Technical Report Documentation Page

1. Report No. FHWA/TX-07/5-4713-01-1	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Estimating Truck Travel Data for Texas		5. Report Date August 31, 2006
		6. Performing Organization Code
7. Author(s) Jolanda Prozzi, Peter Siegesmund, and Rob Harrison		8. Performing Organization Report No. 5-4713-01-1
9. Performing Organization Name and Address Center for Transportation Research The University of Texas at Austin 3208 Red River, Suite 200 Austin, TX 78705-2650		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. 5-4713-01
12. Sponsoring Agency Name and Address Texas Department of Transportation Research and Technology Implementation Office P.O. Box 5080 Austin, TX 78763-5080		13. Type of Report and Period Covered Technical Report
		14. Sponsoring Agency Code
15. Supplementary Notes Project performed in cooperation with the Texas Department of Transportation and the Federal Highway Administration.		
16. Abstract		

16. Abstract

In 2004, the Texas Department of Transportation (TxDOT) awarded the Center for Transportation Research at The University of Texas at Austin an implementation project that was to explain how to use the calibrated multinomial logit (MNL) models to generate county-level truck flows for Texas and to present a detailed explanation of the required steps to calibrate the MNL models in the future. MNL models were developed as part of a TxDOT research project 0-4713 entitled *Development of Truck Travel Database in Texas: Identifying Sources and Methodology.* Specifically, the research team was tasked with (a) embedding the MNL models into Microsoft Excel, (b) using the newly released 2002 Commodity Flow Survey (CFS) data for Texas to populate a truck travel database, (c) developing step-by-step instruction materials on using the MNL models embedded in Excel, and (d) piloting the instruction materials by providing training to district and Transportation Planning and Programming (TPP) personnel on its use. The instruction materials were piloted during a one-day workshop on August 25, 2006, at TxDOT's Riverside facility. This report documents the information presented to the workshop participants.

17. Key Words Texas truck flows, Commodity Flow Survey, Multinomial logit models, Statewide Analysis Model, Truck Travel Database		18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161; www.ntis.gov.
19. Security Classif. (of report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of pages 30	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Estimating Truck Travel Data for Texas

Jolanda Prozzi
Peter Siegesmund
Rob Harrison

Center for Transportation Research
The University of Texas at Austin
3208 Red River
Austin, TX 78705

www.utexas.edu/research/ctr

Disclaimers

Author's Disclaimer: The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official view or policies of the Federal Highway Administration or the Texas Department of Transportation (TxDOT). This report does not constitute a standard, specification, or regulation.

Patent Disclaimer: There was no invention or discovery conceived or first actually reduced to practice in the course of or under this contract, including any art, method, process, machine manufacture, design or composition of matter, or any new useful improvement thereof, or any variety of plant, which is or may be patentable under the patent laws of the United States of America or any foreign country.

Engineering Disclaimer

NOT INTENDED FOR CONSTRUCTION, BIDDING, OR PERMIT PURPOSES.

Research Supervisor: Robert Harrison

Acknowledgments

Table of Contents

List of Tables

1. Introduction

Truck data is critical to transportation planning in any region. Inter-city and interstate truck flows have an important impact on traffic volumes, the mix of traffic, and levels of congestion experienced on state roadways. In Technical Report 0-4713-01, entitled *Development of Sources and Methods for Securing Truck Travel Data in Texas,* a multinomial logit (MNL) approach was proposed to estimate county-level truck travel data from the publicly available 1997 Commodity Flow Survey (CFS) and IMPLAN data over the short term. Although not a required research product, the modeling approach was considered very useful to the Texas Department of Transportation (TxDOT).

In 2004, TxDOT thus awarded the Center for Transportation Research at The University of Texas at Austin with an implementation project to explain how to use the calibrated MNL models to generate county-level truck flows for Texas and to present a detailed explanation of the required steps to calibrate the MNL models in the future. Specifically, the research team was tasked with (a) embedding the MNL models into Microsoft Excel, (b) populating a truck travel database with the newly released 2002 Commodity Flow Survey (CFS) data for Texas, (c) developing step-by-step instruction materials on how to use the MNL models embedded in Excel, and (d) piloting the instruction materials by providing training on their use to district and Transportation Planning and Programming (TPP) personnel. The instruction materials were piloted during a one-day workshop on August 25, 2006, at TxDOT's Riverside facility. This report documents the information presented to the workshop participants. For detailed step-by-step instructions on how to apply and calibrate the MNL models developed, the reader is referred to TxDOT product 5-4713-P3, entitled *Manual for the Computation of Disaggregate County-Level Truck Flows and Explanation of Model Calibration*.

2. Overview and Background

TxDOT has funded the development of the Statewide Analysis Model (SAM) to assess the flows of passengers and freight on the state-maintained roadways. The objective of the SAM is to provide a regional model for the state of Texas that focuses on intercounty travel patterns. Specifically, the freight component aims to:

- provide a clear picture of freight movements on Texas's transportation system;
- determine the impact of freight on Texas's road infrastructure (e.g., bridges and pavements) and the implications in terms of funding;
- evaluate strategies for improving freight mobility;
- forecast system performance; and
- improve the safety and security performance of the road network.

