Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA/TX-10/0-6583-1		

4. Title and Subtitle	5. Report Date
ALTERNATIVE METHODS FOR DEVELOPING EXTERNAL TRAVEL SURVEY DATA	Published: October 2010
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Stephen Farnsworth, Eric Talbot, Praprut Songchitruksa, Phillip Reeder, and David Pearson	Report 0-6583-1
9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
Texas Transportation Institute The Texas A & M University System College Station, Texas 77843-3135	11. Contract or Grant No. Project Number 0-6583
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Texas Department of Transportation Research and Technology Implementation Office P.O. Box 5080 Austin, Texas 78763-5080	Final: March 2009 – February 2010
	14. Sponsoring Agency Code

15. Supplementary Notes

Project performed in cooperation with the Texas Department of Transportation and the Federal Highway Administration.
Project Title: Methods for Developing External Travel Survey Data
URL: http://tti.tamu.edu/documents/0-6583-1.pdf

16. Abstract

The Texas Department of Transportation (TxDOT) has a comprehensive on-going travel survey program that supports the travel demand models being developed for transportation planning efforts in urban areas throughout Texas. One component of the survey program is the external travel survey. External travel surveys provide data on travel movements into, out of, and through urban areas. In recent years, there has been a heightened sensitivity to the methods used to collect external survey data as well as the type of data that is collected. This research examines alternative methods for collecting data on external travel movements and evaluates the potential for synthesizing/modeling external travel in lieu of conducting external surveys. The research will provide recommendations to TxDOT on the most viable methods to estimate external travel movements for use in travel demand models in urban areas in Texas.

17. Key Words	18. Distribution Statement
External Survey, Transportation Planning, External Local, External Through, Trip Tables	No restrictions. This document is available to the public through NTIS: National Technical Information Service Springfield, Virginia 22161 http://www.ntis.gov

19. Security Classif.(of this report)	20. Security Classif.(of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	126	

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

ALTERNATIVE METHODS FOR DEVELOPING EXTERNAL TRAVEL SURVEY DATA

by

Stephen Farnsworth
Associate Research Scientist
Texas Transportation Institute

Eric Talbot
Graduate Assistant Researcher
Texas Transportation Institute

Praprut Songchitruksa
Associate Research Scientist
Texas Transportation Institute

Phillip Reeder
Program Manager, Travel Forecasting Program
Texas Transportation Institute

and

David Pearson
Research Scientist
Texas Transportation Institute

Report 0-6583-1
Project 0-6583
Project Title: Methods for Developing External Travel Survey Data

Prepared in cooperation with the
Texas Department of Transportation
and the
Federal Highway Administration

Published: October 2010

TEXAS TRANSPORTATION INSTITUTE
The Texas A&M University System
College Station, Texas 77843-3135

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 FHWA or TxDOT. This report does not constitute a standard, specification, or regulation. The research supervisor in charge of this project was Stephen Farnsworth.

ACKNOWLEDGMENTS

This project was conducted in cooperation with the Texas Department of Transportation (TxDOT) and the Federal Highway Administration (FHWA). The research reported herein was performed by the Texas Transportation Institute (TTI). Mr. Bill Knowles, Director of the Traffic Analysis Section of the Transportation Planning and Programming Division, served as the research project director.

The authors would like to thank the members of the Project Monitoring Committee for their advisement and guidance:

- Mr. Charlie Hall, Transportation Planning and Programming Division, Texas Department of Transportation;
- Ms. Janie Temple, Transportation Planning and Programming Division, Texas Department of Transportation;
- Mr. Greg Lancaster, Transportation Planning and Programming Division, Texas Department of Transportation;
- Mr. Paul Tiley, Transportation Planning and Programming Division, Texas Department of Transportation;
- Mr. Eduardo Calvo, El Paso District, Texas Department of Transportation.

The authors would like to provide special thanks to Duncan Stewart, Frank Espinosa, and Sylvia Medina of TxDOT's Research and Technology Implementation (RTI) Division for their assistance in this project.

The authors would like to thank the following individuals from TTI for their assistance:

- Mr. Ed Hard – project administration assistance,
- Ms. Debbie Spillane – literature review assistance,
- Mr. Jason Beesinger – data analysis assistance,
- Mr. Gary Lobaugh – review and formatting assistance,
- Mr. Justin Malnar – contract assistance, and
- Editors in TTI's Communications Division.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION

The importance of well developed transportation networks is often overlooked by the public. However, the reality is that transportation networks are crucial to the long-term economic vitality of the United States. The ability to effectively and efficiently move people and goods around the country has a direct correlation to the quality of people's lives. As transportation systems have developed on national, state, and local levels, methods to measure how effectively the networks are performing have also been devised. In addition to measuring the operational characteristics of highway systems, particular interest has been devoted to planning for future growth and development.

Adding capacity to transportation networks, whether it is adding lanes to existing roadways or building new roads altogether, is an expensive process. Right-of-way acquisition, utility relocation, and design/engineering work all contribute to the ever-increasing cost of expanding transportation systems. Urbanized areas typically have many more transportation improvement projects identified than they have the funding to actually undertake. As a result, processes have been developed to help planners and city officials prioritize which projects provide the greatest benefit at the local and/or regional level. An accepted and widely implemented practice is the use of travel demand models (TDM) to forecast and estimate traffic and development patterns. Travel surveys serve as the primary inputs for TDMs in Texas. Having good travel survey data results in more accurate estimates for use in travel demand models, thus resulting in better information for decision makers to utilize in managing transportation system investments.

PROJECT PURPOSE

The Texas Department of Transportation (TxDOT) administers a robust travel survey program that collects travel data in all of the urbanized areas in the state. The travel survey program includes household, workplace, commercial vehicle, and external travel surveys. Data extracted from these surveys serve as inputs to TDMs used by TxDOT in transportation planning and policy analyses. The purpose of this project was to examine alternative methods for collecting data on external travel movements and evaluate the potential for synthesizing external travel data in lieu of conducting external travel surveys.

