FREIGHT AND COMMERCIAL VEHICLE MODEL DEVELOPMENT

MODEL DESIGN

RSG
the science of insight

TASK 3
11.5.2015

PREPARED FOR:
BALTIMORE METROPOLITAN COUNCIL (BMC), MARYLAND STATE HIGHWAY ADMINISTRATION (SHA)

SUBMITTED BY:
RSG

55 Railroad Row
White River Junction, VT 05001
802.295.4999
www.rsginc.com

IN COOPERATION WITH:
UNIVERSITY OF MARYLAND—THE NATIONAL CENTER FOR SMART GROWTH
VISION ENGINEERING & PLANNING

CONTENTS

1.0 INTRODUCTION

This model design memo is a partial deliverable for Task 3 of the Freight and Commercial Vehicle Model Development project. The memo describes the proposed freight and commercial vehicle model ("the model") design developed by the RSG team in collaboration with BMC and SHA staff.

The memo documents the planned design, specification, and performance of the model, and also results from the development of some elements of the model design, for example the estimation results of the commercial vehicle models, which the RSG team has completed estimating. Aspects of the model design are expected to be revisited by the RSG team, BMC, and SHA, in conjunction with other parallel efforts to implement and update other aspects of the model systems into which the freight models are being integrated.

The following sections of the memo cover:

- The purpose of the SHRP 2 C20 project to develop a freight and commercial vehicle model
- Use cases for the model
- The system architecture which describes the high level model entities, how they interface with each other, and the software architecture of the model
- The model design of the national supply chain model and the details of its components
- The model design of the freight truck touring model and the details of its components
- The model design of the commercial services touring model, including the details of its components and the results of the model estimation work that developed the model

This memo is expected to form a core element of the model documentation that will be delivered at the end of the project, and as such will be updated and added to as the model design described here is implemented and refined during the model development, model calibration and validation, and sensitivity testing tasks.

2.0 PURPOSE OF THE FREIGHT AND COMMERCIAL VEHICLE MODEL

2.1 | BMC AND SHA NEEDS TO BE ADDRESSED THROUGH THIS PROJECT

The agencies' application for the SHRP 2 C20 grant[1] outlines their goals for the project and their planning needs that the model is intended to support. The following list summarizes several key aspects of the application that have informed the model design described in the memo:

- Agency needs for information on transportation system performance with particular emphasis on the role of freight movement and its role in the economic growth in region, state, and nation.
- The complexity of the regional transportation system, with Maryland's position in the densely population eastern seaboard and the presence of the congested Baltimore-Washington metropolitan region, means that it is challenging to model freight movement.
- Current tools in place at the agencies have limitations, for example in their ability to model long distance truck movements, empty truck movements, local freight distribution, and the impacts of port expansions and improvements to intermodal facilities.
- Trucks contribute a significant portion of the transportation systems emission, including greenhouse gases, and the agencies require tools that capture the effects of transportation system performance on truck travel.
- The impacts of development that supports freight movements on communities can be significant, and the model outputs will be used to support analysis of these impacts.
- The project, including its data development and model development efforts, represent an important step to reaching the overall modeling goals of the agencies, but are taking place in parallel with other improvement efforts and so must successfully integrate with and support those efforts, and must also results in a roadmap for future improvements to take place after the current project.

2.2 | HOW THE MODEL WILL ADDRESS THESE NEEDS

2.2.1 BEHAVIORAL IMPROVEMENTS

The proposed model is a multi-layer model where freight flows and the resulting vehicles movements at the local level are informed by and sensitive to changes at the upper level. For example, changes over time in the freight flows to and from a region influence the demand for long distance truck travel, but also the need for additional local truck movements to facilitate local deliveries and pickups of shipments.

The model represents the complete supply chain of good movements that involve Maryland. The model includes global suppliers and buyers who produce imports and consume exports, as well as business establishments across the country. That structure supports scenario testing to understand the impacts of changes in the economy over time, including different patterns of long haul domestic flows and imports and exports as the domestic and international trading partners of Maryland businesses change.

[1] MD_SHRP2_C20_Application_Final.PDF, Maryland SHA (2014)

The model also moves from a traditional trip-based approach to a more realistic tour-based approach, where the travel patterns of trucks are modeled to mimic the routings that dispatchers and drivers develop. These more realistic travel patterns are inherently more responsive to the effects of congestion in the transportation system.

2.2.2 MARKET SEGMENTATION SCOPE

The model represents the two major market segments in terms of demand for truck travel: freight movement and non-freight commercial vehicle movement to provide services. This complete treatment of the demand for truck travel means that the drivers of demand can be explicitly represented in the model and the effects of differential changes in those drivers understood. For example, freight movement is a function aspects of the regional economy such as manufacturing and other industry, as well as consumption by the regional population, while changes in many segments of the service sector that use commercial vehicles are more closely aligned with regional population growth.

Within the two major market segments, there is significant detail at business establishment and commodity level within the freight movement model and at the business establishment level within the non-freight commercial vehicle movement model to allow for different behaviors to be represented adequately.

2.2.3 GEOGRAPHIC SCOPE

The supply chain model, which will be integrated with the Maryland Statewide Model (MSTM), has a global geographic scope. At its broadest, it represents import and export movements to and from eight international zones. Domestic movements are represented to and from the rest of the Unites States, and the model design includes simplified transportation networks that cover the entire continental U.S. The model uses the MSTM TAZ system which is more spatially detailed in portions of the states surrounding Maryland, and then again more detailed within the state of Maryland. The national transportation networks are connected to more detailed transportation networks covering Maryland.

The regional freight and commercial vehicle touring models cover the BMC modeling region and will be integrated with the BMC's InSITE model. They will use the same TAZ and networks as the InSITE model.

2.2.4 TEMPORAL SCOPE AND GRANULARITY

The supply chain model covers a one year time frame, representing annual commodity flow movements from business to business. Within that annual time frame, commodity flows are disaggregated to individual shipments to support the development of a daily sample of shipment movements that can be simulated in the truck touring model.

The truck touring models construct tours in a such a way that they produce a trip list with trip timings reported on a minute by minute basis to flexibly support more or less detailed assignment approaches. The tour sequencing algorithms are designed to take into account multiple time periods for travel impedance skims from the highway networks, which is important as plans are made to integrate the regional model with a DTA model with many time periods.

3.0 MODEL USERS, USE CASES, AND MANAGEMENT

This chapter of the memo discusses expected stakeholder and model users at BMC, SHA, and other agencies, use cases for the models, and logistical considerations for the models such as where they will be housed. This chapter summarizes written responses from BMC and SHA staff to questions developed by RSG in March, 2015, follow up discussions at an in person meeting held at BMC, and is also informed by the workshop that took place in November 2014 and other conversations that have been take place to date during the project. The complete set of questions of responses is included as Appendix A for reference, which the November 2014 workshop is documented in the Task 1.2 memo, "Freight Demand Modeling and Data Improvement Wotkshop".

3.1 | LIST OF ACTORS/USERS AT BMC, SHA, AND OTHER AGENCIES

3.1.1 MODEL USERS AND ROLES

At BMC, agency technical staff will be the users of the freight modeling system who are responsible for model maintenance and updates, model applications, analysis of model results, and documenting model output and findings. Technical work is also expected to be performed by modeling consultants under contract to state or local agencies.

At SHA, SHA-Travel Forecasting staff, university researchers and consultants working for SHA and MDOT will be the users of this freight modeling system. The SHA-Travel Forecasting unit will use the freight models for developing and updating the MD Statewide Transportation Model and other sub-area models as needed. University and/or consultant teams may support model updates. Most day to day model use by SHA Travel Forecasting staff is likely to focus on application of the freight models to support agency planning needs.

3.1.2 MODEL OUTPUT CONSUMERS AND PERFORMANCE MEASURES

At BMC, consumers of the output will be members of the BRTB and related committees. The freight model output will be used in regional and corridor transportation decision making by those committees. At the regional level, BMC staff uses model simulation to support long-range plan development and federal transportation conformity determination. Model outputs are used in regional transportation system performance measures of mobility, accessibility, equity, and environmental impacts presented to committees.

At the corridor level performance measures of link level of service, travel speeds and cost are used at the project planning level. Sensitivity of these performance measures to existing and emerging transportation planning scenarios such as additional travel lanes, new interchanges, truck prohibitions, managed/electronic toll lanes and freight generation scenarios such as port expansion or new intermodal transfer facilities are important.

At SHA, consumers of the model outputs will include several groups involved in different aspects of the agency's planning work:

- Engineers and planners working on corridor and project planning studies

- SHA/MDOT environmental teams and MDE staff interested in understanding, freight-specific emissions impacts
- SHA/MDOT/DBED policy analysts looking at the regional economic impacts of freight movement and impacts of freight infrastructure investments on local/regional and state economy
- SHA/MDOT staff involved in performance reporting (MAP-21, Business Plan, MFR)
- MDOT and SHA Freight Offices trying to understand policy impacts e.g restricting peak hour truck operations on certain facilities, tolling scenarios for trucks etc.
- Other stakeholders involved in various levels of decision making who might receive processed results in printed or electronic format.

3.2 | LIST OF USE CASES

3.2.1 LONG-RANGE PLANNING

At the MPO level, a primary use of the regional travel model, including the freight model, is long-range plan project prioritization. This process includes the use of freight performance measures such as highway link simulated truck percentages, and freight mobility and accessibility measures. An additional aspect of long range planning is freight scenario planning to test transportation capacity and policies (lane restrictions, tolling and managed lanes), economic location and concentration scenarios, and intermodal terminal changes.

A key application from a state perspective will be to understand short and long-term freight infrastructure needs and the effect of investments under various scenarios. From a highway design standpoint, SHA will use outputs of the model to develop truck traffic projections by vehicle classes (light/medium/heavy) for corridor segments. The vehicle class based projections will be used to develop ESAL forecasts for pavement design needs. SHA is also interested in using the model outputs to inform operational strategies in a near term (5-10 year planning). Understanding freight markets, flows, diversion potential (mode/route/time-of-day) should help SHA implement policies, select projects/alternatives that balance person and freight movement, and help the state's economic competitiveness. Two other long term planning needs that the model might support are planning for truck parking in the future, e.g., rest area along the state's highways, and overweight truck planning

3.2.2 CORRIDOR STUDIES

From a corridor planning perspective, the model will support understanding freight activity on specific corridors – such as commodity type and value – and how that might change under different corridor plan scenarios or investment alternatives.

3.2.3 CONFORMITY ANALYSIS

The freight modeling system will support the preparation of EPA MOVES software's transportation related assumptions (VMT by source type, road type, hour, and travel speed). Freight travel, especially diesel vehicles, is a major source of NOx mobile source emissions.

3.2.4 REGIONAL ECONOMIC IMPACTS

Regional assessments of the economic impacts of transportation investments, with support from the freight model to understand impacts and relationships between economic activity, freight movement, and the transportation systems. Performance measures such as the economic activity in the region supported by the freight movement system can potentially be developed.

3.2.5 OTHER USES

In addition to the agencies' planning needs, the results from models are used to support the development of publications, such as annual reports that include freight facts, figures, and trends.

3.3 | INTER-AGENCY LOGISTICS AND OPERATIONAL CONSIDERATIONS

3.3.1 CUSTODIANSHIP

BMC will be responsible for the maintenance, update and versioning of the urban freight and commercial vehicle model implemented within the InSITE modeling framework, while SHA will be responsible for maintenance, update and versioning of the statewide supply chain model implemented within the MSTM. It is expected that integration between the two models will be via an extraction of simulated shipments from the supply chain model within the MSTM long distance freight for input to the urban freight and commercial vehicle model within the Baltimore region InSITE modeling framework.

BMC plans to investigate the application of the supply chain model in supporting Baltimore region economic planning studies. Through this use, BMC's intension is to coordinate supply chain model maintenance, update and versioning with the SHA model framework management.

3.3.2 PROJECT COLLABORATION

BMC and SHA plan to continue their existing collaboration on projects using the model system. Both agencies have collaborated on projects, sharing modeling setups and inputs. In addition, collaboration occurs in the application of the model with the simulation of non-capacity enhancement policies such as pricing and managed lanes.

The SHA is engaged in a wide variety of different scales of projects including intersection/interchange design, corridor alternative analysis, and special projects such as transit oriented development, toll, and managed lane analysis requiring travel forecasting and analysis. With the implementation of this freight model and other cooperative state/regional model development projects, an evolving and expanding corridor and special project collaboration is envisioned by the agencies. The freight modeling system will provide both BMC and SHA additional collaboration opportunities with MDOT and other modal (water and airport) operators in developing statewide freight plans and port related freight scenario analysis.

Protocols for database or model component sharing between agencies are in place. SHA relies on BMC to provide the latest endorsed locally developed demographic data and the MPO policy board approved long-range plan for initial base and horizon year alternatives.

3.3.3 SOFTWARE HOSTING, DATA WAREHOUSING, AND SHARING PROTOCOL

The official urban modeling framework (activity based passenger model and urban freight commercial vehicle model) will reside on the BMC network. Model input data will be developed and maintained in spreadsheets, databases, and GIS layers and transferred to a compatible file format prior to model execution. BMC staff anticipate using a file structure similar to the current trip based method (Figure 1). Upon request, for state and local agencies, BMC staff distributes the current validated model using the trip based file structure. Recipients only need to copy the file structure to their hard drive and call the job script from within Cube. Cube's application manager is used to enter replacement variables such as horizon year.

FIGURE 1: FILE STRUCTURE FOR SCENARIO MANAGEMENT

Scenario Name (TIP/Conformity)

- **Model** (job code and model variables identical for all horizon years)
- **Horizon year** (2010)
  - **Inputs** (horizon year specific assumptions – transportation networks, demographics,…)
- **Horizon year** (2020)
  - **Inputs**

The official version of the MSTM resides at SHA-Office of Planning and Preliminary Engineering. The current version 1.0.XX uses a base year of 2007 and forecast year of 2030. SHA will be adopting a version 1.1 in the summer of 2015 with a base year of 2012 and a forecast year of 2040. The zone system of the MSTM is developed from the MPO TAZs and census tracts and the socio-economic data and networks are developed from the MPO inputs.

4.0 HIGH-LEVEL MODEL ENTITIES (SYSTEM ARCHITECTURE)

4.1 | MODEL FRAMEWORK

This chapter of the memo introduces the model's system architecture and describes its high-level model entities – the national supply chain model (NSCM) and the two components of the urban truck-touring model, the freight truck-touring model (FTTM) for freight and the commercial services touring model (CSTM) for non-freight carrying trucks – and how they interface with each other. For each of these main model entities, the memo describes the analytical engine, the input and output databases, and the integration of the models into the larger modeling systems that they will form part of, specifically, SHA's MSTM model and BMC's InSITE model.

Figure 2 shows the overall model system used by SHA and BMC. Each model systems contains passenger travel demand models (gray boxes) that are used to estimate personal travel by auto and other modes. The freight and commercial vehicle travel demand models being developed in this SHRP 2 C20 project are shown in green.

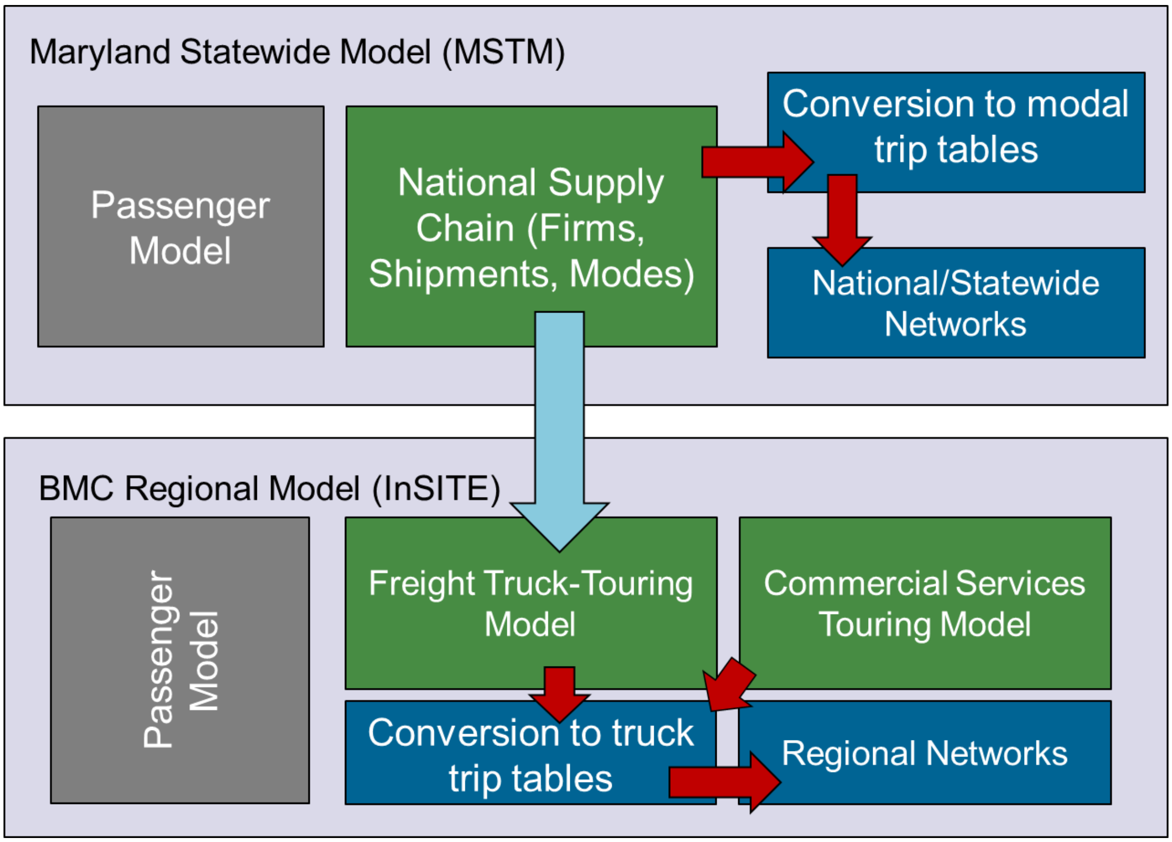
The **national supply chain model** (NSCM) simulates the transport of freight between supplier and buyer businesses in the United States, in this case focusing on movements that involve Maryland. This model component will be integrated within the MSTM. Its output, a list of commodity shipments by mode, is used in two ways. First, in the MSTM, a model component connected to the NSCM converts the annual shipment flows to daily vehicle trip tables (first blue box) that can be assigned to the national and statewide networks (second blue box) in the MSTM along with trips tables from the passenger model. Secondly, as indicated by the blue arrow, the list of commodity shipments by mode is extracted from the MSTM and used an input to the FTTM in BMC's activity-based passenger model, InSITE.

The **freight truck-touring model** (FTTM) simulates truck movements within the Baltimore region that deliver and pick up freight shipments at business establishments. This model component will be integrated within the InSITE model. The model is a tour-based model, and builds a set of truck tours including transfer points at which the shipment is handled before delivery/after pickup for shipments with a more complex supply chain (i.e., a warehouse, distribution center, or consolidation center) and the suppliers and buyer of shipments where those are within the model region. The shipment list from the NSCM is used as the demand input for the FTTM and describes the magnitude and location of delivery and pick up activity in the region that must be connected by truck movements. The model will generate trip lists by commercial vehicle type and time of day so that the outputs from this model can be combined with the outputs from the CSTM and appropriate trip tables from InSITE for highway assignment.

The **commercial services touring model** (CSTM) simulates the remainder of the travel of light, medium, and heavy trucks that is for commercial purposes, i.e., providing services and goods delivery to households and services to businesses. This model component will also be integrated within the InSITE model. As with the FTTM, the CSTM is a tour-based model, but this time demand is derived from the characteristics of the business establishments and households in the region and as such is not affected by the national supply chain model. That is, while the FTTM simulates truck tours based on commodity flows, the CSTM generates and simulates truck and light-duty vehicle movements based on demand for services and goods from certain industries.

In a similar way to the approach used in the MSTM, the outputs from the both the FTTM and the CSTM are converted to trip tables. In this case, a trip list from each model with trip start and end location and trip timing information is aggregated into zone to zone trips by time period (first blue box) that can be assigned to the regional highway networks (second blue box) in the InSITE model along with trips tables from the passenger model.

FIGURE 2: MODEL SYSTEM – TWO-AGENCY IMPLEMENTATION



4.2 | NATIONAL SUPPLY CHAIN MODEL

The NSCM is a simulation of a multimodal supply-chain, connecting the state with the rest of the United States and the world. The most recent version of the NSCM is to be transferred from the Florida DOT statewide freight model to Maryland for this project. The NSCM was originally developed in RSG's FHWA funded work in Chicago and has been extended and refined for a statewide modeling context in Florida.

The NSCM includes components that synthesize business locations, trading relationships between businesses, and the resulting commodity flows, distribution channel, shipment size, and mode and path choices for each shipment made annually. A flow chart of the NSCM is shown in Figure 3. The components shown in the flow chart perform the following model steps:

Firm Synthesis

1. **Business Establishment Synthesis:** Synthesizes all business establishments in the United States and a sample of international business establishments using input employment data, which is spatially detailed for the modeled region and less so for the rest of the United States, and growth factors for future year business synthesis.
2. **Annual Production** by business establishments is characterized based on its industrial classification
3. **Annual Consumption** by business establishments is characterized using relationships described in the input output data that describe the input commodities required to produce commodities.

Commodity Flows

4. **Buyer Seller Matching:** Selects supplier firms for each buyer firm by type, choosing one for each commodity input required by the buyer firm.
5. **Commodity Flow Allocation:** Predicts the annual demand in tonnage for shipments of each commodity type between each firm in the United States, using input commodity flow forecasts.

Transport/Logistics Chain Models

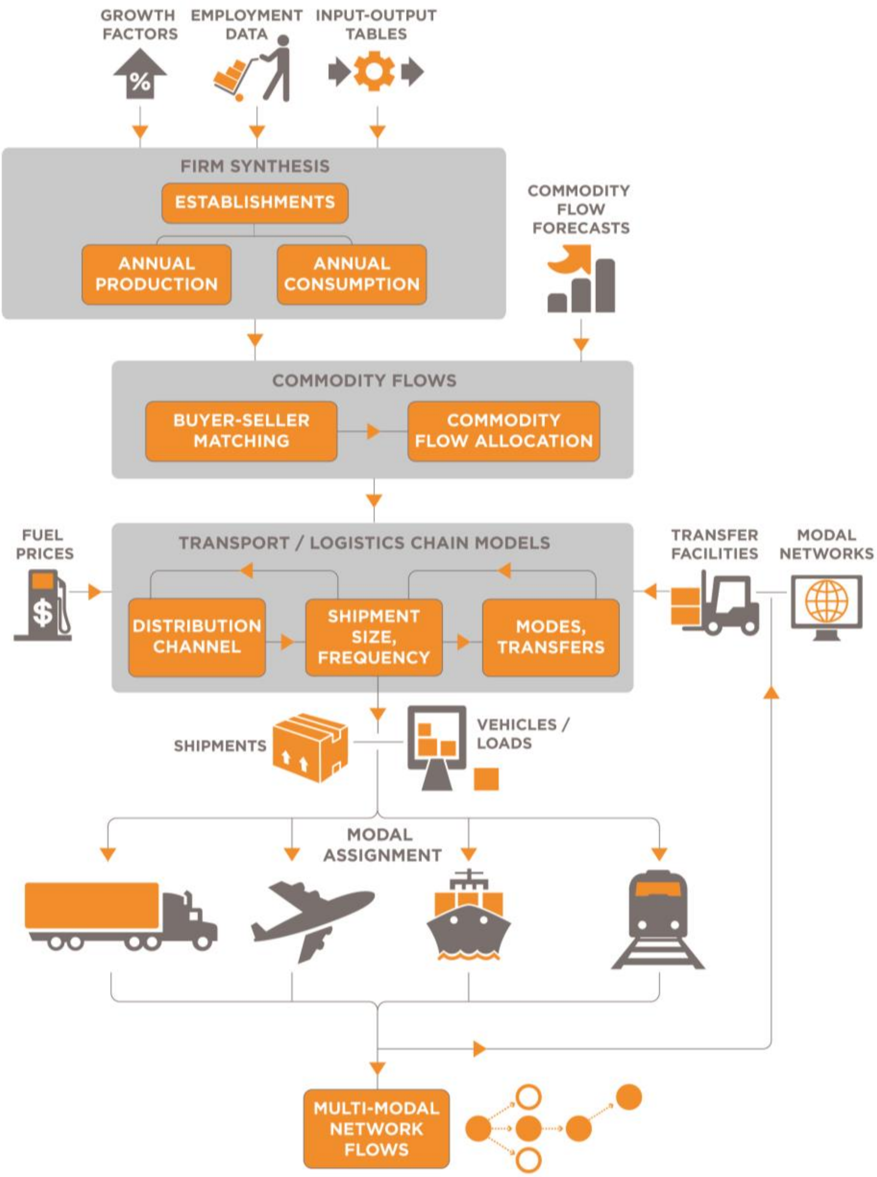
6. **Distribution Channels.** Predicts the level of complexity of the supply chain (e.g., whether it is shipped directly or whether it passes through one or more warehouses, intermodal centers, distribution centers, or consolidation centers).
7. **Shipment Size and Frequency.** Estimates discrete shipments delivered from the supplier to the buyer.
8. **Modes and Transfers.** Predicts four primary modes (road, rail, air, and waterway) and transfer locations for shipments with complex supply chains using inputs from modal networks including descriptions of transfer facilities.

