SANDAG Commercial Vehicle Model and Heavy Truck Model Update

Model Estimation and Calibration

prepared for

San Diego Association of Governments

prepared by

Cambridge Systematics, Inc.

December 2024

www.camsys.com

SANDAG Commercial Vehicle Model and Heavy Truck Model Update

Model Estimation and Calibration
Draft Technical Memorandum

prepared for

San Diego Association of Governments

prepared by

Cambridge Systematics, Inc.
101 Station Landing, Suite 410
Medford, MA 02155

date

December 2024

Table of Contents

List of Tables

List of Figures

1.0 Introduction

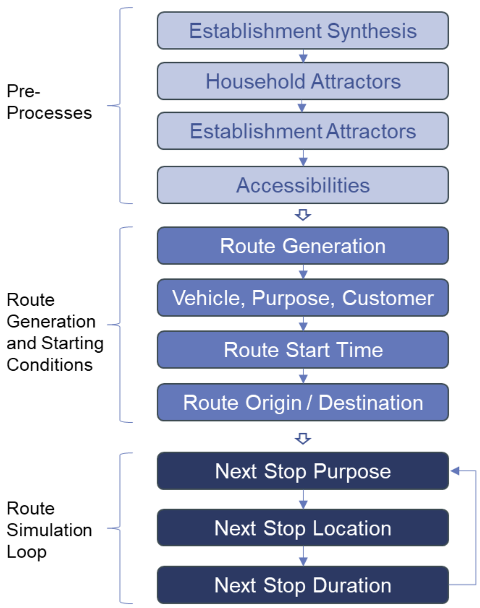
The SANDAG Commercial Vehicle Model (CVM) simulates the weekday demand patterns of commercial vehicle movements throughout the San Diego region. The CVM is an important part of the complete travel demand modeling system for the region, representing a market of travel that dominates the middle part of most weekdays and has been steadily growing as consumer demands for home deliveries and personal services have increased. The design of the CVM is covered in a separate model design technical report, CVM Model Design document, **CVM Model Design** . This report also describes improvements made to the Heavy Truck Model (HTM).

The focus areas of this technical report are the estimation of CVM model components and the subsequent calibration of those components after being implemented in the application software. Readers who would like additional information to better understand the behavioral aspects of model components are encouraged to refer to the model design document. In addition, the last section of this report covers the steps taken to update the HTM.

2.0 CVM Model Component Estimation

The components of the CVM were estimated using the 2022 SANDAG Commercial Vehicle Survey, creating a completely new specification using a consistent data set for all model components. The model components discussed in this section were specified based on the CVM Model Design document. The ordering of components follows from the model design flow diagram in Figure 2.1-1.

Figure 2.1-1 CVM Components



2.1 Establishment Synthesis

There are no components of the Establishment Synthesizer that require estimation. The model is self-calibrating, and validation results are provided in 3.1.

2.2 Household Attraction Generation

Household Attraction Generation model predicts which households in ABM3 will attract commercial vehicle trips for food deliveries (restaurant, groceries), package deliveries (all types), and for service stops (e.g., plumbers, landscapers, healthcare services, etc.).

2.2.1 Dataset

The Household Attraction Generation model is the only model component that does not use the Commercial Vehicle Survey. This model was estimated using SANDAG's 2022 Resident Travel Survey, which asked respondents whether they received deliveries of food or other packages, or whether they received a service stop on the survey day. The dependent variable is formulated as a binary choice and segmented by each of the three commercial vehicle stop types.

2.2.2 Variable Specification

A separate model was estimated for each of the three purposes. The final set of estimated parameters for the Household Attraction Generation model are shown in Table 2.2-1. Only statistically significant parameters were retained in the final specification were retained. The specification testing process considered many different variables which did not make it into the final model.

Table 2.2-1 Household Attractions Model Binary Logit Estimation by Attraction Type: Probability of > 0 Attractions

Commented [JG1]: This table has been updated. The parameter names have been rewritten to be clearer. In addition, there were some missing rows for Service Stops which have been added for Weekday and Age of Oldest HH Member.

	Food Delivery		Package Delivery		Service Stops	
Parameter Name	Coefficient	t-Stat.	Coefficient	t-Stat.	Coefficient	t-Stat.
Constant: Stops	-3.334	-44.56	-2.136	-18.27	-5.569	-15.88
Is Weekday (Mo-Th)			0.468	6.35	1.402	5.51
Is Friday			0.522	5.27	1.092	3.51
Household Size			0.177	6.62		
Oldest HH Member						
Age 55 to 74					0.740	4.48
Age 75-plus					1.569	6.14
Household Income						
$50k to $99k			0.292	2.91	0.508	1.71
$100k to $199k			0.619	6.58	0.776	2.76
$200k-plus			0.676	5.99	1.306	4.39

The following observations are relevant to this specification:

- Food delivery has a smaller number of observations than the other types (4.8% of households reported a food delivery). In particular, the number of people in the household, zero-auto, auto sufficiency, and income segmentation all showed weak significance for explaining variation. In addition, the number of homebased tours made by the household was tested in various combinations—total, for shopping or dining—and were found to be nonsignificant.
- Parameter estimates for package deliveries (41.8% of households) and for service stops (6.8% of households) were more robust and permitted additional segmentation. The estimation data set include all days of the week; therefore, parameter were estimated for Weekdays (Monday-Thursday) and Fridays, which proved significantly different from weekends and were retained to control for differences between days of the week. In the calibrated production version of the model, only the weekday parameter is used.
- Package delivery is significantly positively correlated with more household members.
- For both package delivery and for service stops, the parameter estimates show greater numbers of stops with higher levels of household income. For service stops, there was a significant positive correlation with older households more likely to receive services.

Commented [BS2]: Is this turned off for regular application of the model?

Commented [JG3R2]: correct

Commented [JG4]: Revised descriptions.

2.3 Establishment Attraction Generation

Establishment Attraction Generation model predicts which establishments in the CVM will attract commercial vehicle trips for either goods pickup/delivery or for service calls and how many visits they will attract. As noted in the CVM Model Design document, this is a two-stage model which first predicts whether an establishment will generate at least one attraction, followed by a second model which predicts how many attractions will be generated given there will be at least one.

2.3.1 Dataset

The Establishment Attraction Generation model was estimated using Commercial Vehicle Survey, which asked respondents how many commercial vehicle visits they receive on an average business day for either cargo or services.

2.3.2 Variable Specification

The final set of estimated parameters for the first stage of the Establishment Attraction Generation model are shown in Table 2.3-1. The primary attributes for this model were establishments' industry groups and number of employees. Based on a preliminary analysis of the similarities of attraction generation responses, four groups were formed and separate models estimated for each. Only statistically significant parameters were retained in the final specification were retained.

Table 2.3-1 Establishment Attractions Probability of > 0 Attractions - Binary Logit

Group 1 - Manufacturing, Wholesale, Retail		
Parameter Name	**Coefficient**	**t-Stat.**
employment	0.022	2.75
constant	1.755	12.09

Group 2 - Utilities, Leisure/Food		
Parameter Name	**Coefficient**	**t-Stat.**
employment	0.017	1.86
constant	1.375	6.69

Group 3 - Agriculture/Mining, Construction, Transportation, Education/ Public/Other, Health/Medical		
Parameter Name	**Coefficient**	**t-Stat.**
employment	0.032	4.99
constant	0.920	10.14

Group 4 - Info/FIRE/Professional Services		
Parameter Name	**Coefficient**	**t-Stat.**
employment	0.027	3.09
constant	0.592	4.52

The final set of estimated parameters for the second stage of the Establishment Attraction Generation model are shown in Table 2.3-2 and Table 2.3-3. The two parts to this model were developed to capture both the non-linear relationship between employment and attractions generated as well as variation across industries. Several complex model formulations were tested. Ultimately, it was decided that breaking the model into two steps would be easier to estimate and calibrate to achieve the desired behavior:

1. Nonlinear effects of total employment: $z_n = \alpha X_n$

Where z_n is a preliminary estimate of attraction stops for establishment *n*; X_n is the square-root of employment at the establishment; and α is the estimated coefficient. Table 2.3-2 show the results for a simple regression of the number attraction stops on the square root of employment of the establishment. Various nonlinear forms were tested, and the square root formulation provided both a good fit and a reasonable distribution of predictions.

2. Industry group factors as a modifier: $y_n = \beta_g z_n$

Where y_n is the final estimate of attraction stops for establishment *n*; z_n is the preliminary estimate of attraction stops from Step 1; and β_g is an estimated industry-group specific factor.

Commented [JG5]: New text

Table 2.3-2 Establishment Attractions Square Root of Employment - OLS

R-squared (uncentered):	0.460	
Adj. R-squared (uncentered):	0.459	
Parameter (α)	**Coefficient**	**t-Stat.**
square root of employment	0.705	42.91

Table 2.3-3 shows the estimation results for the industry group factors, which are multiplied times the results from the calculation on square root of employment. Application of these factors provides attraction generation estimates that vary by industry.

Commented [BS6]: The connection between models in 2.32 and 2.33 is a bit unclear. OLS was first estimated at the establishment level to get the coeff on sqrt of employment irrespective of industry? Did that value enter the OLS in 2.33 to estimate industry-specific coefficients? I believe equations here or in the design document would help. I know there are equations in the design document but I am unclear on exactly how many models were estimated here?

Commented [JG7R6]: I hope the above equations make this clearer.

Table 2.3-3 Establishment Attractions Multiplicative Factors by Industry Group - OLS

R-squared (uncentered):	0.473	
Adj. R-squared (uncentered):	0.470	
Parameter Name (β_g)	**Coefficient**	**t-Stat.**
Agriculture/Mining	1.140	7.61
Construction	0.819	9.74
Education/Public/Other Services	0.826	13.68
Industrial/Utilities	1.056	6.33
Info/FIRE/Prof. Services	0.980	14.81
Leisure/Accom./Food	0.934	13.57
Manufacturing	0.950	15.71
Medical/Health Services	1.026	16.48
Retail	1.055	16.53
Transportation	1.633	12.50
Wholesale	1.370	14.23

2.4 Accessibility Variables Impedance Function

The CVM Model Design document, **CVM Model Design** technical memo describes a method used to calculate the impedance curves representing the rate of decay in attractiveness as a function of generalized travel time. As discussed in that report, generalized travel time is composed of congested travel time and toll costs which have been converted to equivalent in-vehicle time minutes using assumed values of time. The calculation details of this method are described below.

2.4.1 Dataset

The primary data source used to develop the impedance curves is the weighted trip records from the Commercial Vehicle Survey. The data were conditioned to produce observations of the travel times and costs between every establishment location and every stop location made by a commercial vehicle belonging to that establishment. The motivation for this model element is to measure the spatial relationship between an establishment and its customers, i.e., how far away are the customer served by that establishment as measured in generalized travel time units and what is the sensitivity to this generalized travel time.

2.4.2 Variable Specification

As discussed in the **CVM Model Design** technical memo, the method uses the following formula:

$$Accessibility_i^k = ln\left(\sum_{\forall j \in J} Attractors_j^k \times \exp\left(\theta_k * (\max(0, GenTravelTime_{ij} - DecayOnsetTime_k)\right)\right)$$

in which the accessibility of zone *i* to attractors of type *k* is measured by the sum of the attractors in all zones j, weighted by the impedance of generalized travel time between zone *i* and each zone *j*.

The decay onset is an assumed amount of travel time to which the decision maker (establishment or driver) is insensitive. For example, a 4-minute decay onset says that drivers see no diminishment of attractiveness in serving customers within the 4-minute time threshold and only begin to consider travel effects for lengthier times. Given an assumed decay onset time, which is a judgement call, a quadratic formula is used to solve for the decay parameter θ as follows:

Let $X = Generalized\ Travel\ Time$

Let $D = Decay\ Onset\ (minutes\ assumed)$

$$a = \frac{X * D - D^2}{2}$$

$$b = X - D$$

$$\theta = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

The results of these calculations are shown below in Table 2.4-1, Figure 2.4-1, and Figure 2.4-2. Initially, different curves were estimated for various combinations of industry groups and attraction variables. However, subsequent implementation of the accessibility variables in the ActivitySim software environment revealed that the computational requirements for creating and using a large number of accessibility variables were prohibitive. In addition, the estimated impedance curves were quite similar. Therefore, a more parsimonious assumption of just two accessibility curves was adopted—one for regular establishments and one for TNCs.

Commented [BS8]: What is the full list of attraction variables? Please specify here.

Commented [JG9R8]: Shown in new table 2.4-2 below.

Commented [BS10]: Can we interpret "Establishments" and "TNCs" as collapsed industry groupings? What attractor variables are in each of the two equations?

Commented [JG11R10]: Yes, please see below in Table 2.4-2.

Table 2.4-1 Accessibility Impedance Curve Parameters

	Establishments	TNCs
N Observations	6,379	3,613
Wtd Observations	999,368	3,613
Avg. Gen Travel Time (unweighted)	20.23	19.75
Avg. Gen Travel Time (weighted)	18.78	19.75
Parameters		
decay onset	4	4
a	67.1	71.0
b	14.8	15.8
c	-1	-1
Theta	**0.0543**	**0.0515**

Figure 2.4-1 Calculated Accessibility Impedance Curve for Establishments

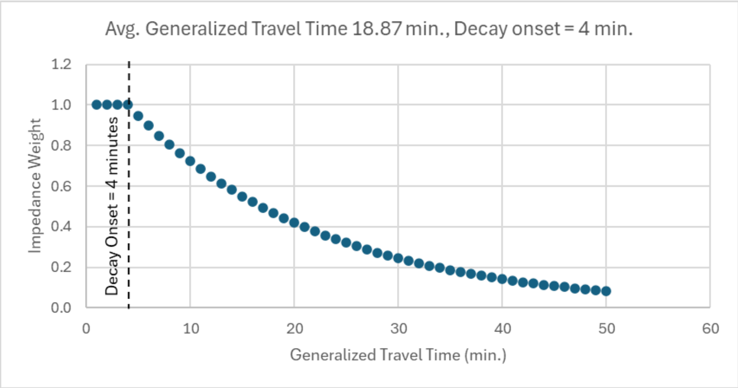
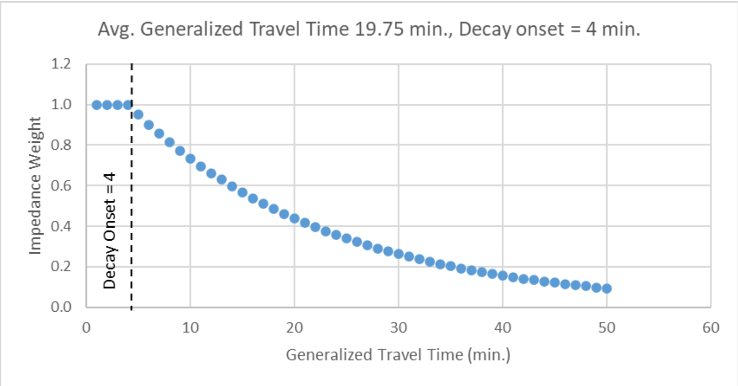


Figure 2.4-2 Calculated Accessibility Impedance Curve for TNCs



A complete list of the attractions variables available to the model are shown in Table 2.4-2, which also indicates how these variables were combined into industry groupings. Using the accessibility variable equation discussed above, the impedance curve for regular establishments was combined with the attraction variables shown in the table to create four accessibility variables for each of four establishment groups, which correspond to the groups used in the establishment route generation model. The impedance curve for regular establishments was also combined with household attractor variables to create three accessibility variables corresponding to household food, package, and service stops.

For TNCs, the impedance curve for TNCs was combined with the attraction variables shown in the lower part of Table 2.4-2 to create accessibility variables representing access to establishments and access to household food and package stops.

Table 2.4-2 Attraction Variables Used for Accessibility Calculations

Establishment Accessibility to:	Attraction Variables
Establishments_Group1	estb_attractions_retail + estb_attractions_wholesale + estb_attractions_manufacuring
Establishments_Group2	estb_attractions_industrial/utilities + estb_attractions_leis/accomd/food
Establishments_Group3	estb_attractions_agric/mining + estb_attractions_construction + estb_attractions_educ/public/other serv + estb_attractions_medical/health + estb_attractions_transport/whsng
Establishments_Group4	estb_attractions_info/FIRE/prof serv
Households_Food	num_hh_food_delivery
Households_Package	num_hh_package_delivery
Households_Service	num_hh_service
TNC Accessibility to:	**Attraction Variables**
Establishments_All_Groups	estb_attractions (sum of all industries)
Households_Food	num_hh_food_delivery
Households_Package	num_hh_package_delivery

During subsequent estimation of the Route Generation model, the four industry-group accessibility variables were combined to form a single variable representing accessibility to establishments of all types. The three household accessibility variables were likewise combined to form a single variable for accessibility to households for all purposes. As the original accessibility variables were created in log form, the new variables were added together in the following form for mathematical consistency:

$$Access\_Estb\_All = ln\begin{pmatrix} \exp(Access\_Estb\_Group1) \\ +\exp(Access\_Estb\_Group2) \\ +\exp(Access\_Estb\_Group3) \\ +\exp(Access\_Estb\_Group4) \end{pmatrix}$$

$$Access\_Hhds\_All = ln\begin{pmatrix} \exp(Access\_Hhds\_Food) \\ +\exp(Access\_Hhds\_Package) \\ +\exp(Access\_Hhds\_Service) \end{pmatrix}$$

Commented [JG12]: New text to address comment.

2.5 Route Generation – Establishments

The Route Generation model predicts which establishments in the CVM will generate commercial vehicle routes and how many. As noted in the CVM Model Design document, this is a two-stage model which first predicts whether an establishment will generate at least one route, followed by a second model which predicts how many routes given at least one.

2.5.1 Dataset

The Commercial Vehicle Survey asked each establishment how many outgoing cargo or service trips they make on an average workday. The responses to these questions were used to estimate the initial model. The survey also includes driver travel diary records from which routes may also be observed; however, because only about half of the establishments in the survey participate in the driver travel diaries, the establishment questions were viewed as a more reliable indicator of which establishments actually generate commercial vehicle routes. The weight-expanded route information from the driver diaries is used for estimating and calibrating other downstream models and for calibrating the route generation model (see Section 3.4).

2.5.2 Variable Specification

The final set of estimated parameters for the first stage of the Route Generation model are shown in Table 2.5-1. The primary attributes for this model were establishments' industry groups and number of employees. Based on a preliminary analysis of the similarities of route generation responses, four groups were formed and separate models estimated for each. Only statistically significant parameters were retained in the final specification were retained.