This project examined experiences and results for highways within and outside Texas, both from existing information and through case studies of selected Texas highways. The research used cause and effect relationships between various policies, actions, and practices and the resulting functionality over the life cycle of highways.

OVERVIEW OF TRAVEL SURVEYS IN TEXAS

Origin-destination travel surveys were first used in Texas in the 1950s to develop trip tables of zone to zone trip movements. In the 1960s, they served as the foundation for early travel models used in transportation planning and programming. Essentially no new large travel surveys were performed during the 1970s and early 1980s. By the mid

1980s, there was a push to revive travel survey data collection efforts using small sample techniques. In 1989–1990, TxDOT initiated several major travel surveys in urban areas to provide information to update their travel demand models. This effort has since evolved into TxDOT's current-day travel survey program (TSP) that represents one of the most comprehensive continuing data collection efforts in the nation.

The Transportation Planning and Programming (TPP) Division of TxDOT funds and manages the TSP, which collects data on travel in all major urban areas in the state of Texas. The TSP coordinates the conduct of travel surveys on a recurring basis in all of the state's 25 Metropolitan Planning Organizations (MPOs). The TSP consolidates the MPOs into 14 travel survey regions in order to combine areas with similar travel characteristics and survey them in a systematic and efficient manner. Figure 1 shows the travel survey regions in Texas.

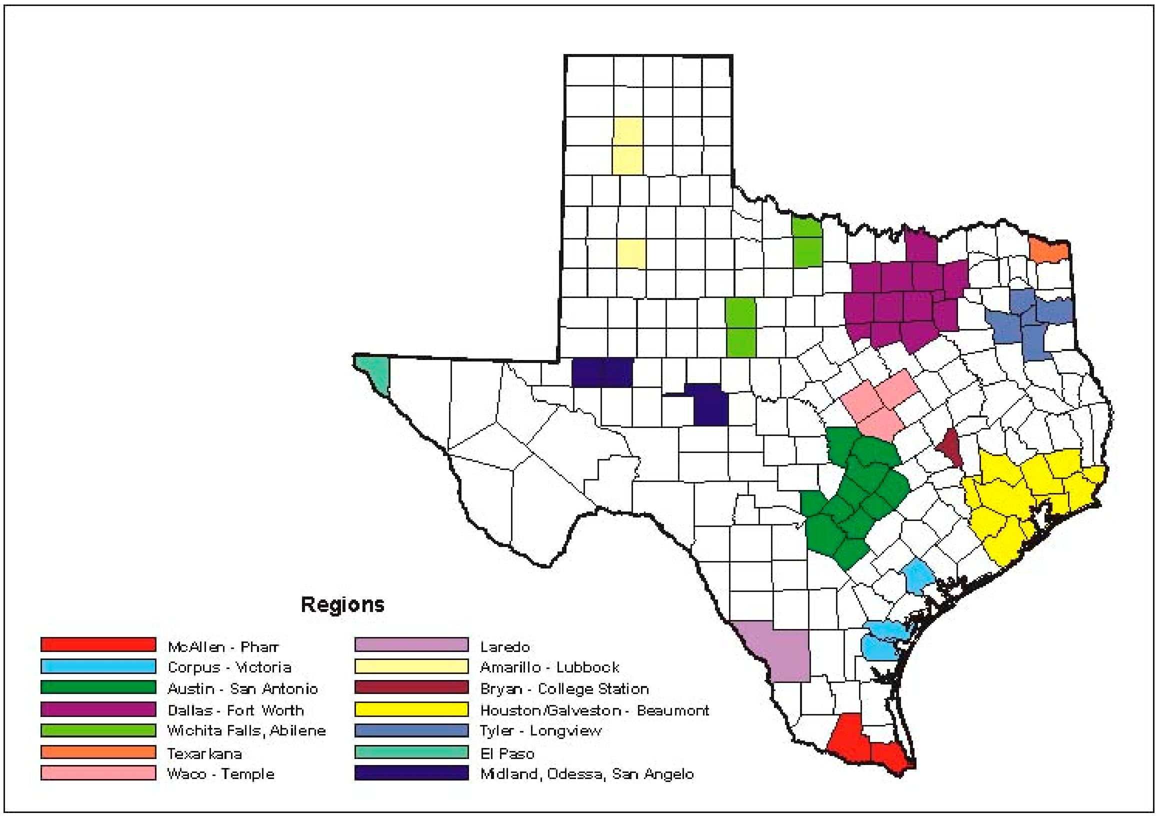

Figure 1. Travel Survey Regions in Texas.

The TSP includes household, workplace, commercial vehicle, and external surveys. Data from these surveys serve as inputs to the travel demand models (TDM) used by TxDOT in transportation planning and policy analyses. These data ensure the travel demand models accurately reflect travel behavior in forecasting travel demand. This is critical to making sound transportation investments.

In concert with travel surveys, TPP also provides crucial support and assistance to MPOs throughout the state in the development and calibration of travel models. As part of this assistance, TPP funds and coordinates travel surveys to collect data on local travel demand, patterns, and characteristics for use in area models. The surveys provide a base level of up-to-date 'real world' travel data that is needed to develop and calibrate models representative of each local area. External travel surveys play an essential role in supporting these models by capturing information on movements into, out of, and through urban areas.

The external travel survey, often termed 'roadside' or 'intercept' survey, is an essential component of TxDOT's travel survey program. External surveys are typically conducted at or near the boundaries of urbanized areas to collect information on the amount and characteristics of vehicles traveling into, out of, and through a defined study area. More specifically, the surveys collect data on internal-external (local) trips and external-external (through) trips by non-commercial and commercial vehicles.

External station surveys typically query motorists as they leave a defined study area. In order to obtain motorist travel information, a traffic control plan is set up on the outbound lane or lanes of roadways where they traverse the study area boundary. Vehicles exiting the study area are directed out of the main flow of traffic into a survey "station" by a trained individual. Once in the survey station, drivers are interviewed by a trained surveyor. Both commercial and non-commercial vehicles are surveyed during the conduct of external surveys, but the two vehicle groups are surveyed using different survey instruments. Figure 2 illustrates a typical setup and survey at an external station.