Modal Assignment

9. **Multimodal Network Flows**: The transport and logistics chain models produce a **list of shipments** that are ready for extraction for use in the regional truck-touring model. This output is also converted into number of **vehicles and loads** ready for **modal assignment** (in this case, assignment is done for all truck trips to the highway network, to produce outputs of trucks volumes on the highway network including to and from transfer facilities).

The NSCM's spatial resolution, multimodal transportation networks that provides supply side inputs, and the approach to integrating it the other high level model components are discussed in the later sections of this chapter. The individual model components, including their specifications, inputs, and outputs are described in Chapter 5.0.

FIGURE 3: NATIONAL SUPPLY-CHAIN MODEL STRUCTURE



4.3 | FREIGHT TRUCK TOURING MODEL

The FTTM simulates truck movements within the Baltimore region that deliver and pick up freight shipments at business establishments. The most recent version of the FTTM is to be transferred from the CMAP mesoscale freight model, where the model was developed and implemented as part of research funded by FHWA to simulate the local delivery and pick-up of shipments in the Chicago region.

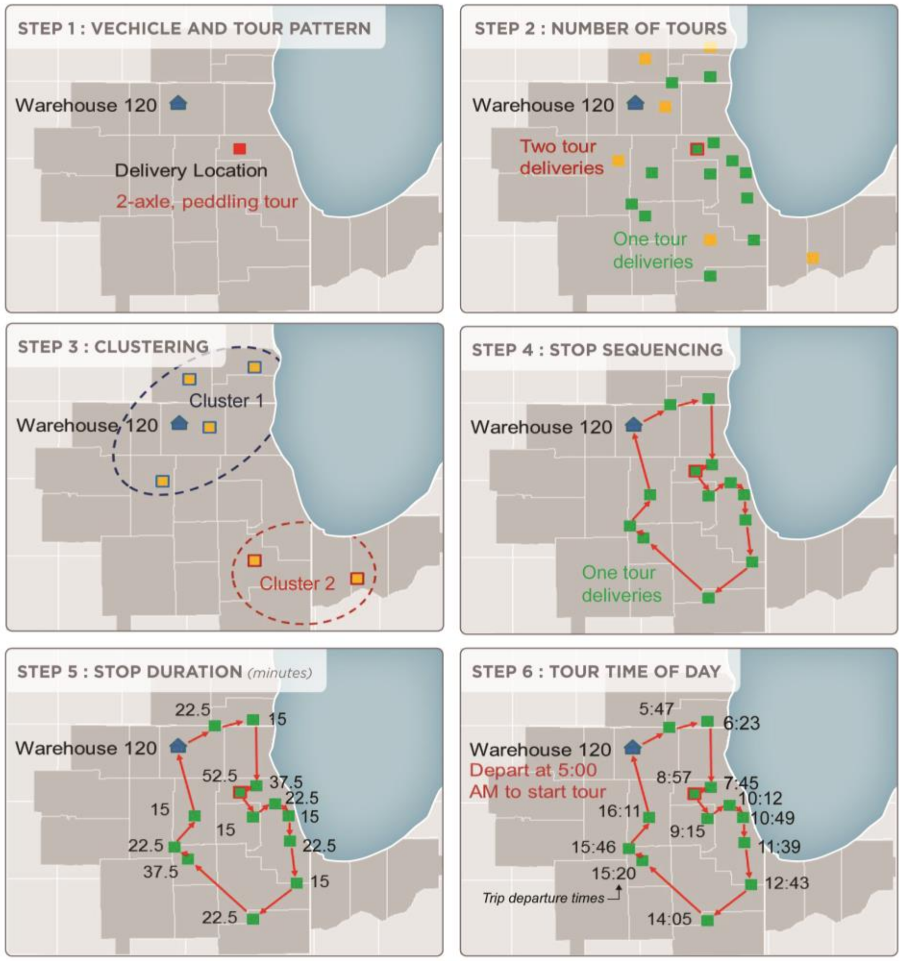
The FTTM picks up where the NSCM leaves off, with a list of zone-to-zone shipments, some of which are allocated to warehouse and distribution centers. The FTTM includes components that sequentially group shipments into tours, and add details to the descriptions of the tours such as stop sequence, stop duration, and tour start time. Figure 4 illustrates how the model build tours, and each step is described briefly here:

- **Vehicle and Tour-Pattern Choice**. The vehicle and tour-pattern choice model (Step 1 in Figure 4) is a multinomial logit model that predicts the joint choice of tour type and vehicle type and size. The tour type choice is whether a shipment will be delivered on a direct tour to the delivery location, or on a multi-stop tour in which the truck makes multiple deliveries or pick-ups. The size of the vehicle is determined jointly with tour pattern (direct or multi-stop), because certain vehicles support different types of tours.
- **Number of Tours** (Step 2 in Figure 4). This model is a multinomial logit model that predicts the complexity of the multi-stop tour in which a shipment stop is contained. For example, a truck might return to the start of its tour (e.g. origin distribution center) after one large loop, or it might break its delivery/stop schedule into two, three, or more tours.
- **Stop Clustering** (Step 3 in Figure 4). This model uses hierarchical clustering to divide stops into spatially co-located groups that can be reasonably served by the same truck during a tour.
- **Stop Sequence** (Step 4 in Figure 4). For multi-stop tours, this model uses a greedy algorithm to sequence the stops in a reasonably efficient sequence, but not necessarily the shortest path.
- **Stop Duration** (Step 5 in Figure 4). This model applies a multinomial logit model to predict the amount of time taken at each stop based on the size and shipment commodity.
- **Delivery Time-of-Day** (Step 6 in Figure 4).This is a multinomial logit model that predicts the departure time of the truck at the beginning of its tour or tours. Combining the chosen departure time with the travel time of each trip and the predicted duration of each delivery or service stop, all of the trips on the tour can be associated with a time period for assignment purposes. At this point, an iterative process is used to identify tours that are too long; overlong tours are then split into tours that meet time constraints (i.e., driver shift-length limits).

In addition to the above model components, intermediate stop models have been developed as part of the design and estimation of the CSTM. The intermediate stop models are being added to the FTTM after the delivery time of day model.

The FTTM's spatial resolution, transportation networks that provides supply side inputs, and the approach to integrating it the other high level model components are discussed in the later sections of this chapter. The individual model components, including their specifications, inputs, and outputs are described in Chapter 0.

FIGURE 4: MODEL COMPONENTS IN THE TOUR-BASED TRUCK MODEL



4.4 | COMMERCIAL VEHICLE TOURING MODEL

The CSTM simulates the remainder of the travel of light, medium, and heavy trucks that is for commercial purposes, i.e., providing services and goods delivery to households and services to businesses. This model component is newly developed for this project, and the work to develop an estimation dataset and estimate the model components is described later in this memo.

As with the FTTM, the CSTM is a tour-based model, but this time demand is derived from the characteristics of the business establishments and households in the region. The CSTM generates and simulates truck and light-duty vehicle movements based on demand for services and goods from certain industries. The CSTM includes components that generate demand for services and goods and then develops truck tours to facilitate the provision of those services and the (residential) delivery of goods. Each step in the CSTM is described briefly here:

- **Establishment Type:** The establishment type model tags each synthetic establishment from the firm synthesis model in NSCM in the modeling region with an industry type and a label that indicates whether the establishment is a goods producer, service provider, or it does both. The model applies a

Monte Carlo simulation method to draw from observed distributions of establishments by industry type and stop type, constructed using truck diary data.

- **Stop Generation:** The stop generation model predicts one day's worth of scheduled stops for each establishment by TAZ, using a count model formulation. Count models predict positive integer values for the frequency of an event. Scheduled stops are grouped in to three market segments: goods stops, service stops, and meeting stops.
- **Vehicle Assignment:** For each stop, the vehicle assignment model assigns one of three commercial vehicle types. These correspond to light (i.e., car, van, and pickup), medium/single-unit, and heavy/multi-unit truck types. The model is formulated as a multinomial logit model and predicts vehicle type as a function of the establishment's industry type, distance between establishment and stops to be served, and the stop's purpose—goods, services, or meeting.
- **Expected Stop Duration:** The expected stop duration model is applied to scheduled stops generated by the stop generation model. For each stop, expected stop duration is drawn from a smoothed, empirical distribution of observed stop distributions for each stop type, based on truck diary data.
- **Stop Clustering:** For each establishment, the stop clustering model groups scheduled goods, service, and meeting stops into feasible commercial vehicle tours, based on spatial proximity, vehicle type, total travel time, and expected stop duration.
- **Arrival Time at First Stop:** For each cluster of stops forming a tour, the arrival time at the first stop on the tour is predicted using a multinomial logit model. Based on an analysis of truck diary data, morning and afternoon arrival times (05:01am to 05:00pm) are modeled at half-hour interval, while evening and night arrival times (05:01pm to 05:00am) are modeled at one-hour interval.
- **Routing Sequence:** Given a set of scheduled stops and their locations on a tour, establishment location, and the time of the day, the model uses a "traveling salesman problem" algorithm to determine the sequence of stops on the tour.
- **Intermediate Stop Choice** This component predicts whether there are any intermediate stops between scheduled stops on a tour. The model simulates whether the driver makes one or more intermediate stops prior to each scheduled goods or tour stop, or prior to returning to the establishment to complete the tour. Purposes for intermediate stops include breaks/meals, vehicle service/refueling, and personal business/other.
- **Intermediate Stop Destination** For each intermediate stop on a tour, this model predicts a destination TAZ. Specifically, for each intermediate stop, the model selects a set of "eligible" TAZ based on attraction factor(s), such as retail employment for break/meal stops, and an impedance factor that accounts for travel distances. This model uses a "rubber banding" method that considers the travel distance from the current stop to each alternative destination and from each alternative destination to the next scheduled stop or returning to the establishment. The idea is to minimize deviations from direct paths between scheduled stop locations.

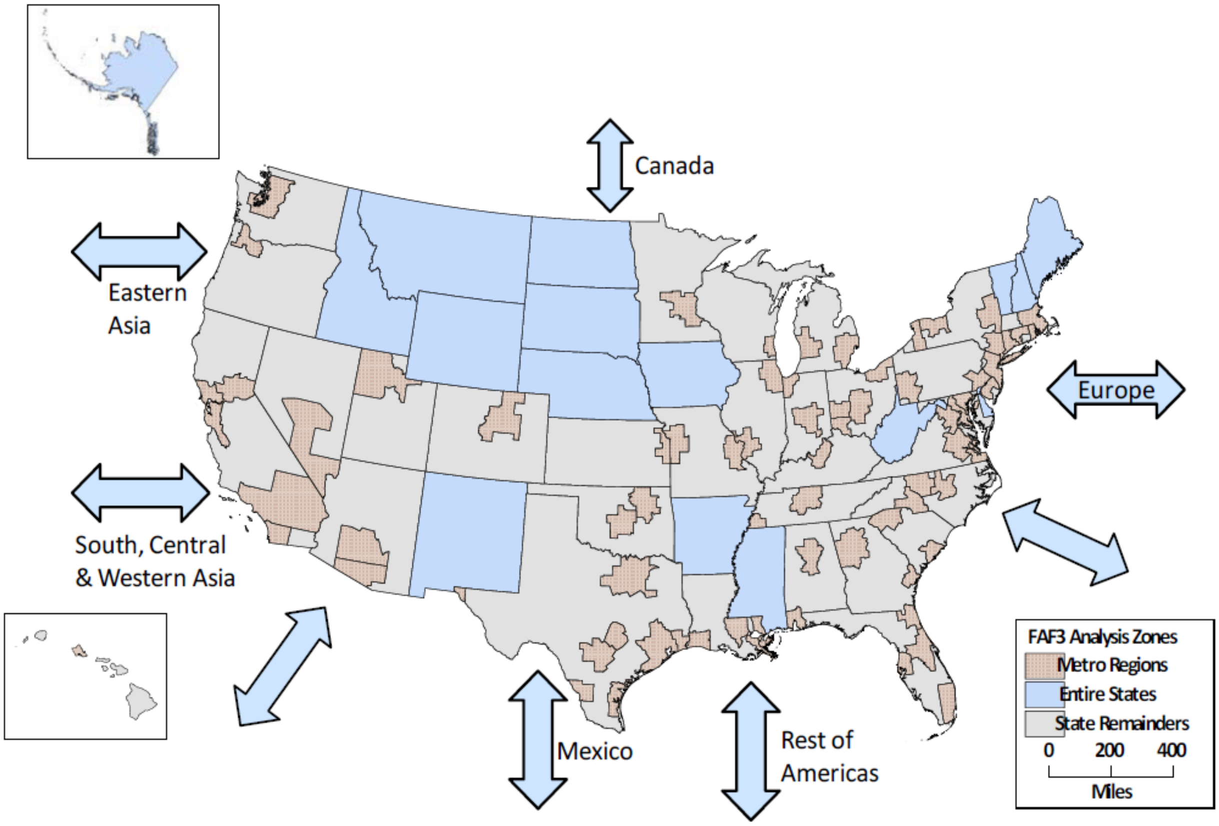
The CSTM's spatial resolution, transportation networks that provides supply side inputs, and the approach to integrating it the other high level model components are discussed in the later sections of this chapter. The individual model components, including their specifications, inputs, and outputs are described in Chapter 7.0.

4.5 | SPATIAL RESOLUTION

There are three levels of spatial resolution used in the model system's zones.

1. **National and International Zones**. This is the broadest zone system, and it is comprised of domestic and international zones from the Freight Analysis Framework, Version 3 (FAF3). The FAF3 was developed by the FHWA to evaluate commodity flow data (see Figure 5**)**. These zones are used to represent all states except Maryland and the surrounding "buffer" states. There are eight international zones used for imports and exports.
2. **MSTM TAZs.** The NSCM is compatible with the MSTM's spatial resolution and zone system, with more detailed traffic analysis zones (TAZs) internal to the modeling region and in adjacent "buffer" states. This zone system is used in several model processes, including firm synthesis and supplier selection.
3. **BMC TAZs**. The regional model's TAZs consists of TAZs that are smaller in size within the Baltimore metropolitan region. This zone system is used for both the tour-based truck models. In areas where BMC TAZs are smaller than the MSTM TAZs, business establishments are allocated to the BMC TAZs as well as to the MSTM TAZs

FIGURE 5. FAF3 GEOGRAPHY[2]



4.6 | MULTIMODAL NETWORKS

A multimodal transportation network is used to facilitate the multimodal logistics costs for alternative shipment paths to be calculated during the mode and transfer model and to allow trips to correctly partition into their separate modal elements. This model covers the global cargo flows into and out of Maryland. Four modes of transportation are included in the network: highway, railroad, air, and waterway. The network has three levels of spatial resolution: rail links, highway links, waterway links, and major logistics nodes are present both nationally and regionally.

The highway network used in the MSTM will be used in association with the non-highway networks for the mode and transfer model. The more detailed highway network in the BMC InSITE model will be used by the truck-touring models.

In order to develop a rail network, the Center for Transportation Analysis (CTA) railroad network was used, which is a representation of the North American railroad system that contains every railroad route in the United States, Canada, and Mexico that has been active since 1993[3]. The latest version—when network

[2] http://faf.ornl.gov/fafweb/Data/FAF3ODDoc611.pdf
[3] http://cta.ornl.gov/transnet/RailRoads.html

development began—was qc15n. It only contains currently operating lines and some interlines to maintain network connectivity. The raw network includes 20,624 nodes and 23,921 links.

Airport locations are from the Federal Aviation Administration (FAA) GIS-airport locations National Transportation Atlas database. This database is a geographic point database of the 19,949 aircraft landing facilities in the United States and its territories. The geospatial data is derived from the FAA's National Airspace System Resource Aeronautical Data Product. It is made public through the 2011 National Transportation Atlas Database by BTS. For foreign airports, the international FAF zones they are located in is identified and used as a tabular input to the model rather than including a network representation.

The US Army Corps of Engineers' Navigable Waterway Network, which is also made public through the 2011 National Transportation Atlas Database, is a comprehensive network database of the nation's navigable waterways. The database contains data on over 40,000 port and waterway facilities in the United States and covers waterway links between the 48 contiguous states plus Hawaii, Alaska, and Puerto Rico. After simplifying the network, a network with 6,249 nodes and 6,464 links was produced. Port point locations obtained from the same data source; the point database contains physical information on the 9,094 commercial facilities at the principal ports in the United States (covering coastal, great lakes, and inland ports).

Cargo may transfer from one mode to another at certain intermodal transfer facilities. The locations of intermodal terminal facilities are also published in the National Transportation Atlas Database. It is a point database of the 3,280 facilities in United States with information about the supported modes. In addition, the database provides a table of the supported cargo types for each facility (e.g. bulk commodities, containerized goods).

4.7 | MODEL INTEGRATION AND SOFTWARE DESIGN

4.7.1 INTEGRATION WITH STATE AND REGIONAL MODELS

The overall model framework and required model component integrations is introduced above (Figure 2): NCSM will become a component of the MSTM while the FTTM and CSTM will become components of InSITE. A relatively loose integration between the models is envisioned, with SHA assuming responsibility for the NSCM that is integrated within their MSTM, and BMC assuming responsibility for the FTTM and CSTM that is integrated within InSITE. In this loose integration, the two models will operate independently. However, the outputs from the NSCM (shipments traveling into, out, and through the BMC region, and the synthetic firms located within the model BMC region) will be provided as inputs to the urban freight modeling system within InSITE.

The BMC region is currently developing InSITE, an activity based model of passenger travel. The validated (2012) InSITE model is scheduled for completion June 2016. The activity based model will disaggregate traditional four step model person trips into each household member's daily activity pattern. Inputs will include: individual daily activity pattern, household value of time, school location, usual workplace location, vehicle availability, transit pass ownership, toll pass ownership and household classification. Output: tour generation, tour choice, tour destination, school escorting, fully joint travel, stop generation, stop purpose and tour mode.

The MD Statewide Transportation Model (MSTM) is currently being updated by SHA. SHA will adopt MSTM Version 1.1 in Summer 2015 with a base year of 2012 and forecast year of 2040.The MSTM uses FAF 3.4 based commodity flows to develop long distance X-X, I-X and I-I flows with "standard" disaggregation processes. MSTM uses a QRFM-II based process to develop zonal truck flows within the statewide model region. The national and regional flows are assigned to a time-of-day network using multi-user class equilibrium. The model runs on Cube Voyager. SHA is also developing a statewide DTA model using DTALite as part of it's SHRP2 C10 and other efforts.

4.7.2 SOFTWARE DESIGN

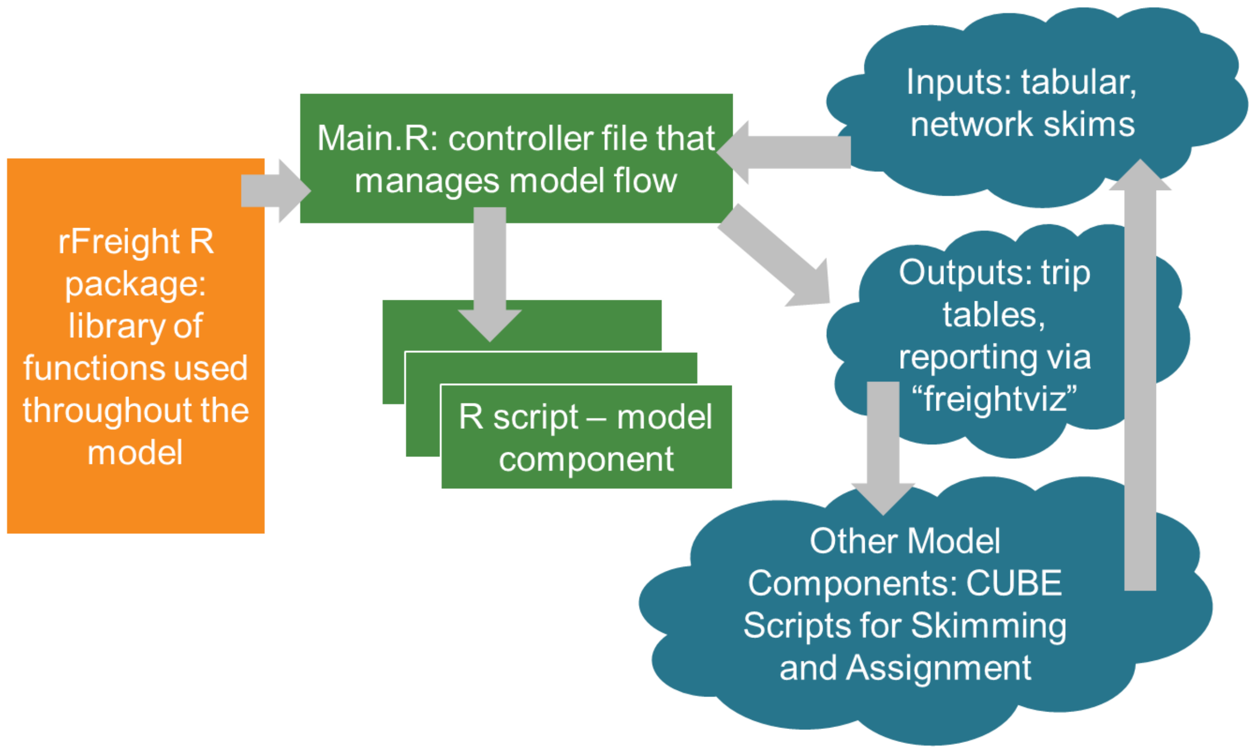
The freight model components are built using an open-source code base initially developed for FHWA. This code has continued to be developed for FDOT and CMAP. This code base covers the simulation models that encompass the NSCM and the FTTM, and the CSTM is being implemented using the same code base. It is written using the R software platform. R is a powerful, open-source, and flexible scripting language used by many in statistical fields of study. It is also increasing in popularity in the transportation field; this increased adoption is due to its ease of use and calculation speed. Enhancements made to be the freight model code during this project will be written to leverage the open-source nature of R, allowing the model to be transferred to other modeling systems without infringing on proprietary software licensing.

Figure 6 presents the general software architecture of the freight models; each of the three freight model components conforms to this structure. The R code components are in two sections. First, the model includes a function library in the form of an R package called rFreight (orange box), which RSG has developed to provide rigorous code management and documentation of the core functions required to implement the freight models. Secondly, there is a set of scripts that call functions from the rFreight library to implement each of the model components (green boxes). A master controller file acts as the main interface for the models, initiating the model run and sequentially sourcing the script files that apply each of the individual model component. The individual scripts are consistently structured to apply a single model component (such as firm synthesis): they load inputs, carry out any required data processing, apply the model, save outputs, and produce summaries from that model component before passing control back to the controller file. In this way, the model code is modularized and individual model components can be updated in future phases of model development.

The R code includes functions to load model inputs (blue box), such as tabular inputs describing the number of business establishments or travel time skims, from the file system. Each model component's R script includes functions to produce summary tabulations of results. The overall model system also includes a component called "freightviz", a further set of functions coded in R to produce customizable reports of model outputs in the form of charts and maps grouped together in a PDF document.

