most purposes, separate decision trees were estimated for each market segment. The market segments are:

* Zero vehicle (v0)
* Vehicle insufficient (vi)
* Vehicle sufficient (vs)

## W\_HBW

The home-based work purpose is sensitive to the usual variables like employment status of the person, their age, and their households income per capita (income divided by size). The models are also sensitive to various measure of accessibility:

* Transit accessibility (t\_access)
* General/auto accessibility (g\_access)
* Walk accessibility (w\_access)

### v0

A diagram of a number

Description automatically generated

### vi

A diagram of a number

Description automatically generated

### vs

A diagram of a number

Description automatically generated

## W\_HBO

This purpose is sensitive to variables like the presence of seniors, gender, size, income, and age.

### v0

A diagram of a number

Description automatically generated

### vi

A diagram of a number of individuals

Description automatically generated

### vs

A diagram of a number of individuals

Description automatically generated

# N\_HBSR

The social/recreational models are sensitive to variables like accessibility, employment, per capita income, age, and presence of children.

### v0

A diagram of a number

Description automatically generated

### vi

A diagram of a number of adults

Description automatically generated

### vs

A screenshot of a computer screen

Description automatically generated

## N\_HBO

These models are sensitive to variables like age, gender, accessibility, and employment status.

### v0

A diagram of a number of adults

Description automatically generated with medium confidence

### vi

A diagram of a number of people

Description automatically generated

### vs

A diagram of a family tree

Description automatically generated

## N\_HBSCH

The most important predictor for these model is the age of the person, but other factors like accessibility and income do influence the number of school trips made. This purpose is not stratified by market segment.

A diagram of a number

Description automatically generated

## N\_HBSHP

Shopping trips are influenced by employment status, accessibility, age, income, along with the presence of kids and seniors.

### v0

A diagram of a number of individuals

Description automatically generated

### vi

A diagram of a number of individuals

Description automatically generated

### vs

A diagram of a tree

Description automatically generated

## Calibration

Calibrating the production rates is done to ensure that the final model is producing the same number of trips per person on average as the survey. One complicating factor in this comparison is that the survey and model socio-economic data have different total populations.

* Survey: 476,187
* Model: 523,542

As a consequence, the survey trip totals are increased by 1.1 and then compared to model results. The table below shows the production model results compared back to the total trips in the survey (factored up). While some purposes are higher or lower than the survey, the model does a good job of predicing total trips even before calibration.

A screenshot of a data table

Description automatically generated

The calibration factors below are the ratio of the observed to modeled trips. These are applied by trip type and auto sufficiency segment to ensure total trip making matches the survey (based on trip weight).

A screenshot of a screen

Description automatically generated

# Non-motorized Split

After resident productions are estimated, the model removes non-motorized trips (bike and walk) using a binary logit model. For each trip purpose, the estimated coefficients are shown along with the adjusted rho-squared and the calibration constant applied to match survey shares of non-motorized trips.

## W\_HBW

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Coefficient** | **t-Statistic** | **Alternatives** | |
| **moto** | **non-moto** |
| HH Kids | 0.066 | 2.8107 |  | X |
| HH Adults | 0.047 | -1.3555 |  | X |
| Vehicle per Adult | 0.254 | -11.657 |  | X |
| Walk Access | 0.073 | 1.2553 |  | X |
| Constant | 0.270 | -3.7151 |  | X |
| **Rho^2** | **0.66** |  |  |  |
|  |  |  |  |  |
| Calibration Constant | 0.60 |  |  | X |

## W\_HBO

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Coefficient** | **t-Statistic** | **Alternatives** | |
| **moto** | **non-moto** |
| HH Kids | -0.290 | -2.213 |  | X |
| HH Adults | 0.149 | 2.2412 |  | X |
| Vehicle per Adult | -3.807 | -7.922 |  | X |
| Walk Access | 0.093 | 0.7413 |  | X |
| Constant | -1.156 | -2.528 |  | X |
| **Rho^2** | **0.69** |  |  |  |
|  |  |  |  |  |
| Calibration Constant | 1.03 |  |  | X |