Figure 2. External Survey on Low-volume Roadway.

In addition to the survey, directional vehicle classification counts are performed at or near the survey site on the same day as the survey. The classification data are aggregated into 15-minute increments for a 24-hour period. External surveys are typically performed only during daylight hours, primarily for safety reasons. Therefore, the classification counts are used to expand the survey data to represent a 24-hour day.

The surveys are designed and implemented following accepted state-of-the-practice methods for roadside intercept travel surveys. While these methods may be the accepted state of the practice, they do not alleviate concerns that can be, and have been, raised by the public.

LEGAL CHALLENGES AND FOUNDATIONS

Travel surveys were originally conducted primarily through face-to-face interviews, typically in respondents' homes or at intercept points along major roadways and transit routes. Urban lifestyles have significantly reduced the ability to conduct face-to-face survey methods due to expenses and breadth of data required. Lack of participation by survey respondents is considered one of the most significant problems of conducting high quality travel surveys. Griffiths, Richardson, and Lee-Gosselin (*1*) researched this basic conflict between the need for increasingly detailed and frequent data on daily travel patterns and the growing difficulty in contacting and interviewing persons about their travel.

In recent years, critics of external surveys contend that establishing traffic control plans for the purpose of collecting travel survey data are analogous to unreasonable search and seizure, which claims to be in violation of the Fourth Amendment to the Constitution. The legal issue related to intercepting vehicles on roadways for the purpose of conducting travel surveys is whether such 'seizure' is unreasonable. In other words, does the government's (state department of transportation) intrusion (intercept of vehicle for travel survey data collection purposes) outweigh the motorist's reasonable expectation of privacy under the Fourth Amendment? In *Delaware v. Prouse* 440 U.S. 648, 653, 59 L.Ed.2d 660, 99 S.Ct. 1391 (1979), the court found that discretionary checks of automobiles (and their drivers) for the sole purpose of validating license and registration information made only marginal contributions to roadway safety, and less intrusive means could have been used to serve the same end. The court of record concluded, "There can be no question that the stopping of a vehicle and the detention of its occupants constitutes a 'seizure' within the meaning of the Fourth Amendment, even though the purpose of the stop is limited and the resulting detention quite brief."

In 1994, the Attorney General (AG) for the State of Kentucky concluded that external (intercept) surveys are illegal (1994 Ky. AG LEXIS 37, OAG 94-26). The Kentucky AG's opinion held that the governmental units involved in the proposed survey could not show a legitimate governmental interest that was sufficient to overcome motorist's legitimate expectation of privacy in their vehicles. It was the AG's opinion that the information sought in the survey could be obtained through less intrusive means.

In some instances, policy changes are made at the institutional level before legal challenges are brought forth. In 2002, the Florida High-Speed Rail Authority sponsored an intercept survey intended to gauge motorists' interest in a high-speed rail network. Approximately 7,000 intercept surveys were administered and nearly 100 individuals complained about being pulled over to be surveyed. Despite the small number of complaints (approximately one percent of the total people surveyed), the Florida Department of Transportation put the practice of intercept surveys on hold "indefinitely" (*2*). By 2006, that position had been modified to allow intercept surveys on facilities other than the state's limited access facilities (*3*).

While there are a few states that have determined that intercept type surveys are illegal, this methodology is considered an acceptable and legal process for collecting survey data in most parts of the United States, as well as in other countries. The justification for the collection of survey data is provided in the Code of Federal Regulations (CFR). In Texas, travel surveys are conducted under a cooperative agreement between TxDOT and MPOs within the state. As part of this agreement, TxDOT conducts travel surveys, performs traffic counts, and maintains travel demand models for the MPOs within the state. The data collected as part of travel surveys and traffic counts serve as inputs to local and regional travel demand models. As specified in 23 CFR 500.202, traffic data are defined as:

> "…data used to develop estimates of the amount of person or vehicular travel, vehicle usage, or vehicle characteristics associated with a system of highways or with a particular location on a highway."

The travel demand models are utilized, in part, to meet local and state planning requirements specified in the CFR. Planning requirements at the local level are specified in 23 CFR 134, and state level requirements are provided in 23 CFR 135. Under Title 23 of the CFR, metropolitan and statewide planning are required to "encourage and promote the safe and efficient management, operation, and development of surface transportation systems that will serve the mobility needs of people and freight and foster economic growth and development within and through urbanized areas, while minimizing transportation-related fuel consumption and air pollution." Therefore, the traffic data that are collected as part of the TSP assists in meeting the planning needs and requirements at both the state and local level.

UNDERSTANDING THE PROBLEM

In order to address the issues presented in the research, it is important to understand the function that various components of an external survey serve in the TDM process. Table 1 provides an overview of the data elements of a typical external survey in Texas and the function, if any, that they serve in the travel demand modeling process.

Without the conduct of external travel surveys, the only means of deriving base year external travel matrices for travel modeling purposes would be to synthetically derive the external trip tables. Unlike other travel model steps such as trip generation or mode choice where it is feasible, though not recommended, to transfer data from a similar urban area, the external travel patterns (i.e., the trip distribution of external trips) are unique for any given area. This is a function of a variety of factors but primarily the number of entry/exit points, the functional classification of and volume on each roadway entering the urban area, the probable combination of interchanges between each external roadway, and the size and urban form of the area itself. Consequently, external travel survey data are a critical component in calibrating and validating base year travel models.

Table 1. Survey Data Element Functions in the Modeling Process.