The freight model uses functionality of the Cube models with which it is to be integrated to skim transportation networks, exported those from matrix format to list format, and to convert model outputs from trip lists into truck trip matrices that can be loaded onto the transportation networks.

FIGURE 6: SOFTWARE ARCHITECTURE



5.0 NATIONAL SUPPLY CHAIN MODEL

5.1 | FUNCTION WITHIN THE MODEL SYSTEM

The NSCM simulates the transport of freight between supplier and buyer businesses in the United States, in this case focusing on movements that involve Maryland. The NSCM includes components that synthesize business locations, trading relationships between businesses, and the resulting commodity flows, distribution channel, shipment size, and mode and path choices for each shipment made annually.

5.2 | INTERFACES WITH OTHER MODEL COMPONENTS

The NSCM requires land use data describing business establishment locations and industry types, input-output data describing the production and consumption of commodities by industry type, and commodity flow data that describes the aggregate, annual movement of commodities to, from, and within the United States. Travel time skims from multimodal networks are used in the mode choice component to assign shipment flows to modes.

The NSCM produces a set of shipments that can either be converted to trip based zone to zone truck volumes for assignment to the statewide model networks or passed to the FTTM for truck-touring simulation. The synthesized firms within the BMC model region are also used as an input to the CSTM.

5.3 | MODEL COMPONENTS

An overview of the NSCM components is provided here, with each component described in detail below.

Firm Synthesis. As shown in Figure 7, the NSCM begins with a synthesis of firms to create individual producers and consumers of commodities segmented by industry (NAICS codes) and firm size in order to capture the primary drivers of commercial vehicle travel. Firms are situated spatially within the zone systems described above. Within the state and in adjacent states within the model area, the firms are completely disaggregate. Outside of the region and in foreign countries, symbolic firms are used to represent a specific industry and size of firm.

Buyer-Supplier Selection. Producers and consumer interests are linked through production relationships represented in make-and-use tables from the US Bureau of Economic Analysis (BEA) Input-Output (IO) accounts. Producers of one commodity seek to satisfy production inputs through consumption of other commodities, which they purchase from firms that produce the needed inputs. In the proposed model structure, buyers choose suppliers using a multinomial logit model, which matches firms based on compatible firms sizes while minimizing spatial separation. Once the spatial relationships between buyers and suppliers are established, the amount of commodity shipped on an annual basis between each pair of firms is apportioned from FAF commodity flows, based on the buying firm's size and industry type. For purposes of calculating shipment sizes and costs, commodities are aggregated into the 43 groups defined by the Standard Classification of Transported Goods (SCTG).

Distribution Channel Selection. The next step is to determine the complexity of the distribution channel used in the supply-chain, a concept pioneered by Wisetjindawat et al. (2003)[4], and adapted by RSG for use in Chicago. Multinomial choice models are used to determine the distribution channel based on buyer-supplier pair characteristics and characteristics of the industry; however, the data to estimate models of these distribution channels are quite limited. The models estimated by RSG have been restricted to alternatives describing the complexity of the distribution channel as represented by combinations of intermodal terminals, warehouses, consolidation centers, and distribution centers (i.e., direct shipments, one type of intermediate stop, two types of intermediate stops, or three or more types of intermediate stops).

Mode-Path Choice. The mode-path choice model considers all feasible combinations of truck, rail, waterborne, and airborne carriage. The model in use in Chicago and Florida was adapted from research conducted by de Jong and Ben-Akiva (2007)[5] to predict the mode and path of long-haul movements of freight in Norway and Sweden, based on a comprehensive accounting of transport and logistics costs. These costs include transport and intermediate handling, inventory, deterioration and damage, pipeline, ordering and stock-out. The mode-path model is applied separately for each type of distribution channel, allowing different parameters for complex supply chains and direct movements. Detailed networks of road and rail in the United States are used, with simpler functions of distance and the value of goods being transported representing air and water modes. The modes and transfer locations on the shipment paths are determined based on the travel time, cost, characteristics of the shipment (e.g., perishable, expedited, containerized), and characteristics of the distribution channel (e.g., whether the shipment is routed via a warehouse, consolidation, or distribution center, and whether the shipment includes an intermodal transfer).

Shipment Size Selection. Shipment size is estimated using a multinomial logit model, considering a variety of firm, commodity, and travel characteristics. At this step, units of analysis change from annual commodity flows between pairs of firms to discrete shipments delivered from the supplier to the buyer.

Daily Shipments and Warehouse Selection. The list of shipments is converted from all annual shipments to a daily sample to represent the day being modeled once the modes and intermodal transfers have been assigned. This component of the model is adjustable to allow for seasonal variations in commodity flows. This component of the model also assigns shipments to specific warehouse, distribution, and consolidation centers if the shipment passes through one of these locations.

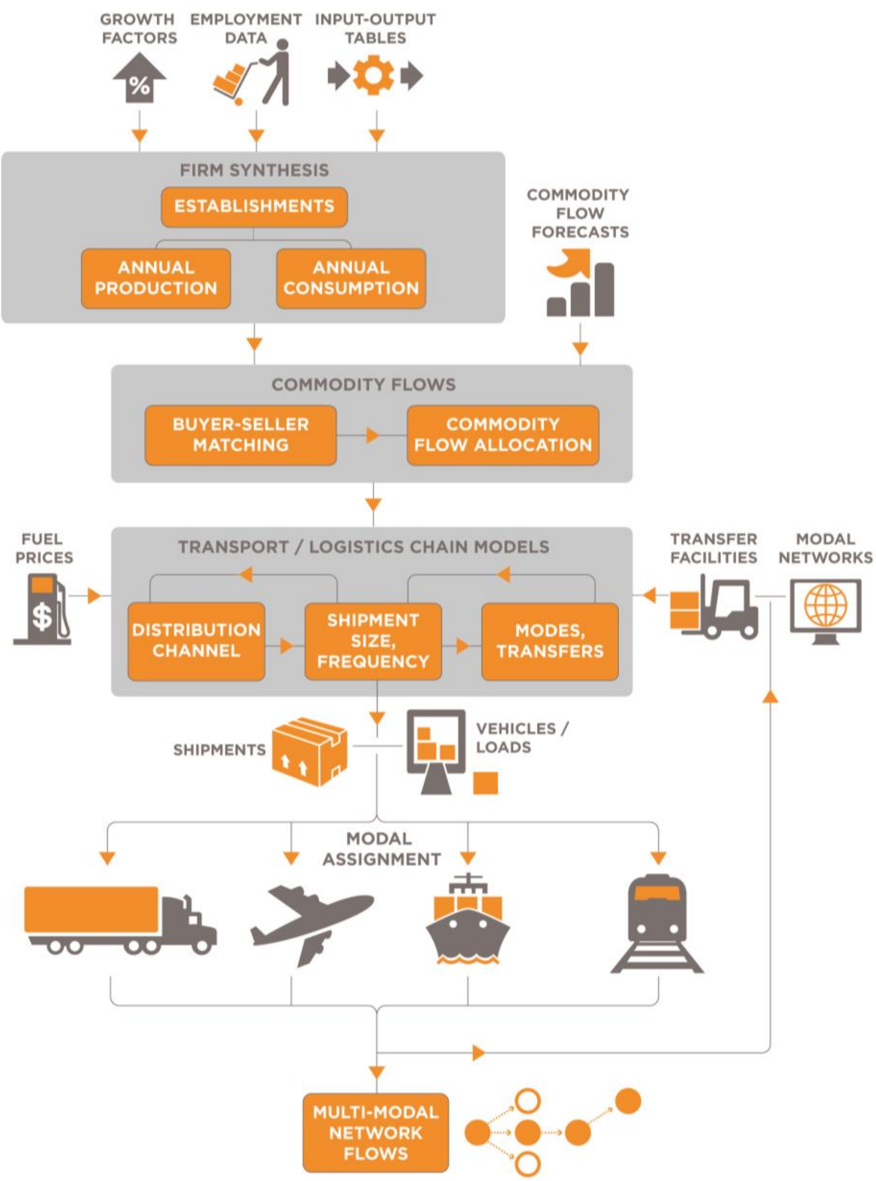
Trip Tables and Truck Assignment. In the statewide modeling context, the daily modal shipment volumes are converted to modal trip tables in this step. The model incorporates a multimodal transportation network that provides supply-side information to the model, including costs for different paths by different modes (or combinations of modes) and truck assignment.

---

[4] Wisetjindawat, W. and Sano, K. 2003. A behavioral modeling in micro-simulation for urban freight transportation, J. East. Asia Soc.Transport. Stud., 5, pp. 2193–2208.

[5] Jong, G.C. de and M. Ben-Akiva. 2007. A micro-simulation model of shipment size and transport chain choice, Special issue on freight transport of Transportation Research B, 41, pp. 950–965.

FIGURE 7: NATIONAL SUPPLY CHAIN MODEL COMPONENTS



5.3.1 FIRM SYNTHESIS

The initial element of the model synthesizes all firms in the United States, and a sample of international firms, in order to capture long-haul freight movements. Each firm has individual characteristics that identify the following:

- Where are they located?
- How large is the firm?

- What industry do they operate in?
- Which commodities do they consume?
- Which commodities do they produce?

The geography within the region of interest (in this case, Maryland and surrounding buffer areas where MSTM TAZs are divisions of States) is divided into counties. The geography outside the region is defined by FAF3 zones. This model synthesizes firms by industry category and by size category to capture the primary drivers of commercial vehicle travel. Firm synthesis is controlled by regional, county, and state control totals.

Data Sources and Model Development

Firms by size and type are allocated to TAZ using available observed data sources on employment by type, consistent with the data used in the passenger travel demand forecasting models. These employment data are primarily from County Business Patterns data collected by the Census Bureau.

Input output data from the BEA are used to describe what each industry produces (makes) and consumes (uses). These relationships are known as make and use tables. When multiple commodities are made or used, then the data represents a proportional value.

The models for firm synthesis and business location were developed originally by the University of Chicago (Samimi et al., 2010) for the Chicago Mesoscale Freight Model (Cambridge Systematics, 2011) and were translated into R by the project team for the development of the FHWA Freight Forecasting Framework project (RSG, 2012). The model for firm synthesis is a direct enumeration process of the firms based on the employment totals. Firms are enumerated by two attributes: 1) industry (NAICS); and 2) the employee size category for each geographic unit.

Model Application

Figure 8 shows a schematic of the firm synthesis process applied. The firm synthesis process enumerates lists of firms, then allocates these firms to zones and identifies them as either producers (suppliers) or consumers (buyers). Figure 9 is a dot density plot by county in Florida from the Florida Statewide Model application of this model showing the density of large and small firms (note that the points indicate overall density by county and not actual firm locations within the county).

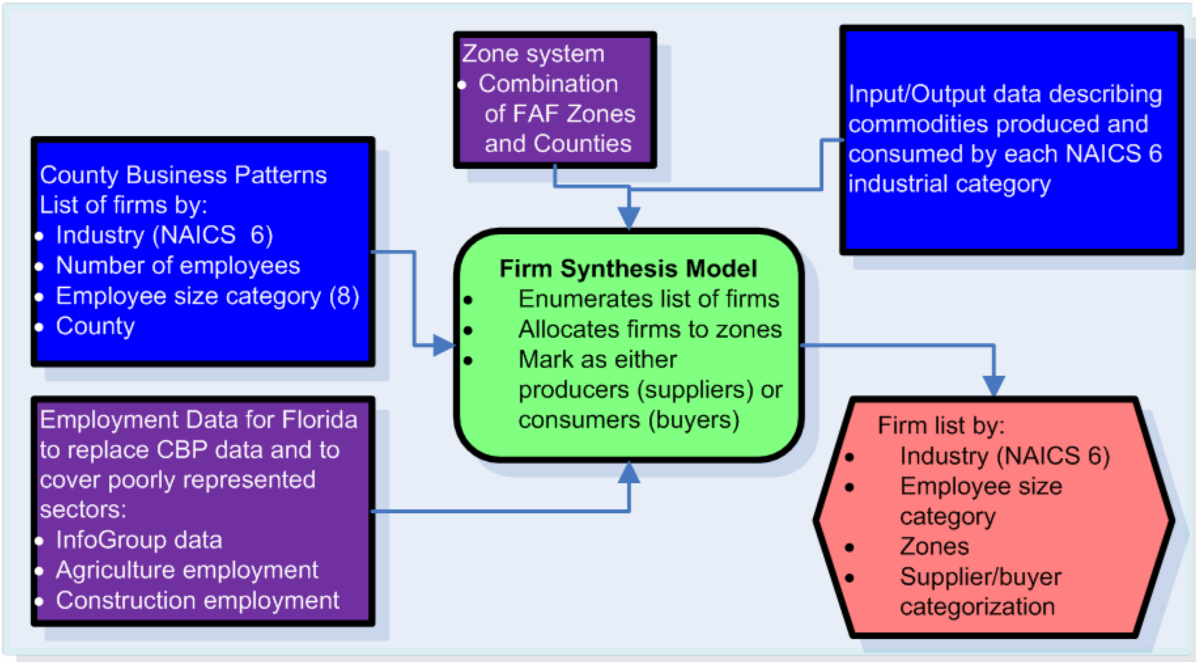


FIGURE 8. FIRM SYNTHESIS PROCESS.

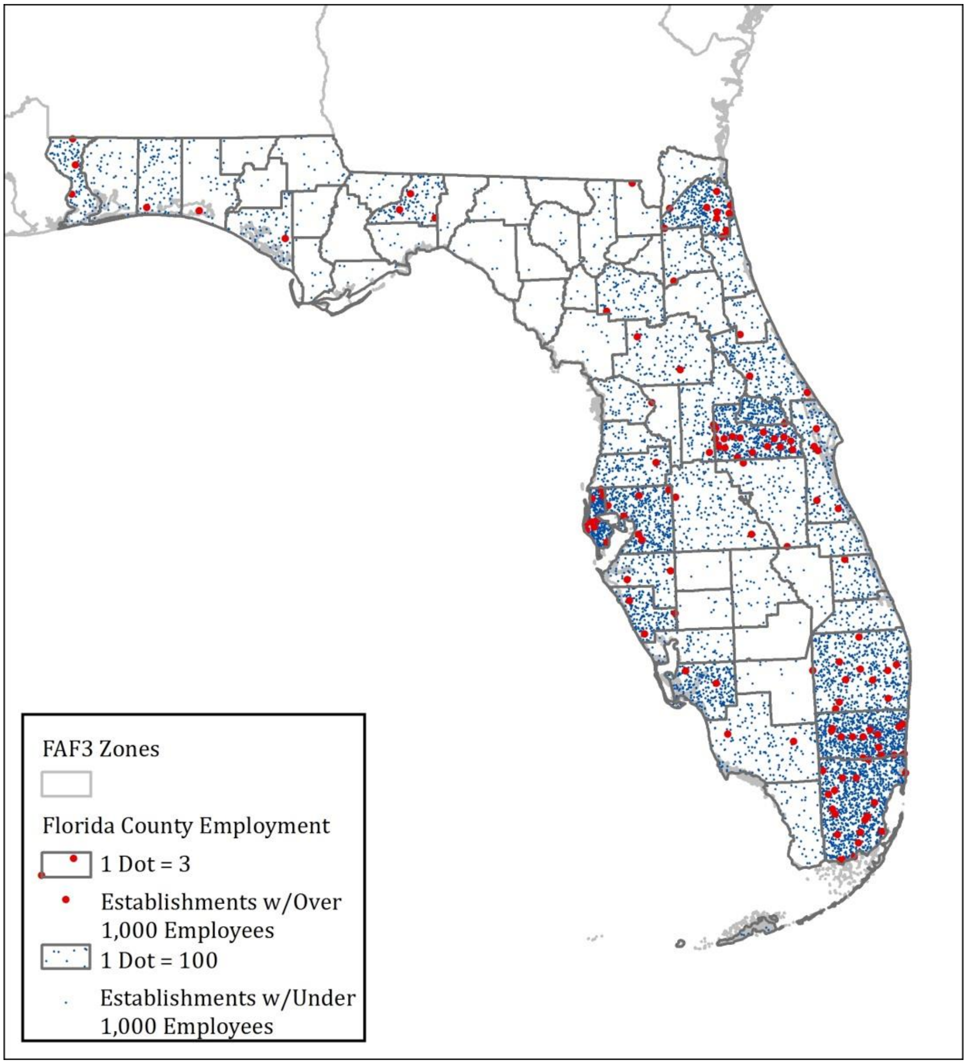


FIGURE 9. FIRMS IN FLORIDA BY SIZE.

A correspondence between NAICS and SCTG is used to attach a unique SCTG commodity to each of the enumerated firms. The firms are then split into producers and consumers based on whether or not they produce any commodities. Some industries produce more than one commodity (such as wholesale). To account for this, the commodity that each firm produces is simulated based on probabilities of the multiple commodities that it could produce. A producer firm-types database is also created, which is used later in the supplier firm selection model. A firm type is defined by a unique combination of industry NAICS, commodity SCTG, and the geographic ID of a firm.

A consumers/users database is created next. This consists of all the firms in the firm database merged with processed Input-Output data based on the NAICS industry code of the consumer. The processed Input-Output data identifies the SCTG commodities (in terms of value) consumed and the corresponding supplier NAICS for each consumer industry NAICS. A limit of 90% of the value consumed is used in order to decrease computational burden by focusing on the largest inputs to each industry. The SCTG commodity for suppliers who could produce more than one SCTG commodities is simulated using probability thresholds. This is done for all the suppliers that are being considered. It is assumed that a certain percentage (in this case, 30%) of consumers would work with a wholesaler instead of directly with a producer. Therefore, some suppliers to consumers, who themselves are not wholesalers, are probabilistically mutated to an appropriate wholesale supplier (NAICS) based on the SCTG commodity being consumed.

Finally, two datasets are created in this step. A makers/suppliers dataset is created from the producers firm type database and consists of firms that are located in a Maryland or buffer zone, all firms that have at least 1,000 employees and one firm by each unique industry-commodity type for each geography. A users/buyers dataset is created from the consumers database and consists of firms that are located in a Maryland or buffer zone, firms that have more than 500 employees, one firm by each unique industry-employee size category for each geography, and a set of 5% of randomly selected consumers.

5.3.2 SUPPLIER FIRM SELECTION

The next element pairs up buyers and suppliers among the firms that have been synthesized in the previous step based on the size of each firm, their industry, and the distance between them.

Data Sources and Model Development

The model for supplier firm selection is based on earlier freight modeling work for the Chicago Metropolitan Agency for Planning (Cambridge Systematics, 2011) and University of Illinois in Chicago (Samimi et al., 2010). For each buyer/consumer firm, a supplier is selected from the suppliers/makers dataset. The selection of a supplier does not mean the selection of an exact business, but that of a firm type (a combination of industry NAICS, commodity SCTG, and the geographic ID of a firm). The exact firm is determined after the next step of firm allocation (of each firm to a TAZ) is done.

The probability of a supplier being paired with a buyer firm (type) depends on the employment sizes of both the firms and the geographic distance between them. Table 1 presents the coefficients used. These coefficients are asserted and not estimated due to the unavailability of data of this nature. In general, the probability of a supplier firm being chosen increases with its employee size and its proximity to the buyer firm. The distances between the firms are GCD values obtained from ORNL county-to-county skims.

TABLE 1. SUPPLIER SELECTION PARAMETERS.

Consumer Business Size (Number of Employees)	Coefficient							
	Producer Business Size (Number of Employees)			**Great Circle Distance between Consumer and Producer (Miles)**				
	1 to 99	**100 to 499**	**500+**	**Intra-county**	**1 to 149**	**150 to 595**	**596 to 1,509**	**Over 1,509**
1 to 99	0.2	0.2	0.4	0.1	0	-0.2	-0.3	-0.4
100 to 499	0.2	0.6	0.6	0.1	0	-0.05	-0.1	-0.2
500+	0.4	0.6	0.6	0.1	0	0	-0.05	-0.1

Source: Cambridge Systematics, 2011, Table 2.8, page 2-23

Model Application

Figure 10 shows a schematic of the supplier selection model. A choice set of suppliers is created for each buyer firm based on the top five commodities it requires and the corresponding NAICS of the suppliers. A supplier firm is excluded from the choice set if no flows for the commodity being traded are observed in the FAF3 dataset between the relevant FAF zones. The GCD values are merged based on the buyer and supplier zones (which are TAZs in Maryland and the buffer area; FAF zones elsewhere). A score for each buyer and potential supplier pair is then calculated using the attested coefficients and adding a random value for stochasticity. For each buyer firm, the supplier firm with the best (highest) score is selected.

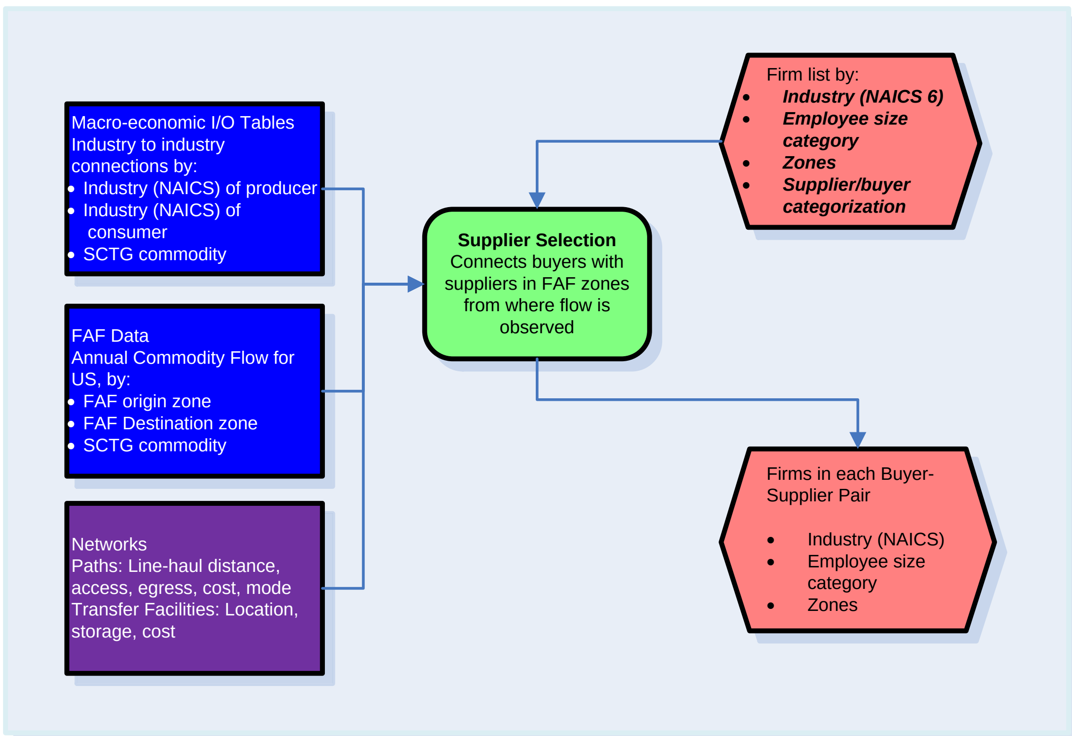


FIGURE 10. SUPPLIER FIRM SELECTION MODEL PROCESS.

5.3.3 GOODS DEMAND

This element predicts the demand in tonnage for shipments of each commodity type by each firm in each industry. The demand is developed to represent the goods produced by each firm and the goods consumed by each firm (and household) in the United States.

Data Sources and Model Development

The goods demand model relies primarily on the FAF3 freight flows and the buyer-supplier pairs estimated in the supplier firm selection model. The model also incorporates input-output tables to determine the allocations between industry types. The amount of commodity shipped on an annual basis between each pair of firms is apportioned based on the number of employees at the buyer and their industry so that observed commodity flows are matched.

Model Application

Figure 11 presents the goods demand model process. Once buyer and supplier pairs have been established, the annual flow between each of the pairs is estimated. The FAF3 flows dataset is used to apportion goods demand to each buyer supplier pair based on the size of the buyer firm. An estimate of consumption (of the commodity being consumed) by a buyer firm is calculated based on the value (in dollars) consumed per employee, which is obtained using Input-Output (Input-Output or make-use) economic tables. The values consumed per employee are calculated for each combination of supplier-buyer industry NAICS from the Input-Output tables.