## N\_HBO

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Coefficient** | **t-Statistic** | **Alternatives** | |
| **moto** | **non-moto** |
| Vehicle per Adult | -0.490 | -2.9348 |  | X |
| Walk Access | 0.470 | 5.9330 |  | X |
| Senior | -0.942 | -2.5293 |  | X |
| Constant | -2.786 | -11.476 |  | X |
| **Rho^2** | **0.56** |  |  |  |
|  |  |  |  |  |
| Calibration Constant | 0.31 |  |  | X |

## N\_HBSCH

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Coefficient** | **t-Statistic** | **Alternatives** | |
| **moto** | **non-moto** |
| HH Kids | 0.320 | 4.4758 |  | X |
| HH Adults | -0.527 | -3.325 |  | X |
| Vehicle per Adult | -1.060 | -3.888 |  | X |
| Walk Access | 0.609 | 4.2519 |  | X |
| Constant | -1.843 | -3.051 |  | X |
| **Rho^2** | **0.48** |  |  |  |
|  |  |  |  |  |
| Calibration Constant | 0.19 |  |  | X |

## N\_HBSHP

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Coefficient** | **t-Statistic** | **Alternatives** | |
| **moto** | **non-moto** |
| Vehicle per Adult | -3.054 | -10.159 |  | X |
| Walk Access | 0.253 | 2.6459 |  | X |
| Senior | -0.661 | -2.0033 |  | X |
| Constant | -1.315 | -5.0432 |  | X |
| **Rho^2** | **0.65** |  |  |  |
|  |  |  |  |  |
| Calibration Constant | 0.66 |  |  | X |

## N\_HBSR

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Coefficient** | **t-Statistic** | **Alternatives** | |
| **moto** | **non-moto** |
| HH Kids | 0.369 | 6.6602 |  | X |
| HH Adults | -0.117 | -1.481 |  | X |
| Vehicle per Adult | -2.378 | -9.723 |  | X |
| Walk Access | 0.453 | 4.8056 |  | X |
| Senior | -0.612 | -1.515 |  |  |
| Constant | -1.816 | -5.326 |  | X |
| **Rho^2** | **0.63** |  |  |  |
|  |  |  |  |  |
| Calibration Constant | 0.46 |  |  | X |

# Time of Day

Peak period determination is based on an analysis of the trips in motion throughout the day. First, the day is broken up into 15-minute increments. A trip from the household survey is said to be “in motion” if any portion of the trip occurs within the 15-minute bin. As a result, a single trip can be counted in multiple bins. Determining the peak period considers the distribution of all trips as well as the distribution of trips on work tours individually. The chart below shows these two distributions, which look as expected with the AM peak being shorter and more condensed than the PM.

A graph of blue and orange lines

Description automatically generated

The next step is to determine the peak hour for all trips and those on work tours. In the tables below, the AM and PM peak hours by type are presented. At a minimum, the AM and PM peak periods should contain the respective peak hour for work trips and all trips.

A screenshot of a computer

Description automatically generated

A graph of different colored lines

Description automatically generated

Based on the trips in motion profile, there are four distinct periods of the day. AM and PM peaks have the highest intensity of trip making, followed by mid-day, with the overnight period containing the fewest trips in motion. Boundaries for these periods are defined such that the variance of trips in motion within periods is smallest while variance between periods is largest.

The final period definitions are shown in the table below including the mid-day (MD) period. The remaining hours of the day are captured in the night (NT) period. In the chart, the final period is represented by the gray rectangle.

A screenshot of a schedule

Description automatically generated

A graph of different colored lines

Description automatically generated

# Destination Choice

Rama/Srini

# Mode Choice

Rama/Srini

# Non-Home-Based

There are many problems related to non-home-based trips in traditional trip-based models arising from the fact that they are disconnected from the home-based trips with which they comprise complete tours. In order to properly represent non-home-based trips, two spatial distribution or destination/spatial choice models are required to account for both the trip’s origin location and destination location. The four-step model architecture is fundamentally flawed because it produces non-home-based trips from only one trip distribution or spatial choice model.