Table 2.5-1 Route Generation Probability of > 0 Routes - Binary Logit

Group 1 – Agriculture/Mining, Manufacturing, Wholesale		
Parameter Name	**Coefficient**	**t-Stat.**
employment	0.0104	2.84
constant	0.2114	2.08

Group 2 - Construction, Utilities, Retail, Transportation		
Parameter Name	**Coefficient**	**t-Stat.**
employment	0.0063	2.60
constant	0	-

Group 3 – Info/FIRE/Professional Services, Leisure/Food		
Parameter Name	**Coefficient**	**t-Stat.**
employment	0.0005	0.23
constant	-0.2428	-2.81

Group 4 – Education/Public/Other, Health/Medical		
Parameter Name	**Coefficient**	**t-Stat.**
employment	0.0064	3.21
constant	-0.8350	-9.40

The final set of estimated parameters for the second stage of the Route Generation model are shown in Table 2.5-2 and Table 2.5-3. The two parts to this model were developed to capture both the non-linear relationship between employment and routes generated as well as variation across industries and the influence of accessibility to customers. Several complex model formulations were tested. Ultimately, it was decided that breaking the model into two steps would be easier to estimate and calibrate to achieve the desired behavior:

1. Nonlinear effects of total employment: $z_n = \alpha X_n$

   Where z_n is a preliminary estimate of routes generated for establishment *n*; X_n is the square-root of employment at the establishment; and α is the estimated coefficient. Table 2.5-2 shows the results for a simple regression of the number attraction stops on the square root of employment of the establishment. Various nonlinear forms were tested, and the square root formulation provided both a good fit and a reasonable distribution of predictions.

2. Industry group factors as a modifier: $y_n = \beta_g z_n(Access\_Estb\_All + Access\_Hhds\_All)$

   Where y_n is the final estimate of routes generated for establishment *n*; z_n is the preliminary estimate of routes generated from Step 1; $Access\_Estb\_All$ and $Access\_Hhds\_All$ are the accessibilities variables created by the Accessibility model (see 2.4); and β_g is an estimated industry-group specific factor.

Commented [JG13]: New text

Table 2.5-2 Route Generation Square Root of Employment

R-squared (uncentered):	0.573		
Adj. R-squared (uncentered):	0.572		
Parameter (α)	**Coefficient**	**Std Err.**	**t-Stat.**
square root of employment	0.958	0.039	24.48

Table 2.5-3 shows the estimation results for the industry group and accessibility factors, which are multiplied times the results from the calculation on square root of employment. Application of these factors provides route generation estimates that vary by industry and by accessibility (potential customer demand). In this specification, this accessibility variable was the sum of the household and establishment accessibility variables described in Section 2.4.2.

Commented [BS14]: Similar comment as above. Were 2.52 and 2.53 estimated as part of a single OLS or separate?

Commented [JG15R14]: Please see above description addressing this. The model was broken into two steps (separate regressions).

Table 2.5-3 Route Generation Multiplicative Factors by Industry Group and Accessibilities – OLS

R-squared (uncentered):	0.622	
Adj. R-squared (uncentered):	0.613	
Parameter (β_g)	**Coefficient**	**t-Stat.**
Agriculture/Mining x Accessibilities	0.0358	3.643
Construction x Accessibilities	0.0541	11.467
Education/Public/Other Services x Accessibilities	0.0512	9.668
Industrial/Utilities x Accessibilities	0.0662	9.208
Info/FIRE/Prof. Services x Accessibilities	0.0367	8.548
Leisure/Accom./Food x Accessibilities	0.0251	3.660

R-squared (uncentered):	0.622	
Adj. R-squared (uncentered):	0.613	
Parameter (β_g)	**Coefficient**	**t-Stat.**
Manufacturing x Accessibilities	0.0304	5.585
Medical/Health Services x Accessibilities	0.0327	6.452
Retail x Accessibilities	0.0302	5.679
Transportation x Accessibilities	0.0696	11.254
Wholesale x Accessibilities	0.0477	8.876

2.6 Route Generation -- TNCs

The CVM Route Generation Model for TNCs involves the application of simple generation rates based on a more complex set of assumptions and calculations, which are described in the **CVM Model Design** technical memo. The table which includes the final rates, which are to be applied to employment by land use zone (LUZ), is repeated below:

Table 2.6-1 Calculation of TNC Route Generation Rates

TNC Customer Type	TNC Weighted Routes	Establishment Employment	TNC Routes Generated Per Job
Restaurant	2,844	140,348	0.020264
Retail	1,816	143,944	0.012484
Non Restaurant or Retail[1]	359	1,271,676	0.000277

2.7 Route Vehicle, Purpose, and Customer Type Assignment

For each route generated, several attributes are created through a series of sub-models. The first sub-model in the sequence assigns to each route a vehicle type, primary purpose, and customer type ("VPC Model"). As noted in the CVM Model Design document, this is formulated as a multinomial logit model. Note that this model is only applied to regular establishments. It is not applied to TNCs, which are assumed to have a single vehicle, purpose, and customer type.

2.7.1 Dataset

The Commercial Vehicle Survey was processed to create route records. Each route record represents one day of travel for a single commercial vehicle and is created from the survey's trip travel diaries. Route records include the vehicle type used on the route, a primary purpose for each route, and a customer type for each route. These route attributes are defined as follows:

[1] Non-restaurant or retail employment includes all employment types in the MGRA file other than food (restaurant) and retail. Also not included are military employment and the non-wage-and salary employment categories.

- **Vehicle Type**:
  - Light Commercial Vehicle (LCV)
  - Single-Unit Truck (SUT)
  - Multi-Unit Truck (MUT)
- **Purpose Type**:
  - Goods: customer-oriented; may include Services and other stop types
  - Service: customer-oriented; does not include Goods stops; may include other stop types
  - Maintenance/Other: refueling, driver breaks, vehicle/equipment repositioning, buying supplies, other; does not include Goods or Services stops; not customer oriented
- **Customer Type**:
  - Residential Only: households, including multi-family buildings (Goods or Services purposes)
  - Non-residential Only: commercial, public/government (Goods or Services purposes)
  - Mixed Residential and Non-residential: (Goods or Services purposes)
  - No Customer: (Maintenance/Other purpose)

The alternatives in the VPC model represent the 21 feasible combinations of vehicle, purpose, and customer type as defined above. The model is formulated as a joint choice in which all three dimensions are represented at the same level (no nesting).

2.7.2 Variable Specification

The final set of estimated parameters for the first stage of the Route Generation model are shown in Table 2.7-1. The alternative-specific constants in the model correspond to each of the feasible combinations of purpose, customer, and vehicle. Also included in the model are various industry-specific effects interacted with one dimension at a time. For example, the transportation industry is negatively associated with service provision (being goods oriented) and positively associate with use of the MUT vehicle type. Only statistically significant parameters were retained in the final specification were retained.

Table 2.7-1 Route Choice of Vehicle, Purpose, and Customer Type -- Multinomial Logit

Parameter Name	Coefficient	t-Stat.
Constants (purpose x customer x vehicle type)		
Goods x Mixed x LCV (reference case)	0	--
Goods x Mixed x MUT	-2.552	-7.49
Goods x Mixed x SUT	-1.775	-5.65
Goods x Non-residential x LCV	-0.321	-1.66
Goods x Non-residential x MUT	-0.434	-2.12
Goods x Non-residential x SUT	-1.003	-4.00
Goods x Residential x LCV	-0.953	-4.38
Goods x Residential x MUT	-3.322	-7.47
Goods x Residential x SUT	-2.044	-5.71

Parameter Name	Coefficient	t-Stat.
Maintenance x LCV	0.144	0.84
Maintenance x MUT	-2.261	-7.71
Maintenance x SUT	-1.585	-5.45
Service x Mixed x LCV	1.277	7.52
Service x Mixed x MUT	-2.507	-7.58
Service x Mixed x SUT	-0.570	-2.55
Service x Non-residential x LCV	2.041	12.09
Service x Non-residential x MUT	-1.027	-4.24
Service x Non-residential x SUT	0.286	1.44
Service x Residential x LCV	0.264	1.41
Service x Residential x MUT	-4.415	-7.07
Service x Residential x SUT	-1.738	-5.78
Industry Effects		
Transportation -> Service Purp	-0.255	-1.74
Transportation -> MUT Vehicle	1.821	9.44
Wholesale -> Non-residential Cust	0.610	3.30
Wholesale -> Service Purp	-0.441	-2.31
Construction -> Non-residential Cust	-1.076	-5.66
Construction -> Mixed Cust	-0.783	-4.09
Construction -> Service Purp	0.876	4.98
Construction -> MUT Vehicle	1.395	6.34
Education/Public/Other Services -> Non-residential Cust	-0.969	-5.11
Education/Public/Other Services -> Service Purp	0.903	5.72
Education/Public/Other Services -> SUT Vehicle	-0.399	-1.68
Industrial/Utilities -> Non-residential Cust	-1.245	-4.55
Industrial/Utilities -> Service Purp	0.958	4.93
Info/FIRE/Prof. services -> Service Purp	0.519	3.80
Info/FIRE/Prof. services -> Maintenance Purp	0.625	2.62
Info/FIRE/Prof. services -> SUT Vehicle	-1.076	-4.05
Info/FIRE/Prof. services -> MUT Vehicle	-1.918	-3.22
Leisure/Accommodations and Food -> Non-residential Cust	0.850	4.12
Leisure/Accommodations and Food -> Mixed Cust	-1.017	-2.36
Leisure/Accommodations and Food -> Maintenance Purp	1.105	3.35
Leisure/Accommodations and Food -> SUT Vehicle	-2.933	-2.88
Medical/Health Services -> Non-residential Cust	0.424	1.86
Medical/Health Services -> Mixed Cust	0.882	3.97
Medical/Health Services -> Service Purp	-1.075	-4.35
Retail -> Non-residential Cust	0.675	3.35
Retail -> Service Purp	-0.498	-2.38

2.8 Route Start Time

The CVM Route Start Time model draws a route starting time from an empirical distribution. Development of these distributions is described in the **CVM Model Design** technical memo. As these empirical distributions were developed from weighted Commercial Vehicle Survey observations, no further model estimation or calibration is required.

2.9 Route Origin and Termination Location Choices – Establishments

For each route generated, once the vehicle type, primary purpose, and customer type attributes have been assigned, and the route is assigned a start time, the next step is to choose an origin location for the route and a termination location. As discussed in the CVM Model Design document, the model design allows for truck routes that begin or end at locations other than the driver's official establishment location. The Route Origin and Termination Location Choice step includes the following sub models:

- **Route Origin Location Choice**
    a. Choice of Location Type – multinomial logit
    b. Given Location Type, Choice of Zone -- destination choice model
- **Route Termination Location Choice**
    a. Choice of Location Type – multinomial logit
    b. Given Location Type, Choice of Zone -- destination choice model

Note that this model set is only applied to regular establishments. The origin and termination locations for TNCs are determined by a separate model, described in Section 0.

2.9.1 Dataset

The Commercial Vehicle Survey for regular establishments was processed to create route records. Each route record represents one day of travel for a single commercial vehicle and is created from the survey's trip travel diaries. The starting point of the route and the location of the last stop on the route were identified by zone and were labeled based on the land use and customer type at the zone. Consistent with the CVM Model Design, the following origin and destination types were defined in the travel diary data:

- **Base** – the establishment of the driver
- **Warehouse** – warehouse or distribution center that is not the establishment of the driver and identified in the survey travel diary as being a "warehouse" place type
- **Transport Node** – zones designated as marine ports, airports, or intermodal facilities and not the establishment of the driver. Overrides the warehouse designation.
- **Residential** – zones with households that were identified in the trip records of the travel diary as a "residential" place type or reported as the location of the driver's home (activity type).

- **Other Commercial** – other zones with employment that do not fall into one of the above categories.

In addition to the survey observations, the zone location choice models required that travel time and costs skims be considered for each zone alternative. The start time of the route is known prior to the destination choice, but the route ending time is not known as it will be determined through the dynamic simulation loop Travel time and cost skims for the start time period were used for the route origin location choice. For the route termination location choice, skims were used for the first time period after the route origin starting time period. Generalized travel times from the base to each alternative zone were calculated as the congested travel time plus any toll costs converted to equivalent in-vehicle minutes. The assumed values of time are shown in Table 2.9-1 and were based on ABM3 highway assignment assumptions for LHDT, MHDT, and HHDT vehicles, respectively. In addition, zone-level data, such as employment and households, were joined from the MGRA land use file.

Table 2.9-1 Value of Time Assumptions

Vehicle Type	$/hour
Light (LCV)	$67
Medium (SUT)	$68
Heavy (MUT)	$89

2.9.2 Variable Specification

Route Origin / Termination Location Types

The final set of estimated parameters for the first stage of the Route Origin Location Type model are shown in Table 2.9-2. The specification includes not only constants but also bias parameters for specific industries, vehicle types, the primary route purpose, the customer type, and the time period in which the route begins.

Table 2.9-2 Route Origin Location Type MNL Choice

Parameter Name	Coefficient	t-Stat.
Alternative Specific Constants		
Base	0.000	NA
Warehouse	-4.003	-13.65
Transport Node[2]	-7.112	-7.11
Residential	-2.278	-14.62
Other Commercial	-1.695	-13.24
Industry Effects (-> Origin Type)		
Estab Industry is Construction -> Origin Warehouse	1.815	4.56
Estab Industry is Info/FIRE/Prof -> Origin Residential	0.689	3.30
Estab Industry is Wholesale -> Origin Warehouse	1.733	3.55
Estab Industry is Retail -> Origin Other Commercial	-2.050	-2.85
Estab Industry is Industrial/Utilities -> Origin Other Commercial	-1.432	-2.39
Vehicle Effects (-> Origin Type)		

[2] Although the transport node (port, intermodal facility location type was included in the original estimation, it was dropped from the final model in the calibration stage due to too few observations for route origin types.

Parameter Name	Coefficient	t-Stat.
Vehicle Type is SUT -> Origin Residential	-0.922	-2.71
Vehicle Type is SUT -> Origin Other Commercial	-1.065	-2.99
Vehicle Type is MUT -> Origin Residential	-0.902	-2.21
Route Purpose Effects (-> Origin Type)		
Route Purpose is Goods -> Origin Warehouse	-2.406	-2.35
Route Purpose is Goods -> Origin Residential	-0.664	-2.27
Route Purpose is Goods -> Origin Other Commercial	-1.148	-3.95
Route Purpose is Maintenance -> Origin Residential	0.655	2.51
Customer Effects (-> Origin Type)		
Customer Type is Residential -> Origin Warehouse	-1.384	-1.87
Customer Type is Mixed -> Origin Other Commercial	-1.814	-3.92
Time of Day Effects (-> Origin Type)		
TOD is AM -> Origin Residential	0.590	3.27
TOD is AM -> Origin Other Commercial	0.450	2.62
TOD is PM -> Origin Residential	-1.918	-1.87

A similar specification for the Route Termination Location type is shown in Table 2.9-3. The termination type also includes parameters representing the origin type choice, which is assumed to be know prior to choosing the termination location. For example, if the origin type is base (establishment), then it is less likely that the termination will be any of the other types, meaning the route is more likely to terminate at the establishment. If the origin is another commercial location, the route is most likely to terminate at another commercial location and least likely to terminate at a residential type.

Table 2.9-3 Route Termination Location Type MNL Choice

Parameter Name	Coefficient	t-Stat.
Alternative Specific Constants		
Base	0.000	NA
Warehouse	-2.173	-6.33
Transport Node[3]	-5.395	-9.33
Residential	0.785	4.06
Other Commercial	0.065	0.28
Route Origin Type Effects (-> Termination Type)		
Route Origin is Base -> Termination Warehouse	-1.273	-3.60
Route Origin is Base -> Termination Residential	-1.841	-9.33
Route Origin is Base -> Termination Other Commercial	-1.400	-5.89
Route Origin is Other Commercial -> Termination Residential	-1.135	-3.85
Route Origin is Other Commercial -> Termination Other Commercial	0.683	2.40
Industry Effects (-> Termination Type)		
Estab Industry is Construction -> Termination Residential	0.547	3.27
Estab Industry is Construction -> Termination Other Commercial	-0.484	-2.19

[3] Although the transport node (port, intermodal facility location type was included in the original estimation, it was dropped from the final model in the calibration stage due to too few observations for route termination types.

Parameter Name	Coefficient	t-Stat.
Estab Industry is Medical -> Termination Other Commercial	0.819	3.72
Estab Industry is Info/FIRE/Prof -> Termination Warehouse	-1.760	-1.73
Estab Industry is Info/FIRE/Prof -> Termination Residential	0.305	1.76
Estab Industry is Manufacturing -> Termination Warehouse	1.342	3.90
Estab Industry is Retail -> Termination Other Commercial	-1.115	-2.94
Estab Industry is Leis/Accom/Food -> Termination Other Commercial	-0.794	-1.98
Vehicle Effects (-> Termination Type)		
Vehicle Type is SUT -> Termination Residential	-0.766	-3.49
Vehicle Type is MUT -> Termination Residential	-1.565	-5.31
Route Purpose Effects (-> Termination Type)		
Route Purpose is Goods -> Termination Warehouse	0.592	1.90
Route Purpose is Goods -> Termination Transport Node	1.551	2.02
Customer Effects (-> Termination Type)		
Customer Type is Residential -> Termination Other Commercial	-0.702	-2.80
Customer Type is Mixed -> Termination Warehouse	0.534	1.80

Route Origin / Location Choice of Zone

The final set of estimated parameters for the Route Origin Location Choice of Zone are shown in Table 2.9-4. As discussed in the CVM Model Design document, these models are formulated as destination choice models which are composed of a size function and an impedance function. The size function includes parameters related to quantities. The table shows parameters specific to each of the route origin types—commercial, residential, and warehouse interacted with zonal land use variables relevant to that type. For example, for a residential location type the number of households and employment in non-wage-and-salary jobs-working from home are quantity predictions.

Note that parameter identification requires that at least one size parameter be fixed, which was done for each location type. Size parameters are required to be positive valued and are actually applied within an exponential function. For example, a coefficient of zero when exponentiated becomes a parameter of 1.0.

Table 2.9-4 Route Origin Location Choice of Zone

Parameter Name	Coefficient	t-Stat.
Size Parameters (γ)		
Commercial Origin Type		
Employment Total Non-WS/Transportation*	0.000	NA
Employment in Accommodations	2.788	2.47
Employment in Construction	2.177	3.42
Employment in Health Services	2.486	4.26
Employment in Manufacturing	2.392	3.43
Employment in Other Services	3.553	7.00
Employment in Wholesale	4.197	9.57
Residential Origin Type		
Non-Wage Salary Work From Home Employment	1.831	6.12

Commented [BS16]: Please confirm that size coefficients are gammas and impedance coefficients are betas here so that we can easily relate these to equations in the design memo.

Commented [JG17R16]: Correct. Please see improved labeling in tables.

Parameter Name	Coefficient	t-Stat.
Households	0.000	NA
Warehouse Origin Type		
Employment in Transportation/Warehousing	0.000	NA
Employment in Wholesale	1.245	2.45
Theta	1.000	NA
Impedance Parameters (β)		
Generalized Travel Time: All	-0.022	-4.63
Generalized Travel Time x Industry		
Agriculture & Mining	-0.045	-2.57
Construction	-0.026	-2.36
Education, Public, Other Services	-0.035	-3.04
Industrial, Utilities	-0.079	-3.74

* Employment Total Non-WS/Transportation = Total Employment – Employment in Transportation/Warehousing – Non-Wage Salary Employment

Commented [BS16]: Please confirm that size coefficients are gammas and impedance coefficients are betas here so that we can easily relate these to equations in the design memo.