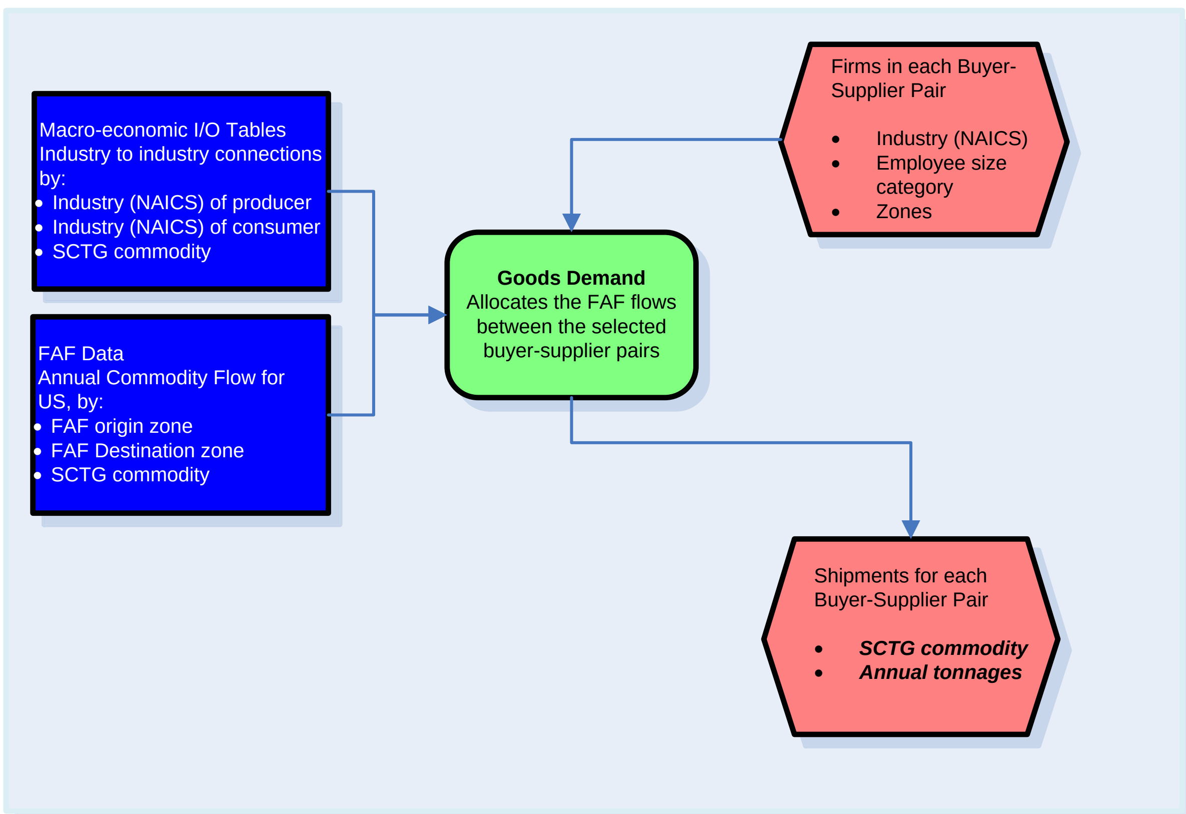


FIGURE 11. GOODS DEMAND MODEL PROCESS.

The values consumed per employee are used to calculate a consumption estimate (in dollars) for each buyer firm. A share of consumption for each firm in a particular zone is then calculated based on the consumption estimate. These shares are used to apportion freight flows for each commodity into a zone for individual buyer firms. This results in an estimate of annual goods demand between each of the buyer-supplier pairs.

5.3.4 DISTRIBUTION CHANNEL

This model selects the distribution channel for the shipment, a key element of the framework that represents an important business decision made by shippers. A distribution channel refers to the supply chain a shipment follows from the supplier to the consumer/buyer and it is critical to business freight related operations. The supplier firms may use their own transportation resources or send shipments to the buyer using third-party logistics firms. The distribution channel might affect the cost, shipment size, and frequency of shipments between a buyer-supplier firm pair.

In this framework, the transfer facilities are represented in the supply chain rather than including all establishments that goods move through as they travel from the producer to the consumer; this is because of limited data for these detailed supply chains. National supply chain data, when they become available, could improve this element. The distribution channel model uses discrete choice methods to identify the unique aspects of the supply chain.

Data Sources and Model Development

An establishment survey conducted by University of Illinois at Chicago was used to represent the elements of the supply chain, which contained data on whether the goods went through a consolidation center, a distribution center, and/or a warehouse. Other aspects of the supply chain were not possible with this dataset, and other datasets did not have the national coverage or details about the supply chain for this purpose. This survey was a small sample (570) of shipments across the United States and a diverse range of industry types. In this survey, there were 47% of shipments with no intermediate stops, 38% with one intermediate transfer location, and the remaining 15% of shipments with two or more stops. As expected, shorter trips tend to have fewer transfers.

The concept of distribution channel was further simplified to obtain a reasonable sample for model estimation (Figure 12). Four alternatives for distribution channels were identified: direct, one-stop type, and two-stop type; and three-stop types, where stop type is a warehouse, distribution center, or consolidation center. Distribution channels that involved only one warehouse stop, or only one distribution center stop, were considered the same.

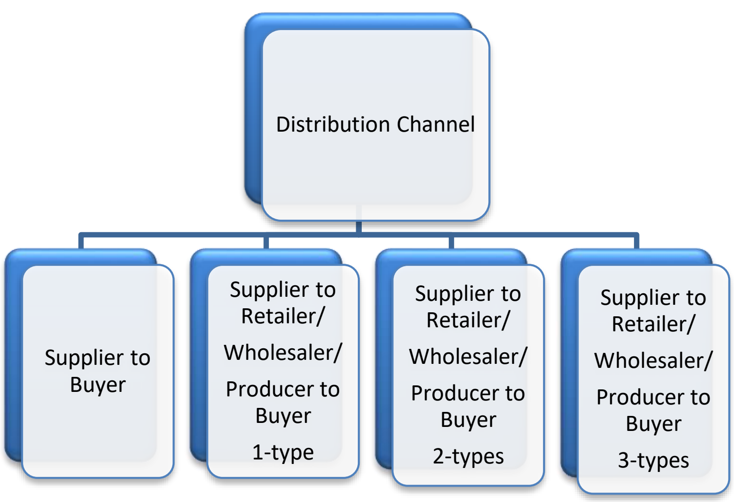


FIGURE 12. DISTRIBUTION CHANNELS.

Table 2 shows the multinomial logit (MNL) model estimated for food products from the FAME survey. The direct distribution channel is the mode preferred, everything else remaining constant. The other variables that affect the choice of distribution channel are firm size and the industry type of the firms involved. It is true that the estimated model does not have a rich specification, but this is reasonable given the data constraints. Table 3 shows the distribution channel MNL model estimated for manufactured goods. The explanatory variable types are the same as those in the food products model.

TABLE 2. DISTRIBUTION CHANNEL MODEL SPECIFICATION FOR FOOD PRODUCTS.

Choices	Variable Description	Variable Name	Coefficient	t-stat
Direct	Alternative Specific Constant	ASC_V1	0 (fixed)	
1-Type Used	Alternative Specific Constant	ASC_V2	-0.932	-2.47
2-Types Used	Alternative Specific Constant	ASC_V3	-3.32	-3.20
3-Types Used	Alternative Specific Constant	ASC_V4	-52.5	-3.11
Direct	49 or less employees firm involved	EMP49_1	0.907	2.03
1-Type Used	Manufacturing industry firm involved	MFGIND2	1.94	3.48
2-Types Used	Transportation\warehousing or wholesale trade firm involved	TRWIND3	3.49	3.23
3-Types Used	Transportation\warehousing or wholesale trade firm involved	TRWIND4	51.4	3.05
3-Types Used	Great circle distance between buyer and supplier zones	DIST1	0.000559	1.14
Number of Observations	**Final Log Likelihood**	**Rho-squared**		
106	-85.326	0.419		

TABLE 3. DISTRIBUTION CHANNEL MODEL SPECIFICATION FOR MANUFACTURED PRODUCTS.

Choices	Variable Description	Variable Name	Coefficient	t-stat
Direct	Alternative Specific Constant	ASC_V1	0 (fixed)	
1-Type Used	Alternative Specific Constant	ASC_V2	-1.96	-5.25
2-Types Used	Alternative Specific Constant	ASC_V3	-2.68	-6.53
3-Types Used	Alternative Specific Constant	ASC_V4	-3.58	-6.04
3-Types Used	50 to 199 employees firm involved	EMP3_4	1.32	2.06

1-Type Used	200 or more employees firm involved	EMP4_2	0.698	1.91
1-Type Used	Transportation\warehousing or wholesale trade firm involved	WHIND2	1.88	4.8
2-Types Used	Transportation\warehousing or wholesale trade firm involved	WHIND4	1.5	3.16
Number of Observations	**Final Log Likelihood**	**Rho-squared**		
182	-176.182	0.302		

Model Application and Results

Figure 13 shows a schematic of the distribution channel model. The distribution channel choice simulated shipments between all the buyer-supplier pairs based on the type of commodity. The manufactured goods model was applied for all commodities other than food. At this stage in the framework, the unit of analysis is shipments by all modes; therefore, the distribution channels are not mode specific and may be completed by a single mode or be multimodal (the process of selecting modes for movement of each shipment takes place in step seven of the model, mode and transfers).

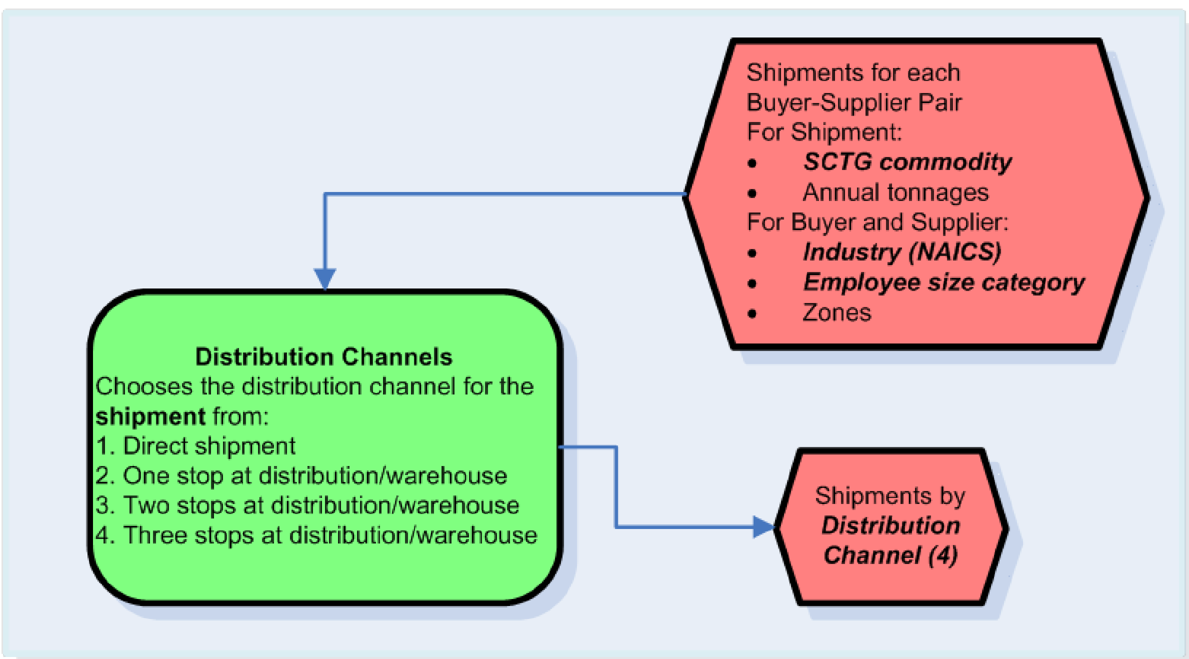


FIGURE 13. DISTRIBUTION CHANNEL MODEL PROCESS.

The final step during application of the model is to adjust the alternative specific constants in the model by commodity groups. In Pouraddollahi (2013), the authors group the surveyed shipments into 11 commodity classes, which are aggregations of the 43 SCTG commodity groups, and present the proportion of shipments by distribution channel for each of these groups. These data were used as calibration targets and the alternative specific constants are iteratively adjusted such that the output from the distribution channel model matches those targets. Figure 14 presents the average number of stops derived from the distribution channels for aggregate groupings of commodities.

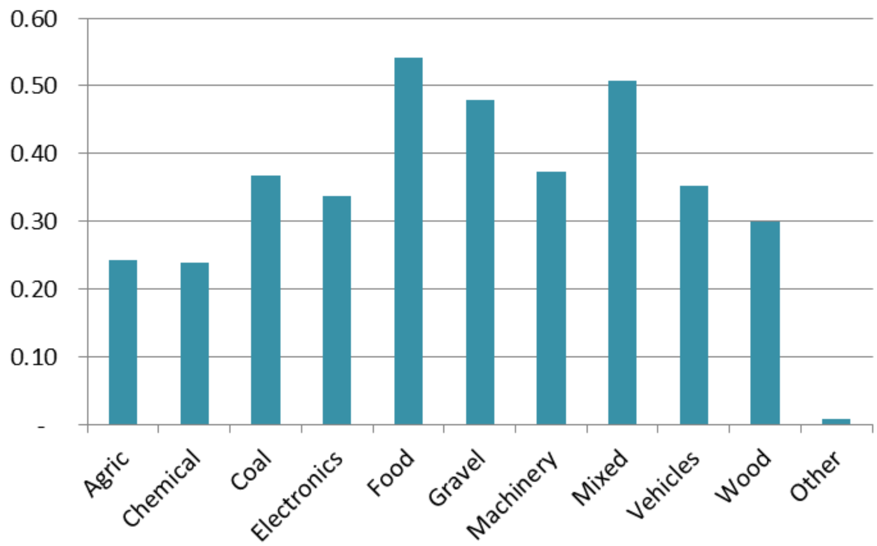


FIGURE 14. AVERAGE NUMBER OF STOPS BY COMMODITY GROUP.

5.3.5 SHIPMENT SIZE AND FREQUENCY

In this step, the annual goods flow between buyer-supplier firms pairs are broken down into individual shipments. The shipment size (weight) and the corresponding number of shipments per year are determined. Shipment size affects the mode used to transport the shipment. This framework is not designed to optimize the shipments or identify the logistics of how shipments may be combined to make a truckload or rail delivery.

Data Sources and Model Development

An MNL model is estimated for choice of shipment size. The Texas commercial vehicle survey dataset was used for estimating the discrete choice model due to its relatively high sample size. This dataset is not ideal for the shipment size model because the shipments represented in the dataset are likely to include many within an urban area. However, this dataset is most appropriate considering the sample sizes in other datasets.

Based on the distribution, three alternatives were selected to form the choice set: less than or equal to 999 lbs., 1000-9999 lbs., and greater than 9999 lbs. It was hypothesized that the distribution channel would influence the choice of shipment size. The distribution channel was not directly available in the Texas dataset. The stop-level data were transformed into tour-level data and the distribution channel was assigned based on the stops made by the truck at ports, intermodal facilities, warehouses, and distribution centers. Two models were estimated initially: food and manufactured products. The coefficients from the manufactured products model were used for other commodities.

Table 4 shows the shipment-size choice model results for food products. It appears that the shipment size between 1000 and 9999 lbs. is the most preferred for food products, everything else being equal. A multi-stop distribution channel in which the shipment stops at three types of facilities seems to positively influence the highest shipment size category (>= 10,000 lbs.). The other explanatory variables in the model specification are: trip length until current shipment stop, from the base location; and industry types at the stop location. Longer trip lengths seem to be associated with shipments greater than 10,000 lbs.

TABLE 4. SHIPMENT-SIZE MODEL SPECIFICATIONS FOR FOOD PRODUCTS.

Choices	Utility Equations
<= 999 lbs.	ASC_V1 * one + SIC11 * SIC1 + SIC21 * SIC2
1000-9999 lbs.	ASC_V2 * one + DISTCHAN12 * DISTCHAN
>=10000 lbs.	ASC_V3 * one + SIC23 * SIC2 + DISTCHAN32 * DISTCHAN_2 + Cost2 * Cost

Choices	Variable Description	Variable Name	Coefficient	t-stat
<= 999 lbs.	Alternative Specific Constant	ASC_V1	0 (fixed)	
1000-9999 lbs.	Alternative Specific Constant	ASC_V2	0.546	3.85
>=10000 lbs.	Alternative Specific Constant	ASC_V3	-1.71	-5.98
>=10000 lbs.	Trip Length (Cost)	Cost2	0.245	2.48
1000-9999 lbs.	Distribution Channel (DISTCHAN) with 1-Type Used	DISTCHAN12	-0.788	-3.58
>=10000 lbs.	Distribution Channel (DISTCHAN) with 3-Types Used	DISTCHAN32	0.759	3.05
<= 999 lbs.	Service Industry (SIC1)	SIC11	5.84	5.77
<= 999 lbs.	Transportation/Construction Industry (SIC2)	SIC21	0.975	3.57
>=10000 lbs.	Transportation/Construction Industry (SIC2)	SIC23	2.88	9.9

Number of Observations	Final Log Likelihood	Rho-squared
738	-554.922	0.316

Table 5 shows the shipment-size choice model results for manufactured products. The explanatory variables in this model are similar to those in the food products model. Shipments less than or equal to 999 lbs. seem to be the most preferable, everything else being equal. Longer trip lengths seem to be associated with shipment sizes less than or equal to 999 lbs.

TABLE 5. SHIPMENT SIZE MODEL SPECIFICATION FOR MANUFACTURED PRODUCTS.

Choices	Utility Equations			
<= 999 lbs.	ASC_V1 * one + cost1 * cost + SIC11 * SIC1 + SIC31 * SIC3			
1000-9999 lbs.	ASC_V2 * one + SIC32 * SIC3 + DISTCHAN12 * DISTCHAN			
>=10000 lbs.	ASC_V3 * one + DISTCHAN33 * DISTCHAN_3			
Choices	**Variable Description**	**Variable Name**	**Coefficient**	**t-stat**
<= 999 lbs.	Alternative Specific Constant	ASC_V1	0	
1000-9999 lbs.	Alternative Specific Constant	ASC_V2	-0.107	-0.5
>=10000 lbs.	Alternative Specific Constant	ASC_V3	-0.349	-1.63
<= 999 lbs.	Trip Length	cost1	0.15	1.69
1000-9999 lbs.	Distribution Channel with 1-Type Used	DISTCHAN12	-0.911	-3.65
>=10000 lbs.	Distribution Channel with 3-Types Used	DISTCHAN33	-1.35	-2.77
<= 999 lbs.	Service Industry	SIC11	2.27	4.16
<= 999 lbs.	Manufacture/Retail/Wholesale/Mining Industry	SIC31	1.98	6.3
1000-9999 lbs.	Manufacture/Retail/Wholesale/Mining Industry	SIC32	1.03	2.86
Number of Observations	**Final Log Likelihood**	**Rho-squared**		
552	-431.443	0.289		

Model Application

Figure 15 illustrates the shipment size and frequency model. The shipment size choice is simulated for all the buyer-supplier firm pairs using the estimated models. The manufactured goods model is applied to all commodities other than food.

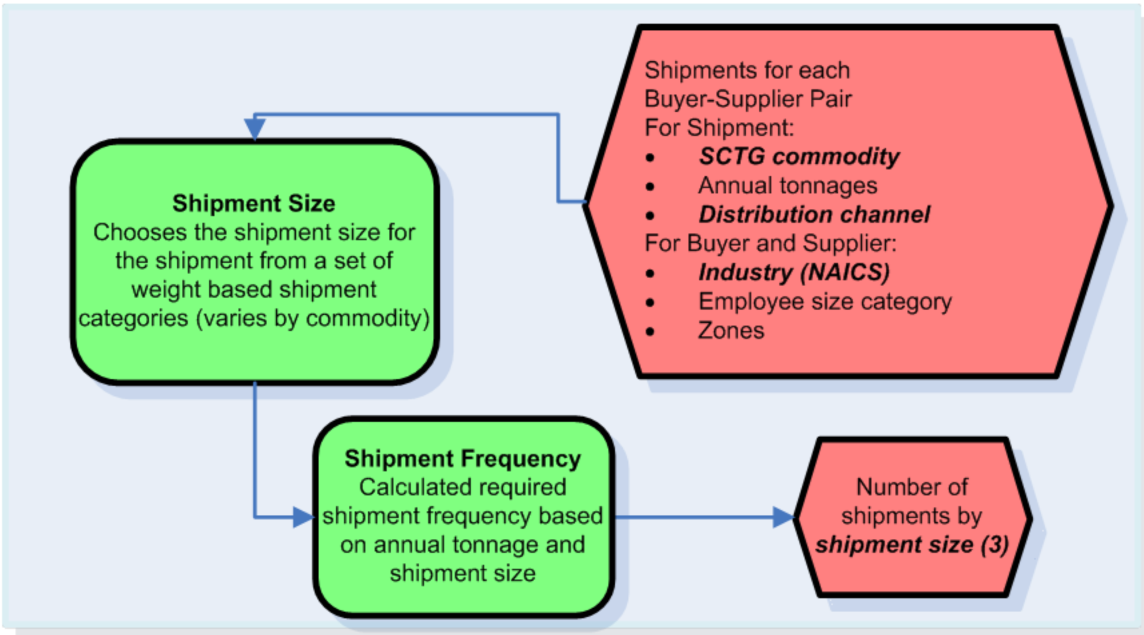


FIGURE 15. SHIPMENT SIZE AND FREQUENCY MODEL PROCESS.

The final step during application of the model is to adjust the alternative specific constants in the model by commodity group. CFS 2012 data showing the distribution of shipment sizes by SCTG commodity group are used as calibration targets and the alternative specific constants are iteratively adjusted such that the output from the shipment size channel model matches those targets.

Simulation of shipment size results in the assignment of a shipment to one of the three broad shipment categories. To obtain a more accurate shipment size, each size category was split into bins and probability thresholds for a shipment being in one of those bins were calculated from a combination of the CFS 2012 by commodity and the Texas survey data. The probability of a shipment falling into each bin was computed using the observed distributions from the survey data. All the modeled shipments were then assigned to one of the finer shipment-size categories using Monte Carlo simulation. An annual delivery frequency was calculated using the annual commodity flow (in tons) and the individual shipment size for all the buyer-supplier firms, for use later in the model system to select a sample of daily activity from the annual shipment flows.

5.3.6 MODE AND TRANSFER CHOICE

This step assigns a mode for shipments transported between each buyer-supplier pair. There are four primary modes (road, rail, air, and water) included in the mode choice model. Networks of all four modes (i.e., road, rail, water, air) for the United States are used.

The modes and transfer locations on the shipment paths are determined based on the travel time, cost, characteristics of the shipment (e.g., bulk natural resources, finished goods), characteristics of the distribution channel (e.g., whether the shipment is routed via a warehouse, consolidation, or distribution center), and whether the shipment includes an intermodal transfer (e.g. truck-rail-truck). A mode and path (from a set of feasible modes and paths) is chosen, one that would have the least annual transport and logistics cost using a two-step process:

- First, a set of feasible paths between each O-D pair is enumerated.
- Second, a reasonable set of parameters is applied to the path skims to generate total annual transport and logistics costs for each combination of path and mode.

In calculating the total annual costs for each pair of seller and buyer, supply chain and inventory control costs are considered and incorporated to account for the inventory-associated costs.

Data Sources and Model Development

Methods developed by De Jong and Ben-Akiva (2007) are used to predict the path and mode of long-haul movements of freight. The path includes identifying the location of intermodal transfer facilities, distribution centers, or warehouses where shipments are consolidated or de-consolidated. Detailed networks of road and rail for the United States were used, in addition to networks describing airport and port locations, domestic waterway connections, and finally GCD distances between airports and between ports and international destinations.