2.1 What Truck Travel Data Does TxDOT Need for the Statewide Analysis Model?

The SAM requires commodity truck tonnage (i.e., weight), flow (e.g., number of loads), and value for 11 aggregated commodity groups (see Table 2.1), as well as for empty trucks, for the following origins and destinations:

- Texas counties-to-states (internal–external),
- Texas county exports-to-Mexican/Canadian provinces (internal–external),
- States-to-Texas counties (external–internal),
- Texas county imports-from-Mexican/Canadian provinces (external–internal),
- Texas county-to-county flows (internal–internal), and
- Texas through flows (external–external).

Table 2.1   Aggregated Commodity Categories Included in SAM

Commodity Group	Commodity Categories
Agriculture	Live animals and live fish; cereal grains; other agricultural products; animal feed and products of animal origin, n.e.c.
Food	Meat, fish, seafood, and their preparations; milled grain products and preparations, and bakery products; other prepared foodstuffs and fats and oils; alcoholic beverages; tobacco products
Building Materials	Monumental or building stone; nonmetallic mineral products; base metal in primary or semifinished forms and in finished basic shapes; articles of base metal
Raw Material	Natural sands; gravel and crushed stone; nonmetallic minerals n.e.c.; metallic ores and concentrates; coal
Chemicals/Petroleum	Gasoline and aviation turbine fuel; fuel oils; coal and petroleum products, n.e.c.; basic chemicals; pharmaceutical products; fertilizers; chemical products and preparations, n.e.c.
Wood	Logs and other wood in the rough; wood products; pulp, newsprint, paper, and paperboard; paper or paperboard articles; printed products; furniture, mattresses and mattress supports, lamps, lighting fittings
Textiles	Plastics and rubber; textiles, leather, and articles of textiles or leather
Machinery	Machinery; electronic and other electrical equipment, components and office equipment; motorized and other vehicles (including parts); transportation equipment, n.e.c.; precision instruments and apparatus; miscellaneous manufactured products
Miscellaneous	Waste and scrap; mixed freight
Secondary	Warehouse and distribution; truck intermodal drayage; truck air drayage
Hazardous	Waste hazardous materials; hazardous materials and substances

SAM can display the statewide truck traffic flows for 11 commodity categories for a base and forecasted year. An embedded TransCAD function assigns the truck tonnage data to the network. However, ultimately the effective use of SAM is critically contingent on the availability of accurate data. The objective of TxDOT research project 0-4713 (*Development of Truck Travel Database in Texas: Identifying Sources and Methodology)* was to recommend a robust methodology to TxDOT planners for collecting and maintaining intercounty and interstate truck travel data in a format that can be used in the SAM, and to advise TxDOT on the cost-effectiveness of acquiring truck travel data for the SAM.

2.2 How Can TxDOT Collect Truck Travel Data?

This research culminated in the recommendation of various approaches available to TxDOT planners for obtaining and maintaining county-level truck travel data in the format required for the SAM. In the short term (i.e., the next one to three years), TxDOT can either continue to purchase the Reebie TRANSEARCH database or use the calibrated MNL models developed in this research project to estimate county-level truck data. Over the medium term (i.e., the next three to five years), reliable truck data for Texas can be collected through an extensive program of truck intercept surveys or a data sharing initiative with trucking companies. In the intermediate long term (i.e., the next five to ten years), a number of national trade and ITS

initiatives could potentially result in more robust freight data for transportation planning. At a minimum, TxDOT should stay abreast of these initiatives. Ideally, however, the agency should participate in any workshops hosted as part of the initiatives to ensure that robust truck travel data are collected and made available for state transportation planning.

2.2.1 Short-term Options

The Reebie TRANSEARCH database is a unique source of detailed freight data that is available for purchase. It is the most often used database for statewide analysis of freight movements—11 out of 19 U.S. state departments of transportation interviewed during the research project used the Reebie TRANSEARCH database to conduct statewide freight modeling. The Reebie TRANSEARCH database is a relatively inexpensive source of detailed freight data. The data sources used to compile the database are proprietary, and the assumptions used to estimate and forecast the data are not disclosed. It is thus not possible to easily verify the accuracy and reliability of the data. This lack of transparency regarding the sources of the data, the methodology used, and the assumptions made have raised several questions about the validity and reliability of the Reebie TRANSEARCH database.

The alternative in the short term is for TxDOT to use the MNL models developed during the research to estimate county-level truck data from publicly available data sources. Of the more than fifty private and public freight data sources reviewed as part of the literature review, only six capture some of the variables of interest to TxDOT—such as truck volume (tonnage, value, and number of loads), origins and destinations, commodities, truckload or less-than-truckload—at various levels of detail. These data sources are as follows: Commodity Flow Survey (CFS), Reebie Associates' TRANSEARCH, Freight Transportation and Logistics Service, Transborder Surface Freight Database, North American Trucking Survey (NATS), and Vehicle Inventory and Use Survey (VIUS).

Among the publicly available freight data published on a continual basis, the Commodity Flow Survey (see text box) was considered the most useful for generating disaggregate truck flows in Texas, pending a decision from TxDOT not to purchase the Reebie TRANSEARCH database.