To address these problems, the model adopts an alternative approach with a simple change to the structure of the trip-based model, running the non-home-based model components after and conditional on the home-based model components instead of in parallel and independently of them as in the traditional four-step model. This relatively simple structural change significantly improves the model’s ability to represent non-home-based trips and their response to land use changes and transportation infrastructure investments. Running a NHB distribution or destination choice model after and conditional on home-based destination choices in this approach, provides the required second spatial distribution model to properly model both the origin and destination of NHB trips.

In this approach, NHB trips are generated separately by mode based on home-based mode choices. This essentially provides information about whether a traveler has a car with them and allows the model, despite its trip-based form, to ensure a reasonable consistency of modes on tour.

Available modes include:

* SOV
* HOV2
* HOV3+
* Auto Pay
* Transit
* Non-motorized

The model coefficients below are the result of multiple linear regression with a forced intercept at zero. In addition, the model estimation will be scaled up to a predicted regional total. As a result of these two factors, the displayed r-squared values are not as meaningful. Instead, the value of the coefficients is in determining the relative effect that various home-based trip types have on non-home-based trip generation.

When reviewing the coefficients below, note their logical consistency: SOV NHB trips are most likely to result when the HB trip is SOV or HOV. NHB walk trips can be made when a person drives from home, but is more much more likely if they walk. These results greatly improve the NHB models compared to traditional trip-based construction.

In the tables below, any “alpha” and “gamma” terms refer to boosting coefficients used. See the section on boosting for an explanation of those terms.

## W\_NHBO

These are non-home-based trips made on a work tour not related to work. In other words, neither end of the trip is work or home.

### SOV

A screenshot of a computer

Description automatically generated

### HOV2

A screenshot of a computer

Description automatically generated

### HOV3+

A screenshot of a computer

Description automatically generated

### Auto Pay

No significant presence of this trip type in the survey, and it is not an important source of travel demand. No model was estimated.

### Transit

A screenshot of a computer

Description automatically generated

### Non-motorized

A screenshot of a phone

Description automatically generated

## W\_NHBW

These are non-home-based trips made on a work tour where one trip end is work or work-related.

### SOV

A screenshot of a computer

Description automatically generated

### HOV2

A screenshot of a computer

Description automatically generated

### HOV3+

A screenshot of a computer

Description automatically generated

### Auto Pay

A screenshot of a calculator

Description automatically generated

### Transit

A screenshot of a calculator

Description automatically generated

### Non-motorized

A screenshot of a computer

Description automatically generated

## N\_NHBSHP

These are non-home-based trips made on a non-work tour with one trip end for shopping.

### SOV

A table of numbers and letters

Description automatically generated

### HOV2

A screenshot of a computer

Description automatically generated

### HOV3+

A screenshot of a computer

Description automatically generated

### Auto Pay

A screenshot of a computer

Description automatically generated

### Transit

A screenshot of a calculator

Description automatically generated

### Non-motorized

A screenshot of a phone

Description automatically generated

## N\_NHBO

These non-home-based trips are those where neither end is shopping. In effect, these are all other non-home-based trips made on non-work tours. Visiting friends or family at their home is an example of this trip type.

### SOV

A screenshot of a table

Description automatically generated

### HOV2

A table of numbers and letters

Description automatically generated

### HOV3+

A table of numbers and letters

Description automatically generated

### Auto Pay

A screenshot of a computer

Description automatically generated

### Transit

A screenshot of a computer

Description automatically generated

### Non-motorized

A screenshot of a calculator

Description automatically generated

## Boosting

“Boosting” is an approach borrowed from machine learning where the errors of a previous model are used to estimate a second model. The models shown above illustrate the link between home- and non-home-based trips by mode, but intuition and experience tells us that accessibility should also influence NHB trip making. Travel to a central business district is more likely to lead to further trips compared to traveling to a rural zone.