Data Element(s)	Function in Modeling Process
Time/date/location	Administrative and statistical summary purposes only
Occupancy	Estimates of number of persons traveling in urban area
Vehicle information/classification	Distribution of vehicle fleet mix operating in urban area
Residence location	Distribution of trips by residents vs. visitors
Overnight information	Statistical purposes
Out of state information	Statistical purposes
Location of trip origin	Average trip length, trip length frequency distribution, trip table development
Time left origin	Estimate of travel times
Type of place at origin	Attraction models
Purpose for being at origin	Disaggregation of trips by purpose (home-based work, home-based non-work, non-home based)
Local trip indicator	Trip table development
Through trip indicator	Trip table development
Location of trip destination	Average trip length, trip length frequency distribution, trip table development
Purpose for traveling to destination	Disaggregation of trips by purpose (home-based work, home-based non-work, non-home based)
Information on travel out of state	Statistical purposes
Information on travel in state	Statistical purposes
Information on trips made prior to being surveyed	Estimate of the number of non-resident internal trips
(Commercial vehicle survey only)	
Cargo being carried	Commodity freight model
Cargo weight	Commodity freight model
Type of container	Statewide Analysis Model (SAM)
Mexican cargo indicator	SAM
Cargo pickup location information	SAM
Cargo drop-off location information	SAM
Location vehicle traveling from	SAM and attraction models
Information on location being in/out of Texas	SAM

External travel surveys provide the following key components to the external distribution portion of the modeling process:

- the apportionment of external local and external through trips for each external station,
- the allocation between non-commercial and commercial vehicle trips,
- an estimate of non-resident travel,
- the average trip length of external local and external through trips, and
- survey expanded trip tables for external local and external through trip purposes.

Following is a brief overview of each of these model input components provided by the external travel survey.

External Local and External Through Trip Apportionment

In general, higher functionally classified roadways tend to have a higher percentage of external through trips than lower functional class roadways (*4*). Apart from that general observation, however, it is not feasible to correctly estimate the percentage of external local and external through traffic at an external roadway based solely on its functional classification. Given that external through trips add an inordinate amount of vehicle miles traveled (VMT) in relation to the actual number of external through trips, it is imperative that an accurate division of local and through trips are calculated for each external station. External surveys monitor and collect whether each external trip exiting an urban area is a local or though trip and, thus, provide the necessary information to accurately apportion external local and through trips at each external station during the base year model development process.

Non-Commercial and Commercial Vehicle Trip Allocation

Similar to the apportionment of local and through trips at each external station, the information on whether each local or through trip is a non-commercial or commercial vehicle trip is also indispensable data in the model development process. The distribution of external local non-commercial trips differs in comparison to the distribution of external local commercial vehicle trips and, consequently, has implications for proper model validation.

Non-Resident Travel

External local non-resident travelers typically make two to four trips within an urban area prior to exiting the urban area. An improper estimate of non-resident travel within an urban area can lead to an inaccurate total VMT estimate from the model. External surveys provide an inventory of non-resident travel that can be incorporated into the base year model and properly forecast for future year model applications.

Average Trip Lengths

The validation of trip distribution models for each trip purpose relies to a great extent on replicating an observed average trip length for each trip purpose. The lack of reasonable modeled trip lengths will impede the model validation process and can only undermine the soundness of the overall model validity.

Survey Expanded Trip Tables

Comparable to census journey-to-work trip tables, external surveys provide expanded local and through trip tables than can be used to benchmark the accuracy of the external trip distribution models. Moreover, without the availability of an expanded through trip table to use as a seed matrix, it would not be feasible to develop a through trip matrix except by synthetically deriving the matrix.

METHODS TO ESTIMATE EXTERNAL DATA

Estimating or having an appropriate measure of an urban area's external travel is a critical component of urban travel models. This is due to the extent to which external travel (i.e., external local and external through trips) comprises a portion of an urban area's total vehicle miles traveled. External travel can constitute 10 to 30 percent of an urban area's VMT and, depending on the study area, in some instances can contribute to more than half of an urban area's VMT. Table 2 provides several examples in Texas of the percentage of urban area external VMT and underscores the need to correctly estimate the distribution of external local and through travel in order to have a defensible model.

Table 2. External VMT Examples from Selected Texas Cities.

Urban Area (*)	Total VMT	External VMT	Percent of Total VMT
Amarillo (2005)	6,177,500	1,794,000	29.0%
Austin (2005)	39,869,000	6,566,100	16.5%
Dallas-Fort Worth (2005)	168,650,000	14,601,700	8.7%
Killeen-Temple (2006)	7,338,100	2,391,000	32.6%
Rio Grande Valley (2004)	18,860,000	2,702,500	14.3%
San Antonio (2005)	48,893,900	7,172,800	14.7%
Tyler (2004)	6,207,000	3,084,600	49.7%
Waco (2006)	6,918,300	3,061,300	44.2%

* Year of most recent external survey

Various methods for estimating origin-destination traffic flows have been developed. One particular method for developing trip tables utilizes network link traffic volumes (*5*). While this methodology has been developed and tested on internal trip movements, the applicability to external and through trips is not known.

As a result, there are essentially three primary means of estimating urban area external travel in lieu of not having relevant external survey data upon which to base the development of modeled external trip matrices. They are as follows:

- developing synthetic external models,
- drawing on data from other surveys, and
- extrapolating data from a statewide model.

Following is a brief discussion of the three approaches including the advantages and disadvantages of each approach.

Developing Synthetic External Models

Though the research and data are rather dated, there have been several previous research efforts conducted on synthetic procedures to estimate external travel, and the results indicate some success in replicating observed data. In the 1970s, Modlin's (*4*) paper documented a process for estimating external through trips at each external station and subsequently distributing the through trips. Building on Modlin's research efforts, Pigman (*6*) transferred the approach to Kentucky and focused on the relevance of urban area population, external station traffic volumes, and the commercial vehicle percentages as key variables for determining the number of through trips at each external station. Chatterjee and Raja (*7*) tested the Modlin and Pigman models in their 1987 paper and presented a two-stage approach for estimating external through trips. In 1993, Reeder (*8*) tested Modlin's and Pigman's through trip estimation process against recent TxDOT external survey results and concluded that the process was not applicable for larger urban areas though it could have some merit in small- and medium-sized Texas area applications. Finally, in 1998, NCHRP Report 365, "Travel Estimation Techniques for Urban Planning" (*9*), summarized Modlin's methodology and provided it as a means of developing external travel estimates in the absence of external survey data.