Commented [JG17R16]: Correct. Please see improved labeling in tables.

The impedance parameter in Table 2.9-4 focus on generalized travel time. There is a generic generalized travel time parameter which applies to all cases, plus for additive factors for specific industries, which, in these cases indicate a greater sensitivity to travel time relative to the establishment for agriculture, construction, education/public/other services, and utilities. Similar specifications for Route Termination Location Choice of Zone are shown in Table 2.9-5.

Table 2.9-5 Route Termination Location Choice of Zone

Parameter Name	Coefficient	t-Stat.
Size Parameters (γ)		
Commercial Destination Type		
Employment Total Non-WS/Transportation*	0.000	NA
Employment in Other Services x Maintenance Route Purpose	3.181	8.030
Employment in Construction	1.508	2.990
Employment in Government	1.646	3.230
Employment in Health Services	1.686	2.580
Employment in Manufacturing	3.081	8.780
Employment in Transportation/Warehousing	3.074	7.000
Employment in Wholesale	3.814	11.230
Residential Destination Type		
Non-Wage Salary Work From Home Employment	2.372	14.060
Households	0.000	NA
Warehouse Destination Type		
Employment in Transportation/Warehousing	0.000	NA
Employment in Wholesale	2.096	9.510
Theta	1.000	NA
Impedance Parameters (β)		
Generalized Travel Time: All	-0.015	-4.60
Generalized Travel Time x Industry		

Education, Public, Other	-0.049	-5.49
Leisure, Accommodations, Food	-0.039	-2.62
Transportation/Warehousing	-0.032	-2.90
Generalized Travel Time x Multi-Unit Truck Type	-0.027	-2.29

* Employment Total Non-WS/Transportation = Total Employment – Employment in Transportation/Warehousing – Non-Wage Salary Employment

2.10 Route Origin and Termination Location Choice – TNCs

For each route generated, once the vehicle type, primary purpose, and customer type attributes have been assigned, and the route is assigned a start time, the next step is to choose an origin location for the route and a termination location. As discussed in the CVM Model Design document, TNC routes are generated through a different process from regular establishments. TNC routes are assumed to have a single location representing both the origin and the termination point of the route. The form of the model is a simplified destination choice model in which a zone would be chosen from among the zones available within the LUZ which generated the route. The model is formulated as a destination choice model, but one which includes only a size function without an impedance as TNC drivers are not attached to a particular base establishment.

2.10.1 Dataset

The Commercial Vehicle Survey of TNC drivers was processed to create route records. Each route record represents one day of travel for a single commercial vehicle and is created from the survey's trip travel diaries. The starting point of the route and the location of the last stop on the route were identified by zone. In addition to the survey observations, zone-level data, such as employment and households, were joined from the MGRA land use file.

2.10.2 Variable Specification

The final set of estimated parameters for the TNC Route Origin and Termination Location Choice of Zone are shown in Table 2.10-1. The model is composed entirely of size parameters related to quantities. The specification includes parameters related to total employment, employment in non-wage-and salary employment working-from-home, total households, and establishment attraction variables related to leisure/accommodations and food services (restaurants) and retail attractions. The households and employment in non-wage-and salary employment working-from-home variable is intended to capture locations where TNC drivers are more likely to reside, whereas the other variables are intended to capture where they are more likely to operate.

Note that parameter identification requires that at least one size parameter be fixed, which was done for each location type. Size parameters are required to be positive valued and are actually applied within an exponential function. For example, a coefficient of zero when exponentiated becomes a parameter of 1.0.

Table 2.10-1 TNC Route Origin and Termination Location Choice of Zone

Parameter Name	Coefficient	t-Stat.
Size Parameters **(γ)**		
Total Employment	0.000	NA
Total Households	0.000	NA
Non-Wage Salary Work From Home Employment	3.247	5.78
Establishment Attractions-Leisure/Accomd/Food	3.292	3.20
Establishment Attractions-Retail	4.165	8.07
Theta	1.000	NA

2.11 Next Stop Purpose Choice

The Next Stop Purposes Choice model predicts the primary purpose for each stop on a commercial vehicle route. The model is formulated as a multinomial logit. As discussed in the CVM Model Design document this is the first model in the dynamic simulation loop, which also includes the choices of the stop location and stop duration. (See the flowchart in Figure 2.1-1 for reference.) Separate yet similarly structured models were developed for regular establishments and for TNCs.

2.11.1 Dataset

The Commercial Vehicle Survey driver diary trip records were processed to identify the primary purpose for each stop consistent with the CVM Model Design. Definitions of stop purposes are as follows:

- **Goods Pickup** – a stop to pick up goods
- **Goods Delivery** – a stop to deliver goods to a customer
- **Service** – a stop to provide a professional service to a customer
- **Maintenance/Other** – a stop for either vehicle maintenance/refueling or driver breaks
- **Base Establishment** – a trip the establishment where the vehicle is based, without ending the route
- **Home** – a trip to the driver's home, without ending the route
- **Terminate Route** – a trip to the final stop on the route where the vehicle will be parked, ending the route.

2.11.2 Variable Specification

The final set of estimated parameters for ~~the first stage of~~ the Next Stop Purpose model are shown in Table 2.11-1 for regular establishments. Attributes include alternative-specific constants and various effects on the propensity to choose each purpose for the next stop, including the effects of accessibility, industry group, vehicle type, route purpose, route customer type, time of day, route elapsed time, and the current stop purpose (purpose at origin of trip). Numerous parameter combinations were tested. Only statistically significant parameters were retained in the final model.

Commented [BS18]: Does this model have two stages too?

Commented [JG19R18]: No. copy-paste error. Corrected text.

Table 2.11-1 Next Stop Purpose MNL Choice -- Establishments

Parameter Name	Coefficient	t-Stat.
Alternative Specific Constants		
Base	0.000	NA
Goods Pickup	0.794	4.65
Goods Delivery	-5.720	-3.69
Service	3.677	30.34
Maintenance	3.584	33.52
Home	-1.371	-7.10
Terminal Destination	2.154	18.86
Accessibility Effects (-> Next Stop Purpose)		
Estab Access to Households Package -> Next Stop Goods Delivery	0.698	5.39
Industry Effects (-> Next Stop Purpose)		
Estab Industry is Construction -> Next Stop Goods Delivery	-1.220	-8.38
Estab Industry is Construction -> Next Stop Service	-0.301	-4.20
Estab Industry is Educ/Public/Other -> Next Stop Goods Delivery	-2.610	-5.09
Estab Industry is Medical -> Next Stop Service	-0.513	-4.74
Estab Industry is Medical -> Next Stop Home	-1.428	-1.98
Estab Industry is Info/FIRE/Prof -> Next Stop Goods Delivery	-2.771	-7.39
Estab Industry is Info/FIRE/Prof -> Next Stop Maintenance	0.359	4.53
Estab Industry is Info/FIRE/Prof -> Next Stop Terminal Destination	0.278	2.87
Estab Industry is Wholesale -> Next Stop Maintenance	0.293	3.24
Estab Industry is Wholesale -> Next Stop Terminal Destination	0.362	3.30
Estab Industry is Retail -> Next Stop Goods Delivery	-0.632	-4.79
Estab Industry is Retail -> Next Stop Service	-0.435	-4.49
Estab Industry is Indus/Utilities -> Next Stop Maintenance	-0.727	-5.47
Estab Industry is Indus/Utilities -> Next Stop Home	-1.038	-1.75
Estab Industry is Agriculture -> Next Stop Service	0.764	2.93
Estab Industry is Agriculture -> Next Stop Maintenance	0.633	2.35
Estab Industry is Agriculture -> Next Stop Terminal Destination	1.492	5.50
Vehicle Effects (-> Next Stop Purpose)		
Vehicle Type is SUT -> Next Stop Goods Delivery	-0.432	-3.79
Vehicle Type is MUT -> Next Stop Goods Pickup	1.044	7.03
Route Purpose Effects (-> Next Stop Purpose)		
Route Purpose is Goods -> Next Stop Service	-3.283	-30.10
Route Purpose is Goods -> Next Stop Maintenance	-2.176	-21.95
Route Purpose is Goods -> Next Stop Home	-2.242	-7.34
Route Purpose is Goods -> Next Stop Terminal Destination	-1.973	-19.29
Customer Type Effects (-> Next Stop Purpose)		
Customer Type is Non-residential -> Next Stop Service	1.016	12.44
Customer Type is Mixed -> Next Stop Service	1.527	17.62
Customer Type is Mixed -> Next Stop Home	0.438	2.02
Customer Type is Mixed -> Next Stop Terminal Destination	-0.764	-10.48
Elapsed Time Effects (-> Next Stop Purpose)		

Parameter Name	Coefficient	t-Stat.
Elapsed Time (min/100) -> Next Stop Service	0.099	3.86
Elapsed Time (min/100) -> Next Stop Maintenance	0.197	7.91
Elapsed Time (min/100) -> Next Stop Home	0.117	1.77
Elapsed Time (min/100) -> Next Stop Terminal Destination	0.360	13.54
Time of Day Effects (-> Next Stop Purpose)		
TOD is AM -> Next Stop Terminal Destination	-1.685	-10.91
TOD is MD -> Next Stop Goods Pickup	-0.513	-3.31
TOD is MD -> Next Stop Terminal Destination	-0.915	-11.40
TOD is PM -> Next Stop Goods Pickup	-0.943	-2.45
TOD is PM -> Next Stop Goods Delivery	-0.382	-2.18
TOD is PM -> Next Stop Home	1.184	4.29
TOD is EV -> Next Stop Goods Pickup	0.686	1.87
TOD is EV -> Next Stop Maintenance	0.424	2.42
TOD is EV -> Next Stop Home	1.676	3.74
TOD is EV -> Next Stop Terminal Destination	0.627	3.75
Origin Stop Purpose Effects (-> Next Stop Purpose)		
Origin Purp is Goods Pickup -> Next Stop Service	-1.414	-5.72
Origin Purp is Goods Pickup -> Next Stop Maintenance	-1.684	-6.82
Origin Purp is Goods Delivery -> Next Stop Goods Pickup	-1.472	-8.08
Origin Purp is Goods Delivery -> Next Stop Service	-2.799	-14.35
Origin Purp is Goods Delivery -> Next Stop Maintenance	-1.603	-12.51
Origin Purp is Service -> Next Stop Goods Pickup	-2.557	-10.23
Origin Purp is Service -> Next Stop Goods Delivery	-3.598	-18.61
Origin Purp is Service -> Next Stop Service	-1.985	-18.85
Origin Purp is Service -> Next Stop Maintenance	-2.119	-19.34
Origin Purp is Maintenance -> Next Stop Goods Pickup	-1.607	-8.33
Origin Purp is Maintenance -> Next Stop Goods Delivery	-2.136	-17.55
Origin Purp is Maintenance -> Next Stop Service	-2.093	-19.14
Origin Purp is Maintenance -> Next Stop Maintenance	-1.460	-13.42
Origin Purp is Home -> Next Stop Goods Pickup	-3.509	-4.77
Origin Purp is Home -> Next Stop Goods Delivery	-3.087	-8.09
Origin Purp is Home -> Next Stop Service	-3.283	-14.85
Origin Purp is Home -> Next Stop Maintenance	-2.792	-12.48
Origin Purp is Home -> Next Stop Terminal Destination	-1.952	-6.71

Commented [BS20]: Just confirming this term means the impact of home origin purpose on maintenance destination purpose.

Commented [JG21R20]: Yes. Please see new labeling

The final set of estimated parameters for the the Next Stop Purpose model are shown in Table 2.11-2 for TNCs. There were many fewer statistically significant parameters leading to a more parsimonious model, compared with regular establishments. Attributes include alternative-specific constants and effects on the propensity to choose each purpose for the next stop, including the effects of: TNC client industry group, customer effects, route elapsed time, and the current stop purpose (purpose at origin of trip).

Table 2.11-2 Next Stop Purpose MNL Choice -- TNCs

Parameter Name	Coefficient	t-Stat.
Alternative Specific Constants		
Base	0.000	NA
Goods Pickup	2.894	21.46
Goods Delivery	1.968	11.39
Service	0.318	1.77
Maintenance	2.906	21.36
Home	0.202	1.14
Terminal Destination	1.532	8.33
Industry Effects		
Client is Non-Restaurant/Retail -> Next Stop Goods Delivery	1.154	15.42
Customer Effects		
Customer Type is Mixed -> Next Stop Terminal Destination	-0.701	-5.56
Elapsed Time Effects		
Elapsed Time (min/100) -> Next Stop Terminal Destination	0.330	14.27
Origin Stop Purpose Effects		
Origin Purp is Goods Pickup -> Next Stop Goods Delivery	2.730	19.24
Origin Purp is Goods Pickup -> Next Stop Service	-1.498	-2.07
Origin Purp is Goods Pickup -> Next Stop Maintenance	-0.887	-5.29
Origin Purp is Goods Delivery -> Next Stop Goods Delivery	1.886	15.80
Origin Purp is Service -> Next Stop Service	1.729	6.95
Origin Purp is Service -> Next Stop Maintenance	-0.441	-2.35
Origin Purp is Maintenance -> Next Stop Goods Delivery	-0.652	-4.20

2.12 Next Stop Location Choice – Establishments

The Next Stop Location Choice model predicts the location type and zone for each stop on a commercial vehicle route. As discussed in the CVM Model Design document this is the second model in the dynamic simulation loop, which also includes the choices of the stop purpose and stop duration. (See the flowchart in Figure 2.1-1 for reference.) Separate yet similarly structured models were developed for regular establishments and for TNCs.

The Next Stop Location Choice model is formulated as a two-level model and consists of the following sub models:

- **Next Stop Location Choice**

    a. Choice of Location Type

    b. Given Location Type, Choice of Zone

2.12.1 Dataset

The Commercial Vehicle Survey establishments driver diary trip records were processed to identify the destination type for each stop consistent with the CVM Model Design. The following destination types were defined in the travel diary data:

- **Warehouse** – warehouse or distribution center that is not the establishment of the driver and identified in the survey travel diary as being a "warehouse" place type
- **Transport Node** – zones designated as marine ports, airports, or intermodal facilities and not the establishment of the driver. Overrides the warehouse designation.
- **Residential** – zones with households that were identified in the trip records of the travel diary as a "residential" place type or reported as the location of the driver's home (activity type).
- **Other Commercial** – other zones with employment that do not fall into one of the above categories.

In addition to the survey observations, the zone location choice models required that travel time and costs skims be considered for each zone alternative. The time period used is determined by the clock time at the start of each trip. Generalized travel times from the base to each alternative zone were calculated as the congested travel time plus any toll costs converted to equivalent in-vehicle minutes. The assumed values of time are shown in Table 2.12-1 and were based on ABM3 highway assignment assumptions for LHDT, MHDT, and HHDT vehicles or Light, Medium, a Heavy CVM vehicle types, respectively. Vehicle operating costs used in the model were derived for the model update and are shown in

Table 2.12-2. In addition, zone-level data, such as employment and households, were joined from the MGRA land use file.

Table 2.12-1 Value of Time Assumptions

Vehicle Type	$/hour
Light (LCV)	$67
Medium (SUT)	$68
Heavy (MUT)	$89

Table 2.12-2 Vehicle Operating Costs

Vehicle Operating Costs	$/mile
Light (LCV)	$0.192
Medium (SUT)	$0.733
Heavy (MUT)	$1.170

2.12.2 Variable Specification

The final set of estimated parameters for the first stage of the Next Stop Location Type model are shown in Table 2.12-3. The specification includes not only constants but also bias parameters for specific industries, destination purposes, the primary route purpose, the customer type, vehicle types, elapsed time on the route at the time the stop choices are being made, and the time period in which the trip would start.

Table 2.12-3 Next Stop Location Type MNL Choice -- Establishments

Parameter Name	Coefficient	t-Stat.
Alternative Specific Constants		
Residential	0.000	NA
Warehouse	-1.060	-10.45
Transport Node	-2.936	-8.60
Other Commercial	1.767	29.00
Industry Effects (-> Next Destination Type)		
Estab Industry is Construction -> Next Destination Transport Node	-3.316	-3.28
Estab Industry is Educ/Public/Other -> Next Destination Warehouse	-1.066	-6.64
Estab Industry is Educ/Public/Other -> Next Destination Transport Node	-2.256	-3.11
Estab Industry is Medical -> Next Destination Warehouse	-2.832	-9.72
Estab Industry is Medical -> Next Destination Other Commercial	-0.276	-3.30
Estab Industry is Info/FIRE/Prof -> Next Destination Warehouse	-1.541	-5.94
Estab Industry is Info/FIRE/Prof -> Next Destination Transport Node	-1.627	-2.70
Estab Industry is Info/FIRE/Prof -> Next Destination Other Commercial	0.512	5.99
Estab Industry is Wholesale -> Next Destination Transport Node	-2.172	-2.98
Estab Industry is Wholesale -> Next Destination Other Commercial	0.285	3.21
Estab Industry is Manufacturing -> Next Destination Warehouse	0.896	6.36
Estab Industry is Manufacturing -> Next Destination Other Commercial	0.533	4.71
Estab Industry is Retail -> Next Destination Transport Node	-1.272	-2.12
Estab Industry is Retail -> Next Destination Other Commercial	0.391	3.96
Estab Industry is Industrial/Utilities -> Next Destination Warehouse	-1.721	-5.96
Estab Industry is Leis/Accom/Food -> Next Destination Warehouse	-1.230	-2.03
Estab Industry is Leis/Accom/Food -> Next Destination Other Commercial	0.747	3.87
Estab Industry is Agriculture -> Next Destination Other Commercial	0.462	2.97
Destination Stop Purpose Effects (-> Next Destination Type)		
Destination Purp is Goods Pickup -> Next Destination Transport Node	1.650	5.54
Destination Purp is Goods Delivery -> Next Destination Transport Node	-1.821	-4.06
Destination Purp is Maintenance -> Next Destination Transport Node	-0.937	-2.73
Route Purpose Effects (-> Next Destination Type)		
Route Purpose is Goods -> Next Destination Warehouse	-0.277	-2.79
Route Purpose is Goods-> Next Destination Other Commercial	-0.547	-8.98
Customer Effects (-> Next Destination Type)		
Customer Type is Non-Residential -> Next Destination Transport Node	2.262	8.34
Customer Type is Residential -> Next Destination Other Commercial	-2.477	-27.72
Customer Type is Mixed -> Next Destination Other Commercial	-1.378	-24.36
Vehicle Effects (-> Next Destination Type)		
Vehicle Type is MUT -> Next Destination Warehouse	1.901	13.32
Vehicle Type is MUT-> Next Destination Transport Node	1.300	4.29
Vehicle Type is MUT -> Next Destination Other Commercial	0.885	7.37

Parameter Name	Coefficient	t-Stat.
Elapsed Time Effects (-> Next Destination Type)		
Elapsed Time (min/100) -> Next Destination Warehouse	-0.201	-7.25
Elapsed Time (min/100) -> Next Destination Transport Node	-0.102	-1.68
Elapsed Time (min/100) -> Next Destination Other Commercial	-0.045	-2.76
Time of Day Effects (-> Next Destination Type)		
TOD is AM -> Next Destination Transport Node	-0.992	-2.63
TOD is MD -> Next Destination Warehouse	0.364	4.01
TOD is MD -> Next Destination Transport Node	-0.685	-2.66
TOD is PM -> Next Destination Other Commercial	-0.395	-4.55
TOD is EV -> Next Destination Other Commercial	-0.706	-4.85

The final set of estimated parameters for the Next Stop Location Choice of Zone are shown in Table 2.12-4. As discussed in the CVM Model Design document, the formulation is a destination choice model, which is composed of a size function and an impedance function. The size function includes parameters related to quantities. The table shows parameters specific to each of the stop location types—commercial, residential, and warehouse interacted with zonal land use variables relevant to that type. For example, for a residential location type the number of households and employment in non-wage-and-salary jobs-working from home are quantity predictions. “Port Transport Node” is a special category that indicates that a zone is designated as either a marine port, airport, or intermodal facility.