The chart below shows how accessibility impacts each trip type and mode combination. The y-axis is a simple factor. When the factor is 1, the trip rates will be the same as displayed in the tables above. A y-value of 0.5 means the trip rates will be reduced by 50 percent. Conversely, a y-value of 1.5 means that trip rates are increased by 50 percent. With this additional factor, the model will understand the role that accessibility plays in NHB trip generation.

A graph of different colored lines

Description automatically generated

## NHB Time of Day Adjustment

The independent feedback by time of day in the model presents a unique challenge for NHB. NHB and HB trips have very different time of day patterns (see Time of Day documentation), but NHB generation is simply rates multiplied by HB trip ends. Without correction, this would mean that the NHB trips would have the same distribution as HB trips. To correct this, Caliper calculated adjustment factors by tour type, mode, and time of day. This was done by comparing raw model outputs back to the observed NHB trips from the survey.

The easiest way to understand the table below is by highlighting one row as an example. For SOV trips on work tours, these factors move NHB trips out of the AM and NT periods and into the MD and PM. Another way of saying it is that, compared to HB trips, NHB trips are less likely in the AM/NT and more likely in the MD/PM. The large mid day factor reflects lunch trips and other mid-day activities. The relative size of the AM and PM periods confirms what we know: that stops are more likely to be made on the way home compared to the way to work (e.g. picking up groceries).

A screenshot of a computer screen

Description automatically generated

# Validation

Once the behavioral models were estimated and calibrated using survey data, Caliper ran the full model and compared outputs to traffic counts (2023 counts). This data provided a check on model performance in the base year and built confidence that the model can be used in the future.

The table below presents the percent difference and percent root mean square error, which are measures of how well the model matches counts in aggregate. The table further breaks down this metric by volume group. The model matches count closely both overall and by volume group.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Volume Group | N | Total Count | Total Volume | % Difference | %RMSE |
| 10000 | 517 | 2052000 | 2147510 | 4.65 | 69.33 |
| 25000 | 199 | 3160300 | 3152887 | -0.23 | 38.24 |
| 50000 | 57 | 1908600 | 1912780 | 0.22 | 18.69 |
| 100000 | 17 | 1199900 | 1285451 | 7.13 | 13.96 |
| 100000+ | 11 | 1509300 | 1527212 | 1.19 | 8.19 |
| All | 801 | 9830100 | 10025839 | 1.99 | 37.01 |

The second table presents the same metrics by HCM type. The model uses independent capacities for each of these roadway types, and this table makes sure each is performing well. Major collectors show large metrics, but with only two counts, this is not cause for concern.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| HCM Type | N | Total Count | Total Volume | % Difference | %RMSE |
| Arterial | 235 | 1865100 | 1807053 | -3.11 | 48.98 |
| Collector | 109 | 276800 | 210763 | -23.86 | 86.79 |
| Freeway | 252 | 4543800 | 4885856 | 7.53 | 25.41 |
| Local | 30 | 56600 | 61305 | 8.31 | 94.52 |
| MajorArterial | 173 | 3087000 | 3059609 | -0.89 | 35.8 |
| MajorCollector | 2 | 800 | 1253 | 56.62 | 148.27 |
| All | 801 | 9830100 | 10025839 | 1.99 | 37.01 |

In addition to aggregate checks, Caliper performed link-level validation using maps like the one shown below. In this map, red colors show where the model is higher than counts while blue show where it is lower. Green indicates that the deviation is within the maximum desirable deviation range. No model with useful sensitivity will achieve a green color for all links. These maps guide the review process by showing where errors are largest and suggesting the presence of network or other errors.

A map of a city

Description automatically generated

## Transit

Transit validation is done by checking total transit ridership to observed boardings and alightings. The table below shows unlinked trips (total boardings) for both model and observed. The observed boardings come from the 2024 APTA report.

Model: 18,490

Observed: 17,700

The model accurately predicts the aggregate level of transit usage.