Total logistics costs that the buyer and supplier encounter is the sum of transport and inventory costs and can be itemized as shown below:

Total Logistics Costs = Transport costs + Inventory costs

Inventory Costs = Ordering + Carrying + Damage + Inventory in-Transit + Safety Inventory

- Ordering = *Order preparation, order transmission, production setup if appropriate*
- Carrying = *Cost of money, obsolescence, insurance, property taxes, and storage costs*
- Damage = *Order lost or damaged*
- Inventory in-transit = *Inventory between shipment origin and delivery location*
- Safety Inventory = *Lost sales cost, backorder cost (Demand and Lead-time uncertainty)*

This formulation models logistics decisions in a joint fashion by capturing transport and logistics costs in a single equation. This effectively reflects the real-world decision-making of freight movers by accounting for different components of costs. These models for mode choice and intermodal transfers were based on the formulation developed by de Jong and Ben-Akiva (2007):

$$G_{mnql} = \beta_{0ql} + \beta_1 \times \left(\frac{Q}{q}\right) + T_{mnql} \times \beta_2 \times j \times v \times Q + \beta_3 \times t_{mnl} \times v \times \frac{Q}{365} + (\beta_4 + \beta_5 \times v)\left(\frac{q}{2}\right) + \beta_5 \times v \times a * \sqrt{(LT \times \sigma_Q^2) + (Q^2 \times \sigma_{LT}^2)}$$

Table 6 provides descriptions of variables and parameter notations. A low (0.01), medium (0.05), or high (0.25) discount rate was used based on the type of commodity being transported. Bulk natural resources have a low discount rate. Animals and intermediate processed goods have a medium discount rate. A high discount rate is applied for finished goods.

TABLE 6. MODE CHOICE AND INTERMODAL TRANSFER MODEL PARAMETERS.

Parameter	Description	Source
G_{mnql}	Logistics cost between shipper m and receiver n with shipment size q and logistics chain l	Calculated in the mode choice model
β_{0ql}	Alternative-specific constant	Asserted values based on commodity category
Q	Annual flow in tons	From goods demand model
q	Shipment size in tons	From shipment size model
β_1	Fixed cost per order	From research (Dominic 2009)
T	Transport and intermediate handling costs	From network skims
β_2	Discount rate	0.01/0.05/0.25 based on commodity
j	Fraction of shipment that is lost or damaged	0.01 assumed
v	Value of goods (per ton)	From FAF flow apportionment
β_3	Discount rate of goods in transit	0.01/0.05/0.25 based on commodity
t	Average transport time (days)	From network skims
β_4	Storage costs per cubic meter per year	From research (Colonial 2009)
β_5	Discount rate of goods in storage	0.01/0.05/0.25 based on commodity
LT	Expected lead time (days)	10 assumed
sdLT	Standard deviation in lead time (days)	1 assumed

Source: de Jong and Ben-Akiva (2007)

Estimation of these parameters was not possible without new data collection, but additional research on a few parameters led to revised assumptions for the Florida application:

Fixed cost per order. This value of $100 per order was obtained from Dominic (2009), who quoted the Supply Management Handbook of $100 of administrative expenses to generate a purchase order. Additional sources confirmed that this average was reasonable (APQC Performance Benchmarks, 2006) with an average $36 for top performers, $162 for median performers, and $507 for bottom performers.

- Storage costs per unit per year. This value of $2,000 per cubic meter per year depends on the physical properties of the commodity. In practice, this is not so much dependent on the weight of the goods but on their volume. This value was based on the assumption provided in the Colonial Diversified (2009) rate schedule of $6 per cubic meter per day after the first seven days, and $3 per cubic meter per day for first seven days, for a range of $1,095 to $2,190 per year.

Transportation and intermediate handling cost (T_{mnql}) is one of the main components of the logistics costs. In order to assess the parameters used in the equation, accurate values were identified in the literature. The transportation and intermediate handling cost in the logistics cost equation equals "annual flows" in tons multiplied by "transportation rate" in $/ton and can be shown as below:

$T_{mnql} = Q \times C$ = annual flows (tons) × transportation rate ($/ton)

The transportation rate term includes line-haul transportation rate and handling, lifting, warehouse/DC, or transload charges. As shown in Table 7, there are different line-haul costs by mode in the literature compared to the set of rates used for in application.

TABLE 7. TRANSPORTATION COST PARAMETERS.

Cost parameter by mode	Sources	Value ($/ton-mile)
Truck	Leachman (2005), EIA	0.080-0.100
Rail	Leachman (2005), EIA, CSX, UP	0.030
Air	Leachman (2005), UPS	3.750
Water	Leachman (2005), EIA	0.005

Table 8 presents a list of the level-of-service parameters. These are consistent with the original research presented by Leachman (2005), except for rail speeds, which have been reduced from 30 miles per hour to 22.5 miles per hour, based on research from the Bureau of Transportation Statistics (2008) and CSX (2013).

TABLE 8. LEVEL OF SERVICE PARAMETERS.

PARAMETER	DESCRIPTION	VALUE
BULKHANDFEE	HANDLING CHARGE FOR BULK GOODS ($ PER TON)	1.00
WDCHANDFEE	WAREHOUSE/DC HANDLING CHARGE ($ PER TON)	15.00
IMXHANDFEE	INTERMODAL LIFT CHARGE ($ PER TON)	15.00
TLOADHANDFEE	TRANSLOAD CHARGE ($ PER TON; AT INTERNATIONAL PORTS ONLY)	10.00
AIRHANDFEE	AIR CARGO HANDLING CHARGE ($ PER TON)	20.00
WATERRATE	LINE-HAUL CHARGE, WATER ($ PER TON-MILE)	0.005
CARLOADRATE	LINE-HAUL CHARGE, CARLOAD ($ PER TON-MILE)	0.03
IMXRATE	LINE-HAUL CHARGE, INTERMODAL ($ PER TON-MILE)	0.04
AIRRATE	LINE-HAUL CHARGE, AIR ($ PER TON-MILE)	3.75
LTL53RATE	LINE-HAUL CHARGE, 53 FEET LTL ($ PER TON-MILE)	0.08

FTL53RATE	LINE-HAUL CHARGE, 53 FEET FTL ($ PER TON-MILE)	0.08
LTL40RATE	LINE-HAUL CHARGE, 40 FEET LTL ($ PER TON-MILE)	0.10
FTL40RATE	LINE-HAUL CHARGE, 40 FEET FTL ($ PER TON-MILE)	0.10
WATERMPH	WATER SPEED (MPH)	5.00
RAILMPH	RAIL SPEED (MPH)	22.50
LHTRUCKMPH	LINE-HAUL TRUCK SPEED (MPH)	60.00
DRAYTRUCKMPH	DRAYAGE TRUCK SPEED (MPH)	45.00
AIRMPH	AIR SPEED (MPH)	500.00
EXPRESSSURCHARGE	SURCHARGE FOR DIRECT/EXPRESS TRANSPORT (FACTOR)	1.50
BULKTIME	HANDLING TIME AT BULK HANDLING FACILITIES (HOURS)	72.00
WDCTIME	HANDLING TIME AT WAREHOUSE/DCS (HOURS)	24.00
IMXTIME	HANDLING TIME AT INTERMODAL YARDS (HOURS)	24.00
TLOADTIME	HANDLING TIME AT TRANSLOAD FACILITIES (HOURS)	12.00
AIRTIME	HANDLING TIME AT AIR TERMINALS (HOURS)	1.00

Table 9 shows the path cost parameters used in the mode and transfer model. The initial application in Chicago was a demonstration and these parameters were asserted (Cambridge Systematics, 2011) rather than estimated. Further research and consideration resulted in recommended values for several parameters.

TABLE 9. PATH COST PARAMETERS RECOMMENDED FOR FLORIDA.

Parameter	Description	Initial Value	Recommended Value
a	Safety stock constant	0.50	0.5 to 2.33 varies by product type
sdQ	Standard deviation in annual flow	1	0.03 to 0.09 times the annual flow varies by product type
CAP1FTL	Truckload capacity (tons)	30	20-40 (assume 30)
CAP1Carload	Carload capacity (tons)	32	70-100 (assume 85)
CAP1Airplane	Air cargo hold capacity (tons)	1	10-40 (assume 25)

- **Safety Stock Constant.** This is a constant used to set the safety stock service level by assuming a fixed probability of not running out of stock. Details on inventory patterns and on product types are provided in Appendix B. The Safety Stock Constant depends on product type, supply chain type and service level, and product demand patterns, and it varies by commodity type (e.g., functional vs. innovative) as follows:
    - Low Multiplier of 0.5 for functional products based on 69% probability of not running out of stock.
    - Medium Multiplier of 1.0 for functional/innovative products based on 84% probability of not running out of stock.
    - High Multiplier of 2.33 for innovative products based on 99% probability of not running out of stock.

  Functional and innovative products are defined in Table 10.

TABLE 10. FUNCTIONAL AND INNOVATIVE PRODUCT DEFINITIONS.

FUNCTIONAL PRODUCTS	INNOVATIVE PRODUCTS
MATURE PRODUCT	EARLY LIFECYCLE STAGE
LOW PRODUCT VARIETY	HIGH PRODUCT VARIETY
PREDICTABLE DEMAND	UNPREDICTABLE DEMAND
MINIMIZE INVENTORY	DEPLOY SIGNIFICANT BUFFER STOCKS
GREATER RELIANCE ON LOW-COST MODES	GREATER RELIANCE ON FAST AND RELIABLE MODES

- **Standard Deviation in Annual Flow.** This also depends on product type, supply chain type and service level, and product demand patterns, and varies by commodity type (e.g., functional vs. innovative). Assuming a normal distribution for demand during lead time (Notteboom, 2011), the parameters are as follows:
    - Low Variability of 0.03 for functional products.
    - Medium Variability of 0.06 for functional/innovative products.
    - High Variability of 0.09 for innovative products.

  Assuming a perfect normal distribution of demand (with 1,000 tons demand per year, going as low as 700 and as high as 1300 tons), the standard deviation would be 150, or it can be about 0.1 to 0.33 of the average for low- and high-demand variable commodities. Table 10 defines functional and innovative products..
- **Lead time.** This is the expected lead time between ordering and replenishment and it varies by mode, which is assumed based on the commodity type, assuming a normal distribution (Leachman, 2005). For Florida, truck modes will be 1-4 days (assume 2.5 days), rail will be 3-10 days (assume 6.5 days), air will be 1-2 days (assume 1.5 days), and waterway will be 30-60 days (assume 45 days).

- **Standard deviation in lead time.** This varies by mode, which is assumed based on the commodity type, assuming a normal distribution (Leachman, 2005). For Florida, truck mode is 2 days, rail is 5 days, air is 1 day, and waterway is 20 days.
- **Truckload Capacity (tons).** The original assumption of 30 tons per truckload, found in the Leachman report (2005), was confirmed to be within the range of 20-40 tons per truckload expected, and it was kept for the Florida model.
- **Carload Capacity (tons).** This was increased from 32 tons per carload to 85 tons per carload to reflect larger carrying capacities of carloads. CSX (2012) reports boxcar and covered hopper capacity ranging from 70-100 tons.
- **Air Capacity (tons).** The original assumption of 1 tons per plane was increased to 25 tons per plane based on an analysis of technical specifications of Boeing aircraft (2013).

Model Application

Figure 16 shows a schematic of the mode and transfer choice model. The buyer-supplier pairs dataset now has information on buyer firm ID, supplier firm type ID, commodity type (SCTG), annual flow in tons and dollars, distribution channel, and the shipment size. An actual business is identified by randomly assigning a business from the pool of businesses that belong to the same firm type. Both the buyer and supplier TAZs are merged from the output of the firm location model. Modal skims developed are merged into the buyer-supplier pairs dataset.

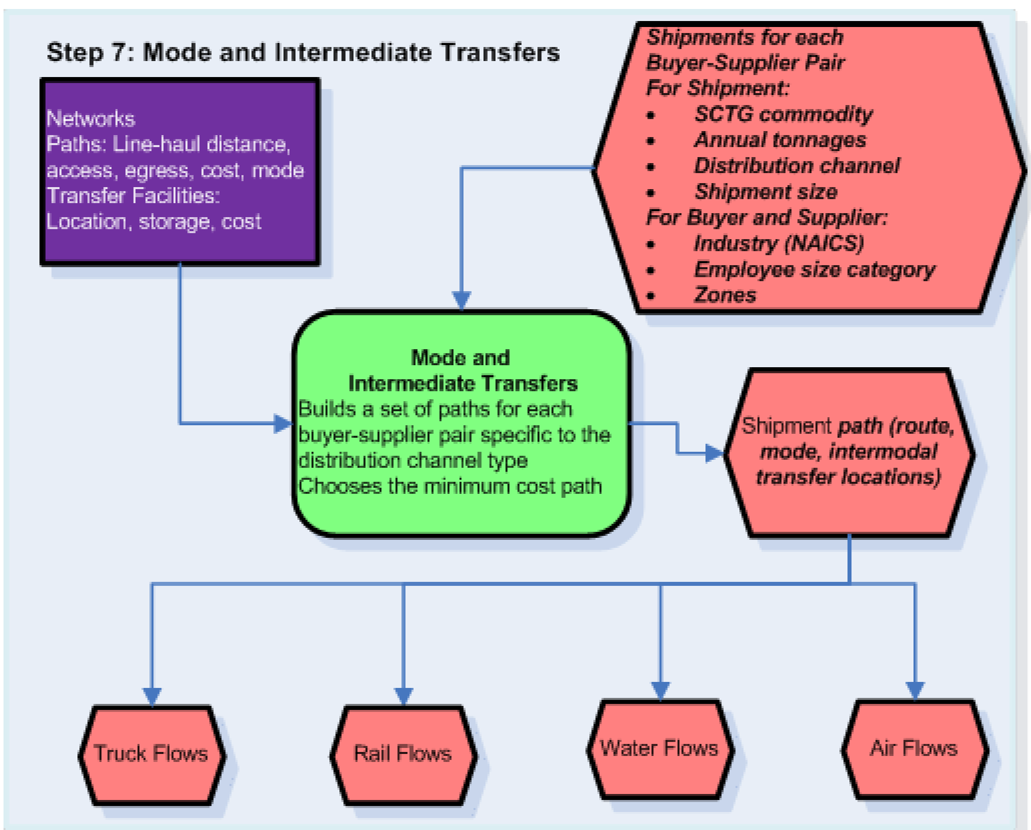


FIGURE 16. MODE AND INTERMEDIATE TRANSFER MODEL PROCESS.

5.3.7 TRIP TABLE CONVERSION

The mode and transfer model is the final freight demand component in the NSCM. This is the point in the model system where the standalone statewide model deviates from the complete integrated model system where truck trips are simulated using the FTTM.

The final part of the NSCM model focus on grouping shipments by mode from the results of the mode and transfer model and aggregating them into zone to zone movements. In the case of shipments moved by truck, those shipment movements are converted to zone to zone truck trips that can be assigned to the highway network.

The trip table conversion component follows a conventional approach to converting shipment flows to truck trips. The steps followed are:

- **Separate full shipment routing from mode and transfer model in to separate zone to zone trips**. In this step, the shipments by mode combinations like truck-rail-truck are split into separate modal trips, i.e. truck and rail trips in the case of truck-rail-truck. This requires using the zone that the intermodal transfer(s) occur in as new trip origins and destinations for the modal trips. The output is a set of shipment trips from an origin TAZ to a destination, described with variables including mode and the shipment's characteristics such as commodity, weight, and value.
- **Convert shipments trip to vehicle trips by size.** Payload factors that are based on distance and SCTG commodity are applied to the shipments to identify the number of truck trips that are required to deliver the shipment
- **Divide truck trips into vehicle classes.** The observed distribution of medium and heavy trucks in the truck counts is used to simulate a truck size for each truck trip
- **Add empty truck trips.** An allowance is made for some return empty trips based on empty factors that vary by SCTG commodity group. These are applied using simulation; if a trip is selected to have a return empty trip then this trip is added to the trip list as the reverse trips with zero weight and value and then origin and destination transposed.
- **Sampling from the annual trips to create a daily sample.** Up until this point, the trip list being manipulated is a representation of a full year's commodity movements and truck travel. A sample that represents an average day is created from the annual trip list. For annual frequency conversion to daily, a factor of 310 is used as recommended in NCFRP Report 8 (Cambridge Systematics, 2010). For daily frequencies of more than one (i.e. 310 or more trips per year), the daily frequency is rounded and otherwise sampling is used to identify whether or not the trip should be included in the daily sample.
- **Aggregate trips into trip tables.** The trip list (a database of individual trips, with one row per trip) is aggregated into a trip table, a simpler table with one row for each pair of TAZs and truck type, with the sum of the corresponding number of truck trips as the only data item. The model exports both the complete trip list as well as the trip table, which is at this point ready to be converted by Cube to a matrix and used in the trip assignment stage.

6.0 FREIGHT TRUCK TOURING MODEL

6.1 | FUNCTION WITHIN THE MODEL SYSTEM

The FTTM provides freight truck vehicle movements within the entire model region, providing the same spatial and temporal coverage as the CSTM, but covering a different market segment of commercial vehicle movements – those which involve the pick up and delivery of shipments at business establishments. In terms of market coverage, the FTTM should not overlap with the CSTM

6.2 | INTERFACES WITH OTHER MODEL COMPONENTS

The FTTM uses the synthetic establishments generated by the NSCM and also used by the CSTM as simulation agents. Travel time skims produced by the network assignment models will be used to inform some intermediate stop choices and routing choices. The shipment list produced by the NSCM provides the demand for vehicles movements used by the FTTM.

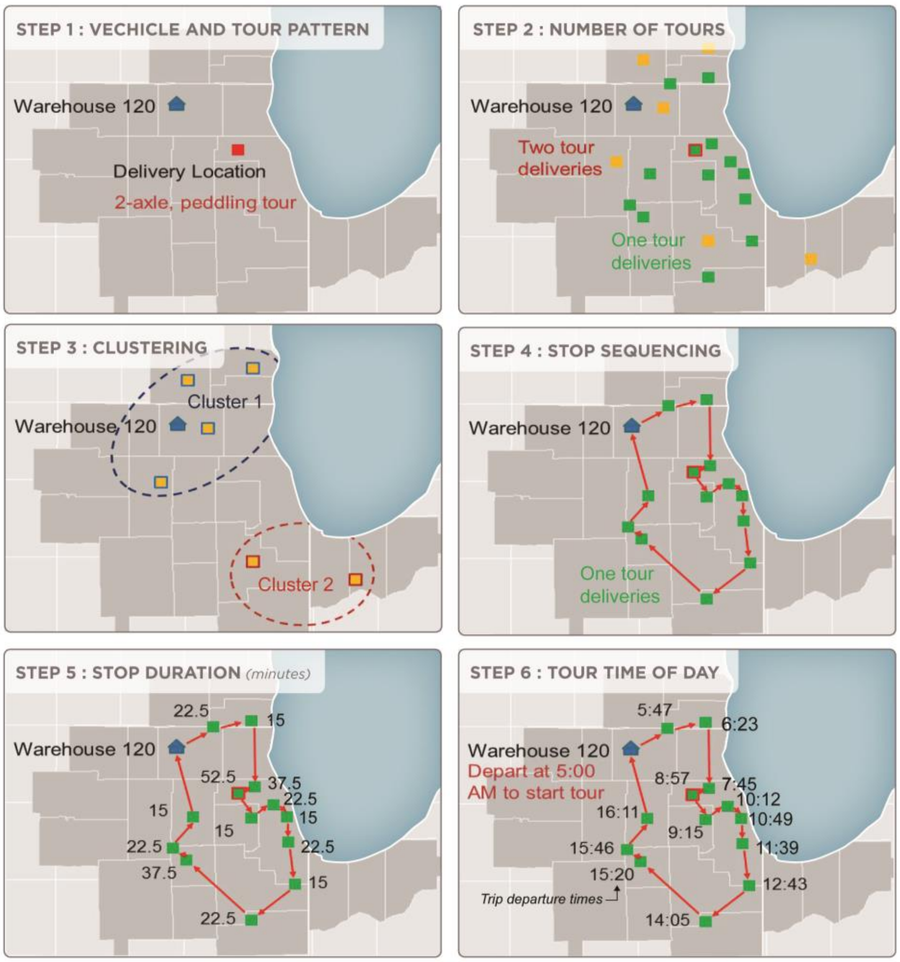
The FTTM will generate commercial vehicle trip lists in the same format as the CSTM. These trip lists subsequently are aggregated into time-period vehicle-specific trip tables for highway assignment. The trip tables generated by the FTTM could be combined with those generated by the CSTM and InSITE trip tables for the same vehicle types and time periods, or they may be kept as separate vehicle classes to reflect differences in value-of-time or other factors, should computing resources allow it.

6.3 | MODEL COMPONENTS

The FTTM is comprised of 6 main model components, shown in Figure 17. The intermediate stop models developed as part of the CSTM are being added to the FTTM (see 7.4.7). The FTTM picks up where the NSCM leaves off, with a list of zone-to-zone shipments, some of which are allocated to warehouse and distribution centers. The FTTM includes components that sequentially group shipments into tours, and add details to the descriptions of the tours such as stop sequence, stop duration, and tour start time. Each sub-model is described in detail below.

Vehicle and Tour-Pattern Choice. The vehicle and tour-pattern choice model (Step 1) is a multinomial logit model that predicts the joint choice of tour type and vehicle type and size. The tour type choice is whether a shipment will be delivered on a direct tour to the delivery location, or on a multi-stop tour in which the truck makes multiple deliveries or pick-ups. The size of the vehicle is determined jointly with tour pattern (direct or multi-stop), because certain vehicles support different types of tours. Commercial service vehicles, making service stops, would be handled similarly.

FIGURE 17: MODEL COMPONENTS IN THE FTTM



Number of Tours and Clustering (Step 2 and 3). This model is composed of two elements, the first being a multinomial logit model that predicts the complexity of the multi-stop tour in which a shipment or service stop is contained. For example, a truck might return to the start of its tour (e.g. origin distribution centre) after one large loop, or it might break its delivery/stop schedule into two, three, or more tours. The second element uses hierarchical clustering to divide the shipments or service stops into spatially co-located groups that can be reasonably served by the same truck during a tour.

Stop Sequence and Stop Duration (Step 4 and 5). For multi-stop tours, this model uses a greedy algorithm to sequence the stops in a reasonably efficient sequence, but not necessarily the shortest path. Next, a multinomial logit model predicts the amount of time taken at each stop based on the size and shipment commodity or, if the purpose is a service stop, by truck type.

Delivery Time-of-Day (Step 6).This is a multinomial logit model that predicts the departure time of the truck at the beginning of its tour or tours. Combining the chosen departure time with the travel time of each trip and the predicted duration of each delivery or service stop, all of the trips on the tour can be associated with a time period for assignment purposes. At this point, an iterative process is used to identify tours that are too long; overlong tours are then split into tours that meet time constraints (i.e., driver shift-length limits). After correcting for overlong tours, the vehicle tours can be converted to an aggregate zone-to-zone trip table that can be assigned using a network model.

6.3.1 DAILY SAMPLE

In this component, the transition is made to the regional touring model. The model takes a sample from the overall table of annual shipments for local BMC region, daily simulation.

Table 11 shows the input parameter that is used in the daily sample step of the model. There are no new inputs used in this model component.

TABLE 11: PARAMETERS USED IN THE DAILY SAMPLE STEP

PARAMETER	BASE VALUE	DESCRIPTION
annualfactor	310	Annualization factor to select a daily sample from the annual shipments

Model Steps

The daily sample component of the model applies the following steps:

1. A shipments database is created for just shipments to, from, or within the region.
2. Daily shipment frequency is calculated as the annual purchase quantity divided by the shipment weight, with the conversion to daily frequency made using the *annualization* parameter.
3. Not all of the firm pairs simulated in the NSCM have daily deliveriesand/or pickups. A random number between 0 and 1 is generated and firm pairs for which the random number is less than the daily frequency are assumed to have deliveries on the day being simulated.
4. Firm pairs with a daily frequency greater than one have daily deliveries and/or pickups.
5. From this point onwards only the selected firm pairs and corresponding shipments are simulated to represent a typical day scenario for deliveries and/or pickups in the region.

6.3.2 WAREHOUSE ALLOCATION

Once the daily sample has been taken, for those shipments that follow an indirect distribution channel via a distribution center, warehouse, or consolidation center the model assigns transfer points in the BMC region.

Model Steps

The warehouse allocation component of the model applies the following steps:

1. A local warehouse (within the modeling area) is randomly selected for indirect distribution channel shipments. This warehouse would represent the last transfer stop for incoming shipments and first transfer stop for outgoing shipments.
2. Based on the mesozone of the warehouse and whether or not a shipment is incoming or outgoing, the origin and destination mesozones of all the shipments that are to be simulated in a day are determined.