Commodity Flow Survey (CFS) Data

The CFS captures shipment data from a sample of businesses in the mining, manufacturing, wholesale trade, and certain retail industries in the fifty U.S. states and the District of Columbia. The budget for the 2002 CFS survey cycle is estimated at $13 million. It is by far the most comprehensive freight flow database in the U.S., providing freight data in terms of tons, ton-miles, value, shipment distance, commodity, and weight for all major transportation modes (i.e., air, truck, rail, water, and pipeline), as well as intermodal combinations. The CFS is conducted every five years in years ending in 2 and 7. It has been conducted in 1993, 1997 and 2002. Unfortunately, budget constraints have resulted in the sample size of the CFS survey being reduced from 200,000 (1993) to 100,000 (1997), and 50,000 (2002)—the latter represents approximately 7 percent of the 750,000 industries registered in the Census Bureau's Business Register. In addition, concerns exist about the adequate coverage of agricultural shipments (i.e., the movement from the farm to the first point of assembly), shipments from transportation and service businesses, and most of the retail businesses (National Research Council of the National Academies, 2003). Finally, the CFS database provides limited information on foreign shipments, because only domestic shippers are sampled. For example, imports are captured only when transported from the importer's location to a destination elsewhere in the U.S. Thus, the movement between the port of entry and the importer's location is not captured in the CFS database. Similarly for exports, the export destination is provided as the port of exit in the U.S. (National Research Council of the National Academies, 2003). The CFS is available on a CD-ROM free of charge from the Census Bureau (see www.census.gov). The CD-ROM version of the CFS contains a database and a software application that can be used to access, sort, and export the data. Using this software, the CFS data can be sorted by mode, commodity, origin, and/or destination.

The CTR research team calibrated two MNL models for nine of the eleven commodity groups included in the SAM from the publicly available CFS data. The calibrated MNL models were subsequently used to estimate county-level truck flows for Texas (see Chapter 3). Because one of the primary data sources in compiling the Reebie TRANSEARCH database is the CFS data, the research team compared the CTR model estimates with the TRANSEARCH data for Texas. The paired sample t test was used to determine whether there was any statistically significant difference between the model estimates and the Reebie data used in the SAM. A confidence interval of 95 percent was specified. For six of the nine commodity groups, the mean of the model estimates for Texas county-to-state truck flows was statistically similar to the mean of the TRANSEARCH data for these flows. In the case of state-to-Texas county truck flows, the model estimates were statistically similar to the TRANSEARCH data for four of the nine major commodity groups, and for Texas county-to-county truck flows, the model estimates were statistically similar to the TRANSEARCH data for five of the nine commodity groups. It is thus believed that the MNL approach proposed by the research team provides a cost-effective alternative for obtaining Texas county-level truck flows for at least some of the commodity groups in the short term.

2.2.2 Medium-term Options

The research team detailed two data collection approaches—i.e., roadside intercept surveys and truck carrier participation—that showed the most promise of providing TxDOT with the data needed for the SAM over the medium term. Roadside intercept surveys are expected to be the more costly—in excess of $5 million—of the two approaches. However, the total cost will be largely a function of the statistical reliability that is required, which is influenced by the degree of geographical coverage and the number of days surveyed. To ensure geographical coverage of the state of Texas and to account for variations in the characteristics of trucks that use different roadways, it was recommended that a minimum of ninety-six sites be identified: three sites per major highway type (i.e., interstate, state highway, Texas highway, and farm-to-market/ranch-to-market) in each of the eight National Transportation Analysis Regions (Nears) of Texas. To account for time-of-day, day-of-week, and seasonal variation, it would be ideal if each site were surveyed for 24 hours, seven days a week in each of the four seasons, but it is expected that the costs would be prohibitive. In addition, collecting truck travel data through roadside intercept surveys tend to be time consuming and poses concerns about interviewer safety and potential disruptions to traffic flows.

The truck carrier participation approach is based on the hypothesis that a statistically representative sample of truck companies operating in, from, to, and through Texas can be convinced to share a sample of their operational data (e.g., operational data for one week four times a year) with TxDOT. The research team contacted eight trucking companies, which have been exposed to transportation planning through their involvement with the North Central Texas Council of Government's Intermodal Freight and Safety Committee, to determine whether these companies would consider participating in a data sharing initiative with TxDOT and what their conditions for participation would be. All eight representatives indicated a willingness to participate provided that certain conditions can be met. The most often mentioned conditions were that (a) no information about the company will be included in the database compiled and used by TxDOT, (b) the data will not be used for law enforcement or litigation against the company, (c) the Texas Motor Transportation Association (TMTA) will be involved to protect the interests of those that participate, and (d) no severe cost burden will be imposed on the trucking companies in compiling and submitting the data to TxDOT. Given this positive response, the research team was encouraged to believe that statistically reliable and verifiable information can be obtained through a truck carrier participation program. Because no previous precedent exists for collecting statewide truck travel data through such a data sharing partnership, it is very difficult to estimate the costs associated with this approach. It is, however, foreseen that the initial costs associated with recruiting a statistically representative sample of trucking companies could be significant. There would also be initial costs in determining the most appropriate data collection methods (i.e., ranging from electronic reporting options to less sophisticated paper-based systems) for obtaining data from companies that vary not only in size but also in the level of sophistication in how the trucking companies capture this information. The recurring costs would, however, be limited to maintaining a dialogue and relationship with the trucking companies involved and the costs associated with compiling and managing the database.