The principal advantage in using synthetic external models is that it negates the need to conduct external surveys and collect information on local external travel patterns. This would be a cost saving measure for TxDOT and no longer necessitate the need for stopping the traveling public for roadside surveys. The drawback to developing defensible synthetic external models is that it will require further research to assess whether or not previous procedures are relevant and transferable to Texas, and whether or not they can be modified for application to any size Texas urban area. The methodologies cited are now 30 years old and have only been applied in small urban areas (*10*). A useable approach for Texas will need to encompass small- and medium-sized areas as well as large metropolitan areas.

Data from Other Surveys

TxDOT currently supports a robust travel survey program that includes household, workplace, commercial vehicle, and travel time studies in addition to the external travel surveys. Though limited in nature, these other surveys collect some information on external travel. For example, if a person participating in a household survey makes a trip outside an urban area on their designated survey day, that external local trip will be captured. Data such as the examples that are gleaned from the household, workplace, and commercial vehicle surveys could conceivably be used to extrapolate external local travel patterns.

Similar to the synthetic model advantages, travel information assembled from other surveys could replace current external survey data and also negate the need for conducting further external surveys. The limited number of external trip observations that occur within the samples collected from other surveys is the main disadvantage to this approach. Given the potential for collecting a limited amount of external travel information from other surveys challenges the ability to develop a complete and defensible dataset of external local travel patterns. In addition, there would still be the need to collect information on external through movements.

Statewide Model Data Extrapolation

A third option would be to use the Statewide Analysis Model (SAM) to derive external local and through splits for urban area external stations based on SAM modeled data. Likewise, SAM modeled data could also be used as the basis for developing urban area external through trip matrices.

One advantage to this approach is that urban area models would be more closely linked to the SAM model. The primary disadvantage would be that the SAM model would not encompass all the years for which TxDOT-TPP develops base year models. Another disadvantage would be the need to determine how well the SAM model currently replicates external local and through trip patterns within Texas urban areas. Lastly, due to the different zone structures between urban area models and the SAM, a method for estimating the average trip length for external local trips would still need to be developed for use in urban area models.

CHAPTER 2. ALTERNATIVE APPROACHES TO COLLECTING EXTERNAL TRAVEL SURVEY DATA

The initial task in the research project was to examine the state-of-the-practice in external survey data collection methodologies, with particular focus on methods other than the traditional roadside/intercept survey method. Included in the literature review was the identification of research and case studies where external data was synthesized from other sources and external data was transferred from one study area to another. Additionally, this task involved a query of travel demand modeling professionals in order to ascertain how Metropolitan Planning Organizations and state Department of Transportation agencies model external travel when no new survey data are available.

To acquire information for this task, journal articles, dissertations, presentations, and other scholarly texts on this topic were reviewed. Source material for this task was identified through internet searches, consultants, professional contacts, and a listserv query. The findings of the literature review revealed a number of articles that pertain to the subject of this research project.

METHODS FOR COLLECTING NEW TRAVEL DATA

Methods used to collect external-related traffic data vary depending on objectives of the survey and policies set forth by agencies conducting the survey. In Texas, the primary method used to collect external traffic data has been the roadside intercept interview method. However, TxDOT is interested in investigating alternative methods to collect external survey data. A potential data collection method is the use of internet-based travel surveys. In the past, not having easy access to the internet was a huge disadvantage of the methodology due to the sampling bias that it created. However, a study performed in 2006 by the Pew Internet and American Life Project reported that nearly 73 percent of Americans are regular internet users and nearly half (42 percent) had some sort of broadband connection in their home (*11*). These percentages continue to grow as the demand for internet availability increases. Another platform for accessing the internet is cellular phones and other mobile devices. It has been estimated that in 2013 nearly 97 percent of Americans will own some sort of mobile phone and nearly half (44 percent) of those users will use those mobile devices to access the internet (*12*).

In 2006, TxDOT implemented an internet-based travel survey as part of the external travel survey for the Killeen/Temple study area. This particular survey focused on travel patterns of persons entering and exiting the Fort Hood military installation. While the feasibility of a widespread application of web-based travel surveys in Texas has yet to be determined, the methodology was reviewed to determine the advantages and disadvantages that it offers (*13*). The primary advantages of internet based travel surveys include:

- very low operational and maintenance costs,
- interactive geocoding can increase accuracy of survey results,
- increasing or decreasing sample size does not increase cost of survey,

- multimedia and graphical tools can make survey more user-friendly,
- electronic medium allows logic checks and ability to skip questions based on previous responses,
- fast data retrieval, and
- increased efficiency in data analyses.

The primary disadvantages of internet based travel surveys include:

- sample bias due to potential participants not having internet access,
- potential for respondents to enter erroneous data,
- maintaining control over who replies to the survey and preventing an individual from responding multiple times,
- determining the best method for advertising/notifying the public about the survey, and
- traditional external survey results are expanded based on traffic counts collected on the day of the survey. A determination would need to be made about how to reconcile survey data with traffic count data.

While internet-based travel surveys are becoming a more frequently utilized methodology, determining the most effective method for recruiting participants remains a difficult undertaking. One method mentioned in the research was the use of dynamic message signs (DMS), also referred to as changeable message signs (CMS). Of primary concern with regards to the use of DMS is whether or not state policies allow the signs to be used for non-traffic/emergency information. A Federal Highway Administration (FHWA) memorandum dated March 21, 2003, states that they "continue to discourage the display of general public information or other nonessential messages on CMS." The memorandum also adds that FHWA feels that it is unsafe to expect motorists to "write telephone numbers, websites, addresses, or other lengthy information while they are moving." The FHWA position on acceptable uses of DMS serves as a guide and does allow the states to develop their own guidelines and policies.

In a majority of areas that utilize DMS, traffic-related information is the primary data displayed on the signs. However, certain non-traffic related information has become acceptable to display. This includes AMBER alert messages, special event messages, and certain public service announcements (PSA). The decision of whether or not to utilize DMS to notify the public of a survey would need to be a policy dictated by TxDOT.