Note that parameter identification requires that at least one size parameter be fixed, which was done for each location type. Size parameters are required to be positive valued and are actually applied within an exponential function. For example, a coefficient of zero when exponentiated becomes a parameter of 1.0.

The impedance parameter in in Table 2.12-4 focus on generalized travel time. As discussed in the CVM Model Design document, this model considers the travel impedance from the current stop to all alternative next stops as well as the impedance from each alternative stop to the route’s termination location. This may be expressed as:

$$Total\ Impedance\ Alt\ j\ = \beta X_{ij} + \beta X_{jk}$$

Which is the impedance of travel *X* between zone *i* (current zone) and zone *j* (alternative next zone), and the impedance of travel *X* between zone *j* (alternative next zone) and zone *k* (route terminal zone). The estimated parameter *β* measures the disutility of travel times and costs.

The specification includes a generic generalized travel time parameter which applies to all cases, plus for additive factors for specific stop purposes, which indicate a greater sensitivity to travel time for home, maintenance/other (e.g., lunch, fueling), or for service stops. There is also an additive term for the interaction between generalized travel time and the elapsed time on the route which represents a greater sensitivity as the elapsed time increases.

The specification also includes vehicle operating cost parameters segmented by vehicle type, which assumed the costs shown above in Table 2.12-2.

Commented [BS22]: How exactly? Added a new comment in Design document to make it clearer.

Commented [JG23R22]: See new clarifying text. Made similar changes in the Model Design memo.

Commented [JG24]: Clarifying text.

Commented [BS25]: Is this a sum of times from route_origin-to-alternative and alternative-to-route_dest?

Commented [JG26R25]: Yes. Please see clarifying text.

Table 2.12-4 Next Stop Location Choice of Zone -- Establishments

Parameter Name	Coefficient	t-Stat.
Size Parameters (γ)		
Commercial Destination Type		
Employment in Food x Maintenance Stop Purpose	1.983	15.07
Employment in Other Services x Maintenance Stop Purpose	3.358	32.89
Employment in Retail x Maintenance Stop Purpose	3.091	42.65
Employment Total Non-WS/Transportation*	0.000	NA
Establishment Attractions-Agriculture & Mining	4.961	45.79
Establishment Attractions-Construction	2.764	31.79
Establishment Attractions-Education, Public, Other	1.520	3.31
Establishment Attractions-Info/FIRE/Prof. Services	2.077	6.66
Establishment Attractions-Leisure/Accomd./Food	3.331	27.72
Establishment Attractions-Medical/Health	1.576	5.39
Establishment Attractions-Retail	0.657	2.18
Establishment Attractions-Transportation/Warehousing	2.564	12.12
Establishment Attractions-Wholesale	2.298	13.03
Port Destination Type		
Is Port Transport Node	0.000	NA
Residential Destination Type		
Employment in Non-Wage & Salary-Other	3.671	8.29
Employment in Non-Wage & Salary-Work from Home	5.917	16.44
Total Households	0.000	NA
Household Attractions-Package Delivery	1.404	11.13
Household Attractions-Services	0.000	NA
Warehouse Destination Type		
Employment in Transportation/Warehousing	0.000	NA
Establishment Attractions-Retail	2.183	7.99
Establishment Attractions-Transportation	3.130	10.86
Establishment Attractions-Wholesale	3.405	18.40
Theta	1.000	NA
Impedance Parameters (β)		
Generalized Travel Time: All	-0.012	-8.21
Generalized Travel Time x Purpose is Home	-0.388	-22.69
Generalized Travel Time x Purpose is Maintenance	-0.010	-6.10
Generalized Travel Time x Purpose is Service	-0.006	-4.13
Generalized Travel Time x Elapsed Route Minutes/1000	-0.045	-12.10
Operating Costs ($) x Vehicle Type		
LCV	-0.063	-8.02
SUT	-0.024	-6.67
MUT	-0.018	-5.72

* Employment Total Non-WS/Transportation = Total Employment – Employment in Transportation/Warehousing – Non-Wage Salary Employment

2.13 Stop Location Choice – TNCs

The Next Stop Location Choice model predicts the location type and zone for each stop on a commercial vehicle route. As discussed in the CVM Model Design document this is the second model in the dynamic simulation loop, which also includes the choices of the stop purpose and stop duration. (See the flowchart in Figure 2.1-1 for reference.) Separate yet similarly structured models were developed for regular establishments and for regular establishments.

The Next Stop Location Choice model is formulated as a two-level model and consists of the following sub models:

- **Next Stop Location Choice**
    c. Choice of Location Type
    d. Given Location Type, Choice of Zone

2.13.1 Dataset

The Commercial Vehicle Survey TNC driver diary trip records were processed to identify the destination type for each stop consistent with the CVM Model Design. The following destination types were defined in the travel diary data:

- **Warehouse** – warehouse or distribution center that is not the establishment of the driver and identified in the survey travel diary as being a “warehouse” place type
- **Transport Node** – zones designated as marine ports, airports, or intermodal facilities and not the establishment of the driver. Overrides the warehouse designation.
- **Residential** – zones with households that were identified in the trip records of the travel diary as a “residential” place type or reported as the location of the driver’s home (activity type).
- **Other Commercial** – other zones with employment that do not fall into one of the above categories.

In addition to the survey observations, the zone location choice models required that travel time and costs skims be considered for each zone alternative. The time period used is determined by the clock time at the start of each trip. Generalized travel times from the base to each alternative zone were calculated as the congested travel time plus any toll costs converted to equivalent in-vehicle minutes. SOV high-income skims were assumed for TNCs, which include an assumed value of time of $85 per hour. Vehicle operating costs used in the model were assumed to be $0.188 per mile. In addition, zone-level data, such as employment and households, were joined from the MGRA land use file.

2.13.2 Variable Specification

The final set of estimated parameters for the first stage of the Next Stop Location Type model are shown in Table 2.13-1. The specification includes not only constants but also bias parameters for specific industries, destination purposes, customers, the elapsed time on the route at the time the stop choices are being made, and the time period in which the trip would start.

Table 2.13-1 Next Stop Location Type MNL Choice -- TNCs

Parameter Name	Coefficient	t-Stat.
Alternative Specific Constants		
Residential	0.000	NA
Warehouse	-2.374	-9.88
Transport Node	-7.284	-12.14
Other Commercial	0.355	5.43
Industry Effects (-> Next Destination Type)		
Client is Retail -> Next Destination Other Commercial	-0.574	-7.83
Client is Non-Restaurant/Retail -> Next Destination Warehouse	0.741	4.38
Client is Non-Restaurant/Retail -> Next Destination Other Commercial	-0.734	-10.45
Destination Stop Purpose Effects (-> Next Destination Type)		
Destination Purp is Service -> Next Destination Transport Node	2.752	4.46
Customer Effects[4] (-> Next Destination Type)		
Customer Type is Residential -> Next Destination Warehouse	-3.229	-3.17
Customer Type is Residential -> Next Destination Other Commercial	-2.309	-7.85
Customer Type is Mixed -> Next Destination Warehouse	-1.414	-8.55
Elapsed Time Effects (-> Next Destination Type)		
Elapsed Time (min/100) -> Next Destination Transport Node	0.554	5.56
Elapsed Time (min/100) -> Next Destination Other Commercial	-0.053	-3.33
Time of Day Effects (-> Next Destination Type)		
TOD is AM -> Next Destination Warehouse	0.941	2.94
TOD is MD -> Next Destination Warehouse	0.860	3.97
TOD is MD -> Next Destination Other Commercial	0.348	5.45
TOD is EV -> Next Destination Warehouse	-0.711	-1.92

The final set of estimated parameters for the Next Stop Location Choice of Zone for TNCs are shown in Table 2.13-2. As discussed in the CVM Model Design document, the formulation is a destination choice model, which is composed of a size function and an impedance function. The size function includes parameters related to quantities. The table shows parameters specific to each of the stop location types—commercial, residential, and warehouse interacted with zonal land use variables relevant to that type. For example, for a residential location type the number of households and employment in non-wage-and-salary jobs, and household delivery attraction variables for food and packages are quantity predictions. "Port Transport Node" is a special category that indicates that a zone is designated as either a marine port, airport, or intermodal facility.

Note that parameter identification requires that at least one size parameter be fixed, which was done for each location type. Size parameters are required to be positive valued and are actually applied within an exponential function. For example, a coefficient of zero when exponentiated becomes a parameter of 1.0. The impedance parameter in in Table 2.13-2 focus on generalized travel time. As discussed in the CVM Model Design document, this model considers the travel impedance from the current stop to all alternative

[4] Although customer effects were part of the originally estimated model, subsequently TNC routes were simplified to be for a single customer type—mixed residential/non-residential. The estimated parameters related to residential customer types therefore have no effect in the model.

next stops as well as the impedance from each alternative stop to the route termination location. As discussed in the CVM Model Design document, this model considers the travel impedance from the current stop to all alternative next stops as well as the impedance from each alternative stop to the route's termination location. This may be expressed as:

$$Total\ Impedance\ Alt\ j\ = \beta X_{ij} + \beta X_{jk}$$

Which is the impedance of travel *X* between zone *i* (current zone) and zone *j* (alternative next zone), and the impedance of travel *X* between zone *j* (alternative next zone) and zone *k* (route terminal zone). The estimated parameter β measures the disutility of travel times and costs.

Commented [BS27]: similar to the above comment, how exactly?

Commented [JG28R27]: Same as above

Commented [JG29]: Clarifying text.

Table 2.13-2 Next Stop Location Choice of Zone -- TNCs

Parameter Name	Coefficient	t-Stat.
Size Parameters (γ)		
Commercial Destination Type		
Establishment Attractions-All x Goods Delivery Purpose	1.835	17.63
Employment in Business Services x Maintenance Stop Purpose	-3.941	-2.91
Employment in Food x Maintenance Stop Purpose	1.644	7.62
Employment in Retail x Maintenance Stop Purpose	2.305	18.15
Employment in Food x Goods Pickup Purpose	4.776	34.85
Employment in Retail x Goods Pickup Purpose	4.694	18.16
Employment Total Non-WS/Transportation*	0.000	NA
Port Destination Type		
Is Port Transport Node	0.000	NA
Residential Destination Type		
Total Households	0.000	NA
Employment in Non-Wage & Salary-Other	4.782	4.90
Employment in Non-Wage & Salary-Work from Home	7.769	9.35
Household Attractions-Food Delivery	4.931	54.18
Household Attractions-Package Delivery	2.651	26.14
Warehouse Destination Type		
Establishment Attractions-Wholesale x Goods Delivery Purpose	2.925	13.91
Employment in Retail x Goods Pickup Stop Purpose	3.232	7.98
Employment in Wholesale x Goods Pickup Stop Purpose	4.169	22.78
Employment in Transportation/Warehousing	0.000	NA
Theta	1.000	NA
Impedance Parameters (β)		
Generalized Travel Time: All	-0.037	-23.09
Generalized Travel Time x Purpose is Home	-0.261	-13.01
Generalized Travel Time x Purpose is Maintenance	-0.011	-4.88
Generalized Travel Time x Purpose is Service	0.018	3.15
Generalized Travel Time x Elapsed Route Minutes/1000	-0.014	-2.92
Operating Cost for LCV-TNC	-0.085	-8.18

* Employment Total Non-WS/Transportation = Total Employment – Employment in Transportation/Warehousing – Non-Wage Salary Employment

The specification includes a generic generalized travel time parameter which applies to all cases, plus for additive factors for specific stop purposes, which indicate a greater sensitivity to travel time for home, maintenance/other (e.g., lunch, fueling), or for service stops. There is also an additive term for the interaction between generalized travel time and the elapsed time on the route which represents a greater sensitivity as the elapsed time increases. The specification also includes a vehicle operating cost parameter.

2.14 Stop Duration Simulation

As described in the CVM Model Design document, the Stop Duration model is formulated as a random draw from a fitted distribution. After testing numerous model forms and ways of segmenting the observations, a generalized Beta distribution was selected, which allows flexibility through estimated shape parameters and predicts probabilities for continuous variables within a unit interval.

2.14.1 Dataset

The Commercial Vehicle Survey travel diary trip records provide information on the duration of stops, excluding the final stops of each driver's day which indicate the end of the commercial vehicle route. Stops at the establishment ("base") and at the driver's home ("home") are included if the driver was observed to make additional stops afterwards.

2.14.2 Variable Specification

Beta distributions used to describe the distribution of stop durations for regular establishments are shown in Table 2.14-1. In each case, the maximum of the Beta distribution must be rescaled from the default of 1 to "maximum stop duration" (which becomes the effective maximum possible duration) and becomes a multiplicative constant. The estimated shape parameters—the alphas and betas—are estimated based on a normalized distribution, which does not account for maximum duration. The values shown in the table for "maximum stop duration" are rough estimates based on observed maximum durations but are really the starting point for calibrated values (see Section 3.10).

Commented [JG30]: New text.

The distribution was estimated separately for the "Before 12pm" segment of stops and the "After 12pm" segment of stops for different combinations of stop purpose and vehicle type. As discussed in the CVM Model Design document, it was observed that arrival at stops after 12 p.m. are likely to be shorter than stops with morning start times because there is less time left in the workday.

Beta distributions used to describe the distribution of stop durations for TNCs as shown in Table 2.14-2. In each case, the maximum of the Beta distribution must be rescaled from the default of 1 to "maximum stop duration" (which was used as the effective maximum possible duration). The distribution was estimated separately for five stop purposes, segmented by three density variables. These density variables represent the total population + employment density as found on MGRA land use file records. As discussed in the CVM Model Design document, density is a proxy for the amount of time that TNC drivers might spend in parking activities and inside buildings when making pickups and deliveries. Other stop purposes are not subject to density segmentation.

Table 2.14-1 Stop Duration Model for Establishments

Market Segment (Purpose x Vehicle Type)	Beta Distribution Parameters				
	max. stop duration (min)	Before 12pm Segment		After 12pm Segment	
		alpha	beta	alpha	beta
Goods Delivery x [LCV or SUT]	300	0.67	20.93	0.74	22.99
Goods Delivery x MUT	480	0.97	10.53	0.97	15.38
Goods Pickup x [LCV or SUT]	300	0.89	8.56	1.02	11.28
Goods Pickup x MUT	480	0.87	6.46	1.20	23.31
Service x LCV	900	0.47	5.86	0.60	13.79
Service x SUT	900	0.43	8.48	0.67	21.07
Service x MUT	900	0.47	8.33	0.61	15.66
Base	900	0.50	2.78	0.72	9.11
Home	900	0.42	3.50	0.65	8.58
Maintenance / Other	900	0.55	19.81	0.54	20.37

Table 2.14-2 Stop Duration Model for TNCs

Market Segment (Purpose)	Beta Distribution Parameters						
	max. stop duration (min)	Low Density (<15)		Medium Density (15 to 35)		High Density (> 35)	
		alpha	beta	alpha	beta	alpha	beta
Goods Delivery	300	0.59	17.26	0.62	14.90	0.58	10.95
Goods Pickup	300	0.71	10.95	0.71	10.95	0.71	10.95
Base	900	0.52	5.65	0.52	5.65	0.52	5.65
Home	1320	0.53	4.54	0.53	4.54	0.53	4.54
Maintenance / Other	900	0.54	10.42	0.54	10.42	0.54	10.42

3.0 CVM Model Component Calibration

The CVM model component calibration statistics reported below reflect the active calibration of model component parameters through March 12, 2024. Subsequent to that date, SANDAG staff implemented some small changes to model inputs as part of the model validation and preparation for the long range plan analysis. The statistics reported in this memo include results from validation runs that took place after some of these input changes were made, which affected the calibration fit statistics for certain model components as noted below.

3.1 Establishment Synthesis

The Establishment Synthesizer generates synthetic establishments by industry group and establishment size group through an optimization process. The resultant synthetic establishment records are considered to be calibrated as the algorithm works by matching employment by industry group for each MGRA while attempting to satisfy LUZ targets for the distribution of establishments by size class. In satisfying these two conditions, the model also determines the total number of establishments of each type.

Commented [BS31]: Are these based on observed data? Durations as high as 15 hours?

Commented [JG32R31]: Please see enhanced explanation in text above.

Commented [BS33]: 2024?

Commented [JG34R33]: 2024

Commented [SX35]: https://github.com/SANDAG/Data_CVM/blob/9aa6a57fd81c069a3fa9694fc881b207b267acad/Calib_Summary/calibration_targets_data/CVM_Calibration_03122024.xlsx
Is this the latest result?

Commented [JG36R35]: https://github.com/SANDAG/Data_CVM/blob/main/Calib_Summary/CVM_Calibration_FINAL.xlsx

Table 3.1-1 shows the industry group names and definitions used in the CVM and synthetic establishments. The establishment size categories are shown in Table 3.1-2.

Table 3.1-1 CVM Establishment Industry Groups

Code	Industry Group	Abbreviation	MGRA File Fields
1	Agriculture/Mining	AGM	emp_ag_min
2	Manufacturing	MFG	emp_mnf
3	Industrial/Utilities	IUT	emp_utl
4	Retail	RET	emp_ret
5	Wholesale	WHL	emp_whl
6	Construction	CON	emp_con
7	Transportation	TRN	emp_trn_wrh
8	Info/Finance/Insurance/Real Estate/Prof. Services	IFR	emp_fin_res_mgm, emp_bus_svcs
9	Education/Public/Other Services	EPO	emp_educ, emp_gov, emp_oth
10	Medical/Health Services	MHS	emp_hlth
11	Leisure/Accommodations and Food Services	LAF	emp_ent, emp_accm, emp_food
12	Military	MIL	emp_mil

Table 3.1-2 Establishment Size Class Definitions

Code	Class Definition
1	1 to 4 employees
2	5 to 9 employees
3	10 to 19 employees
4	20 to 49 employees
5	50 to 99 employees
6	100 to 249 employees
7	250 to more employees

3.1.1 Target Data

The target data for the Establishment Synthesizer are the two primary inputs to the process:

- Zonal land use file containing employment by industry sector in each zone. These employment values are industry-specific targets for the scenario year in question. The Establishment Synthesizer was specified to synthesize establishments at the MGRA geographic level and used the file **"mgra15_based_input2022.csv"** for model development. Table 3.1-1, above, shows the MGRA employment fields in the rightmost column that were used for each industry group.