3. If the distribution channel is direct, the origin mesozone is the mesozone of the seller; if not, it is the warehouse mesozone (in the case of a drop-off a truck will start from the warehouse with a shipment and in the case of a pick-up, the truck will start from the warehouse without a shipment).
4. If the distribution channel is indirect and the buyer is outside the modeling area, the destination mesozone is the seller mesozone (to simulate a pick-up); if not, it is the buyer mesozone.

6.3.3 VEHICLE CHOICE AND TOUR PATTERN

This model component applies the vehicle choice and tour pattern model. This is a joint model that assigns each shipment to a particular size of vehicle and whether it will be picked up or delivered on a direct tour (i.e. one that includes only a single stop) or a peddling tour (one that includes multiple stops).

Table 12 shows the input tables that are used in the vehicle choice and tour pattern step of the model. The input tables are read into memory at the start of this step of the model and named as shown in the object name column of the table. Table 13 shows the input parameters that are used in the vehicle choice and tour pattern step of the model.

TABLE 12: INPUTS TO THE VEHICLE CHOICE AND TOUR PATTERN STEP

FILENAME	OBJECT NAME	DESCRIPTION
data_emp_cbpzone.csv	emp_cbpzone	Number of employees by zone
data_mesozone_emprankings.csv	mzemp	Employment ranking by industry by county
model_vehicle_tourpattern.csv	vehtourpat	Vehicle and tour pattern model variables and coefficients

TABLE 13: PARAMETERS USED IN THE VEHICLE CHOICE AND TOUR PATTERN STEP

PARAMETER	BASE VALUE	DESCRIPTION
wgtmax_2axl	35000	Maximum load weight of light duty truck in pounds
wgtmax_3axl	65000	Maximum load weight of medium duty truck in pounds
wgtmax_semi	100000	Maximum load weight of heavy duty truck in pounds

Model Steps

The vehicle choice and tour pattern component of the model applies the following steps:

1. There are six alternatives for this model, which are combinations of two tour patterns (direct and multi-stop) and three vehicle types (2 axle, 3-4 axle, and semi/trailer). The model is influenced by commodity type, destination type, pick-up/drop-off weight and county employment in the destination zone.
2. Additional variables required by the model are created, including dummy variables indicating whether the shipment is a food product or manufacture product, whether the delivery is a pick-up or drop off, the log transformed shipment weight, and dummy variables for the industry of the business establishment where the delivery or pick-up will take place.

3. County employment data is merged with the dataset of daily shipments to be simulated.
4. The logit prediction function is applied to predict the tour pattern and vehicle type choice for each daily shipment.
5. Shipments using a direct distribution channel are forced to have a direct tour pattern keeping the vehicle type choice same as that output by the simulation/prediction.
6. For very large total daily delivery amounts (i.e., the sum of all shipments in the sample day from the supplier to the buyer), the weight of just the shipments could exceed the capacity of the vehicle assigned by the mode. In those cases, the vehicle choice is moved up to the next vehicle size category that could successfully carry the load.
7. For even larger total daily delivery amounts, the weight of just the deliveries could exceed the capacity of any single truck to make in one delivery. In those cases, the shipments are split into separate shipments for individual delivery and are simulated separately in subsequent model steps.

Outputs

Table 14 below, show the outputs that are produced by the vehicle choice and tour pattern step of the model.

TABLE 14: MODEL OUTPUTS FROM THE VEHICLE CHOICE AND TOURS PATTERN STEP

FILENAME	TYPE OF OUTPUT	DESCRIPTION
vehtourpat_allcommodities.csv	CSV Table	Frequency of each alternative (combination of tour patterns and vehicle types) for all commodities
vehtourpat_food.csv	CSV Table	Frequency of each alternative (combination of tour patterns and vehicle types) for food commodity group
vehtourpat_mfg.csv	CSV Table	Frequency of each alternative (combination of tour patterns and vehicle types) for manufacturing commodity group

6.3.4 STOP SEQUENCE

In this component, additional detail is added to the structure of the delivery tours. The model applies the number of tours model, the clustering model, and the stop sequence model to identify the full set of stops that comprise each tour and the order in which they are made.

Table 15 shows the input tables that are used in the stop sequence step of the model. The input tables are read into memory at the start of this step of the model and named as shown in the object name column of the table. There are no new parameters used in this model component.

TABLE 15: INPUTS TO THE STOP SEQUENCE STEP

FILENAME	OBJECT NAME	DESCRIPTION
data_mesozone_centroids.csv	mz_centroids	Zone centroid coordinates
data_mesozone_skims.csv	mz_skims	Skims for origins and destinations with available path costs and times
model_numberoftours.csv	numberoftours	Number of tours model variables and coefficients

Model Steps

The stop sequence component of the model applies the following steps to determine the number of truck tours for each shipment, cluster the indirect shipments according to the number of truck tours category, and sequencing all the stops within a tour:

1. There are three alternatives for the category of number of truck tours: all (stops) in one tour, two tours, three tours, and four tours. All in one tour means the shipment belongs to a tour category in which a truck covers all the stops assigned to it in a single tour. The purpose of this is to be able to determine the number of tours for all the stops serviced from a particular warehouse.
2. The additional required variables to apply the number of tours model are created. The model requires pick-up/drop-off weights, buyer and supplier industry categories, etc.
3. The model is applied to the daily shipments dataset.
4. The shipment stops are now clustered based on the number of truck tours category predicted. This is done only for indirect shipments since direct shipments are serviced by a direct tour with a single stop.
5. A separate vehicle type variable is created from the tour pattern vehicle type choice.
6. The count of shipment stops is obtained by warehouse, number of tours category, and vehicle type. Vehicle type is also used here in the aggregation so that stops grouped/clustered together have the same vehicle type since they are assumed to be serviced by the same vehicle.
7. It is ensured that the count of shipment stops is at least as much as the number of tours in the corresponding category. For example, if a particular shipment falls in a 4-tour category, there should be three more shipments in the 4-tour category that are assigned to the same warehouse for consistency. If that is not the case, the number of tours category is modified to be consistent.
8. The count of shipment stops is obtained again by warehouse, number of tours category (modified), and vehicle type.
9. The X and Y coordinates for each mesozone are merged into the shipments data by destination mesozone (stop mesozone).
10. A function clusters the shipments for each warehouse by the number of tours category and vehicle type. For example, if a warehouse has five indirect shipment stops assigned to it and all of them fall in the two tours category using 2-axle truck, the clustering function would cluster the five stops in two clusters using the Euclidean method.
11. After all the shipments stops are clustered into specific tours, unique tour and trip IDs are assigned to all the records.

12. For sequencing all the stops within a tour, a greedy algorithm is applied. It assigns the first stop as the one that is closest to the base (warehouse) based on the travel times in the mesozone skims. It then keeps adding the next closest stop from the previous stop to the trip sequence until all of the stops are sequenced. After the final stop, the truck returns to the base as the last trip of the tour.
13. The truck load at each point during the tour is calculated based on the sequence of deliveries and the shipment weights.

Outputs

Table 16 below, show the outputs that are produced by stop sequence step of the model.

TABLE 16: MODEL OUTPUTS FROM THE STOP SEQUENCE STEP

FILENAME	TYPE OF OUTPUT	DESCRIPTION
stopseq_numstopspertour.csv	CSV Table	Frequency of each stop count category
stopseq_tourcatbyshipsize.csv	CSV Table	Frequency of tour category by shipment size
stopseq_stopcluster_sample.png	PNG File	Sample chart of stop clustering for each tour type category

6.3.5 STOP DURATION

Description

This model component applies the stop duration model to estimate the time spent at each delivery or pick-up activity.

Table 17 shows the input table that is used in the stop duration step of the model. The input table is read into memory at the start of this step of the model and named as shown in the object name column of the table. There are no new parameters used in this model component.

TABLE 17: INPUTS TO THE STOP DURATION STEP

FILENAME	OBJECT NAME	DESCRIPTION
model_stopduration.csv	stopduration	Stop duration model variables and coefficients

Model Steps

The stop duration component of the model applies the following steps:

1. There are six alternative stop durations in the stop durations model: less than or equal to 15 minutes, 15-30, 30-45, 45-60, 60-75, and greater than 75 minutes.
2. Additional variables required to apply the model are created in the shipments datasets.
3. The logit simulation function is used to predict the stop duration categories for both direct and indirect shipment datasets.
4. The exact stop duration in both datasets is assigned as the mid-point of the respective stop duration category:

a. 0-15 min = 0.25 hrs.
b. 15-30 min = 0.375 hrs.
c. 30-45 min = 0.625 hrs.
d. 45-60 min = 0.875 hrs.
e. 60-75 min = 1.125 hrs.
f. >75 min = 2 hrs.

Outputs

Table 18 below, show the outputs that are produced by stop duration step of the model.

TABLE 18: MODEL OUTPUTS FROM THE STOP DURATION STEP

FILENAME	TYPE OF OUTPUT	DESCRIPTION
stopdur_durationbytourtype.csv	CSV Table	Frequency of stop duration categories by tour types

6.3.6 TOUR TIME OF DAY

Description

During this step, the model applies the time of day model, which estimates the times that tours begin. Each individual trip in the tour are then given a start and end time.

Table 19 shows the input table that is used in the tour time of day step of the model. The input table is read into memory at the start of this step of the model and named as shown in the object name column of the table. There are no new parameters used in this model component.

TABLE 19: INPUTS TO THE TOUR TIME OF DAY STEP

FILENAME	OBJECT NAME	DESCRIPTION
model_timeofday.csv	tod	Time of day model variables and coefficients

Model Steps

The tour time of day component of the model applies the following steps:

1. The tour time of day multinomial logit model simulates the time of departure of the first trip in a particular truck tour. There are five alternatives for this – before 6 AM, 6-8 AM, 8-9 AM, 9-10 AM, and after 10 AM.
2. Additional variables required to apply the model, such as total tour and stop durations, are calculated for both direct and indirect tours.
3. Time of day categories are simulated for direct and indirect tour datasets. Based on the category of time of day an exact time is assigned to the tour start as follows:
   a. Before 6 AM = 5
   b. 6-8 AM = 7

c. 8-9 AM = 8.5
d. 9-10 AM = 9.5
e. After 10 AM = 10.5

4. For indirect/multi-stop tours, the start times for each of the individual trips are calculated by adding the travel duration to the stop duration at the current stop to the start time of the previous trip. Since the tour start time or the start time for the first trip is already simulated, start times for all other trips in the tour can be calculated.
5. The final portion of this model component reviews the tours against two constraints and split tours where they do not meet the constraint:
    a. Vehicle loading: at no point during the tour should the load on the vehicle exceed the capacity of the vehicle. Tours are divided to meet the vehicle's weight capacity
    b. Tour start time: the threshold for the latest trip start time was set as 10 PM. Shipments flowing in and out of the modeling area had been randomly allocated to the warehouses in the area. In certain cases, due to the number of shipments to be delivered/picked up and/or due to the locations of shipments, it is found that certain tours had trips starting very late in the day. Since it is unlikely that such trips occur, these tours are broken down further to bring all the trip start times within a reasonable range.
6. The number of new tours into which these tours are broken down is determined by the greater of the number of 8-hours periods within the total tour time of the tours and the ratio of the maximum load on the vehicle during a tour to its capacity. For example, if the total tour time for a particular tour (which has a trip starting later than 10 PM) is 12 hours, it is split into two tours and a 19-hour tour is split into three tours. Similarly, a tour with a maximum load of 50,000lbs using a vehicle with a capacity of 35,000lbs will be split into two tours. The split tours could be interpreted as a new driver/vehicle added to the set of tours from a warehouse.
7. The vehicle type is kept the same as in the original/parent tour. All the stops are then clustered using the clustering algorithm into the number of sets that the tour is to be split into. Each cluster of stops is then sequenced using the greedy algorithm. The tour start times of the split tours are kept the same as that of the original tour. The individual trip start times are recalculated to reflect the new split tour start times.
8. Steps 5 through 7 are repeated until all trips start before 10 PM and at all points on all tours, vehicle are loaded at or below the vehicle capacity.

Outputs

Table 20 below, show the outputs that are produced by tour time of day step of the model.

TABLE 20: MODEL OUTPUTS FROM THE TOUR TIME OF DAY STEP

FILENAME	TYPE OF OUTPUT	DESCRIPTION
tod_todbytourtype.csv	CSV Table	Frequency of time of day categories by tour types

6.3.7 PREPARE TRIP TABLES

Description

This final component of the model aggregates the individual trips that form the pick-up and delivery tours into zone-to-zone trip tables by time period.

Model Steps

The prepare trip tables component of the model applies the following steps:

1. Text descriptions of vehicle types and times of day are added to the direct and indirect tour datasets
2. The direct and indirect tour datasets are combined together into a single tours dataset.
3. The tours dataset is aggregated by origin TAZ, destination TAZ, vehicle type, and time of day to produce the final trip tables.

Outputs

Table 21 below, show the outputs that are produced by prepare trip table step of the model.

TABLE 21: MODEL OUTPUTS FROM THE PREPARE TRIP TABLE STEP

FILENAME	TYPE OF OUTPUT	DESCRIPTION
trip_table.csv	CSV Table	Number of trips by origin TAZ, destination TAZ, vehicle type, and time of day

7.0 COMMERCIAL SERVICES VEHICLE TOURING MODEL

7.1 | FUNCTION WITHIN THE MODEL SYSTEM

The CSTM provides commercial vehicle movements within the entire model region, providing the same spatial and temporal coverage as the FTTM, but covering a different market segment of commercial vehicle movements. In terms of market coverage, the CSTM should not overlap with either the FTTM or the passenger demand model, InSITE.

- To avoid overlap with the FTTM, the CSTM will generate tours for which the primary activity stops are the provision of commercial services, rather than freight pickup or delivery. In addition, the model will allow stops for the pickup or delivery of goods, such as appliances, materials and equipment, which are typically part of service-oriented businesses such as home remodeling services, HVAC services and landscapers.
- In addition, the CSTM will cover goods pickup and delivery to private residences such as parcel delivery and waste collection, which the FTTM does not cover, and these home-delivery stops may constitute the primary stops on tour. The FTTM provides only for the transport of commodities between businesses.
- The CSTM will generate tours with stops for the purpose of business meetings or sales calls at customer sites. The model was developed using establishment survey data (described below) in which only persons whose jobs involved daily driving were sampled. Some of their stop purposes included business and sales meetings performed in the context of a customer service relationship, such as when a contractor visits a home or other site to make an develop a cost estimate or as a follow up inspection of work already performed. In this context, these stop purposes are unlikely to overlap to a large extent with InSITE's work-related travel purposes.

7.2 | INTERFACES WITH OTHER MODEL COMPONENTS

The CSTM requires land use data—households and employment— at the zonal level, which it uses to generate demand. It also uses the synthetic establishments generated by the NSCM and also used by the FTTM as simulation agents. Travel time skims produced by the network assignment models will be used to inform some intermediate stop choices and routing choices.

The CSTM will generate commercial vehicle trip lists in the same format as the FTTM. These trip lists subsequently are aggregated into time-period vehicle-specific trip tables for highway assignment. The trip tables generated by the CSTM could be combined with those generated by the FTTM and InSITE trip tables for the same vehicle types and time periods, or they may be kept as separate vehicle classes to reflect differences in value-of-time or other factors, should computing resources allow it.

7.3 | KEY ASSUMPTIONS

The first important assumption underlying the model design is that the demand for commercial services is customer-driven. Although the model will utilize synthetic establishments to provide these services and these establishments will have some number of employees, the assumption is that they will staff drivers/technicians to meet the demand for commercial services, both in the base year and the future. Accordingly, the amount of service and related stops generated by each establishment will be partially a function of regional employment and households and partially a function of the number of employees in the synthetic establishment. This approach is an advantage in simulating future-year scenarios, because information on both future service-oriented establishments (supply) and the future customer base (demand) are used to generate commercial vehicle movements.

The second important assumption is that the market segmentation assumptions, described above, will provide full commercial vehicle movement coverage, without overlapping the FTTM or InSITE model coverage. To evaluate how well the model performs in this regard, it will be necessary to validate the CSTM results simultaneously with those of the FTTM and InSITE. Comparisons should be made to traffic counts by vehicle class. Heavy and medium vehicle movements will be represented by both FTTM and CSTM outputs. Light-duty vehicle movements, essentially equivalent to passenger cars and light trucks, will be generated by both CSTM and InSITE.

A third important assumption is that the parameters of the various model components, described below, can be estimated using truck diary data from the 2004 Ohio General Establishment Survey, and that these parameters are transferable to Maryland. This survey has been used to develop a commercial vehicle model for the Ohio Department of Transportation and has good representation of non-freight commercial movements. It is anticipated that the plan for future development of the freight model system will include a local, Maryland-based establishment survey that may be use to refresh the model. The plan for continuing development of the freight model is a deliverable for a later task in this project.

7.4 | MODEL COMPONENTS

Nine sub-models comprise the complete CSTM. The overall model flow is depicted below in Figure 7-1, and each sub-model is described in detail below.

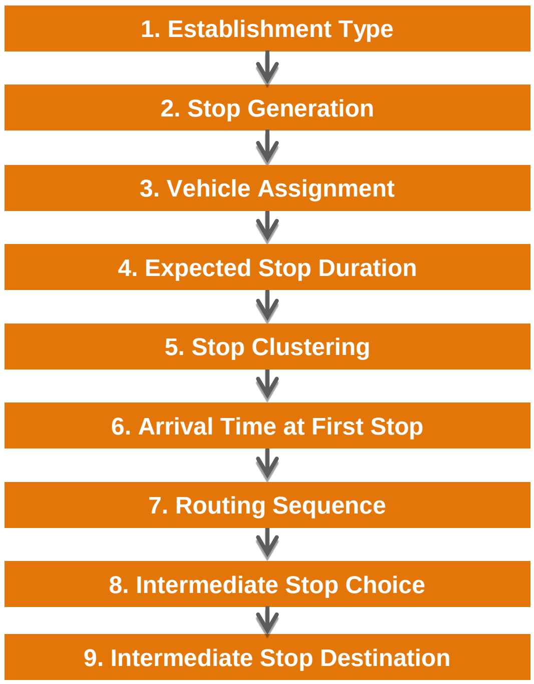


FIGURE 7-1: COMMERCIAL SERVICES TOURING MODEL FLOW DIAGRAM

7.4.1 ESTABLISHMENT TYPE

The establishment type model tags each synthetic establishment in the modeling region with an industry type and a label that indicates whether the establishment is a goods producer, service provider, or do both. The model applies a Monte Carlo simulation method to draw from observed distributions of establishments by industry type and stop type, constructed using truck diary data. In CSTM:

A. Establishments tagged as "goods only" producers are permitted to generate goods delivery and pickup only as a function of residential demand. As mentioned above, the aim is to avoid overlap with the business-to-business freight movements covered by the FTTM. To achieve this, establishments that belong to this category but are part of manufacturing and wholesale industries are filtered out after establishment type model is applied.
B. Establishments tagged as "service only" providers are permitted to generate service stops as the primary stops on tours, for both private residences and businesses. In addition, they are allowed to make goods stops for the purpose of picking up and dropping off materials and equipment that are a necessary part of many service providers, and to make meeting stops, such as when a service makes an initial inspection and cost estimate prior to actually providing a service.
C. Establishments tagged as "goods and service" providers are permitted to generate tours without service stops, if the goods stops are generated by residential demand. In addition, they are permitted to generate service stops as the primary stops on tours, for both private residences and businesses, as well as stops for the purpose of picking up and dropping off materials and equipment and meeting stops

7.4.2 STOP GENERATION

The stop generation model predicts one day's worth of scheduled stops for each establishment by TAZ, using a count model formulation. Count models predict positive integer values for the frequency of an event. For the current stop generation model, based on an analysis of the truck diary data, scheduled stops are grouped in to three market segments: goods stops, service stops, and meeting stops. For each market segment, the stop generation model predicts the number of scheduled stops for establishments as a function of zonal households and travel impedance as well as each establishment's industry type and size. For each TAZ within a specified service radius of an establishment (determined empirically), the model predicts some number of stops of the specified type. For a given TAZ, zero stops could result either because the establishment does not serve that TAZ, or because there is simply no demand for that day. Because many TAZs are likely to generate zero stops, the model uses a count model formulation that accounts for excessive zeros and heterogeneity. Specifically, a two-stage zero-truncated count model of the following form is formulated:

a. A Hurdle model that predicts 0 and 1+ stops

b. For 1+ stops, a zero-truncated count model that predicts the number of stops

The Hurdle models use binomial distributions, while the zero-truncated count model for goods uses a negative binomial distribution and the zero-truncated count model for service and meetings use Poisson distributions, after specification testing (more on this in Model Estimation Results chapter).

7.4.3 VEHICLE ASSIGNMENT

For each stop, the vehicle assignment model assigns one of three commercial vehicle types. As shown below in Figure 7-2, these correspond to light (i.e., car, van, and pickup), medium/single-unit, and heavy/multi-unit truck types. The model is formulated as a multinomial logit model and predicts vehicle type as a function of the establishment's industry type, distance between establishment and stops to be served, and the stop's purpose—goods, services, or meeting, after specification testing.

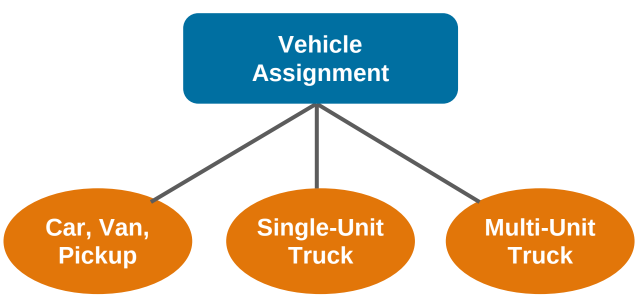


FIGURE 7-2: VEHICLE ASSIGNMENT MODEL STRUCTURE

Expected Stop Duration

The expected stop duration model is applied to scheduled stops generated by the stop generation model as well as any intermediate stops on the tour generated downstream (intermediate stop generation is discussed later in this chapter). For each stop, expected stop duration is drawn from a smoothed, empirical distribution of observed stop distributions for each stop type, based on truck diary data. Duration distributions are segmented by establishment industry type and stop purpose, after specification testing.

7.4.4 STOP CLUSTERING

For each establishment, the stop clustering model groups scheduled goods, service, and meeting stops into feasible commercial vehicle tours, based on spatial proximity, vehicle type, total travel time, and expected stop duration. The spatial proximity-based hierarchical clustering algorithm used in the FTTM has been modified to incorporate these additional criteria.

7.4.5 ARRIVAL TIME AT FIRST STOP

For each cluster of stops forming a tour, the arrival time at the first stop on the tour is predicted using a multinomial logit model. Based on an analysis of truck diary data, morning and afternoon arrival times (05:01am to 05:00pm) are modeled at half-hour interval, while evening and night arrival times (05:01pm to 05:00am) are modeled at one-hour interval, as presented below in Figure 7-3. The model is formulated as a function of the scheduled stop durations, after specification testing. The arrival and departure times at other stops on the tour, including the establishment-departure time for the first trip, are determined after the remainder of the tour is sequenced and any intermediate stops are added, along with their durations. Travel times are determined from skims.

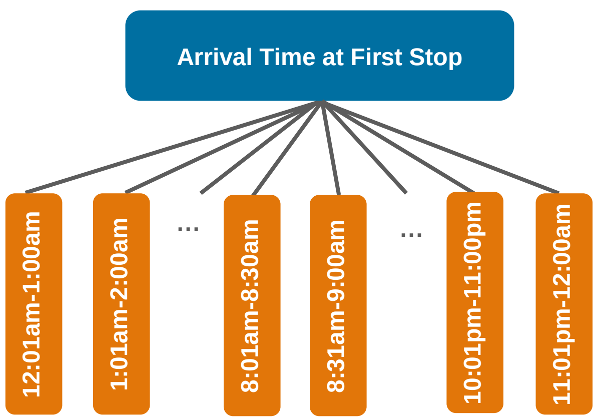


FIGURE 7-3: ARRIVAL TIME AT FIRST STOP MODEL STRUCTURE

7.4.6 ROUTING SEQUENCE

Given a set of scheduled stops and their locations on a tour, establishment location, and the time of the day, the CSTM uses a "traveling salesman problem" algorithm to determine the sequence of stops on the tour. The sequencing accounts for scheduled goods, service, and meeting stops only, not intermediate stops, which are added as subsequent steps. In application, the model is designed in such a way as to select the correct skims for the time period, building from first arrival time and updating clock time using trip travel times and durations.