Another consideration when determining the feasibility of utilizing DMS to publicize travel surveys is what information can be displayed that the traveling public will be able to remember. Either an internet URL or telephone number would appear to be the most likely information needed to make the public aware of a survey effort. Conventional wisdom would indicate that short website addresses and telephone numbers would be easiest to remember. However, no information that directly related to the issue of website addresses on DMS was available. However, research has been performed on the public's recall capabilities for traditional DMS messages and telephone numbers provided in

AMBER alerts. Studies have shown that message length and message familiarity factor into how much of the message a motorist will remember (*14*). A study on telephone numbers in AMBER alert messages determined that it took 3.2 seconds longer to read a DMS that contained a 10-digit telephone number versus a message that did not have a telephone number (*15*). Additionally, only 35 percent of the study participants were able to correctly recall the 10-digit phone number that was provided in the message. That level rose to 60 percent when the message was displayed for 6 seconds. Having a 1-800 or 1-888 prefix did not appear to improve the ability of people to remember the entire phone number. However, almost all of the study participants (99 percent at 2 seconds and 100 percent at 4 seconds) could recall a 3-digit number like 911 or 511.

Yet another potential method for collecting external survey data is the postcard method. With this method, postcards with prepaid postage are mailed or handed out to motorists at selected locations. The postcards typically contain questions on trip origin and destination, trip purpose, trip frequency, and vehicle occupancy. Postcard surveys have been implemented in various parts of the country, including Texas, with good results. However, one of the negative aspects of postcard surveys is a low response rate. It is not uncommon for postcard surveys to have a 10 percent or less response rate. Additionally, some surveys can be returned with incomplete or missing data, thus lowering the rate of usable surveys. Distribution methods for the postcards vary depending on the policies of the entities conducting the survey. Postcard surveys have been conducted in North Carolina using license plates captured at selected locations (*16*). License plates were recorded and transcribed, and postcards were mailed to the motorist's home address. Another method for distributing the postcards is handing them out to motorists at roadside locations. This method was used in Sussex County, New Jersey, in 2002. The postcard also included a website where participants could go to fill out the survey online. This survey effort had a 32 percent response rate, of which 18 percent of the respondents submitted their surveys via the website (*17*). Using this method would require some sort of traffic control plan in order to safely distribute the postcards. Postcard surveys may also be distributed at toll plazas and freeway on-ramps.

Another emerging area to monitor is the use of cellular phone global positioning satellite (GPS) data. In 2009, it was estimated that approximately 89 percent of Americans are mobile subscribers and nearly 20 percent of U.S. households were wireless only (*18*). Given the deep market penetration of this technology, the utility of cellular phones as a means to collect travel information is increasing. While the use of data from GPS enabled cellular phones involves some legal and privacy issues, there are an increasing number of transportation related agencies that have utilized this type of data for various planning purposes. A project called "Mobile Millennium" sponsored by Caltrans uses cellular phone GPS data to estimate real-time traffic conditions (*19*). Other programs, like the "TRAC-IT" program developed at the University of South Florida, involve the use of a software application that is downloaded onto a cellular phone (*20*). The software has a passive mode that records the GPS related data and an active mode that provides the phone owner the ability to provide trip-related information such as trip mode, trip purpose, and vehicle occupancy. While there are many other programs similar to these that are emerging, a salient consideration with the use of this technology is the recruitment process and how the recruitment can minimize any sampling bias that is produced.

METHODS FOR DEVELOPING NEW ESTIMATES USING EXISTING DATA

There are a number of methods for developing new external estimates that have been devised over the years. The following sections provide an overview of the methods for developing both external-external and internal-external trip matrices.

Estimating External-External (E-E) Trip Matrices

Modlin developed a two-phase process to estimate the E-E trip matrix for small cities. The first phase uses regression equations based on the study area population, average daily traffic (ADT) at the external station, and percent of trucks at the external station to predict the percent of trips at each external station that are E-E trips. This model assumes, as do most models, that this percentage is the same for both directions over a 24-hour period. The second phase uses a set of five regression equations to estimate the distribution of E-E trips between external stations. Each of the five equations corresponds to one of five functional classes applied to the origin external station. The independent variables are percent of trips at the destination station that are E-E trips, route continuity between origin or destination (0 or 1), ratio of ADT at destination station to total ADT at all stations, ADT at destination station, and percent trucks at the destination station. Modlin developed these models using data from cities with less than 50,000 people. Pigman developed similar equations.

Anderson developed a regression equation that simultaneously estimates the percent of traffic at each external station that is E-E trips, and the distribution of E-E trips between external stations, for cities with less than 50,000 people. The independent variables for this model include the presence of a major city at the destination station (0 or 1), route continuity between the origin station and the destination station (0 or 1), the ADT at the destination station, and if the external station is used for E-I trips (0 or 1) (*21*).

One of the drawbacks of the Modlin and Anderson Models is that they do not explicitly address the economic context and network configuration of the study area. To address this weakness, Anderson evaluated three spatial economic models, including the simple gravity model, Huff's probability contours, and Zipf's law of spatial interaction. These models consider the effect that neighboring competing communities have on external travel patterns for the study area. In other words, they determine if a person from a neighboring city will patronize businesses within the study area, or travel through the study area to patronized businesses in another city. Anderson applied these models to develop an E-E matrix for a test study area, and concluded that Huff's probability contours provided the best results. Anderson later applied the same model to other cities with populations less than 50,000, and concluded that the model gave good results (*22*, *23*).

Horowitz developed a geometric/geographic model to estimate the distribution of E-E trips between external stations. In his model, each external station has a catchment area outside of the study area. Line segments connecting points within pairs of catchment areas are constructed, some of which cross the study area boundary, and some do not. Catchment area pairs with more line segments that cross the study area boundary are more likely to have E-E trips between their corresponding external stations. In addition, the catchment areas can be constructed to simulate a barrier to travel, such as a body of water. This model generates weighting factors that are then used in the procedure described in the Quick Response Freight Manual. It should be noted that it is assumed that the E-E volume at each external station is known. The model was applied to two medium-sized cities with good results.