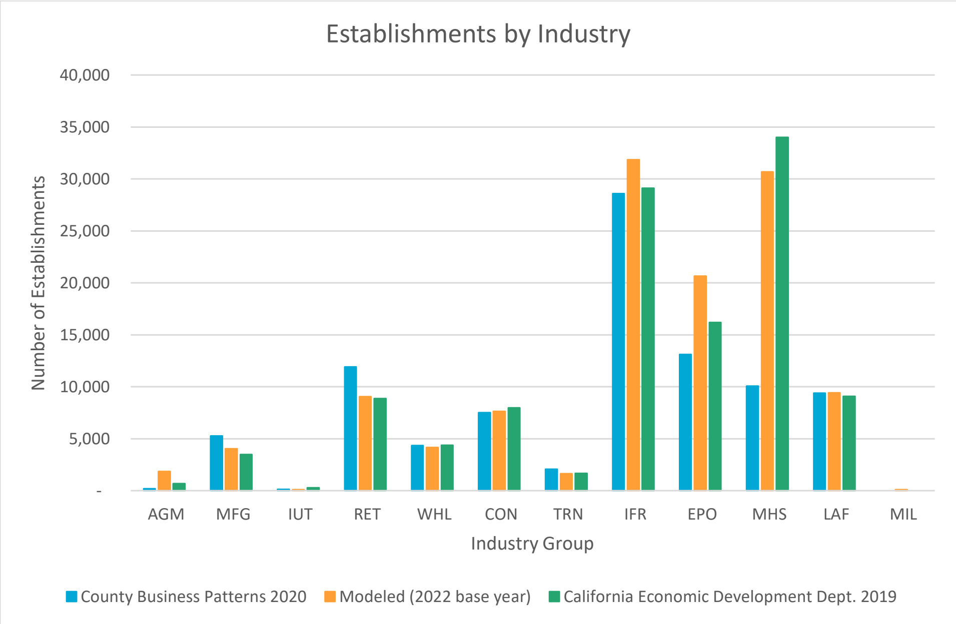
- LUZ data file containing baseline target percentage distributions of establishments by size and industry for each of 245 LUZs covering the SANDAG model area. The Establishment Synthesizer was specified to use the distributions created by SANDAG's land use and economics team who created the input file **"percent of establishments by luz-size-emp_cat.xlsx"**. This file was prepared in August 2023 using the industry groups and size classes shown above in Table 3.1-1 and Table 3.1-2.

3.1.2 Fit Statistics

The Establishment Synthesizer loops through each of the 17 MGRA employment types shown in the rightmost column of Table 3.1-1 and, for each employment category, it loops through each LUZ. In effect, tries to solve a separate optimization problem for each employment type and LUZ (17 x 245 = 4165). In effect, the model tries to solve a separate optimization problem for each employment type and LUZ, skipping LUZs where there is no employment of a particular type. Note that the final set of allocated establishment records used in the CVM use the 12 categories listed in Table 3.1-1, combining some of the MGRA categories as indicated.

As noted above, the Establishment Synthesizer creates synthetic establishment records as an outcome of trying to satisfy industry-specific employment at the MGRA level and establishment size distributions at the LUZ level. Figure 3.1-1 shows the total number of establishments generated by the Establishment Synthesizer centered in between two independent sources of establishment data which were used as references for weighting the 2022 SANDAG commercial vehicle survey (CVS), both representing different years from the 2022 base year represented in the model. The County Business Patterns (CBP) data is pre-pandemic 2019, while the California Economic Development Department (CEDD) data are from 2020, which may be affected by the COVID 19 pandemic, depending on what month of the year they represent. The modeled establishments were derived by the process described above and are rooted in SANDAG's estimates of employment in each MGRA for 2022 and the distribution of establishments by size class for each LUZ. Overall, the modeled data compares most closely to the CEDD establishments data but produces about 5 percent more total establishments across all the categories. Industry groups where the modeled data are notably higher are Agriculture and Mining and the Public Sector, which are not well represented in the CBP data, which also tends to under-represent small businesses, and are often subject to undercounting or differences in classifications by different sources. The model accounts for Military establishments, which are not covered in the other sources. In addition, the model predicts higher numbers of establishments in the group representing information, finance, insurance, real estate, and professional services, which may be attributed to growth in those sectors post pandemic.

Figure 3.1-1 Comparison of Synthetic Establishments vs. Independent Sources



Fit statistics for each of the 17 MGRA employment sectors developed for ABM3 are shown below in Table 3.1-3. For each employment sector, the model reports having matched the total employment for each TAZ, which was a hard constraint for the program, resulting in zero discrepancies. In addition, the model reports the total number of establishments created for each establishment size class and compares the modeled percent distribution to the target percent distribution for the entire model area based on an establishment-weighted average across all LUZs. The total coincidence ratio, the degree of overlap between the modeled and target distributions for the entire industry is also reported. A second coincidence ratio, the LUZ average, is reported as the establishment-weighted average of the coincidence ratios for individual LUZs, an indicator of how well on average the model was able to match individual LUZ targets.

Table 3.1-3 Establishment Synthesis Fit Statistics by ABM3 MGRA Industry Group

MGRA Industry Sector:	**Government**		
MGRAs with employment discrepancies: 0			
Estab. Size Class	Establishments	Model	Target
1	98	7.3%	8.0%
2	67	5.0%	4.8%
3	272	20.3%	20.5%
4	412	30.8%	29.9%
5	156	11.7%	11.3%
6	146	10.9%	11.2%
7	187	14.0%	14.4%
Totals	1,338	100.0%	100.0%
Coincidence Ratio Total:	0.971		
Coincidence Ratio LUZ Avg:	0.868		

MGRA Industry Sector:	**Agriculture and Mining**		
MGRAs with employment discrepancies: 0			
Estab. Size Class	Establishments	Model	Target
1	1,333	72.3%	72.1%
2	238	12.9%	13.3%
3	154	8.4%	8.4%
4	92	5.0%	4.7%
5	16	0.9%	1.0%
6	7	0.4%	0.5%
7	5	0.3%	0.0%
Totals	1,845	100.0%	100.0%
Coincidence Ratio Total:	0.986		
Coincidence Ratio LUZ Avg:	0.949		

MGRA Industry Sector:	**Business Services**		
MGRAs with employment discrepancies: 0			
Estab. Size Class	Establishments	Model	Target
1	13,384	64.1%	64.5%
2	2,629	12.6%	12.6%
3	1,945	9.3%	9.5%
4	1,593	7.6%	7.5%
5	682	3.3%	3.1%
6	448	2.2%	2.0%
7	198	1.0%	0.9%
Totals	20,879	100.0%	100.0%
Coincidence Ratio Total:	0.989		
Coincidence Ratio LUZ Avg:	0.953		

MGRA Industry Sector:	**FIRE, Management Services**		
MGRAs with employment discrepancies: 0			
Estab. Size Class	Establishments	Model	Target
1	7,130	65.1%	65.6%
2	1,780	16.2%	16.6%
3	979	8.9%	8.6%
4	627	5.7%	5.4%
5	230	2.1%	2.1%
6	147	1.3%	1.2%
7	66	0.6%	0.6%
Totals	10,959	100.0%	100.0%
Coincidence Ratio Total:	0.983		
Coincidence Ratio LUZ Avg:	0.950		

MGRA Industry Sector:	**Education**		
MGRAs with employment discrepancies: 0			
Estab. Size Class	Establishments	Model	Target
1	915	30.5%	29.9%
2	367	12.2%	12.3%
3	399	13.3%	13.2%
4	565	18.9%	18.8%
5	468	15.6%	16.2%
6	206	6.9%	7.1%
7	78	2.6%	2.5%
Totals	2,998	100.0%	100.0%
Coincidence Ratio Total:	0.982		
Coincidence Ratio LUZ Avg:	0.863		

MGRA Industry Sector:	**Health and Medical Services**		
MGRAs with employment discrepancies: 0			
Estab. Size Class	Establishments	Model	Target
1	25,222	82.2%	83.2%
2	2,386	7.8%	7.4%
3	1,557	5.1%	4.8%
4	836	2.7%	2.5%
5	326	1.1%	1.1%
6	216	0.7%	0.7%
7	133	0.4%	0.4%
Totals	30,676	100.0%	100.0%
Coincidence Ratio Total:	0.981		
Coincidence Ratio LUZ Avg:	0.963		

MGRA Industry Sector:	**Retail**		
MGRAs with employment discrepancies: 0			
Estab. Size Class	Establishments	Model	Target
1	3,716	41.0%	40.9%
2	2,062	22.8%	23.5%
3	1,537	17.0%	17.2%
4	995	11.0%	10.6%
5	437	4.8%	4.7%
6	248	2.7%	2.6%
7	63	0.7%	0.6%
Totals	9,058	100.0%	100.0%
Coincidence Ratio Total:	0.981		
Coincidence Ratio LUZ Avg:	0.914		

MGRA Industry Sector:	**Transportation and Warehousing**		
MGRAs with employment discrepancies: 0			
Estab. Size Class	Establishments	Model	Target
1	861	52.8%	52.9%
2	238	14.6%	14.7%
3	209	12.8%	13.9%
4	175	10.7%	10.0%
5	77	4.7%	4.2%
6	42	2.6%	2.6%
7	30	1.8%	1.8%
Totals	1,632	100.0%	100.0%
Coincidence Ratio Total:	0.974		
Coincidence Ratio LUZ Avg:	0.887		

MGRA Industry Sector:	**Construction**		
MGRAs with employment discrepancies: 0			
Estab. Size Class	Establishments	Model	Target
1	4,387	57.5%	57.2%
2	1,440	18.9%	19.1%
3	822	10.8%	11.0%
4	637	8.4%	8.5%
5	203	2.7%	2.6%
6	112	1.5%	1.4%
7	29	0.4%	0.3%
Totals	7,630	100.0%	100.0%
Coincidence Ratio Total:	0.989		
Coincidence Ratio LUZ Avg:	0.922		

MGRA Industry Sector:	**Utilities**		
MGRAs with employment discrepancies: 0			
Estab. Size Class	Establishments	Model	Target
1	35	37.2%	39.0%
2	14	14.9%	12.4%
3	12	12.8%	14.0%
4	10	10.6%	11.6%
5	7	7.5%	8.4%
6	8	8.5%	6.7%
7	8	8.5%	7.9%
Totals	94	100.0%	100.0%
Coincidence Ratio Total:	0.906		
Coincidence Ratio LUZ Avg:	0.887		

MGRA Industry Sector:	**Manufacturing**		
MGRAs with employment discrepancies: 0			
Estab. Size Class	Establishments	Model	Target
1	1,554	38.6%	38.7%
2	579	14.4%	14.9%
3	619	15.4%	16.0%
4	582	14.5%	14.1%
5	314	7.8%	7.4%
6	222	5.5%	5.1%
7	154	3.8%	3.9%
Totals	4,024	100.0%	100.0%
Coincidence Ratio Total:	0.976		
Coincidence Ratio LUZ Avg:	0.918		

MGRA Industry Sector:	**Wholesale**		
MGRAs with employment discrepancies: 0			
Estab. Size Class	Establishments	Model	Target
1	2,346	56.3%	56.6%
2	782	18.8%	19.3%
3	506	12.1%	12.7%
4	350	8.4%	7.7%
5	119	2.9%	2.4%
6	54	1.3%	1.0%
7	13	0.3%	0.3%
Totals	4,170	100.0%	100.0%
Coincidence Ratio Total:	0.970		
Coincidence Ratio LUZ Avg:	0.921		

MGRA Industry Sector:	**Entertainment**		
MGRAs with employment discrepancies: 0			
Estab. Size Class	Establishments	Model	Target
1	981	58.2%	58.2%
2	205	12.2%	11.8%
3	210	12.5%	12.8%
4	191	11.3%	11.9%
5	44	2.6%	3.1%
6	32	1.9%	1.5%
7	23	1.4%	0.7%
Totals	1,686	100.0%	100.0%
Coincidence Ratio Total:	0.972		
Coincidence Ratio LUZ Avg:	0.883		

MGRA Industry Sector:	**Accommodations**		
MGRAs with employment discrepancies: 0			
Estab. Size Class	Establishments	Model	Target
1	151	24.6%	25.0%
2	133	21.7%	21.1%
3	126	20.6%	20.8%
4	101	16.5%	16.4%
5	57	9.3%	9.2%
6	28	4.6%	4.9%
7	17	2.8%	2.6%
Totals	613	100.0%	100.0%
Coincidence Ratio Total:	0.982		
Coincidence Ratio LUZ Avg:	0.905		

MGRA Industry Sector:	**Food and Drink**		
MGRAs with employment discrepancies: 0			
Estab. Size Class	Establishments	Model	Target
1	1,516	21.3%	20.1%
2	1,374	19.3%	19.3%
3	1,754	24.7%	26.3%
4	1,760	24.8%	26.2%
5	504	7.1%	6.4%
6	173	2.4%	1.5%
7	28	0.4%	0.2%
Totals	7,109	100.0%	100.0%
Coincidence Ratio Total:	0.940		
Coincidence Ratio LUZ Avg:	0.882		

MGRA Industry Sector:	**Other Services**		
MGRAs with employment discrepancies: 0			
Estab. Size Class	Establishments	Model	Target
1	5,552	66.4%	66.4%
2	1,524	18.2%	18.5%
3	836	10.0%	10.1%
4	333	4.0%	3.7%
5	67	0.8%	0.7%
6	33	0.4%	0.4%
7	13	0.2%	0.2%
Totals	8,358	100.0%	100.0%
Coincidence Ratio Total:	0.992		
Coincidence Ratio LUZ Avg:	0.939		

MGRA Industry Sector:	**Military**		
MGRAs with employment discrepancies: 0			
Estab. Size Class	Establishments	Model	Target
1	-	0.0%	0.0%
2	-	0.0%	0.0%
3	3	3.7%	3.7%
4	-	0.0%	0.0%
5	6	7.4%	6.5%
6	10	12.4%	11.4%
7	62	76.5%	78.4%
Totals	81	100.0%	100.0%
Coincidence Ratio Total:	0.964		
Coincidence Ratio LUZ Avg:	0.964		

3.2 Household Attraction Generation

The Household Attraction Generation model predicts the number of households attracting commercial vehicle trips for each of three purposes--food delivery, package delivery, and service provision. There is a separate model for each purpose. The estimated models, prior to calibration, are discussed above in Section 2.2. Application of the model requires the synthetic households and tour records from the Resident model, which are considered to be calibrated and validated. The binary logit model formulation requires calibration of a single constant. In addition, the models were calibrated to household income group targets for two of the three attraction types.

3.2.1 Target Data

The target data used to calibrate the Household Attraction Generation model come from the expanded 2022 SANDAG household travel survey, which asked respondents whether they received deliveries of food or other packages, or whether they received a service stop on the survey day. The dependent variable is formulated as a binary choice and segmented by each of the three commercial vehicle stop types.

3.2.2 Fit Statistics

Comparisons of model outputs to survey target data are shown in Table 3.2-1, Table 3.2-2, Table 3.2-3, and Table 3.2-4. Due to non-reported income groups, it was not possible to perfectly align the rates by income group with the overall rates. The models were calibrated to the overall rates as shown in Table 3.2-1. Nevertheless, the results are generally in agreement with the targets for the Package and Service attractions. For Food deliveries, due to the small number of observations, the calibration focused only on total deliveries per household although comparisons to income groups are provided in Table 3.2-4.

Table 3.2-1 Percent of All Household Days Attracting Food, Package, or Service Received at Home

All Households	Model	Survey	Difference
Food	4.3%	4.8%	-0.005
Service	6.7%	6.8%	-0.002
Package	38.7%	40.8%	-0.022

Table 3.2-2 Percent of Household Days with Food Delivery by Income Group

Household Income Group	Model	Survey	Difference
Not Reported	---	3.1%	---
Less than $50,000	3.5%	3.2%	0.2%
$50,000-$99,999	4.8%	4.9%	-0.1%
$100,000-$199,999	4.7%	6.3%	-1.6%
$200,000 or more	4.7%	7.9%	-3.2%

Table 3.2-3 Percent of Household Days with Package Delivery by Income Group

Household Income Group	Model	Survey	Difference
Not Reported	---	34.3%	---
Less than $50,000	19.5%	27.7%	-8.1%
$50,000-$99,999	37.0%	38.7%	-1.6%
$100,000-$199,999	52.0%	53.5%	-1.5%
$200,000 or more	55.6%	57.0%	-1.4%

Table 3.2-4 Percent of Household Days with Service Visit by Income Group

Household Income Group	Model	Survey	Difference
Not Reported	---	7.1%	---
Less than $50,000	2.2%	3.8%	-1.6%
$50,000-$99,999	5.3%	5.4%	-0.1%
$100,000-$199,999	6.0%	6.0%	0.0%
$200,000 or more	18.3%	18.3%	0.0%

3.3 Establishment Attraction Generation

Establishment Attraction Generation model predicts which establishments in the CVM will attract commercial vehicle trips for either goods pickup/delivery or for service calls and how many visits they will attract. As noted in the CVM Model Design document, this is a two-stage model which first predicts whether an establishment will generate at least one attraction, followed by a second model which predicts how many attractions will be generated given there will be at least one.

3.3.1 Target Data

The Establishment Attraction Generation model was calibrated using the expanded Commercial Vehicle Survey, which asked respondents how many commercial vehicle visits they receive on an average business day for either cargo or services.

3.3.2 Fit Statistics

Comparisons of model outputs to survey target data are shown in Table 3.3-1 for the probability of at least one attraction, segmented by the aggregate industry groups specified in the model estimation (See Section 2.3). The results show a tight fit to the target proportions.

Table 3.3-1 Establishment Attraction Generation Calibration Proportion of Establishments with at Least 1 Attraction

Industry Group	Model	Survey	% Difference
MFG, WHL, RET	0.8790	0.8756	0.39%
IUT, LAF	0.8396	0.8424	-0.34%
AGM, CON, TRN, EPO, MHS	0.7545	0.7499	0.61%
IFO	0.6820	0.6852	-0.46%

Table 3.3-2 shows the fit statistics for the model that predicts the actual number of establishment attractions, given there is at least one. The match to the target data is generally quite accurate for most industry categories with two notable exceptions, Agriculture and Mining and Industrial and Utilities. This difference is due to a change in a land uses file input that took place after this model component had been calibrated but was part of the validation run use for this report. This discrepancy is not expected to have a meaningful outcome on model behavior but should be recalibrated as part of a future model update.