7.4.7 INTERMEDIATE STOP CHOICE

The intermediate stop choice model predicts whether there are any intermediate stops between scheduled stops on a tour. The model simulates whether the driver makes one or more intermediate stops prior to each scheduled goods or tour stop, or prior to returning to the establishment to complete the tour. Purposes for intermediate stops include breaks/meals, vehicle service/refueling, and personal business/other. Note that tours entirely composed of intermediate stops cannot occur in the current implementation. Intermediate stops are only added to tours with at least one goods, service, or meeting stop.

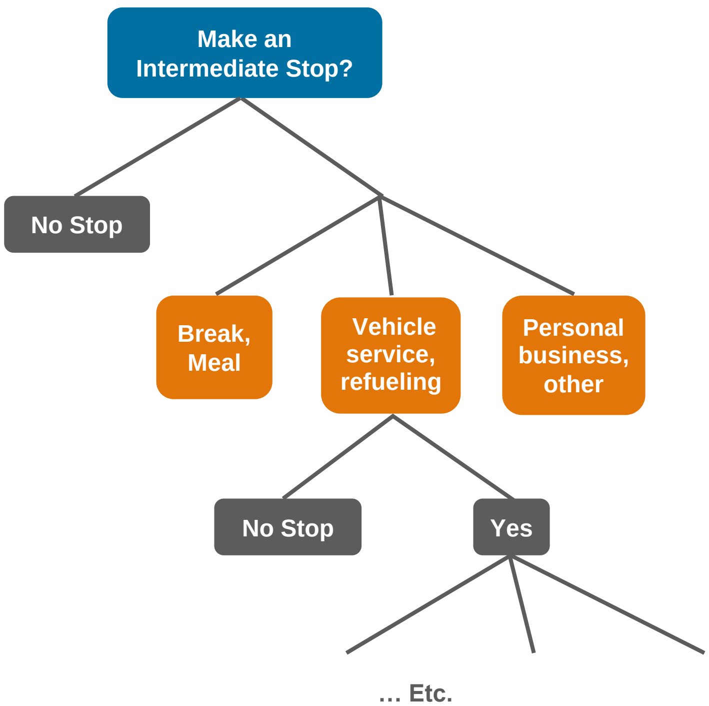


FIGURE 7-4: INTERMEDIATE STOP GENERATION MODEL STRUCTURE

The model structure is shown in Figure 7-4. It is applied to each stop in the order determined by the stop-sequencing model. Given a current location and next scheduled stop, the model predicts whether there will be an intermediate stop and, if so, for which purpose. If the answer is "no stop," then none is added prior to the next scheduled stop. If the answer is "yes," then one of three intermediate stop purposes is chosen and the model is run again, as indicated by the lower set of choices in Figure 7-4. If "no stop" results, then only the first intermediate stop will be added to the tour. If "yes" results, then another stop purpose will be chosen and this stop, too, will be added to the tour. The model continues until no additional stops are added.

Based on truck diary data, it is anticipated that when the model is applied the "no stop" alternative will result in the majority of cases. In addition, the model is history-dependent, making the propensity to add stops to a tour less likely if intermediate stops have already been made. For example, if vehicle service/refueling was chosen on a previous intermediate stop, then this alternative becomes unavailable for the subsequent intermediate stops. Remaining time on the tour, time of day, distance between establishment and the current stop, and a variable that indicates whether the current stop is the establishment and also the first stop of the tour are considered, after specification testing.

7.4.8 INTERMEDIATE STOP DESTINATION

For each intermediate stop on a tour, this model predicts a destination TAZ. Specifically, for each intermediate stop, the model selects a set of "eligible" TAZ based on attraction factor(s) and an impedance factor that accounts for travel distances. As shown below in Figure 7-5, this model uses a "rubber banding" method that considers the travel distance from the current stop to each alternative destination and from each alternative destination to the next scheduled stop or returning to the establishment. The idea is to minimize deviations from direct paths between scheduled stop locations. The model applies a maximum-distance threshold for deviation, determined empirically from truck diary data and considering only those TAZs that fall within the boundary set by the maximum-distance limit. For each intermediate stop, the set of "eligible" TAZs is refined further by retaining only those TAZs that have positive attraction factor(s). In a simple implementation of the model, retail employment may be used as an attraction factor for break/meal and vehicle service/refueling, and total household may be used for personal business/other stop type. The attraction factor associated with each TAZ is treated as a weight and is used to calculate relative weight of each TAZ in the choice set, as presented below in Equation (1). Finally, values of r_k are used to calculate cumulative relative weights and are used in Monte Carlo simulation to select an intermediate stop destination from the choice set.

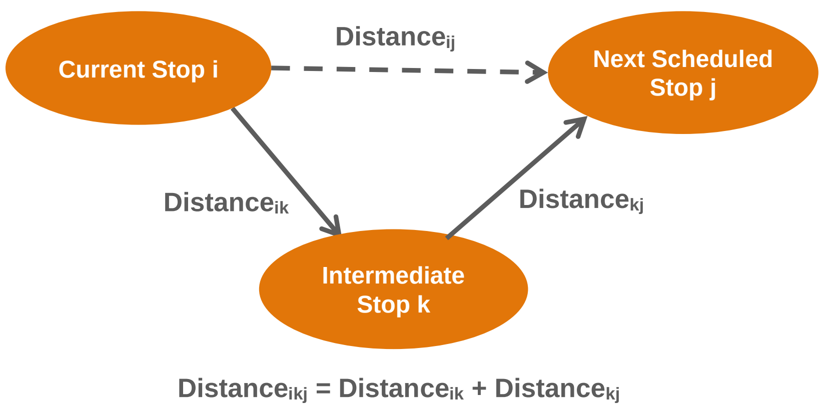


FIGURE 7-5: CALCULATING TRAVEL DISTANCES OF INTERMEDIATE STOPS

Relative weight associated with TAZ *k*, $r_k = \frac{w_k}{\sum_{i=1}^{N} w_i}$ (1)

Where w_k is attraction factor for TAZ *k* and *N* is number of TAZs in the choice set.

Importantly, the departure time needed to arrive at the first scheduled stop on the tour will take into account congested travel times, derived from travel model skims and the expected durations of any intermediate stops, which are determined by upstream model components. To illustrate, let us assume that a driver starts the first tour of the day at the establishment and makes an intermediate stop for vehicle refueling before arriving at the first scheduled stop at 8:00 a.m. Let us further assume that the a.m. peak travel time between the establishment and the first scheduled stop is 30 minutes and intermediate stop duration (predicted by the expected duration model) is 10 minutes. In the current model, the driver takes all these into consideration and adjusts departure time from the establishment to 7:20 a.m. to start the tour.

7.5 | ESTIMATION DATA DEVELOPMENT

The data used for the estimation of CSTM sub-models is drawn from the 2003–2004 Ohio Statewide General Establishment Survey (GES survey hereafter). As part of the GES, 24-hour diaries of sampled employees who drove every day to perform their jobs were collected. Collected information include start location of the first trip of the day, departure time for the first trip of the day, arrival and departure time at each subsequent stops, stop locations, activities at the stops, and vehicle used (car, small van, single-unit, truck, and multi-unit truck). In addition, data on the establishments was also collected including establishment name, location, industry type, primary activity (goods producer and/or service provider), primary commodities, number of employees, number of employees who travel as part of their work, and number and type of vehicles owned by the establishment.

The diary data included travel information at stop-level. Thus, the data needed to be processed in several steps to obtain the sample for the current study. A brief description of the steps involved in the sample formation is provided below.

- Using establishment location as an anchor point, stop-level data for each driver were converted to a sequence of trips/tours undertaken by each driver on the diary day.
- Next, each establishment was identified as either a goods producer, service provider, or both. Industries that are likely to be covered by FTTM model such as metal, light industry, heavy industry, and transportation equipment were combined to a form a single manufacturing industry.
- Activities undertaken at the stops were used to classify each stop as either a service, goods, meeting, break/meal, vehicle service/refueling, or personal business/other stop. If a stop belonged to one of the first three categories then it was labeled as a scheduled stop. Similarly, if a stop belonged to one of the last three categories then it was labeled as an intermediate stop.
- Next, a series of screening and consistency checks were performed and records with missing or inconsistent data were eliminated
- Finally, a number of land use information and appropriate network level of service data were appended to the sample data set.

The final sample includes data from 440 establishments, as presented in Table 7-1. The sample includes establishments from a wide variety of industries. The common industries are agriculture/forestry/fishing, wholesale, construction, government, and other services. Though data from all 440 establishments was used to construct distributions for the establishment type model, the establishment records highlighted in yellow in Table 7-1 were not included in the subsequent models. Table 7-2 provides the number of observations available for each stop type and corresponding average duration. The most common stop type is goods, followed by service and meeting. Less than 15% stops are intermediate stops. As expected, service stops have the longest average duration and goods stops have the shortest average duration among the scheduled stops. Regardless of stop type, the average duration of intermediate stops are significantly shorter than the scheduled stops.

TABLE 7-1 NUMBER OF ESTABLISHMENTS BY INDUSTRY AND PRIMARY ACTIVITY

Industry	Goods producer	Service provider	Both goods producer and service provider	Total
Agriculture/Forestry/Fishing	8	24	26	58
Manufacturing (metal, light, heavy, and transportation equipment)	23	1	13	37
Wholesale	25	6	37	68
Retail	11	15	11	37
Hotel & Real Estate	3	15	3	21
Construction	2	26	26	54
Health	5	27	7	39
Transportation Handling	3	9	5	17
Other Services	3	34	21	58
Government	4	24	23	51
Total	87	181	172	440

TABLE 7-2 STOP TYPE

Stop type	Number of observations	Average activity duration (in minutes)
Service	2384	84
Goods	4106	25
Meeting	1440	61
Break/meal	360	41
Vehicle service/refueling	147	19
Personal business/other	697	60

7.6 | MODEL ESTIMATION RESULTS

7.6.1 ESTABLISHMENT TYPE

This is the first model in the CSTM system and is use to predict the whether the primary activity of the establishment is producing goods, providing service, or both producing goods and providing service.

Table 7-3 presents observed distributions of industry type by the primary activities of the establishments. To ensure that the distributions are constructed using a reasonable number of observations, the metal, light industry, heavy industry, and transportation equipment industries were combined into a single manufacturing industry. As discussed previously, goods producing establishments that belong to manufacturing or wholesale industries are not covered by the CSTM model. The model does not require any estimation and can be implemented in a straightforward manner by using a Monte Carlo simulation method to draw from observed distributions.

TABLE 7-3 DISTRIBUTION OF INDUSTRY BY PRIMARY ACTIVITIES OF ESTABLISHMENTS (SOURCE: 2003–2004 OHIO STATEWIDE GENERAL ESTABLISHMENT SURVEY)

Industry	Goods producer	Service provider	Both goods producer and service provider	Total
Agriculture/Forestry/Fishing	0.1379	0.4138	0.4483	1.0000
Manufacturing	0.6216	0.0270	0.3514	1.0000
Wholesale	0.3676	0.0882	0.5441	1.0000
Retail	0.2973	0.4054	0.2973	1.0000
Hotel & Real Estate	0.1429	0.7143	0.1429	1.0000
Construction	0.0370	0.4815	0.4815	1.0000
Health	0.1282	0.6923	0.1795	1.0000
Transportation Handling	0.1765	0.5294	0.2941	1.0000
Other Services	0.0517	0.5862	0.3621	1.0000
Government	0.0784	0.4706	0.4510	1.0000

7.6.2 STOP GENERATION

The Stop Generation model predicts the number of scheduled stops of a given type (goods, service, and meeting) for each establishment and TAZ combination. A Hurdle model for each type of stop was estimated using the Ohio GES data. The Hurdle model is useful for regressions on count data with an abundance of zeros (Zeileis et al. 2008). The model structure is two-fold: a Stop/No-Stop binary classification and a truncated count model for the number of stops given that a stop was made. The predicted number of stops takes on a positive integer value, or zero if no stops are made. These two sub-models are jointly estimated via maximum likelihood estimation. The assumed distributions are summarized in Table 7-4.

Model validation and sensitivity testing may identify the need to constrain the stop zone consideration sets for establishments. The Ohio GES area is much larger than the Baltimore modeling area, and establishments in the Baltimore modeling area may not consider zones as distant as establishments in the Ohio GES. Establishment and zone pairs that exceed a distance threshold may be removed from the stop generation process to avoid unrealistically long trips. This threshold can be calibrated to validation data for the Baltimore area. Additionally, the Hurdle models themselves can be calibrated to match observed stop distribution data by adjusting the zero and count sub-models' intercept terms.

TABLE 7-4: HURDLE MODEL SPECIFICATION

	Goods Stops	Service Stops	Meeting Stops
Stop/No-Stop Distribution	Binomial	Binomial	Binomial
Stop/No-Stop Link Function	Cauchit	Logit	Logit
Count Model Distribution	Negative Binomial	Poisson	Poisson

TABLE 7-5: ZONE EMPLOYMENT INDUSTRY CATEGORIES

NAICS Zone Employment Categories	Model Zone Employment Categories
Agriculture, Forestry, Fishing and Hunting	Industrial
Mining	Industrial
Utilities	Industrial
Construction	Industrial
Manufacturing	Industrial
Manufacturing	Industrial
Manufacturing	Industrial
Wholesale Trade	Industrial
Transportation	Industrial
Warehousing	Industrial
Retail Trade	Retail
Retail Trade	Retail
Information	Office
Finance and Insurance	Office
Real Estate	Office
Professional, Scientific and Technical	Office
Management of Companies and Enterprises	Office
Government	Office
Educational Services	Education
Health Care	Medical Services
Administrative, Support, Waste Management	Other Services
Arts, Entertainment and Recreation	Other Services
Hotels	Other Services
Other Services	Other Services

Travel time from establishment to zone, establishment employment, zone employment, establishment industry, toll cost, and number of zone households were used in predicting the number of goods, service, and meeting stops between the establishment and zone pair. Table 7-5 shows the employment categories (NAICS) used to map the Ohio GES definitions of employment groups onto the definitions used in BMC's InSITE model. Manufacturing establishments making tours composed entirely of goods-related stops were excluded from the modeling as these are likely freight trips and are already captured in the freight model. However, manufacturing establishments making tours composed of a combination of goods, service, and meeting stops were included. Table 7-6 to Table 7-8 contain the final estimated parameters for the three models and their sub-models.

TABLE 7-6: GOODS HURDLE MODEL RESULTS

Zero Hurdle Variable	Coef.	Z-Stat	Count Model Variable	Coef.	Z-Stat
Intercept	78.54	6.94	Intercept	-15.89	-0.02
Industry			Industry		
Agriculture	0	-	Agriculture	0	-
Manufacturing	0.39	0.19	Manufacturing	0.14	0.48
Wholesale	4.37	4.11	Wholesale	-0.01	-0.05
Retail	-2.87	-0.85	Retail	0.37	1.26
Hotel & Real Estate	-4.34	-0.99	Hotel & Real Estate	-0.06	-0.12
Construction	-1.56	-1.48	Construction	-0.06	-0.23
Health	-2.09	-0.36	Health	0.43	1.12
Transportation Handling	8.59	4.30	Transportation Handling	0.42	1.59
Other Services	0.11	0.09	Other Services	0.04	0.15
Government	-3.86	-2.42	Government	0.85	3.15
Establishment Emp.	0.27	4.79	Establishment Emp.	0.08	16.63
Zone Employment			Zone Employment		
Office + Medical Services (000's)	1.29	2.36	Industrial + Retail (000's)	0.11	5.46
Retail + Industrial + Education (000's)	0.94	2.36	Office + Medical Services + Education (000's)	0.06	1.92
Toll (dollars)	-0.54	-3.91	Log(Travel time from establishment)	-0.75	-13.03
Log(Travel time from establishment)	-16.85	-7.02	Log(Theta[6])	-16.63	-0.02
$LL_0 = -6783$					
$LL_f = -2519$					
$Adjusted\ \rho^2 = 0.624$					

[6] Estimated shape parameter for the negative binomial distribution.

TABLE 7-7: SERVICE HURDLE MODEL RESULTS

Zero Hurdle Variable	Coef.	Z-Stat	Count Model Variable	Coef.	Z-Stat
Zero Hurdle (binomial with logit link)			Intercept	1.51	13.47
Intercept	20.06	9.24	Industry		
Establishment Emp.	0.11	4.31	Agriculture	0	-
Zone Employment			Manufacturing	-0.33	-1.48
Retail (000's)	1.90	2.40	Wholesale	-1.67	-9.46
Office + Medical Services + Education (000's)	0.08	0.44	Retail	-1.68	-4.54
Toll (dollars)	-8.94	-4.36	Hotel & Real Estate	-1.57	-1.60
Log(Travel time from establishment)	-4.49	-9.83	Construction	-0.57	-6.74
			Health	-1.39	-7.05
			Transportation Handling	-0.83	-2.18
			Other Services	0.15	2.00
			Government	-0.48	-5.80
			Retail x Total Households (000's)	0.45	3.76
			Health x Total Households (000's)	0.39	5.72
			Establishment Emp.	0.04	13.37
			Zone Employment		
			Retail (000's)	0.12	4.16
			Office + Medical Services + Education (000's)	0.02	1.82
			Log(Travel time from establishment)	-0.50	-16.64
$LL_0 = -3962$					
$LL_f = -2133$					
$Adjusted\ \rho^2 = 0.456$					

TABLE 7-8: MEETING HURDLE MODEL RESULTS

Zero Hurdle Variable	Coef.	Z-Stat	Count Model Variable	Coef.	Z-Stat
Intercept	20.24	9.42	Intercept	-0.55	-3.19
Industry			Establishment Emp.	0.06	18.79
Agriculture	0	-	Zone Employment		
Manufacturing	1.00	0.93	Total Employment (000's)	0.03	4.28
Wholesale	3.30	4.19	Log(Travel time from establishment)	-0.29	-6.33
Retail	0.86	0.85			
Hotel & Real Estate	-0.46	-0.28			
Construction	1.47	1.79			
Health	0.52	0.42			
Transportation Handling	-2.37	-0.34			
Other Services	2.24	2.71			
Government	0.99	1.18			
Establishment Emp.	0.03	1.87			
Zone Employment					
Total Employment (000's)	0.58	4.95			
Log(Travel time from establishment)	-4.84	-10.54			
$LL_0 = -2631$					
$LL_f = -1027$					
$Adjusted\ \rho^2 = 0.603$					

7.6.3 VEHICLE ASSIGNMENT

The vehicle assignment model assigns either a light (i.e., car, van, pickup), medium, or a heavy truck to each scheduled stop. The final model estimation results, obtained after specification testing, are summarized below in Table 7-9. The base alternative in the model is heavy truck (i.e., coefficient =0), and the coefficients for other alternatives are estimated with respect to this base alternative. The main implications of the estimated coefficients are the following:

- Stops are more likely to be made by light and medium trucks than by heavy trucks.
- Interaction of light and medium trucks and a number of industries are important. For example, establishments in retail, construction, government, other services, and agriculture, forestry, and fishing are likely to make scheduled stops by light or medium trucks than by heavy trucks. Though, establishments in transportation handling industry are likely to make scheduled stops by heavy trucks than by light or medium trucks.
- Meeting stops are more likely to be made by light or medium trucks while goods only stops are less likely to be made by these trucks, relative to heavy trucks.
- Light trucks are likely to serve stops that are less than 20 miles away from the establishment, relative to heavy trucks.
- Medium trucks are likely to serve stops that are 2 to 10 miles away from the establishment, relative to heavy trucks.

Depending on the performance of the upstream models and the overall performance of CSTM as a system, it may be necessary to calibrate[7] the vehicle assignment model before applying the model to the BMC study area. Calibration of the vehicle assignment model will involve changing the alternative-specific constants (i.e., constants associate with light and medium truck) to match observed distribution of vehicle type, generated from GES survey data, with model prediction.

TABLE 7-9 VEHICLE ASSIGNMENT MODEL ESTIMATION RESULTS

Variable	Coefficient	t-statistic
<u>Vehicle type (base: heavy truck)</u>		
Light truck	0.220	1.83
Medium truck	0.274	2.46
<u>Industry type (base: Wholesale and other omitted industries)</u>		
Light truck × Agriculture/Forestry/Fishing	0.878	6.01
Light truck × Construction	1.970	12.60
Light truck × Government	3.870	19.25
Light truck × Manufacturing	-1.100	-5.43
Light truck × Other Services	2.950	12.60
Light truck × Retail	1.920	9.48
Light truck × Transportation Handling	-2.770	-10.27
Medium truck × Agriculture/Forestry/Fishing	1.490	10.56
Medium truck × Construction	1.010	6.07

[7] Model calibration is the process of applying the estimated models, comparing the results to observed values, and adjusting either the model specification or the alternative-specific constants.

Medium truck × Government	1.190	5.28
Medium truck × Other Services	0.598	2.17
Medium truck × Retail	0.762	3.43
Medium truck × Transportation Handling	-1.210	-9.85
Stop type (base: service)		
Light truck × Goods only	-1.530	-14.08
Light truck × Meeting	3.840	11.45
Medium truck × Goods only	-0.935	-8.44
Medium truck × Meeting	1.280	3.62
Distance between the establishment and the stop (in miles) (base: distance ≥ 20)		
Light truck × (0<distance<2)	1.340	11.61
Light truck × (2≤distance<10)	1.530	14.80
Light truck × (10≤distance<20)	0.503	5.84
(base: 0<distance<2 and distance ≥ 10)		
Medium truck × (2≤distance<5)	0.772	6.95
Medium truck × (5≤distance<10)	1.370	14.37
Observations	8,311	
Final log likelihood	-5,306.98	
Adjusted rho-square	0.416	

7.6.4 EXPECTED STOP DURATION

The expected stop duration model is applied to scheduled stops generated by the stop generation model as well as any intermediate stops on the tour. Empirical stop duration distributions from the Ohio GES were segmented by establishment industry type and stop purpose, after specification testing. Log-normal distributions were fitted to each empirical distribution segment to allow for duration draws to be made from a smoothed distribution. For each stop of a particular type made by a driver in a particular industry, an expected stop duration value is drawn from the appropriate distribution segment.

Table 7-10 summarizes the stop duration distribution segments and reports the estimated mean and standard deviation of the underlying normal distributions. Draws from these underlying normal distributions are exponentiated to get the associated log-normally-distributed values. To avoid unrealistically long or short stop durations, only the middle 90% of the probability density function is drawn from. A minimum stop duration for goods or service stops of one minute is also imposed.

TABLE 7-10 EXPECTED STOP DURATION DISTRIBUTIONS SEGMENTS (SOURCE: 2003–2004 OHIO STATEWIDE GENERAL ESTABLISHMENT SURVEY)

Industry	Activity	Observations	Mean of underlying normal distribution (μ)	Standard deviation of underlying normal distribution (σ)

Agriculture, Construction, Government, Health, Hotel & Real Estate, Manufacturing, Other Services, Retail, Transportation Handling, and Wholesale	Break/Meal	360	3.28	0.92
Government	Goods	685	1.08	1.06
Other Services	Service	330	2.09	2.13
Other Services	Goods	217	2.20	1.19
Retail	Goods	198	2.37	1.00
Health	Goods	87	2.62	0.94
Wholesale	Goods	1643	2.73	0.87
Agriculture	Goods	200	2.83	1.09
Hotel & Real Estate	Goods	37	2.84	0.68
Construction	Goods	299	2.91	1.03
Wholesale	Meeting	739	2.99	0.96
Government	Service	454	3.01	1.35
Transportation Handling	Goods	541	3.13	0.97
Agriculture	Service	712	3.26	1.25
Manufacturing	Goods	199	3.30	0.84
Agriculture	Meeting	105	3.43	1.10
Wholesale	Service	164	3.58	1.27
Health	Service	217	3.76	1.30
Hotel & Real Estate and Retail	Meeting and Service	174	3.65	1.25
Construction	Meeting	189	4.14	0.85
Construction	Service	360	4.63	1.33
Agriculture, Construction, Government, Health, Hotel & Real Estate, Manufacturing, Other Services, Retail, Transportation Handling, and Wholesale	Vehicle Service	147	2.52	0.83
Manufacturing and Transportation Handling	Meeting and Service	83	4.27	1.06
Government, Health, and Other Services	Meeting	297	4.16	0.96
Agriculture, Manufacturing, Wholesale, Retail, Hotel & Real Estate, Construction, Health, Transportation Handling, Other Services, and Government	Personal business/other	697	2.83	1.63

7.6.5 STOP CLUSTERING

Once the scheduled stops and their durations have been generated for a given vehicle and establishment, they are clustered into groups to be serviced by a vehicle tour. The dissimilarity matrix of the latitude and longitude coordinates defining the locations of the scheduled stops is calculated using the Euclidean method. This matrix is used in a complete-linkage hierarchical cluster algorithm to agglomerate stops into spatially similar groups. The result is a "tree" of stop clusters where the "leaves" of the tree are individual stop locations.