Han developed regression equations to estimate the percent of trips at external stations that are E-E trips, as well as the distribution of E-E trips between external stations. The equations combined included independent variables similar to those found in the Modlin and Anderson models, with independent variables reflecting results from the Horowitz model and Zipf's law of spatial interaction. The independent variables for the E-E trip split model are the functional classification of the road, whether or not the city is small (less than 50,000 people), whether or not the external station connects to a marginal route, ADT at the external station, population of the study area, percentage of trucks at the external station, area in square miles of the study area, and the employment in the study area. This equation was developed for cities with population less than 200,000. The E-E trip distribution model included two equations: one for cities with population less than 50,000 and one for cities with population between 50,000 and 200,000. The equation for small cities included route continuity, number of lanes at destination, percent trucks at destination, a factor related to Horowitz's model, and a factor related to Zipf's law as independent variables. The independent variables for the medium-sized city equation are route continuity, ratio of ADT at the destination to the sum of all ADTs, a factor related to Horowitz's model, and the percent trucks at the origin external station (*24*).

Estimating Internal-External (I-E/E-I) Trip Matrices

NCHRP Report 365 describes a process to estimate external trips. The process starts with estimating the percent of external trips that are through trips, and distributing the through trips between external stations, using equations developed by Modlin. The split of E-I/I-E trips between purposes (home-based work, home-based other, and non-home-based), is estimated based on possible ranges developed from previous studies and local knowledge. The percent of trips of each purpose with productions outside of the study area is estimated in the same way. For example, the observed A.M. directional split at external stations can be used to estimate the percentage of home-based work trips that are produced outside of the study area. Trips are distributed as usual using the gravity model.

Han et al. presented a method for estimating all travel for small urban areas. This method is unique from the previously described methods in that it provides a way to estimate I-I trips made by non-residents. It recommended that E-E trips be estimated using a software program (SYNTH), or using growth factors from historical data. I-E trips are estimated as a percentage of trips produced within the study area, based on local knowledge. E-I trips are estimated as the ADT minus the E-E trips and the I-E trips. E-I trips are assumed to lead to secondary I-I trips (non-home-based trips). These are estimated to be a certain percentage of E-I trips, based on local knowledge. The E-I trips are distributed according to the percent of employment in each zone. Secondary I-I trips are combined with all other I-I trips and distributed according to households at the origin zone and employment at the destination zone (*25*).

The FHWA Model Validation and Reasonableness Checking Manual also describes a method to estimate I-I trips made by non-residents. The ratio of the number of non-home-based trips to home-based trips is assumed to be the same for non-residents as it is for residents (*26*).

Method for Developing New Estimates Using State Planning Data

The use of statewide planning models has the potential to provide some of the data elements needed for input into urban travel demand models. When querying the Travel Model Improvement Program (TMIP) listserv, a number of respondents indicated that their organizations utilized statewide planning models to develop urban external estimates. For instance, in Oregon an analysis is performed at select link locations (at MPO boundaries) within the statewide model. From that, external-local and external-through trip tables are developed for the MPO area. The Texas Statewide Analysis Model includes a utility for creating urban area trip tables. The process assigns volumes to all external links and then generates trip matrices for E-E and I-E/E-I trips. A limitation in the statewide planning model is that it is limited to the base and forecast years. Therefore, if a different year was needed for an urban area, measures would be required to adjust the estimates. Additionally, the current SAM does not include proposed future facilities.

CHAPTER 3. ANALYSIS OF TRAFFIC CONTROL PLANS

The conduct of external travel surveys in Texas has consisted primarily of a roadside intercept interview method. With this method, contracted personnel establish a traffic control plan (TCP), and then randomly request motorists' participation in the survey. This method has proven to be very effective at capturing high-quality travel data at a reasonable cost. In 2008, TxDOT developed a new set of standardized TCPs to be used during the conduct of external travel surveys. Since no new external surveys have been conducted since the introduction of the new TCPs, an analysis of the impact of the new TCP was performed. As part of a TxDOT research project (*27*), a queuing analysis of two TCP configurations was performed to ascertain the impact of the bottleneck created by the TCP. In order to determine the impact of the changes to the TCP, a comparison of the old and new TCPs was performed. The comparison was limited to 4-lane cross-sections of roads where two lanes of traffic (traveling in the same direction) were merged into one prior to the survey station. Figure 3 provides the old TCP, while Figure 4 and Figure 5 provide the new TCPs without and with shoulders, respectively.