Table 3.3-2 Establishment Attraction Generation Calibration Mean Number of Attractions if at Least 1

Industry	Model	Survey	% Difference
Agriculture Mining	2.622	2.214	18.40%
Manufacturing	3.076	3.093	-0.56%
Industrial Utilities	4.468	4.346	2.82%
Retail	4.531	4.532	-0.03%
Wholesale	3.325	3.334	-0.27%
Construction	3.080	3.109	-0.93%
Transportation	3.683	3.721	-1.03%
Info FIRE Professional Services	2.585	2.583	0.07%
Education, Public, Other Services	3.259	3.257	0.05%
Medical Health	2.237	2.235	0.08%
Leisure, Accommodation, Food	3.269	3.243	0.82%

3.4 Route Generation

The Route Generation model predicts which establishments in the CVM will generate commercial vehicle routes and how many. As noted in the CVM Model Design document, this is a two-stage model which first predicts whether an establishment will generate at least one route, followed by a second model which predicts how many routes given at least one.

3.4.1 Target Data

The establishment Route Generation model was calibrated to targets from the expanded Commercial Vehicle Survey. As discussed in 2.5.1, above, the survey asked each establishment how many outgoing cargo or service trips they make on an average workday. The responses to these questions were used to estimate the initial model. The survey also includes driver travel diary records from which routes may also be observed; however, because only about half of the establishments in the survey participate in the driver travel diaries, the establishment questions were viewed as a more reliable indicator of which establishments actually generate commercial vehicle routes. In the course of validating the model, it was found that the expanded number of routes as aggregated from the driver trip diary portion of the survey were much closer to matching system-wide targets for total commercial vehicle trips and vehicle miles traveled (VMT). Therefore, the weight-expanded route information from the driver diaries were used for calibrating the Route Generation model for establishments.

3.4.2 Fit Statistics

Comparisons between the Route Generation Model and the expanded survey results for driver diaries are shown below in. Overall, the model matches the targets quite well with two exceptions, Agriculture and Mining and Industrial and Utilities. As noted above, this difference is due to a change in a land uses file input that took place after this model component had been calibrated but was part of the validation run use for this report. This discrepancy is not expected to have a meaningful outcome on model behavior but should be recalibrated as part of a future model update.

Table 3.4-1 Establishment Route Generation Calibration

Industry	Model	Survey	% Difference
Agriculture Mining	716	2,327	-69.23%
Manufacturing	14,657	14,639	0.13%
Industrial Utilities	3,540	3,243	9.17%
Retail	13,053	12,893	1.24%
Wholesale	10,920	10,681	2.24%
Construction	31,784	32,129	-1.07%
Transportation	9,254	9,182	0.78%
Info FIRE Professional Services	43,951	43,519	0.99%
Education, Public, Other Services	55,771	54,986	1.43%
Medical Health	31,661	31,416	0.78%
Leisure, Accommodation, Food	7,027	7,218	-2.64%
Total	222,334	222,232	0.05%

Calibration of the TNC Route Generation model is based on adjustments to the rate factors described in the CVM Model Design document, which describes the process of creating expansion weights for the TNC driver survey. This model generates TNC routes for three client groups based on total employment of establishments in each group by LUZ. Overall, the model matches the weighted estimation of the establishment survey well as shown in Table 2.2-1.

Table 3.4-2 TNC Route Generation Calibration

	Total Routes Per Day		
TNC Client Group	**Model**	**Survey**	**Difference**
Not Restaurant or Retail	327	359	-9.0%
Restaurant	2,875	2,844	1.1%
Retail	1,814	1,816	-0.1%
Total	5,016	5,019	-0.1%

Commented [BS37]: The survey numbers in this table differ from those in table 2.61 (and also those in design memo).

Commented [JG38R37]: These are correct and represent a more current version of the processed survey data. Updated numbers in the other tables and the design memo.

3.5 Route Vehicle, Purpose, and Customer Type Assignment

For each route generated, several attributes are created through a series of sub-models. The first sub-model in the sequence assigns to each route a vehicle type, primary purpose, and customer type ("VPC Model"). As noted in the CVM Model Design document, this is formulated as a multinomial logit model.

3.5.1 Target Data

The expanded Commercial Vehicle Survey was processed to create route records and used to calibrate the VPC model. Each route record represents one day of travel for a single commercial vehicle and is created from the survey's trip travel diaries. Route records include the vehicle type used on the route, a primary purpose for each route, and a customer type for each route. These route attributes are defined as follows:

- **Vehicle Type**:
  - Light Commercial Vehicle (LCV)
  - Single-Unit Truck (SUT)
  - Multi-Unit Truck (MUT)
- **Purpose Type**:
  - Goods: customer-oriented; may include Services and other stop types
  - Service: customer-oriented; does not include Goods stops; may include other stop types
  - Maintenance/Other: refueling, driver breaks, vehicle/equipment repositioning, buying supplies, other; does not include Goods or Services stops; not customer oriented
- **Customer Type**:
  - Residential Only: households, including multi-family buildings (Goods or Services purposes)
  - Non-residential Only: commercial, public/government (Goods or Services purposes)
  - Mixed Residential and Non-residential: (Goods or Services purposes)

- No Customer: (Maintenance/Other purpose)

The alternatives in the VPC model represent the 21 feasible combinations of vehicle, purpose, and customer type as defined above.

3.5.2 Fit Statistics

The alternative-specific constants in the VPC model were calibrated to match the expanded survey targets for each of the 21 alternatives in the model. As shown in Table 3.5-1, the model matches the targets well across nearly all alternatives. Marginal results for comparisons by route purpose, customer type, and vehicle type are shown in Table 3.5-2, Table 3.5-3, and Table 3.5-4, respectively.

Table 3.5-1 Route Choice of Primary Purpose, Customer Type, and Vehicle Type Calibration All Joint Alternatives

	Model			Survey			
All Joint Alternatives (purpose x customer x vehicle type)	Routes	% Routes		Routes	% Routes		% Diff
Goods_Mixed_LCV	8,519	3.8%		8,461	3.8%		0.69%
Goods_Mixed_MUT	2,986	1.3%		3,024	1.4%		-1.26%
Goods_Mixed_SUT	4,207	1.9%		4,229	1.9%		-0.53%
Goods_NonRes_LCV	5,012	2.3%		5,045	2.3%		-0.66%
Goods_NonRes_MUT	13,045	5.9%		13,160	5.9%		-0.87%
Goods_NonRes_SUT	4,148	1.9%		4,159	1.9%		-0.27%
Goods_Res_LCV	6,084	2.7%		6,057	2.7%		0.44%
Goods_Res_MUT	1,779	0.8%		1,704	0.8%		4.41%
Goods_Res_SUT	4,701	2.1%		4,776	2.1%		-1.58%
Maintenance_LCV	13,689	6.2%		13,588	6.1%		0.74%
Maintenance_MUT	2,254	1.0%		2,346	1.1%		-3.91%
Maintenance_SUT	5,331	2.4%		5,267	2.4%		1.22%
Service_Mixed_LCV	33,871	15.2%		33,394	15.0%		1.43%
Service_Mixed_MUT	3,458	1.6%		3,454	1.6%		0.12%
Service_Mixed_SUT	13,612	6.1%		13,543	6.1%		0.51%
Service_NonRes_LCV	50,065	22.5%		50,409	22.7%		-0.68%
Service_NonRes_MUT	6,241	2.8%		6,362	2.9%		-1.91%
Service_NonRes_SUT	23,040	10.4%		23,153	10.4%		-0.49%
Service_Res_LCV	16,164	7.3%		15,971	7.2%		1.21%
Service_Res_MUT	594	0.3%		594	0.3%		0.00%
Service_Res_SUT	3,534	1.6%		3,535	1.6%		-0.03%
Total	222,334	100.0%		222,232	100.0%		0.05%

Table 3.5-2 Route Choice of Primary Purpose, Customer Type, and Vehicle Type Calibration by Route Purpose

	Model		Survey		
Route Purpose	**Routes**	**% Routes**	**Routes**	**% Routes**	**% Diff**
Goods	50,481	22.7%	50,617	22.8%	-0.27%
Maintenance	21,274	9.6%	21,201	9.5%	0.35%
Service	150,579	67.7%	150,414	67.7%	0.11%
Total	222,334	100.0%	222,232	100.0%	0.05%

Table 3.5-3 Route Choice of Primary Purpose, Customer Type, and Vehicle Type Calibration by Customer Type

	Model		Survey		
Customer Type	**Routes**	**% Routes**	**Routes**	**% Routes**	**% Diff.**
Mixed	66,653	30.0%	66,105	29.7%	0.83%
Residential	32,856	14.8%	32,637	14.7%	0.67%
Non-Residential	101,551	45.7%	102,289	46.0%	-0.72%
None (Maint./Other)	21,274	9.6%	21,201	9.5%	-1.95%
Total	222,334	100%	222,232	100.0%	0.05%

Table 3.5-4 Route Choice of Primary Purpose, Customer Type, and Vehicle Type Calibration by Vehicle Type

	Model		Survey		
Vehicle Type	**Routes**	**% Routes**	**Routes**	**% Routes**	**% Diff.**
LCV	133,404	60.0%	132,925	59.8%	0.36%
MUT	30,357	13.7%	30,644	13.8%	-0.94%
SUT	58,573	26.3%	58,663	26.4%	-0.15%
Total	222,334	100.0%	222,232	100.0%	0.05%

3.6 Route Start Time Choice

The Route Start Time choice model simulates route start times as draws from empirical distributions. Development of these distributions is described in the **CVM Model Design** technical memo. As these empirical distributions were developed from weighted Commercial Vehicle Survey observations, no further model estimation or calibration is typically required.

3.6.1 Target Data

Weighted observations for route start times from the Commercial Vehicle Survey.

3.6.2 Fit Statistics

Comparisons of the modeled route start times with the weighted survey data are shown in Table 3.6-1 for regular establishments and in Table 3.6-2 for TNCs. Both models match the survey targets will by the percent of routes starting in each time period.

Table 3.6-1 Route Start Time Calibration -- Establishments

Commented [JG39]: These numbers have been updated. Prior draft showed total route start times, which included TNCs. TNCs have been separated out and are shown in Table 3.6.2.

	Model		Survey		
Time Period	Model Routes	Model % Share	Survey - Routes	Survey % Share	Difference
AM	100,343	45.1%	101,824	45.8%	-0.7%
MD	93,503	42.1%	91,342	41.1%	1.0%
PM	7,468	3.4%	7,167	3.2%	0.1%
EV	5,539	2.5%	5,768	2.6%	-0.1%
EA	15,481	7.0%	16,130	7.3%	-0.3%
Total	222,334	100.0%	222,232	100.0%	0.0%

Table 3.6-2 Route Start Time Calibration -- TNCs

Commented [JG40]: These numbers have been updated. Prior draft showed Trip start time numbers (not route).

	Model		Survey		
Time Period	Model Routes	Model % Share	Survey - Routes	Survey % Share	Difference
AM	553	11.0%	588	11.7%	-0.7%
MD	2,951	58.8%	2,992	59.6%	-0.8%
PM	1,013	20.2%	1,017	20.3%	-0.1%
EV	413	8.2%	363	7.2%	1.0%
EA	87	1.7%	59	1.2%	0.6%
Total	5,017	100.0%	5,019	100.0%	0.0%

3.7 Route Origin and Termination Location Choices

As discussed in Section , the Route Origin and Termination Location Choice step includes the following sub models:

- **Route Origin Location Choice**
  a. Choice of Location Type – multinomial logit
  b. Given Location Type, Choice of Zone -- destination choice model
- **Route Termination Location Choice**
  a. Choice of Location Type – multinomial logit
  b. Given Location Type, Choice of Zone -- destination choice model

Calibration of these models involved matching targets for location types and for mean distances between the establishment and the route origin or termination location. Note that this set of models only applies to regular establishments. TNCs route starting and ending locations follow a different process that was not calibrated.

3.7.1 Target Data

The Commercial Vehicle Survey travel diary records were aggregated to routes and the weighted route observations were used to create targets for the percentage of routes with origins and terminations of each type as well as the mean trip distances from the establishment.

3.7.2 Fit Statistics

Comparisons of Origin Type choice and Termination Type choice shares are shown in Table 3.7-1 and Table 3.7-2, respectively. In both cases, the model matches the target shares well.

Table 3.7-1 Route Origin Type Calibration

	Routes		Shares		
Origination Type	**Model**	**Survey**	**Model**	**Survey**	**% Diff**
Base (Establishment)	178,920	177,361	80.47%	80.51%	0.0%
Warehouse	2,896	2,976	1.30%	1.35%	0.0%
Residential	18,643	18,812	8.39%	8.54%	-0.2%
Other Commercial	21,875	21,149	9.84%	9.60%	0.2%
Total	222,334	220,298			

Table 3.7-2 Route Termination Type Calibration

	Routes		Shares		
Termination Type	**Model**	**Survey**	**Model**	**Survey**	**% Diff**
Base	121,402	119,137	54.60%	54.86%	-0.3%
Warehouse	5,345	5,048	2.40%	2.32%	0.1%
Residential	45,212	43,927	20.34%	20.23%	0.1%
Other Commercial	50,375	49,044	22.66%	22.58%	0.1%
Total	222,334	217,156			

Comparisons of the mean distance to the establishment for Origin and Termination Location choices are shown in Table 3.7-3 and Table 3.7-4, respectively. In each case, the model matches the target means well.

Table 3.7-3 Route Origin Calibration Mean Distance (Miles) from Establishment if not Establishment

Mean Distance from Establishment			
Origination Type	**Model**	**Survey**	**% Diff**
Warehouse	8.14	8.60	-5.4%
Residential	11.69	11.57	1.0%
Other Commercial	11.57	12.12	-4.5%

Table 3.7-4 Route Termination Calibration Mean Distance (Miles) from Establishment if not Establishment

Mean Distance from Establishment			
Termination Type	**Model**	**Survey**	**% Diff**
Warehouse	2.81	2.81	0.2%
Residential	3.65	3.69	-1.2%
Other Commercial	5.03	5.18	-2.8%

It is worth noting that the percentage of routes with origins at the establishment is about 80 percent, whereas the percentage of routes that terminate at the establishment is about 55 percent. This difference can be explained by the perception of the driver. A vehicle which is dropped off by a driver based at establishment “A” at another establishment “B” may be driven the following day by a driver based at establishment “B.”

It is also important to consider that the drivers who originate their routes at a non-establishment location are not necessarily the same persons who terminate their routes at a non-establishment location. While there are fewer drivers starting their day at a non-establishment origin, those origins tend to be farther from their base establishment, which makes sense if the first stop on the vehicle route is not close to the establishment and that first reporting to the establishment would be out of the way or inconvenient. At the end of the day, it is more common for drivers to terminate their route at a non-establishment location, but those locations are on average closer to the establishment, which suggests reasons related to end-of-day vehicle positioning for warehouses operations, vehicle maintenance, or because the driver happens to live near the establishment. Table 3.7-5 was created to show that the total extra travel implied by these positioning perspectives balances out to be of similar magnitude for Origins and Terminations when the location type is Other Commercial, but is skewed towards longer travel to Origins when the location type is Warehouse or Residential.

Commented [JG41]: Enhanced explanation.

Table 3.7-5 Route Origin and Termination Total Travel Distances Relative to Establishment if Not Establishment

Implied Total Travel Distances from Establishments				
	Origins		**Terminations**	
Location Type	**Model**	**Survey**	**Model**	**Survey**
Warehouse	23,566	25,583	15,044	14,181
Residential	217,855	217,606	164,992	162,246
Other Commercial	253,028	256,249	253,523	253,887
Total	494,449	499,438	433,559	430,313

Commented [BS42]: The difference between 499k and 430k (69k) represents routes where origin and destination were both at base?

Commented [JG43R42]: No, this table does not show when route origins and destinations are both at the base. I tried to provide a better explanation in the text.

3.8 Next Stop Purpose Choice

The Next Stop Purposes Choice model predicts the primary purpose for each stop on a commercial vehicle route. The model is formulated as a multinomial logit.

3.8.1 Target Data

The Commercial Vehicle Survey driver diary trip records were expanded to create target shares for stop purposes as follows:

- **Goods Pickup** – a stop to pick up goods
- **Goods Delivery** – a stop to deliver goods to a customer
- **Service** – a stop to provide a professional service to a customer
- **Maintenance/Other** – a stop for either vehicle maintenance/refueling or driver breaks
- **Base Establishment** – a trip the establishment where the vehicle is based, without ending the route
- **Home** – a trip to the driver's home, without ending the route
- **Terminate Route** – a trip to the final stop on the route where the vehicle will be parked, ending the route.

Because transitions between an origin purpose and the next stop purpose are important to the overall model design and, the calibration focused on these transition proportions.

3.8.2 Fit Statistics

For regular establishments, the stop purpose transition shares are shown in Table 3.8-1 for the weighted survey proportions and in Table 3.8-2 for the weighted modeled proportions. The differences between the shares are shown in Table 3.8-3. Differences tend to be small and the differences when viewed by row and column marginals are close to one percent.

For TNCs, the stop purpose transition shares are shown in Table 3.8-4 for the weighted survey proportions and in Table 3.8-5 for the weighted modeled proportions. The differences between the shares are shown in Table 3.8-6. Differences tend to be small and the differences when viewed by row and column marginals are all less than one percent.

Table 3.8-1 Next Stop Purpose Transition Proportions Establishments: Survey

Trip Origin Purpose\Trip Destination Purpose	Base	Goods Pickup	Goods Delivery	Service	Maint / Other	Home	Terminal Destination	Marginal
Originate	0.50%	0.59%	1.53%	6.61%	3.61%	0.18%	0.35%	13.36%
Base	0.00%	0.11%	0.79%	1.90%	1.31%	0.05%	0.42%	4.58%
Goods Pickup	0.12%	0.57%	0.79%	0.24%	0.26%	0.00%	0.29%	2.26%
Goods Delivery	0.97%	0.36%	15.11%	0.39%	1.63%	0.05%	1.43%	19.95%
Service	1.94%	0.18%	0.32%	20.91%	6.71%	0.32%	6.37%	36.75%
Maintenance/Other	1.00%	0.46%	1.38%	6.38%	8.38%	0.26%	4.41%	22.26%
Home	0.05%	0.00%	0.03%	0.32%	0.36%	0.00%	0.09%	0.85%
Marginal	4.58%	2.26%	19.95%	36.75%	22.26%	0.85%	13.36%	100.00%

Table 3.8-2 Next Stop Purpose Transition Proportions Establishments: Model

Trip Origin Purpose\Trip Destination Purpose	Base	Goods Pickup	Goods Delivery	Service	Maint / Other	Home	Terminal Destination	Marginal
Originate	0.23%	0.09%	0.36%	10.43%	3.12%	0.00%	0.08%	14.32%
Base	0.00%	0.00%	0.01%	2.94%	1.47%	0.00%	0.01%	4.43%
Goods Pickup	0.08%	0.35%	0.97%	0.26%	0.32%	0.01%	0.20%	2.19%
Goods Delivery	0.47%	0.36%	13.84%	0.33%	1.97%	0.03%	1.45%	18.46%
Service	2.57%	0.71%	1.30%	18.68%	5.64%	0.53%	8.26%	37.68%
Maintenance/Other	0.99%	0.67%	1.95%	4.87%	9.31%	0.18%	4.19%	22.16%
Home	0.08%	0.01%	0.04%	0.17%	0.33%	0.00%	0.13%	0.76%
Marginal	4.43%	2.19%	18.46%	37.68%	22.16%	0.76%	14.32%	100.00%

Commented [SX44]: The number does not match the one on the file: https://github.com/SANDAG/Data_CVM/blob/9aa6a57fd81c069a3fa9694fc881b207b267acad/Calib_Summary/calibration_targets_data/CVM_Calibration_03122024.xlsx

And 'terminal destination' is missing in the trip origin purpose in both tables 3.8-1 and 3.8-2

Commented [JG46R44]: Terminal Destination is never an origin purpose. By definition, it can only be the last purpose on the route. If it were a row, it would be all zeros.