To avoid clustering stops into tours that cannot be realistically served by a single vehicle in a single day, the leaves of the tree are weighted by the associated stop durations. A recursive weighted-branch trimming algorithm then cuts off branches of the tree if the sum of the leaf weights for that branch is less than some threshold. If not, the algorithm continues down to the next split in the branch and evaluates the weights again. The leaves of the trimmed branches define the final stop clustering and are assigned together as a tour. If a single stop location has a duration of more than the weight threshold, the branch trimming process produces a branch with just that stop location, resulting in an out-and-back tour (though this is not the only way such tours can be assigned).

The leaf weight threshold will be calibrated using the distribution of stops observed by vehicle type in the Ohio GES. The threshold can be adjusted such that average number of simulated stops per tour is approximately equal to the average number of observed stops per tour. Alternatively, the threshold could be adjusted to considered the average total stop duration per tour.

This approach has the effect of grouping spatially similar stop locations within the constraints of a working day. Because the tour sequence is not known at this point in the modeling sequence, the tour travel time is left to fall out as-is when the Routing Sequence model is applied. Figure 7-6 visualizes the scheduled stop clustering of a toy example on a generic coordinate plane.

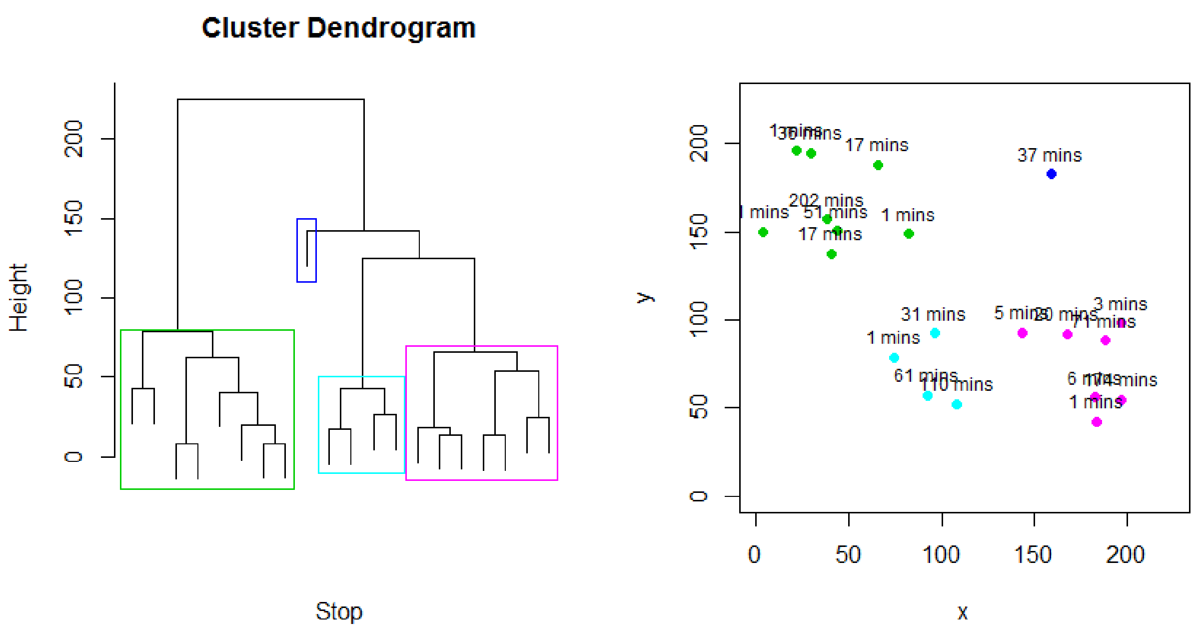


FIGURE 7-6: STOP CLUSTERING AND BRANCH TRIMMING

7.6.6 ARRIVAL TIME AT FIRST STOP

This model predicts arrival time of day at the first scheduled stop on a tour. The final model estimation results, obtained after specification testing, include a number of alternative-specific constants and the effect of summation of duration of scheduled stops in the tour (Table 7-11). The base alternative in the model is 07:31-08:30am. The model estimation results indicate the following:

- In general, arrival at the first stop of a tour is more likely to occur between 12:00am to 4:00 am and 8:31am to 11:59pm, and less likely to occur between 4:01am to 7:30am, relative to the time period between 7:31am to 8:30am. Some arrival time periods with similar likelihood of being chosen are 10:01am to 12:30pm,12:31pm to 13:30pm,1:31pm to 3:30pm, 3:31pm to 6:00pm, 6:01pm to 8:00pm, and 8:00pm to 11:00pm.
- As the duration of scheduled stops in a tour increases so does the likelihood of a tour starting earlier in the day. Figure 7-7 shows the effect of tour duration on the probability of arrival at first stop on a tour. The peak of the 2-hour tour duration graph is located right to the peak of the 8-hour tour duration graph. Also, the 2-hour tour duration graph has a fatter tail than the 8-hour tour duration graph, indicating that tours with short durations are more likely to start later in the day than tours with long durations.
- As noted above in the description of the stop duration models, establishment industry type and stop type are used to determine the duration of each stop and, thus, are indirectly included in this model as they affect expected tour duration.

TABLE 7-11 ARRIVAL TIME AT FIRST STOP MODEL ESTIMATION RESULTS

Variable	Coefficient	t-statistic
Time of day (base: 0731-0830 hour)		
0001-0100 hour	0.802	0.72
0101-0200 hour	0.561	0.77
0201-0300 hour	0.455	0.62
0301-0400 hour	0.897	1.27
0401-0500 hour	-2.080	-2.89
0501-0530 hour	-1.520	-2.19
0531-0600 hour	-1.340	-1.95
0601-0630 hour	-4.960	-4.15
0631-0700 hour	-2.970	-4.70
0701-0730 hour	-1.730	-3.15
0831-0900 hour	0.998	2.53
0901-0930 hour	1.330	3.19
0931-1000 hour	1.810	4.37
1001-1030 hour	2.500	6.90

1031-1100 hour	2.120	5.77
1101-1130 hour	2.250	4.76
1131-1200 hour	2.420	5.05
1201-1230 hour	2.600	6.76
1231-1300 hour	2.870	7.53
1301-1330 hour	2.980	6.59
1331-1400 hour	3.280	7.07
1401-1430 hour	3.830	11.25
1431-1500 hour	3.610	10.45
1501-1530 hour	3.320	9.41
1531-1600 hour	2.670	7.04
1601-1630 hour	2.950	8.05
1631-1700 hour	2.540	6.58
1701-1800 hour	2.850	7.71
1801-1900 hour	2.340	5.87
1901-2000 hour	2.090	4.99
2001-2100 hour	2.990	4.93
2101-2200 hour	2.910	4.75
2201-2300 hour	2.830	4.56
2301-2400 hour	1.210	1.12
Logarithm of summation of duration of scheduled stops in a tour (log(D)) (base: 0731-0830 hour)		
0101-0400 hour × log(D)	-0.805	-5.09
0401-0600 hour × log(D)	-0.166	-1.22
0601-0630 hour × log(D)	0.590	2.71
0631-0700 hour × log(D)	0.462	3.92
0701-0730 hour × log(D)	0.254	2.42
0831-0900 hour × log(D)	-0.187	-2.35
0901-0930 hour × log(D)	-0.321	-3.74
0931-1000 hour × log(D)	-0.443	-5.09
1001-1100 hour × log(D)	-0.631	-8.26
1101-1130 hour × log(D)	-0.678	-6.47
1131-1200 hour × log(D)	-0.741	-6.86
1201-1300 hour × log(D)	-0.801	-9.66
1301-1330 hour × log(D)	-0.854	-8.24
1331-1400 hour × log(D)	-0.972	-8.85
1401-2000 hour × log(D)	-1.120	-15.30
2001-2300 hour × log(D)	-1.410	-9.00
2301-0100 hour × log(D)	-1.110	-4.07
Observations	2,534	
Final log likelihood	-7,713.8	
Adjusted rho-square	0.145	

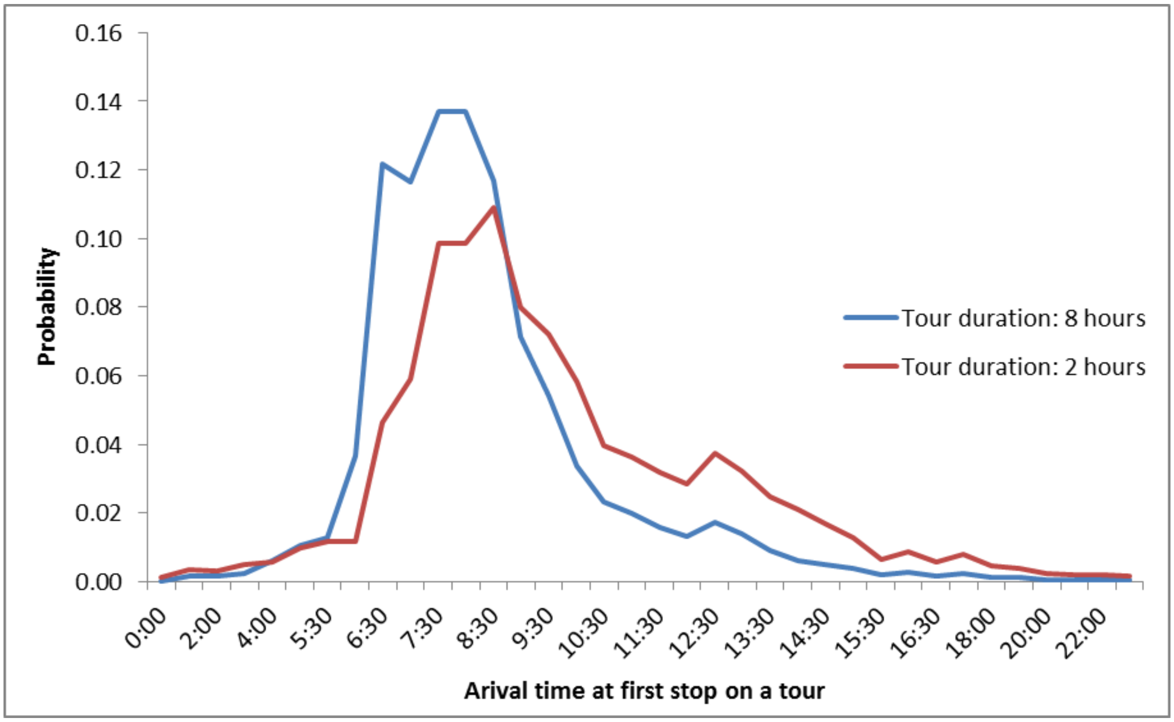


FIGURE 7-7 EFFECT OF TOUR DURATION ON THE PROBABILITY OF ARRIVAL TIME AT FIRST STOP ON A TOUR

7.6.7 ROUTING SEQUENCE

Once the scheduled stops have been identified for a given tour, vehicle, and establishment, they are assigned to a sequence. A table containing the establishment location, scheduled stop locations, and their travel time skim matrix is constructed for use with a path-finding algorithm. Because the order in which stops are made is unknown at this point, and because stop to stop travel time depends on time of day, the travel skim matrix is assumed fixed for a generic time of day.

The TSP R package is used as the path-finding algorithm and contains a number of algorithms for solving symmetric and asymmetric traveling salesman problems. The traveling salesman problem is a class of network path-finding problems where a set of locations in a network must be visited exactly once via the shortest possible route and return to the initial location (a Hamiltonian cycle). The vehicle routing problem is a special case of the traveling salesman problem and determining the optimal solution is very computationally challenging (NP-Hard). Thus, finding an optimal solution is not computationally feasible when the number of locations to visit is more than a handful or so. Commercial vehicle drivers are also not necessarily strict route optimizers and so approximate solutions that can be obtained quickly are doubly preferable. The arbitrary insertion method described by Rosenkrantz et al. (1977) is used to route the tours. Figure 7-8 visualizes the tour sequencing behavior for a toy example on a generic coordinate planes.

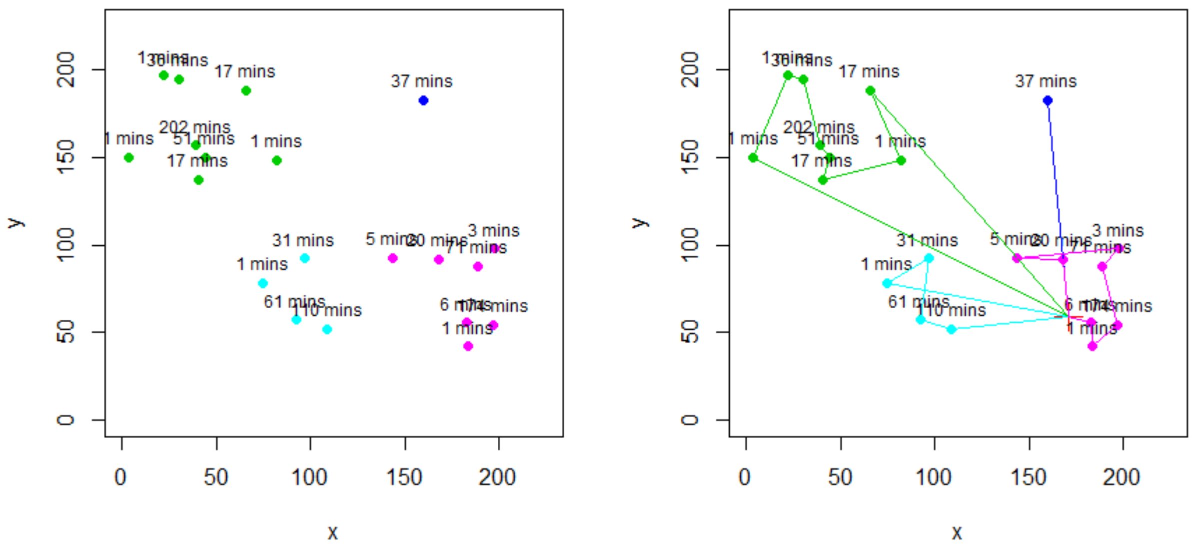


FIGURE 7-8: TOUR SEQUENCING

7.6.8 INTERMEDIATE STOP CHOICE

The model predicts intermediate stops and their purposes between "scheduled" stops. The final model estimation results, obtained after specification testing, are summarized below in Table 7-12. The base alternative in the model is "no break." The main implications of the estimated coefficients are the following:

- Drivers are less likely to take a break between or prior to a scheduled stop.
- The likelihood of a driver taking a break also decreases with an increase in the duration of remaining scheduled stops in the tour.
- Medium-truck drivers are more likely to take vehicle service/refueling break, but less likely to take a break for personal business/other reasons.
- Heavy-truck drivers are less likely to make an intermediate stop for a break/meal or for personal business/other reasons.
- Drivers are more likely to make an intermediate stop for break/meal between 11:30am to 12:30pm, relative to any other time periods of the day.
- If the tour starts at the establishment, then the drivers are more likely to make an intermediate stop for vehicle service/refueling before making the first schedule stop.
- Finally, an increase in distance between the scheduled stops and the establishment is likely to increase the likelihood of making an intermediate stop for vehicle service/refueling.

TABLE 7-12 INTERMEDIATE STOP CHOICE MODEL ESTIMATION RESULTS

Variable	Coefficient	t-statistic
Stop type (base: no break)		
Break/meal	-2.910	-29.88
Vehicle service/refueling	-4.130	-14.70
Personal business/other	-0.962	-18.21
Log of summation of remaining scheduled stops duration (log(D)) (base: no break)		
Break/meal × log(D)	-0.298	-10.48
Vehicle service/refueling × log(D)	-0.391	-9.34
Personal business/other × log(D)	-0.678	-24.50
Vehicle type (base: light truck)		
Vehicle service/refueling × Medium truck	0.396	2.03
Personal business/other × Medium truck	-0.879	-6.70
Break/meal × Heavy truck	-0.593	-3.13
Personal business/other × Heavy truck	-2.310	-8.69
Other variables		
Break/meal × 1130-1230 hour	1.770	14.56
Vehicle service/refueling × tour starts at establishment ×first stop of the tour	1.220	4.65
Vehicle service/refueling × log(distance between establishment and current stop)	0.254	2.88
Observations	10,175	
Final log likelihood	-3,782.20	
Adjusted rho-square	0.731	

7.6.9 INTERMEDIATE STOP DESTINATION

This model is applied to select destination for each intermediate stop. The model does not require any parameter estimation, only selection of a set of eligible TAZ (i.e., choice set) for each intermediate stop, based on a suitable maximum distance threshold for deviation (ΔD_{max}) and attraction factor(s), as discussed previously. Figure 7-9 presents a simple illustration of how the model will be implemented to identify eligible intermediate stop locations/TAZs. TAZ_i and TAZ_j are the locations of the current stop and the next scheduled stop, respective. And, TAZ_k, TAZ_l , TAZ_m, and TAZ_n are four potential intermediate stop locations with positive attraction factor. For each potential intermediate stop location, k, deviation from the direct path will be calculated as following:

Distance deviation for TAZ_k (ΔD_k) = $Distance_{ik}$ (D_{ik}) + $Distance_{kj}$ (D_{kj}) – $Distance_{ij}$ (D_{ij})

Next, ΔD_k will be compared with a predetermined value ΔD_{max}, and TAZ_k will be an eligible TAZ for the intermediate stop if $\Delta D_k \leq \Delta D_{max}$.

To determined appropriate value of ΔD_{max}, observed deviations in distance for intermediate stops (ΔD) were analyzed. Table 7-13 presents summary statistics of observed ΔD, and Figure 7-10 shows distribution of observed ΔD. In lieu of observed data from the BMC/SHA modeling area, when the model is implemented, the mean value of ΔD obtained from Ohio GES survey data will be assumed to be transferable. As part of model calibration, it may be necessary to adjust the implemented value of ΔD until the predicted ΔD distribution matches observed data (presented in Figure 7-10) reasonably well.

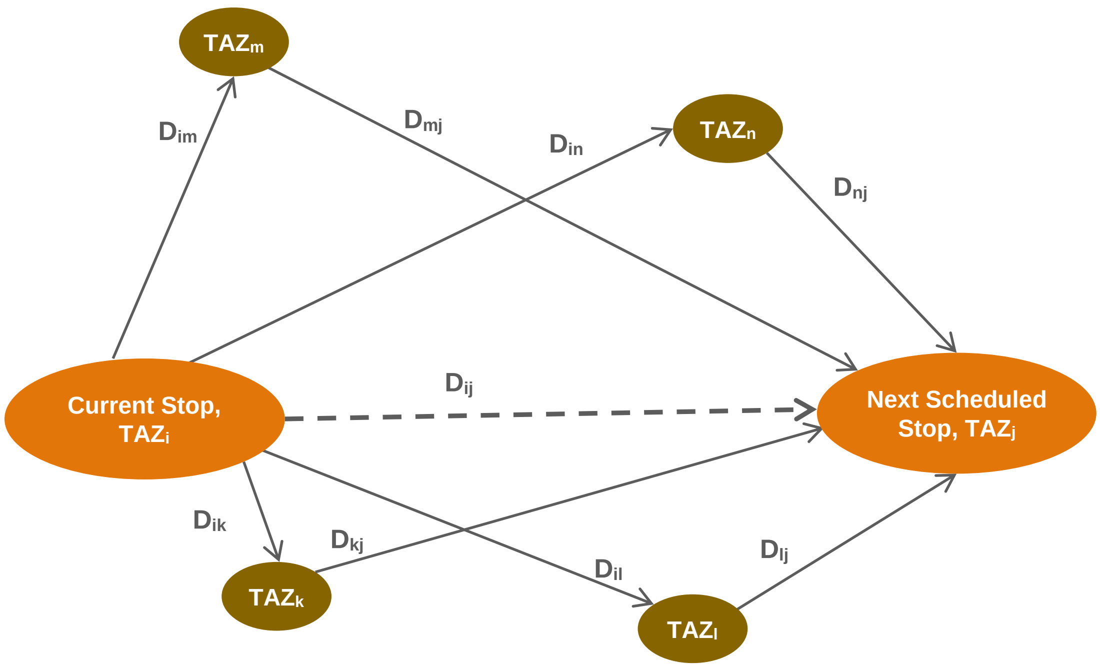


FIGURE 7-9: IDENTIFYING ELIGIBLE TAZS FOR INTERMEDIATE STOP LOCATION CHOICE SET

TABLE 7-13 OBSERVED DEVIATION IN DISTANCE (ΔD) PAPAMETERS

Parameter	Value (in miles)
Mean	3.09
Median	1.37
Minimum	0.10
Maximum	31.01

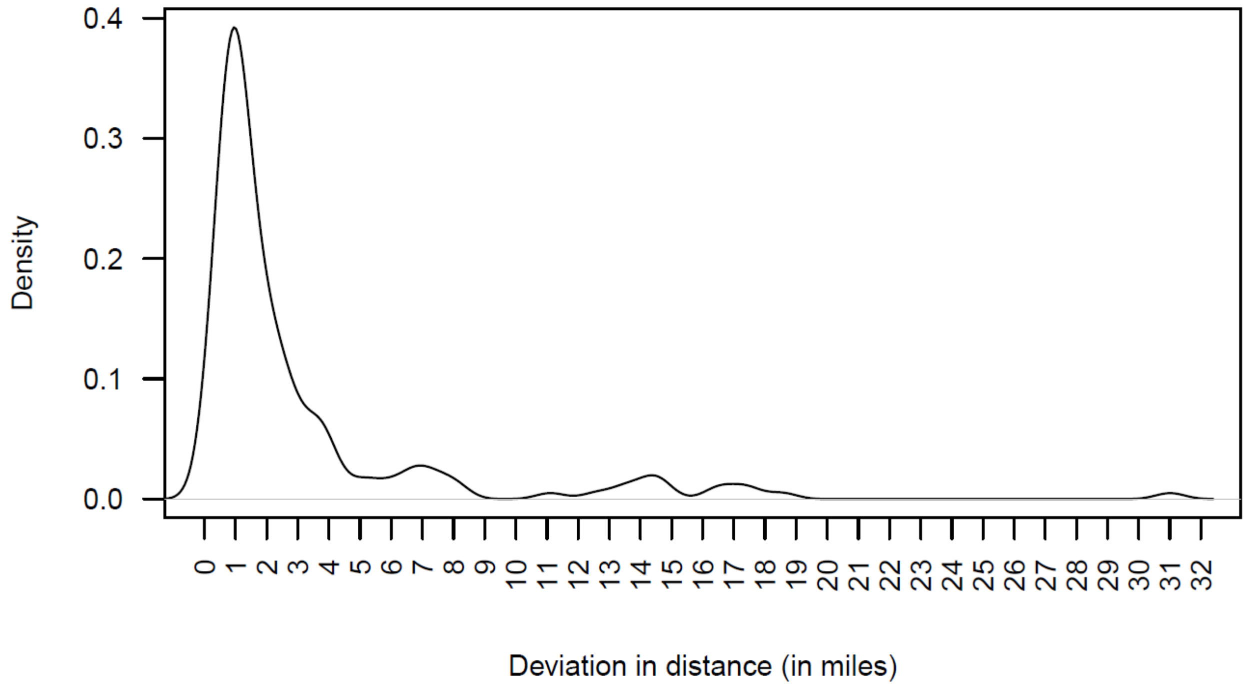


FIGURE 7-10: DISTRIBUTION OF OBSERVED DEVIATION IN DISTANCE (ΔD)

APPENDIX A. FREIGHT MODEL DEVELOPMENT QUESTIONS