2.2.3 Long-term Options

In the intermediate long term, the research team highlighted a number of national initiatives that could potentially result in more robust freight data for transportation planning.

Initiatives such as the Transportation Research Board's proposed national freight data program, the FHWA's nationwide deployment of a universal electronic freight manifest, U.S. Customs and Border Protection's Automated Manifest System, the automated commercial environment, and the proposed International Trade Data System might result in timely and robust freight data that can be invaluable to future statewide transportation planning efforts. Finally, the research team summarized the data attributes captured by current Intelligent Transportation System (ITS) technologies and highlighted a new federal ITS initiative that could potentially provide state and local transportation planners with more robust truck data over the long term. The new ITS initiative—dedicated short range communications (DSRC) devices—is a medium-range communication service intended to support both public safety (e.g., collision avoidance) and licensed private operations by means of roadside-to-vehicle and vehicle-to-vehicle communication. It is foreseen that these DSRC devices could be the underlying technology for electronic waybills that will make the collection of commodity data and truck characteristics feasible.

3. Production and Attraction Flow Distribution Models

TxDOT research project 0-4713 utilizes a statistical model called a *Logit model*. Logit models are a class of econometric models that are used to evaluate the relationship between a set of independent variables and a binary dependent variable. The predicted dependent variable is expressed as a ratio of the probability of an outcome divided by the probability of that outcome not occurring. This is an odds ration, and the natural logarithm of this ration is known as the logistic function, from which the Logit model derives its name.

The Logit Model has the following functional form:

$$\ln(\frac{p}{1-p}) = \alpha + \beta_1 \chi_{1,i} + ... \beta_k \chi_{k,i}$$

MNL models are an expansion of the basic, binary Logit models. They are used to evaluate the relationship between a set of independent variables and a set of dependent variables that represent mutually exclusive, discrete alternatives. These alternatives can be alternate transportation routes, consumer choices, or voting patterns to name just a few.

The MNL has the following functional form:

$$\Pr(y_i = j) = \frac{e^{x_i \beta_j}}{\sum_i^j e^{x_i \beta_j}}$$

3.1 Production and Attraction Flow Distribution Models

In general, the movement of commodity flows between zones can be presented as production flows from an origin zone (e.g., state, county) and attraction flows to a destination zone (e.g., state, county). Specifically, the annual truck flows (tonnage) from each production state to the 50 attraction states for a particular commodity k can be represented as follows:
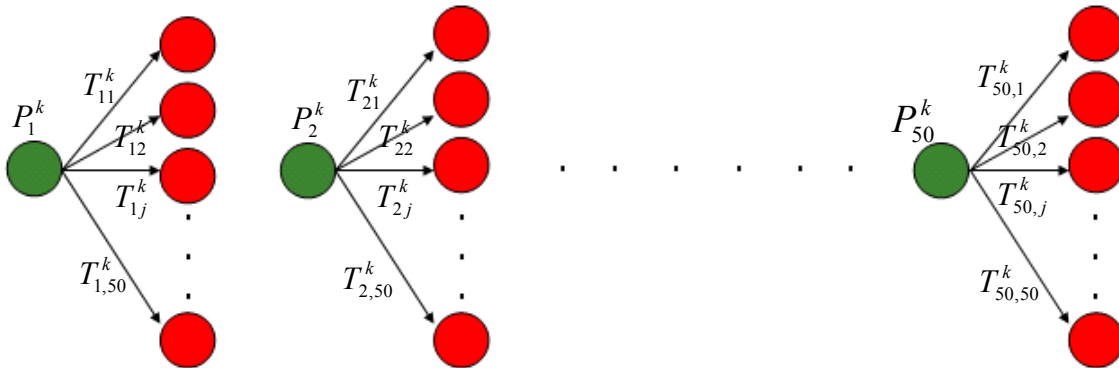
Similarly, the annual truck flows (tonnage) attracted to each state from the 50 production states of a specific commodity k can be illustrated as follows:
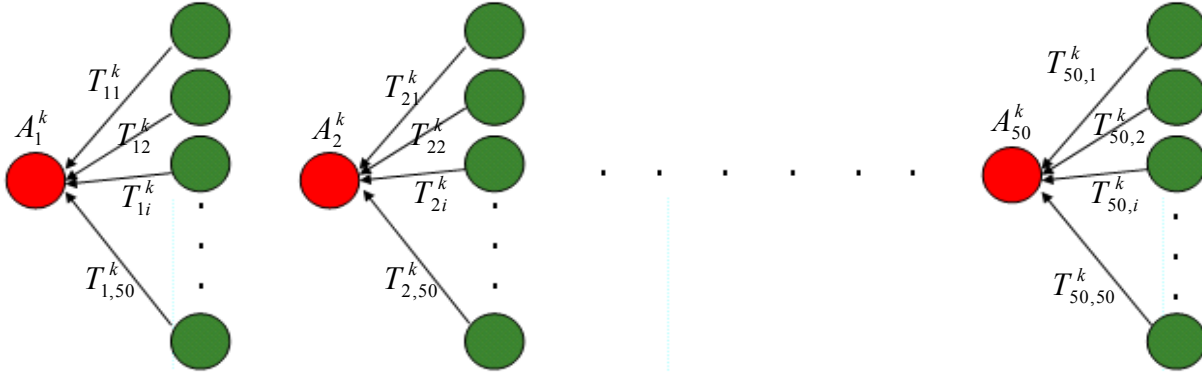
During TxDOT research project 0-4713, two MNL models were calibrated for each of nine commodity categories at the state level to estimate truck flows at the county level. The production flow distribution model models truck flows from each of the production states to each of the 50 attraction states. Similarly, the attraction flow distribution model models truck flows to each attraction state from the 50 production states.