Commented [JG45R44]: https://github.com/SANDAG/Data_CVM/blob/main/Calib_Summary/CVM_Calibration_FINAL.xlsx is where the numbers came from.

Table 3.8-3 Next Stop Purpose Transitions Differences in Transition Proportions (Model-Survey) Establishments

Trip Origin Purpose\Trip Destination Purpose	Base	Goods Pickup	Goods Delivery	Service	Maint / Other	Home	Terminal Destination	Marginal
Originate	-0.27%	-0.50%	-1.18%	3.82%	-0.48%	-0.17%	-0.26%	0.97%
Base	0.00%	-0.11%	-0.78%	1.04%	0.16%	-0.05%	-0.41%	-0.15%
Goods Pickup	-0.04%	-0.22%	0.18%	0.02%	0.06%	0.01%	-0.09%	-0.07%
Goods Delivery	-0.50%	0.00%	-1.27%	-0.06%	0.34%	-0.02%	0.02%	-1.49%
Service	0.62%	0.53%	0.97%	-2.23%	-1.07%	0.22%	1.89%	0.93%
Maintenance/Other	-0.01%	0.22%	0.57%	-1.51%	0.93%	-0.08%	-0.22%	-0.09%
Home	0.03%	0.01%	0.01%	-0.15%	-0.04%	0.00%	0.04%	-0.10%
Marginal	-0.15%	-0.07%	-1.49%	0.93%	-0.09%	-0.10%	0.97%	0.00%

Table 3.8-4 Next Stop Purpose Transition Proportions TNCs: Survey

Trip Origin Purpose\Trip Destination Purpose	Base	Goods Pickup	Goods Delivery	Service	Maint / Other	Home	Terminal Destination	Marginal
Originate	0.05%	3.20%	1.68%	0.33%	2.72%	0.05%	0.19%	8.24%
Base	0.00%	0.44%	0.08%	0.04%	0.56%	0.02%	0.09%	1.23%
Goods Pickup	0.13%	2.52%	16.76%	0.03%	1.05%	0.19%	1.04%	21.72%
Goods Delivery	0.48%	8.99%	24.98%	0.56%	6.50%	0.61%	3.48%	45.59%
Service	0.08%	0.49%	0.16%	0.50%	0.49%	0.01%	0.25%	1.98%
Maintenance/Other	0.49%	5.56%	1.84%	0.50%	7.81%	0.61%	2.97%	19.77%
Home	0.00%	0.51%	0.10%	0.02%	0.63%	0.00%	0.22%	1.48%
Marginal	1.23%	21.72%	45.59%	1.98%	19.77%	1.48%	8.24%	100.00%

Table 3.8-5 Next Stop Purpose Transition Proportions TNCs: Model

Trip Origin Purpose\Trip Destination Purpose	Base	Goods Pickup	Goods Delivery	Service	Maint / Other	Home	Terminal Destination	Marginal
Originate	0.00%	3.53%	1.66%	0.35%	2.59%	0.03%	0.23%	8.39%
Base	0.00%	0.48%	0.09%	0.04%	0.65%	0.00%	0.04%	1.30%
Goods Pickup	0.17%	2.71%	17.46%	0.04%	1.18%	0.22%	0.49%	22.27%
Goods Delivery	0.54%	9.19%	23.52%	0.61%	6.47%	0.60%	3.84%	44.77%
Service	0.08%	0.51%	0.15%	0.45%	0.52%	0.01%	0.29%	2.00%
Maintenance/Other	0.52%	5.33%	1.80%	0.48%	7.77%	0.58%	3.34%	19.82%
Home	0.00%	0.52%	0.09%	0.03%	0.63%	0.00%	0.18%	1.45%
Marginal	1.30%	22.27%	44.77%	2.00%	19.82%	1.45%	8.39%	100.00%

Table 3.8-6 Next Stop Purpose Transitions Differences in Transition Proportions (Model-Survey) TNCs

Trip Origin Purpose\Trip Destination Purpose	Base	Goods Pickup	Goods Delivery	Service	Maint / Other	Home	Terminal Destination	Marginal
Originate	-0.05%	0.33%	-0.03%	0.01%	-0.13%	-0.02%	0.04%	0.15%
Base	0.00%	0.04%	0.02%	0.00%	0.09%	-0.02%	-0.05%	0.07%
Goods Pickup	0.04%	0.19%	0.70%	0.01%	0.14%	0.04%	-0.55%	0.55%
Goods Delivery	0.05%	0.20%	-1.46%	0.06%	-0.03%	0.00%	0.36%	-0.82%
Service	0.00%	0.02%	-0.01%	-0.05%	0.03%	0.00%	0.03%	0.02%
Maintenance/Other	0.03%	-0.23%	-0.03%	-0.02%	-0.04%	-0.02%	0.37%	0.05%
Home	0.00%	0.02%	0.00%	0.01%	-0.01%	0.00%	-0.04%	-0.03%
Marginal	0.07%	0.55%	-0.82%	0.02%	0.05%	-0.03%	0.15%	0.00%

3.9 Next Stop Location Choice

The Next Stop Location Choice model predicts the location type and zone for each stop on a commercial vehicle route. For both regular establishments and TNCs, the Next Stop Location Choice model is formulated as a two-level model and consists of the following sub models:

- **Next Stop Location Choice**
  1. Choice of Location Type
  2. Given Location Type, Choice of Zone

3.9.1 Target Data

The Commercial Vehicle Survey driver diary trip records were expanded to create calibration targets for each stop Location Type and to create a mean trip distance by vehicle type. The following destination types were defined in the travel diary data:

- **Warehouse** – warehouse or distribution center that is not the establishment of the driver and identified in the survey travel diary as being a "warehouse" place type
- **Transport Node** – zones designated as marine ports, airports, or intermodal facilities and not the establishment of the driver. Overrides the warehouse designation.
- **Residential** – zones with households that were identified in the trip records of the travel diary as a "residential" place type or reported as the location of the driver's home (activity type).
- **Other Commercial** – other zones with employment that do not fall into one of the above categories

3.9.2 Fit Statistics

The calibrated shares of stops by Location Type are shown in Table 3.9-1. All differences between the model and survey shares are less than one percent.

Table 3.9-1 Nest Stop Location Type Choice Calibration Shares

Destination Type	Model	Model Share	Survey	Survey Share	Difference
Other Commercial	772,157	49.7%	837,063	50.3%	-0.6%
Residential	519,251	33.4%	554,103	33.3%	0.1%
Base	190,122	12.2%	195,288	11.7%	0.5%
Warehouse	64,662	4.2%	71,130	4.3%	-0.1%
Transport Node	6,382	0.4%	6,426	0.4%	0.0%
Total	1,552,574		1,664,011		

The calibrated mean trip distances for the Next Stop Location Choice model are shown in Table 3.9-2 and Table 3.9-3 for regular establishments and for TNCs, respectively. For regular establishments, the differences are about 2 percent or less. For TNCs, the differences in mean trip lengths are less than one percent.

Table 3.9-2 Average Distance (Miles) by Vehicle Type -- Establishments

Vehicle Type	Model	Survey	% Difference
LCV	7.53	7.47	0.78%
SUT	6.31	6.18	2.11%
MUT	9.57	9.77	-2.06%

Table 3.9-3 Average Distance (Miles) by Vehicle Type -- TNCs

Vehicle Type	Model	Survey	% Difference
LCV	4.23	4.23	-0.14%

3.10 Stop Duration Simulation

The Stop Duration model simulates the dwell time or duration of a commercial vehicle at a stop. As described in Section 2.14, this model is formulated as a fitted Beta distribution. For regular establishments, the models were segmented by stop purpose, vehicle type, and whether the trip to the stop began before or after 12 noon. For TNCs, the model segmentation was by trip purpose and three land use density variables.

The models are most readily calibrated by adjusting the maximum duration parameter, which will change the mean duration, and effectively is a multiplicative constant. Since stop duration affects the total time spent on a commercial vehicle route, along with travel time, it has an impact on the total number of stops that may be accomplished in a day and indirectly has an impact on VMT. For simplicity, it was decided to calibrate stop duration solely on vehicle types.

Commented [BS47]: What exactly needs to be adjusted to calibrate stop durations by vehicle type? New calibration terms were added?

Commented [JG48R47]: The maximum duration is a multiplicative constant that transforms the distribution from being on a 0 to 1 interval to being on an interval between the minimum and maximum values which we set.

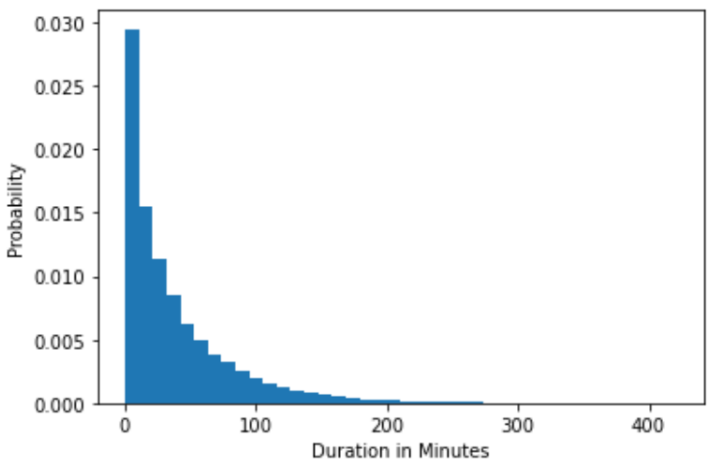
Calibrating the Stop Duration model resulted in a new set of maximum duration parameters. These parameters are shown in Table 3.10-1.

Table 3.10-1 Calibrated Values for Maximum Stop Duration Multiplicative Constants (Minutes)

	Vehicle Types		
Stop Purpose	LCV	SUT	MUT
Base	1043	821	992
Goods delivery	348	274	529
Goods pickup	348	274	529
Home	1043	821	992
Maintenance	1043	821	992
Service	1043	821	992

The shape parameters of the distribution, however, make the likelihood of drawing an extremely large value vanishingly small. For example, shown in is a draw of 50,000 samples from the calibrated simulation model for Base-LCV, which has a mean of 38.2 minutes.

Table 3.10-2 Example Draws from Calibrated Duration Model for Base Purpose, LCV Vehicle Type



3.10.1 Target Data

The Commercial Vehicle Survey travel diary trip records were expanded to create a targets for stop duration by vehicle types.

3.10.2 Fit Statistics

Comparisons of mean stop durations by commercial vehicle types are shown in Table 3.10-3.

Table 3.10-3 Stop Duration Model Calibrated to Mean Dwell Time by Vehicle Type

Vehicle Type	Survey	Model	Difference
LCV	38.5	38.1	-0.9%
SUT	25.7	26.6	3.5%
MUT	40.0	39.5	-1.4%

4.0 HTM Update

4.1 FAF Disaggregation to FAZs

This section briefly explains the Freight Analysis Framework (FAF) data disaggregation for the HTM update to get external freight movements to/from San Diego. It starts with describing FAF data and the basis for disaggregation of FAF freight flows. Subsequently, it explains the steps taken for the FAF disaggregation process.

4.1.1 FAF5 Data

FAF data provides national commodity flows for 43 commodity groups coded with Standard Classification of Transported Goods (SCTG) categories and by 7 modes of transportation, including Truck, Rail, Water, Air (includes truck-air), Multiple modes & mail, Pipeline, Other & unknown. For import and export flows, it includes Port of Entry/Exit (POE). The base-year is 2017 (based on the national commodity flow survey data), and forecasts are provided for the years 2025 to 2050 in five-year increments. Provisional annual data are also available for 2018-2023.

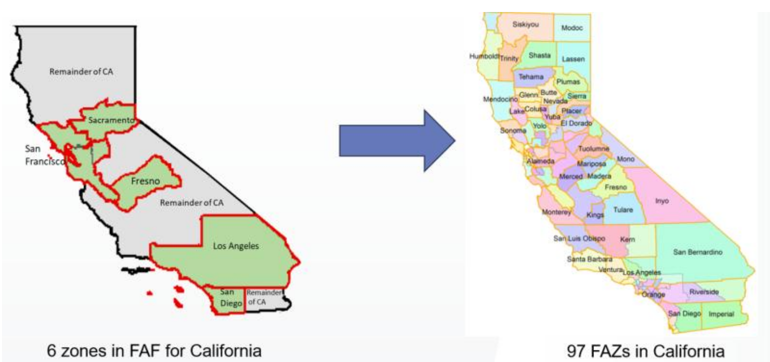
FAF creates a comprehensive picture of freight movement among states and major metropolitan areas by all modes of transportation. FAF integrates data from a variety of sources. Starting with data from the Commodity Flow Survey (CFS) and international trade data from the Census Bureau, FAF incorporates data from agriculture, extraction, utility, construction, service, and other sectors. FAF is produced by the Bureau of Transportation Statistics (BTS) with support from the Federal Highway Administration (FHWA).FAF5.5 (most recent version available) zones within the U.S. (i.e. 132 domestic zones, 6 zones in California, San Diego region as one zone) are too aggregate for estimating regional commodity flows to/from San Diego (see Figure 4.1-1). For example, San Diego County is represented by a single FAF zone, the Los Angeles metropolitan area (including Riverside and San Bernardino County) is represented by a single FAF zone, and the Imperial County is represented as part of the “Remainder of California” FAF zone. Assuming San Diego and Los Angeles FAF zones as origins and destinations (O-D) of commodity flows and assigning truck flows to these zones as single points (zonal centroid) is inadequate for estimating regional movements and results in unrealistic truck travel patterns. Therefore, FAF5 data must be disaggregated for relevant zones near San Diego FAF Zone, using employment data.

Figure 4.1-1 FAF Zones in California

4.1.2 FAF Disaggregation

For the HTM model update, FAF5 flows in California were disaggregated into Freight Analysis Zones (FAZs), movements between FAZs, to/from California (see Figure 4.1-2). This is in line with the California Statewide Freight Forecasting Model (CSFFM) Update approach (currently underway) and SANDAG staff can extract the FAF disaggregation from the statewide model when it is next updated. It also makes the inter-regional movements consistent with the statewide model and better accounts for internal to California movements with more detailed zones. Detailed 3-digit NAICS employment data was used for disaggregation. The disaggregation steps are explained below.

Figure 4.1-2 FAF Zones and Freight Analysis Zones (FAZs) in California



Step 1 Obtain the Most Recent Version of FAF5

FAF5.5 data were downloaded from the FAF website. FAF5 provides estimates for tonnage (unit: thousand tons), value (unit: million dollars), and ton-miles (unit: million ton-miles) by origin-destination pair of FAF zones, commodity type, and mode for the base year (2017), and forecast year estimates (2025-2050).

Step 2. Filter FAF5.5 to Exclude Non-Pass-Through California traffic

FAF is a national database and includes many flows that do not pass through California, (e.g., FAF011-Brimingham, AL to FAF131-Atlanta, GA.). To ensure data relevance for California-specific analyses, the routing table developed for FHWA's Office of Highway Policy Information's (OHPI) Reinventing Conditions and Performance project was used to exclude FAF records that do not utilize corridors in California, thereby improving the accuracy of regional flow estimations.

Step 3. Develop a Crosswalk between FAF Regions and Statewide FAZs

A crosswalk was developed between FAF zones and statewide FAZs. This crosswalk also includes information on whether a FAZ included a Seaport access and whether the county was designated by the Oak Ridge National Laboratories Center of Transportation Analysis (ORNL CTA) as having rail tracks that could be used to provide rail service.

Step 4. Develop Economic Activity Disaggregation Factors for Statewide FAZs

The Bureau of Economic Analysis's (BEA) Make and Use tables were used to link the commodities with the economic sectors producing and consuming them. Claritas employment data was used as the primary employment data at the 3-digit NAICS level while control totals were derived from the American Community Survey (ACS) employment data at the 2-digit NAICS level. Crosswalks were developed to connect the commodities used by the BEA with the Standard Classification of Transported Goods at the 2-digit level (STCG2) classification utilized in FAF. Similarly, crosswalks were developed to align the economic sectors used by the BEA with the North American Industrial Classification System at the 3-digit level (NAICS3). Equations were developed for each STCG2 classification for each origin (make), and destination (use), to reflect the economic sectors contributing more than 5 percent to a commodity's value. Typically, the economic sector making or using an economic sector was a NAICS3 industry. Exceptions included personal consumption, assumed to be correlated with population, and business consumption, assumed to correlate with total employment. The disaggregation factors were developed as relative shares to ensure that the distribution of commodities aligns closely with the underlying economic activities and sectoral contributions.

Step 5 Apply Disaggregation Factors to Filtered FAF5.5

Apply the disaggregation factors developed in Step 4 to the FAF filtered to include only California II, IE, EI, and EE pass through trips developed in Step 2. The consultant team used in-house proprietary scripts for this step.

Step 6 Review and Process Validation of Disaggregated Flows

Various summaries were reviewed, and process validation and cross-checks were performed to ensure the disaggregation process functioned correctly without introducing errors. This review ensured that no tons were added or lost, and that totals were maintained accurately by mode and commodity.

4.1.3 Employment Data and Industries

The previous version of HTM did not account for the most appropriate industries that produce and consume the various commodities in the current FAF data. For example, Manufacturing was too broad to be effectively used with specific commodities. For instance, STCG 02 (Grain) is produced by the agricultural industry, which falls under NAICS code 11, but it is used by both NAICS 11 and NAICS 311 (Food Manufacturing).

For the FAF disaggregation update, detailed 3-digit industry codes (NAICS3) data from Claritas were utilized, as described earlier in this document. By using the 3-digit industry code details, the relationships between industries and commodities were updated and refined accordingly.

4.1.4 Regions Disaggregated

The HTM produces a more accurate assignment and avoids “lumpy loading”, by disaggregating its main input, the FAF data. However, disaggregating FAF for all counties in the U.S. is unnecessary and adds to the model’s computational complexity. For instance, while disaggregating regions to counties improves assignment accuracy on I-95 for commodity movements from New York to Florida, it offers no benefit for San Diego. For this update, only FAF zones with flows to, from, or through San Diego FAF Zone have been disaggregated to counties. This approach reduces the model’s computational details and improves run time without losing any detail related to freight movements associated with the San Diego FAF Zone.

The primary purpose of disaggregating FAF for the HTM is to accurately identify freight flows crossing the SANDAG model boundary. Smaller commodity flows were intentionally neglected for simplicity, as exchanges of less than 10 kilotons per year translate to only about 500 trucks annually, or fewer than 2 trucks per day). When utilized for daily assignment, this minimal flow should not affect the entry or exit points for the SANDAG model, particularly after the annual to daily conversion is done. Consequently, only a subset of the FAF data was disaggregated to maintain focus on the most relevant flows.

Most IE/EI tonnages to or from San Diego FAF Zone are attributed to adjacent FAF zones, primarily consisting of regions in California and neighboring states. It is important to note that import and exports transported by domestic truck through the Port of San Diego, San Diego International Airport, or the Mexican border crossings at Tecate or Otay Mesa are recorded with a San Diego domestic trip end in the FAF.

Table 4.1-1presents the top 10 regions that are IE or EI with respect to San Diego and have a minimum of 20 kilotons per year (highlighted in red) transported by domestic truck for domestic, imports, or exports with San Diego as one end of the trip. It is noted that more than 96% of the IE/EI truck tonnage to or from San Diego is in the top 10 regions. It is noteworthy that none of the tons exported from San Diego, predominantly by water, are delivered by truck; instead, these exports primarily arrive by rail. The tons listed in Table 1 are sorted by total IE/EI tons with one end in San Diego. Only these regions are likely to impact the assignment of flows for San Diego when disaggregated to counties. As anticipated, the major trading partner in FAF5 is the Los Angeles region.

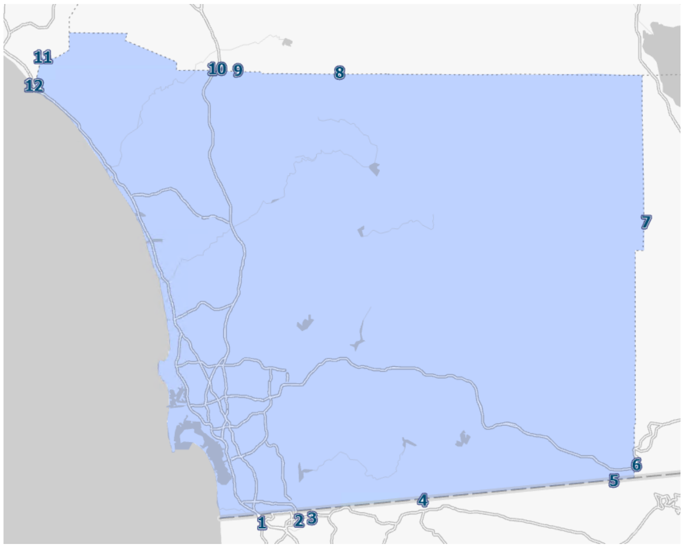
Table 4.1-1 Top 10 FAF Regions IE/EI Flows with Respect to San Diego FAF Zone (2017 kilotons)

Rank	FAF Region	Name	Dms O	Dms D	Import D	Import O	Export D	Export O	Total IE/EI
1	061	Los Angeles CA	8,669.32	11,941.68	3,254.09	2,366.45	#N/A	1,976.76	28,208.30
2	064	San Francisco CA	276.21	378.79	986.32	2,091.38	#N/A	1,350.67	5,083.37
3	069	Rest of CA	1,174.10	899.44	0.12	288.92	#N/A	9.84	2,372.42
4	489	Rest of TX	34.23	1,250.78	48.68	41.76	#N/A	7.82	1,383.27
5	041	Phoenix AZ	118.69	564.37	0	235.35	#N/A	#N/A	918.41
6	062	Sacramento CA	45.85	260.87	#N/A	251.2	#N/A	#N/A	557.92
7	486	Houston TX	37.9	282.44	56.14	88.6	#N/A	45.7	510.78
8	261	Detroit MI	6.74	21.96	70.07	179.88	#N/A	170.74	449.39
9	065	Fresno CA	142.68	192.48	#N/A	87.15	#N/A	#N/A	422.31
10	131	Atlanta GA	31.58	49.04	0	238.5	#N/A	0	319.12

4.1.5 National Flows passing through San Diego

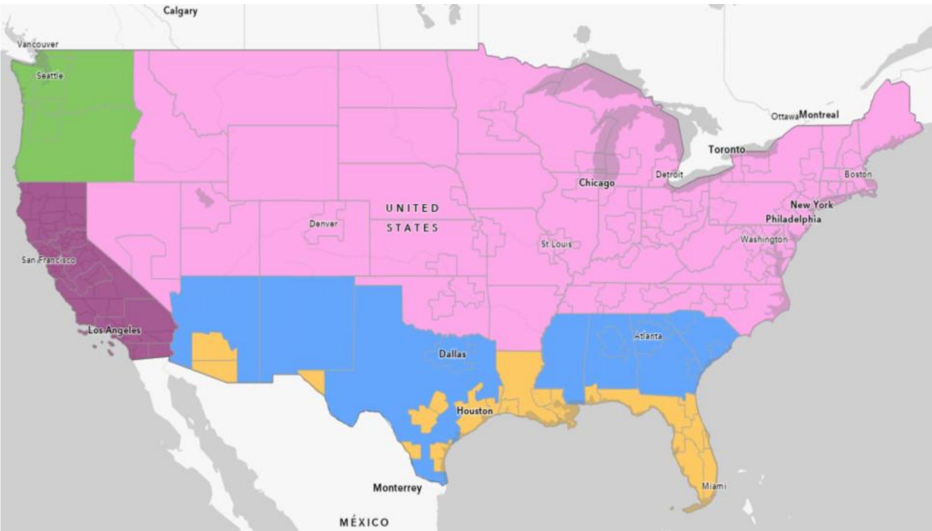
SANDAG CVM/HTM has 12 gateways, although some of them experience very low truck traffic. The primary corridors connecting San Diego County to the rest of the nation include the gateways at I-5 (#12), I-15 (#10), I-8 (#6) and the border crossing at the Mexico-US border. See Figure 4.1-3.

Figure 4.1-3 SANDAG CVM/HTM external Gateways



San Diego County is bordered by the Pacific Ocean to the west and Mexico to the south, which limits the routes from other parts of the U.S. into the San Diego region. The National FAF regions are grouped together, as shown in Figure 4.1-4, based on the potential paths that may cross through San Diego County. This streamlined approach eliminates the need to extract data from the entire national FAF flow using sub-area extraction models, allowing SANDAG to use data directly from the California statewide model instead of relying on a detailed disaggregation of national FAF data. In Figure 4.1-4, truck trips between regions of the same color and San Diego will use the same gateway to pass through San Diego County.

Figure 4.1-4 National FAF Regions Aggregation



To identify the truck routes between each region and San Diego County, Google Maps was to verify the shortest path. Figure 4.1-5 and Figure 4.1-6 show the examples for routing via I-10 and I-8. These figures provide a visual representation of the most efficient routes for truck traffic connecting these regions to San Diego County. Using Google Maps is reasonable for identifying efficient routes for long-haul trucks, primarily because these trucks typically travel on major highways and freeways to pass-through the San Diego region.

For example, flows from the Green region including Washington and Oregon are all allocated to I-5. Flows from the Orange region including most of southern FAF regions are allocated 85% to I-8 and 15% to SR-78. The split percentages are approximately estimated based on truck traffic flows in California Statewide model and calibrated to match the traffic counts.

Figure 4.1-5 Example of Regional Routs, I-10 Corridor

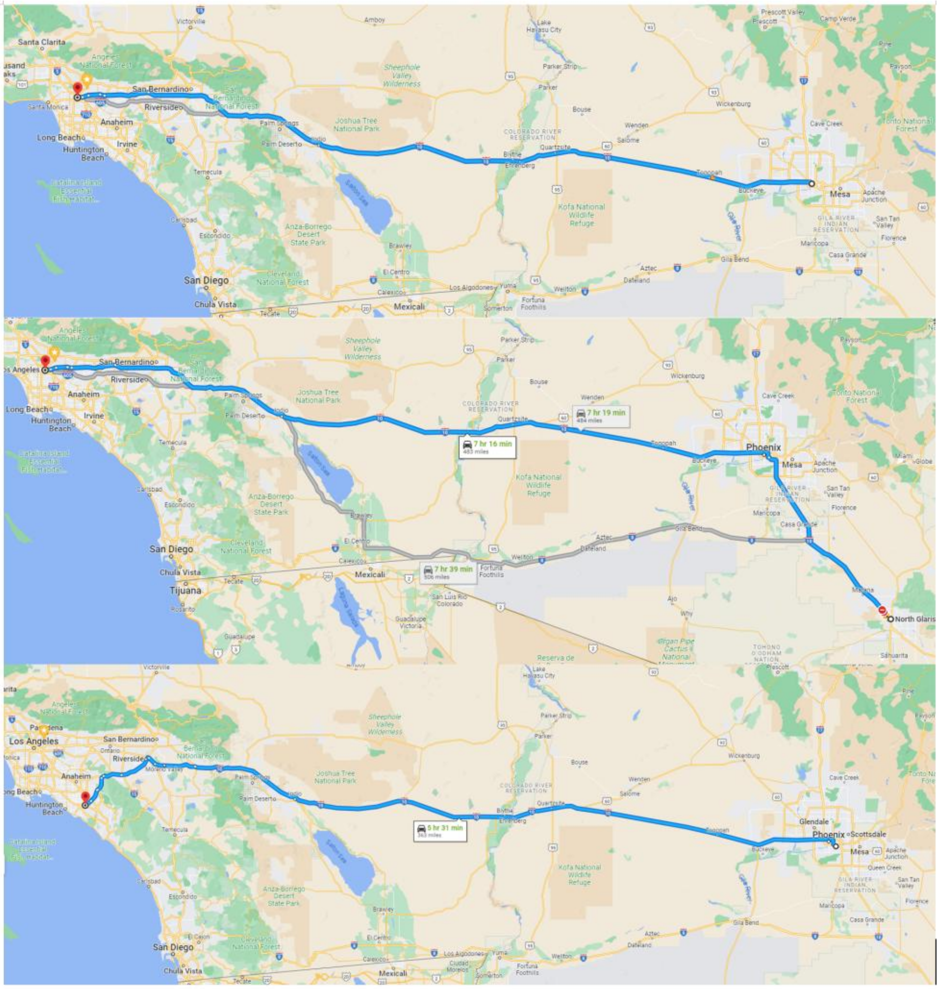
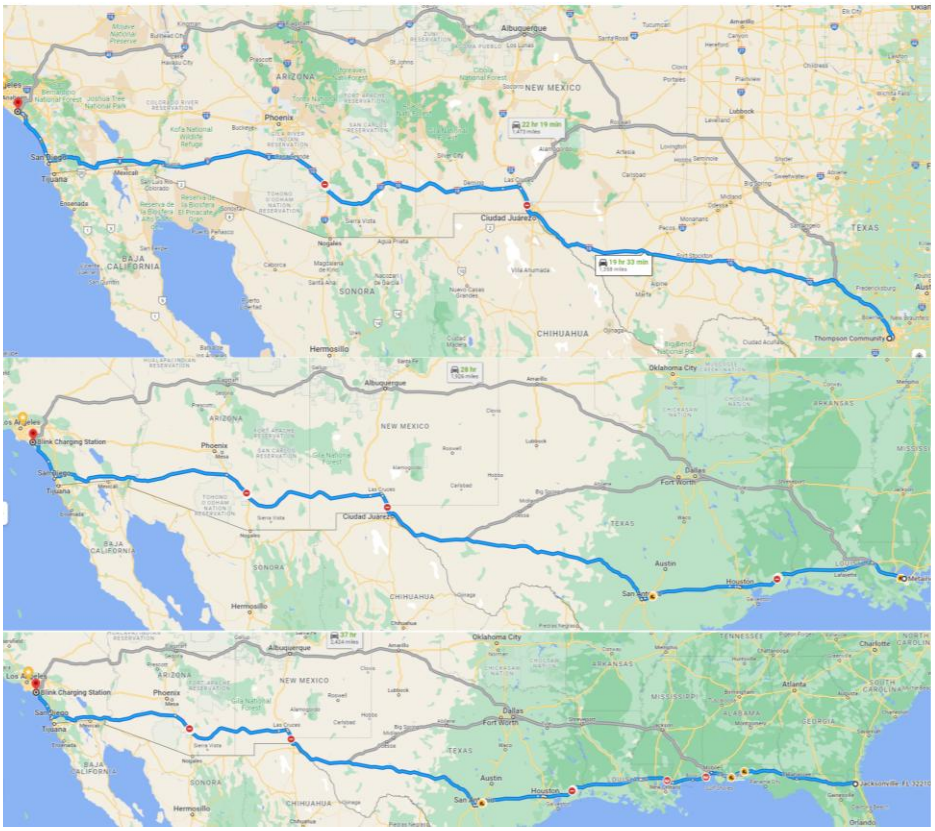


Figure 4.1-6 Example of Regional Routs, I-8 Corridor



A similar process was developed to estimate pass-through flows in San Diego County. Table 4.1-2 shows an aggregated OD (origin-destination) table corresponding to the regions in Figure 4.1-4. The gray cells represent routes that do not pass through San Diego County. For example, truck routes from Oregon (green region) to New York (pink region) do not cross San Diego County. However, for routes from Orange County, California, to Houston, Texas (orange region), there is a small possibility of using I-8 as a potential route. The percentages are defined based on a high-level review of National FAF assignments and calibrated to truck counts.

Table 4.1-2. Distribution of Regional Flows to each SANDAG Gateway

Area Code	Area	Green	Pink	Blue	Yellow	CA North of OC	Orange County	Imperial/ Calexico/ Calexico east/ Andrade	SD/Otay Mesa/ Tecate/ Port of SD/ SD Airport
1	Green	-	-	-	-	-	-	0.5	1
2	Pink	-	-	-	-	-	-	-	1
3	Blue	-	-	-	-	-	-	-	1
4	Yellow	-	-	-	-	-	0.05	-	1
5	CA North of OC	-	-	-	-	-	-	-	1
6	Orange County	-	-	-	0.05	-	-	0.5	1
7	Imperial/ Calexico/ Calexico east/ Andrade	0.5	-	-	-	-	0.5	-	1
8	SD/Otay Mesa/ Tecate/ Port of SD/ SD Airport	1	1	1	1	1	1	1	1

4.1.6 Reference Tables

This section presents the reference tables that are used to develop HTM.

Commodity Groups

Table 4.1-3 shows how the 43 commodities in FAF data are aggregated to 15 groups. These groups are consistent with California Statewide Freight Model commodity groups.

Table 4.1-3. Commodity Groups

SCTG	SCTG Description	CG	CG Description
1	Live Animals and Fish	CG-1	Agriculture products
2	Cereal Grains (including seed)	CG-1	Agriculture products
3	Other Agricultural Products, except for Animal Feed	CG-1	Agriculture products
4	Animal Feed and Products of Animal Origin, not elsewhere classified (n.e.c.)	CG-1	Agriculture products
5	Meat, Fish, and Seafood, and Their Preparations	CG-7	Food, beverage, tobacco products
6	Milled Grain Products and Preparations, and Bakery Products	CG-7	Food, beverage, tobacco products
7	Other Prepared Foodstuffs, and Fats and Oils	CG-7	Food, beverage, tobacco products
8	Alcoholic Beverages	CG-7	Food, beverage, tobacco products
9	Tobacco Products	CG-7	Food, beverage, tobacco products
10	Monumental or Building Stone	CG-5	Gravel/ sand and non-metallic minerals
11	Natural Sands	CG-5	Gravel/ sand and non-metallic minerals
12	Gravel and Crushed Stone	CG-5	Gravel/ sand and non-metallic minerals
13	Non-Metallic Minerals, n.e.c.	CG-5	Gravel/ sand and non-metallic minerals
14	Metallic Ores and Concentrates	CG-6	Coal / metallic minerals
15	Coal n.e.c.	CG-6	Coal / metallic minerals
16	Crude Petroleum Oil	CG-3	Crude petroleum
17	Gasoline and Aviation Turbine Fuel	CG-4	Fuel and oil products
18	Fuel Oils	CG-4	Fuel and oil products
19	Coal and Petroleum Products, n.e.c.	CG-4	Fuel and oil products
20	Basic Chemicals	CG-9	Chemical/ pharmaceutical products
21	Pharmaceutical Products	CG-9	Chemical/ pharmaceutical products
22	Fertilizers	CG-9	Chemical/ pharmaceutical products
23	Chemical Products and Preparations, n.e.c.	CG-9	Chemical/ pharmaceutical products
24	Plastics and Rubber	CG-8	Manufactured products
25	Logs and Other Wood in the Rough	CG-15	Logs
26	Wood Products	CG-2	Wood, printed products
27	Pulp, Newsprint, Paper, and Paperboard	CG-2	Wood, printed products
28	Paper or Paperboard Articles	CG-2	Wood, printed products
29	Printed Products	CG-2	Wood, printed products
30	Textiles, Leather, and Articles of Textiles or Leather	CG-8	Manufactured products
31	Non-Metallic Mineral Products	CG-10	Nonmetal mineral products

SCTG	SCTG Description	CG	CG Description
32	Base Metal in Primary or Semi-Finished Forms and in Finished Basic Shapes	CG-11	Metal manufactured products
33	Articles of Base Metal	CG-11	Metal manufactured products
34	Machinery	CG-11	Metal manufactured products
35	Electronic and Other Electrical Equipment and Components, and Office Equipment	CG-13	Electronics
36	Motorized and Other Vehicles (including parts)	CG-14	Transportation equipment
37	Transportation Equipment, n.e.c.	CG-14	Transportation equipment
38	Precision Instruments and Apparatus	CG-13	Electronics
39	Furniture, Mattresses and Mattress Supports, Lamps, Lighting Fittings, and illuminated Signs	CG-8	Manufactured products
40	Miscellaneous Manufactured Products	CG-8	Manufactured products
41	Waste and Scrap	CG-12	Waste material
43	Mixed Freight	CG-8	Manufactured products

Table 4.1-4. Payloads by Truck Type (Pounds)

Commodity Group	Light Truck	Medium1 Truck	Medium2 Truck	Heavy Truck
CG-1	3320	5902	14618	40653
CG-2	1585	5569	9757	38711
CG-4	1189	4355	10066	36822
CG-5	2235	5679	18394	46112
CG7	2479	5236	12266	38500
CG-8	2361	4711	9354	34209
CG-9	2364	3609	12933	37348
CG-10	2235	4834	18394	46648
CG-11	2822	5489	12980	36338
CG-12	1961	4552	10503	35360
CG-13	995	3334	11126	20792
CG-14	2705	4600	10803	37836
CG-15	2364	5788	12393	41333

Table 4.1-5. Truck Allocation Distribution by Distance Category

Distance Range	Light	Medium1	Medium2	Heavy
less than 50 miles	1.0%	9.0%	8.8%	81.2%
50 to 100 miles	0.5%	5.0%	2.6%	91.9%
100 to 150 miles or One end in Orange County	0.6%	3.0%	2.8%	93.6%
201 miles or more	0.3%	2.0%	2.2%	95.5%
One end in Mexico	0.0%	0.0%	0.0%	100.0%

Table 4.1-6. Time of Day Distribution

Peak Period	Share
EV	0.29
AM	0.13
MD	0.25
PM	0.17
EA	0.16