3.1.1 Production Flow Distribution Model

The MNL production flow distribution model estimates the fraction of the total productions in a state moving to each attraction state by truck as a function of the relative attraction level of the destination states and the generalized cost of transportation. However, due to a lack of generalized cost data, centroidal distances between states along the highway network were employed as the impedance measure affecting freight flow distribution.
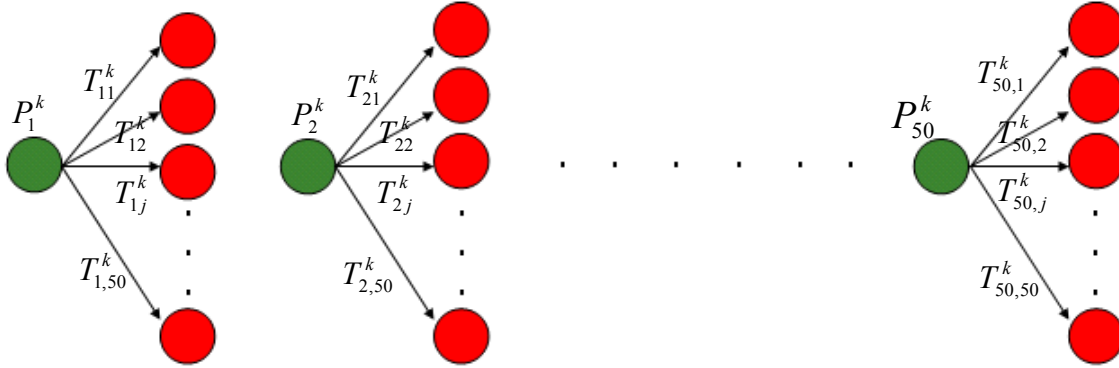
Given the definition of the MNL model, the truck flows from state i to state j of commodity k can be calculated from the following mathematical equation:

$$T_{ij}^{k} = P_{i}^{k} * \frac{e^{V_{ij}^{k}}}{\sum_{j=1}^{n} e^{V_{ij}^{k}}}$$

T_{ij}^{k} = truck flows from state i to state j of commodity k

P_{i}^{k} = total productions of commodity k in state i

V_{ij}^{k}
= utility for flows from production state i to each of the attraction states j

Where,

$$V_{ij}^{k} = \alpha_{0}^{k} + \alpha_{1}^{k} d_{ij} + \alpha_{2}^{k} FA_{j}^{k}$$

$\alpha_{0}^{k}, \alpha_{1}^{k}, \alpha_{2}^{k}$ = the coefficients to be determined during model calibration

d_{ij} = the centroidal distance between states i and j

FA_{j}^{k} = the fractional attraction level of state j for commodity group k

The following data/software was used for calibrating the production flow distribution model:

- The 1997 CFS data was used to obtain (a) the annual truck tonnage from each production state i to each attraction state j for each commodity category k and (b) the fractional attraction level of state j for each commodity category k.
- TransCAD GIS software was used to calculate the interstate centroidal distances as the shortest path distance along the highway network between state centroids.

3.1.2 Attraction Flow Distribution Model

The MNL attraction flow distribution model estimates the fraction of the total attractions destined in a state originating from each of the 50 production states by truck as a function of the relative production level of the origin states and the generalized cost of transportation. As in the case of the production flow distribution model, the centroidal distance between states along the highway network were used as the impedance measure affecting freight flow distribution.
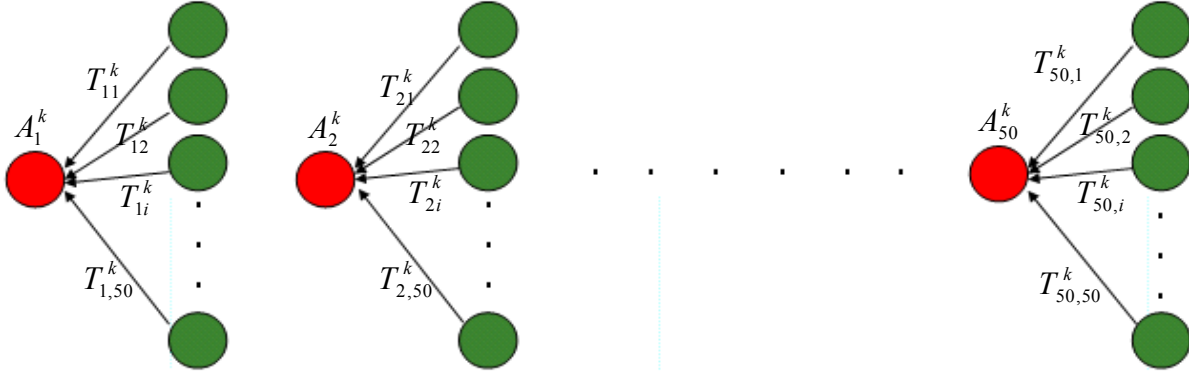
Given the definition of the MNL model, the truck flows of commodity k attracted to state j from state i can be calculated from the following mathematical equation:

$$T_{ji}^k = A_j^k * \frac{e^{V_{ji}^k}}{\sum\limits_{i=1}^{n} e^{V_{ji}^k}}$$

T_{ji}^k = truck flows of commodity k attracted to state j from state i

A_j^k = total attractions of commodity k to state j

V_{ji}^k
= utility for flows to attraction state j from each of the production states i

Where,

$$V_{ji}^k = \beta_0^k + \beta_1^k d_{ji} + \beta_2^k FP_i^k$$

$\beta_0^k, \beta_1^k, \beta_2^k$ = are the coefficients to be determined during model calibration

d_{ji} = the centroidal distance between zones j and i

FP_i^k = the fractional production level in zone i of commodity group k

The following data/software was used for calibrating the attraction flow distribution model:

- The 1997 CFS data was used to obtain (a) the annual truck tonnage to each attraction state j from each production state i for each commodity category k and (b) the fractional production level of state i for each commodity category k.
- TransCAD GIS software was used to calculate the interstate centroidal distances as the shortest path distance along the highway network between state centroids.

3.2 Compute Disaggregate Texas County Truck Flows

The objective of the state-level model calibration is to generate county-level truck flows for Texas. The CFS state-to-state flows, the IMPLAN county production and attraction levels, the calculated inter-county and county-state centroidal distances, and the calibrated state-level MNL models are thus used to estimate:

- Texas county-to-state truck flows of commodity k
- State-to-Texas county truck flows of commodity k
- Texas county-to-county truck flows of commodity k

3.2.1 Model Estimates: Texas County-to-State Truck Flows

Texas county-to-state flows are estimated from the Texas-to-state truck flows (by commodity category) captured in the CFS and by applying the MNL attraction flow distribution model. The truck flow distribution to attraction state j from Texas counties can be illustrated as follows:
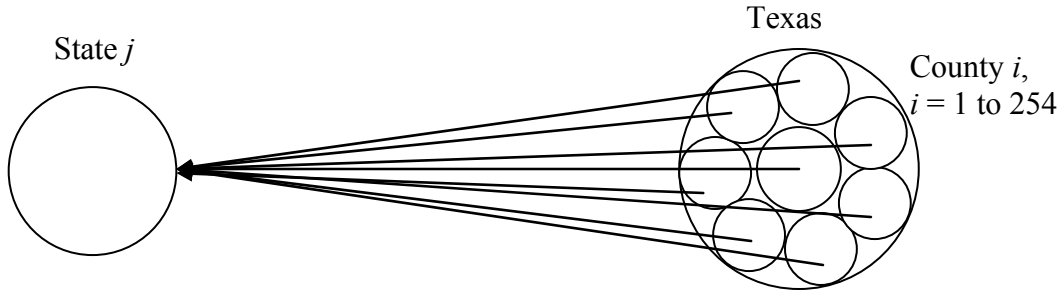
By applying the MNL attraction flow distribution model, the fraction of the total attractions of commodity k in state j from Texas, originating from county i, can be determined as a function of the production level of county i of commodity k and the centroidal distance between county i and state j. The mathematical equation is as follows:

$$T_{jC_i}^k = T_{jT}^k * \frac{e^{\beta_0^k + \beta_1^k d_{jC_i} + \beta_2^k FP_{C_i}^k}}{\sum\limits_{i=1}^{254} e^{\beta_0^k + \beta_1^k d_{jC_i} + \beta_2^k FP_{c_i}^k}}$$

$T_{jC_i}^k$ = Truck flows of commodity k to state j from county i in Texas

T_{jT}^{k} = Truck flows of commodity k to state j from Texas (obtained from CFS Texas-to-state truck flows of commodity k)

$\beta_0^k, \beta_1^k, \beta_2^k$ = Coefficients (obtained from MNL Attraction Flow Distribution Model)

d_{jC_i} = Centroidal distance between state j and Texas county i (obtained from TransCAD)

$FP_{C_i}^{k}$ = Fractional production level of county i for commodity k (Texas county productions by commodity category are derived from the Output field in IMPLAN)

3.2.2 Model Estimates: State-to-Texas County Truck Flows

State-to-Texas county truck flows are estimated from the state-to-Texas truck flows (by commodity category) captured in the CFS and by applying the calibrated MNL production flow distribution model. The truck flow distribution from production state i to Texas county j can be illustrated as follows:
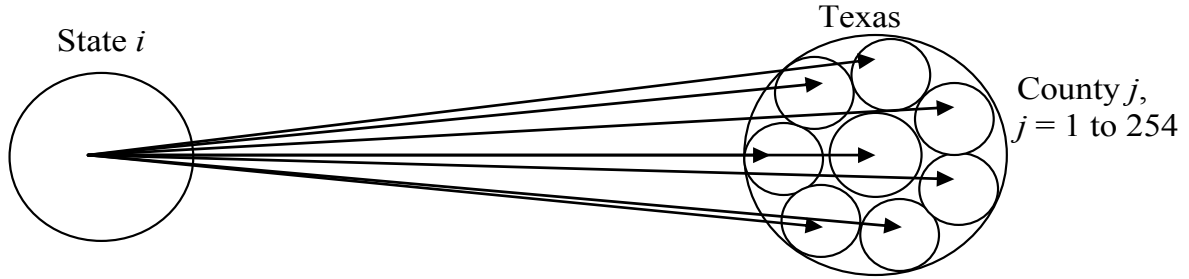
By applying the MNL production flow distribution model, the fraction of the total truck flows of commodity k from origin state i to Texas, destined for county j, can be determined as a function of the attraction level in county j of commodity k and the centroidal distance between state i and county j. The mathematical equation is as follows:

$$T_{iC_j}^{k} = T_{iT}^{k} * \frac{e^{\alpha_0^k + \alpha_1^k d_{iC_j} + \alpha_2^k FA_{C_j}^k}}{\sum_{j=1}^{254} e^{\alpha_0^k + \alpha_1^k d_{iC_j} + \alpha_2^k FA_{c_j}^k}}$$

$T_{iC_j}^{k}$ = Truck flows of commodity k from state i to county j in Texas

T_{iT}^{k} = Truck flows of commodity k from state i to Texas (obtained from CFS state-to-Texas truck flows of commodity k)

$\alpha_0^k, \alpha_1^k, \alpha_2^k$ = Coefficients (obtained from MNL Production Flow Distribution Model)

d_{iC_j} = Centroidal distance between state i and Texas county j (obtained from TransCAD)

$FA_{C_j}^k$ = Fractional attraction level of county j for commodity k (Texas county attractions by commodity category are derived from the institutional commodity demand field in IMPLAN)

3.2.3 Model Estimates: Texas County-to-Texas County Truck Flows

Texas county-to-county truck flows were estimated from the Texas-to-Texas truck flow data (by commodity category) captured in the CFS and the calibrated MNL attraction and production flow distribution models. County-to-county truck flows were thus generated following a two-step procedure. *First (Step 1),* the fraction of the total intrastate Texas attractions originating in each Texas county i is determined by applying the MNL attraction flow distribution model.

The distribution of the intrastate Texas attractions originating in each Texas county i can be illustrated as follows:
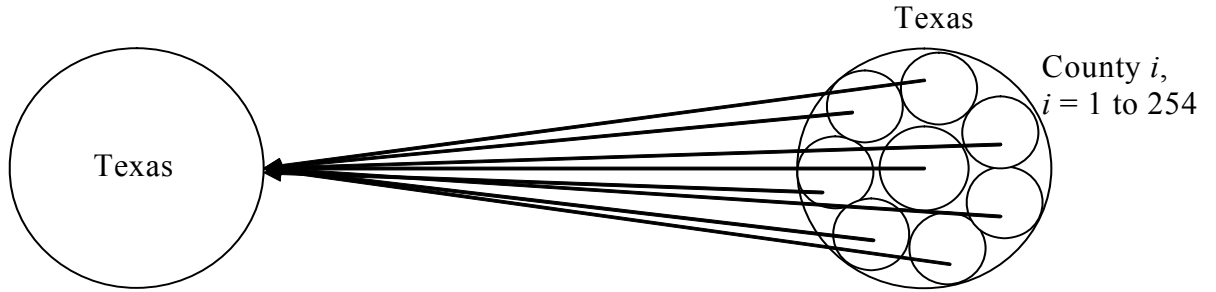
By applying the MNL attraction flow distribution model, intrastate truck flows attracted to Texas from each Texas county can be calculated as follows:

$$T_{TCi}^k = T_{TT}^k * \frac{e^{\beta_0^k + \beta_1^k d_{TCi} + \beta_2^k FP_{Ci}^k}}{\sum_{i=1}^{254} e^{\beta_0^k + \beta_1^k d_{TCi} + \beta_2^k FP_{ci}^k}}$$

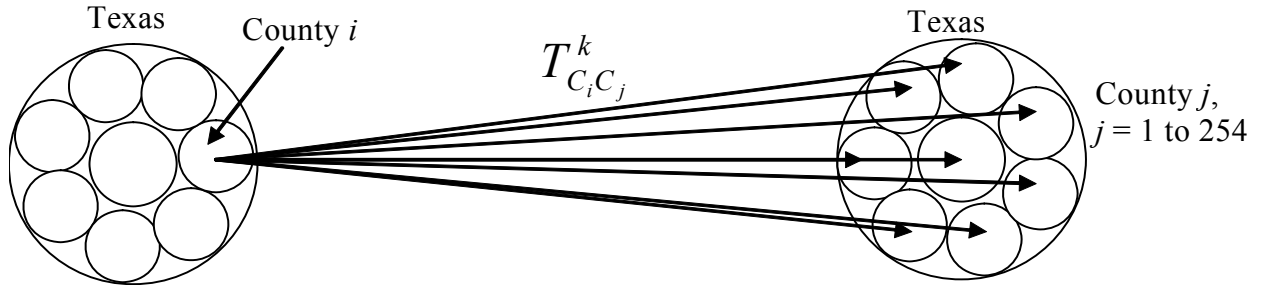T_{TCi}^k = Truck flows of commodity k from Texas county i to Texas

T_{TT}^k = Truck flows of commodity k from Texas to Texas (obtained from CFS Texas-to-Texas truck flows of commodity k)

$\beta_0^k, \beta_1^k, \beta_2^k$ = Coefficients (obtained from MNL Attraction Flow Distribution Model)

d_{TCi} = Centroidal distance between Texas and Texas county i (obtained from TransCAD)

FP_{Ci}^{k} = Fractional production level of county i for commodity k (Texas county productions by commodity category are derived from the Output field in IMPLAN)

Second (Step 2), given the Texas county origins calculated in Step 1, Texas county-to-county truck flows are estimated by distributing the intrastate truck flows originating in each Texas county i to each destination county j using the production flow distribution model. The distribution of intrastate truck flows originating in each Texas county i to each Texas destination county j can be presented as follows:

Applying the MNL production flow distribution model, truck flows of commodity k from county i to county j can be calculated as follows:

$$T_{CiC_j}^{k} = T_{CiT}^{k} * \frac{e^{\alpha_0^k + \alpha_1^k d_{CiC_j} + \alpha_2^k FA_{C_j}^k}}{\sum_{j=1}^{254} e^{\alpha_0^k + \alpha_1^k d_{CiC_j} + \alpha_2^k FA_{c_j}^k}}$$

$T_{CiC_j}^{k}$ = Truck flows of commodity k from Texas county i to Texas county j

T_{CiT}^{k} = Truck flows of commodity k from Texas county i (calculated in Step 1)

$\alpha_0^k + \alpha_1^k + \alpha_2^k$ = Coefficients (obtained from MNL Production Flow Distribution Model)

d_{CiC_j} = Centroidal distance between Texas county i and j (obtained from TransCAD)

$FA_{C_j}^{k}$ = Fractional attraction level of Texas county j for commodity k (Texas county attractions by commodity category are derived from the institutional commodity demand field in IMPLAN)

4. Concluding Remarks

The MNL modeling approach, developed in this research, provides a cost-effective methodology for generating disaggregate truck flow data from the publicly available CFS data. However, even though MNL models are particularly suited for modeling the production and attraction flow distribution of trucks and can be calibrated relatively easily, their application is impacted by the quality and reliability of the data used and the power of the explanatory variables in predicting the distribution of truck flows. The CFS data used for the model calibration has a number of specific limitations. For example, the CFS covers only shipments originating from the manufacturing, mining, wholesale, and selected retail sectors. The farming, forestry, fishing, construction, crude petroleum production, household, government, foreign establishment, and most retail and service businesses are thus not covered (Transportation Research Board, 2003). Inadequate interstate truck flow data for primary agriculture movements, secondary traffic, and hazardous material (HAZMAT) shipments thus have prevented the development of truck flow distribution models for these shipments.

In addition, the MNL model for production (attraction) flows is calibrated assuming that the distribution of truck flows is a function of the interstate centroidal distances and attraction levels (production levels) of states. A number of factors other than distance and the production/attraction levels of states can impact the truck flow distribution between states, including the economy, land use patterns (e.g., the location of industrial centers, intermodal terminals, inland ports, and seaports), supply and logistics strategies, trade and transportation agreements, legislation (environmental, tax, and transportation policies), and a number of mode choice factors. These various factors interact to translate finally in truck volumes on a state's infrastructure. An understanding of these factors, how they change, and how they interact is necessary to model current and estimate future impacts on truck traffic volumes. Because these factors (explanatory variables) have not been considered in the model, it does affect the accuracy of the results. Also, the impedance measure for truck flows included in the model is represented by the interstate centroidal distances. Truck flows, however, typically occur between regions of economic activity, which are not necessarily located at the geometrical centroid of the state. Centroidal distance is thus a crude proxy for the impedance measure for interstate truck flows

To conclude, it was recommended that the calibrated MNL models developed be used by TxDOT to estimate county-level truck data for the SAM in the short term (i.e., next one to three years). Over the medium (i.e., next three to five years) and intermediate long term (i.e., next five to ten years), a number of alternative options were described for obtaining more robust truck travel data for the SAM (see 0-4713-01, entitled *Development of Sources and Methods for Securing Truck Travel Data in Texas.*

5. References

National Research Council of the National Academies (2003). *Measuring Personal Travel and Goods Movement: A Review of the Bureau of Transportation Statistics Surveys.* Transportation Research Board Special Report 277. Washington, DC: Author.

Transportation Research Board (2003). *Special Report 276: A Concept for a National Freight Data Program.* Washington, DC: The Committee on Freight Transportation Data: A Framework for Development.