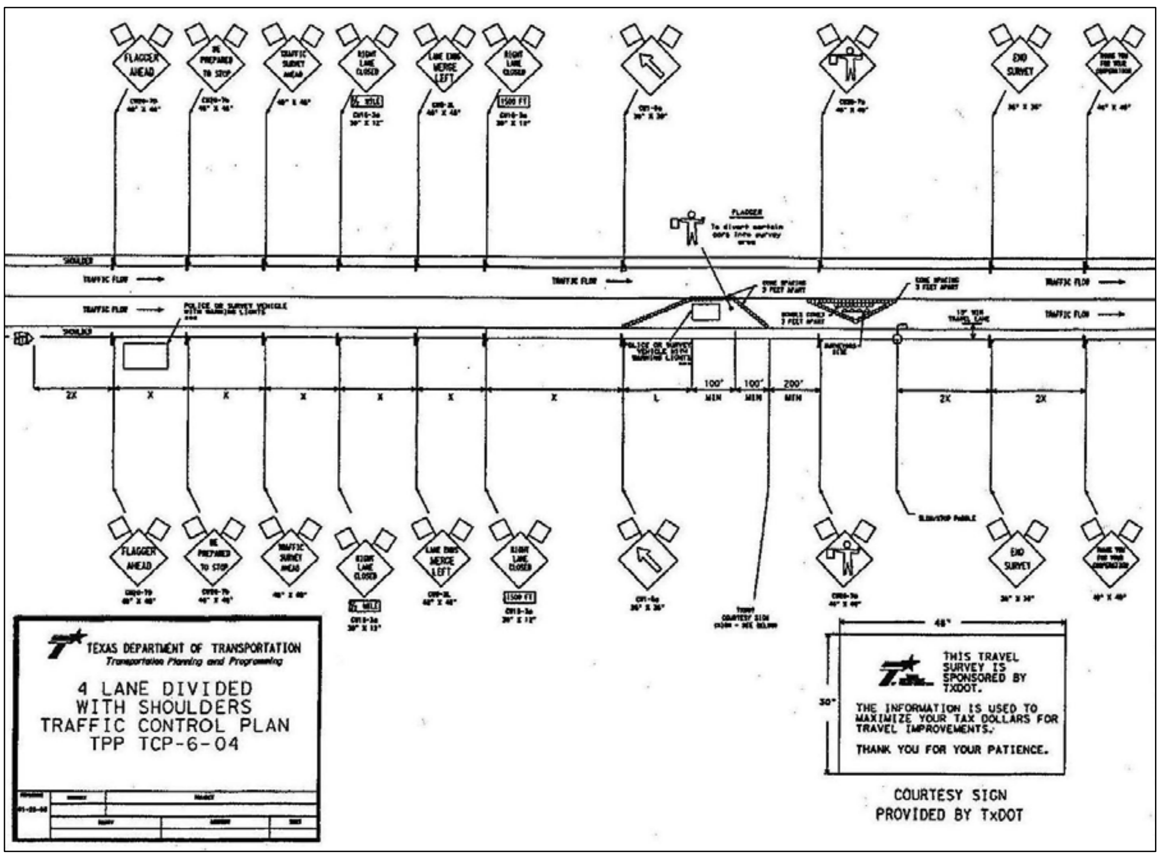

Figure 3. Old TxDOT External Station TCP.

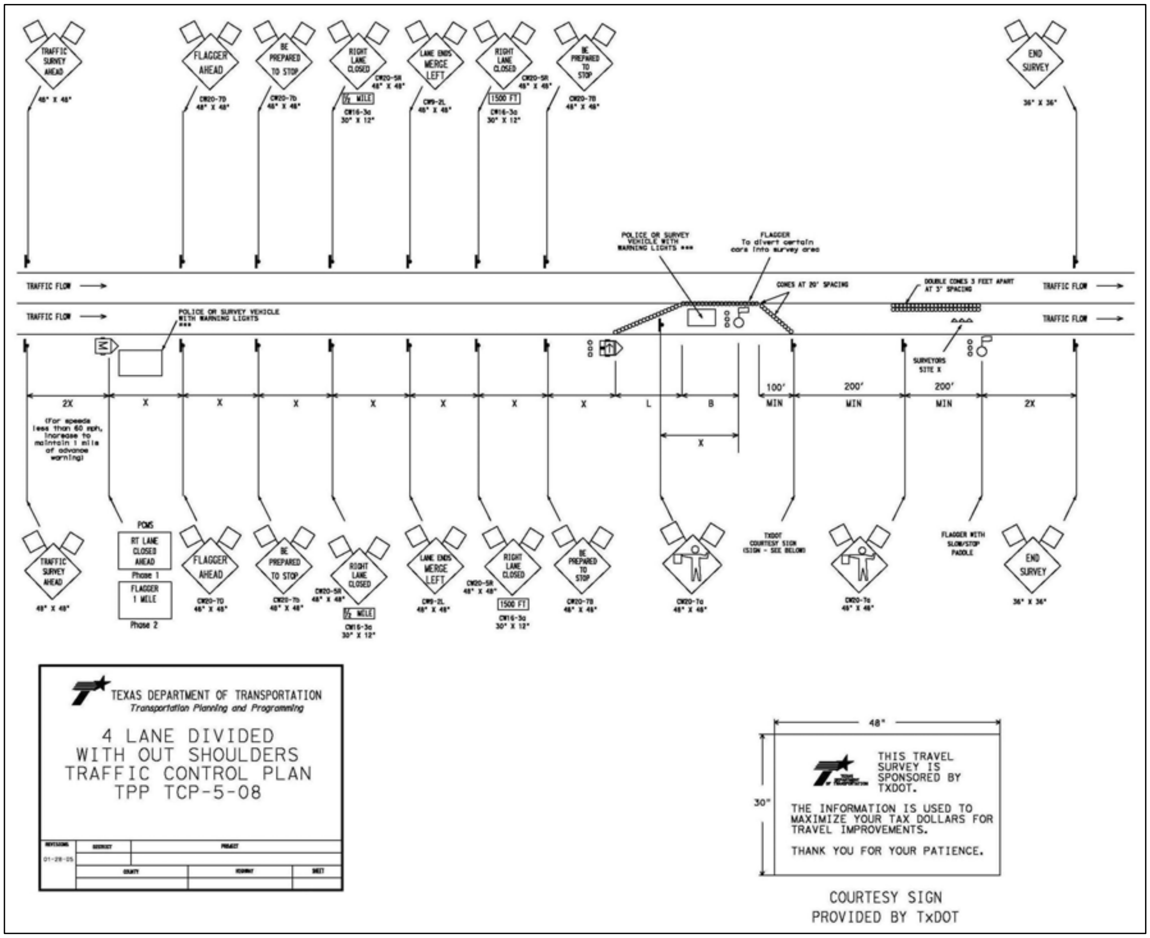

Figure 4. New TxDOT External Station TCP (without shoulders).

All of the scenarios involve tapering of a two-lane roadway into a one-lane roadway as the vehicles approach the survey station. Under the old TCPs, the drivers were slowed down and flagged into the survey bay from the outer main lane. The vehicles flagged into the bay were asked to complete the survey. Due to safety and operational factors, drivers who refused to participate in the survey had to remain in the survey station until all the other survey participants completed the surveys. In other words, only the drivers diverted into the bay were delayed by the survey procedure.

Under the new TCPs, the drivers will be stopped and asked whether they would be willing to participate in the survey. Drivers electing to not participate will be able to continue their trip without being delayed by the survey procedure. However, they may still be delayed by other factors such as the stop-and-ask procedure itself.

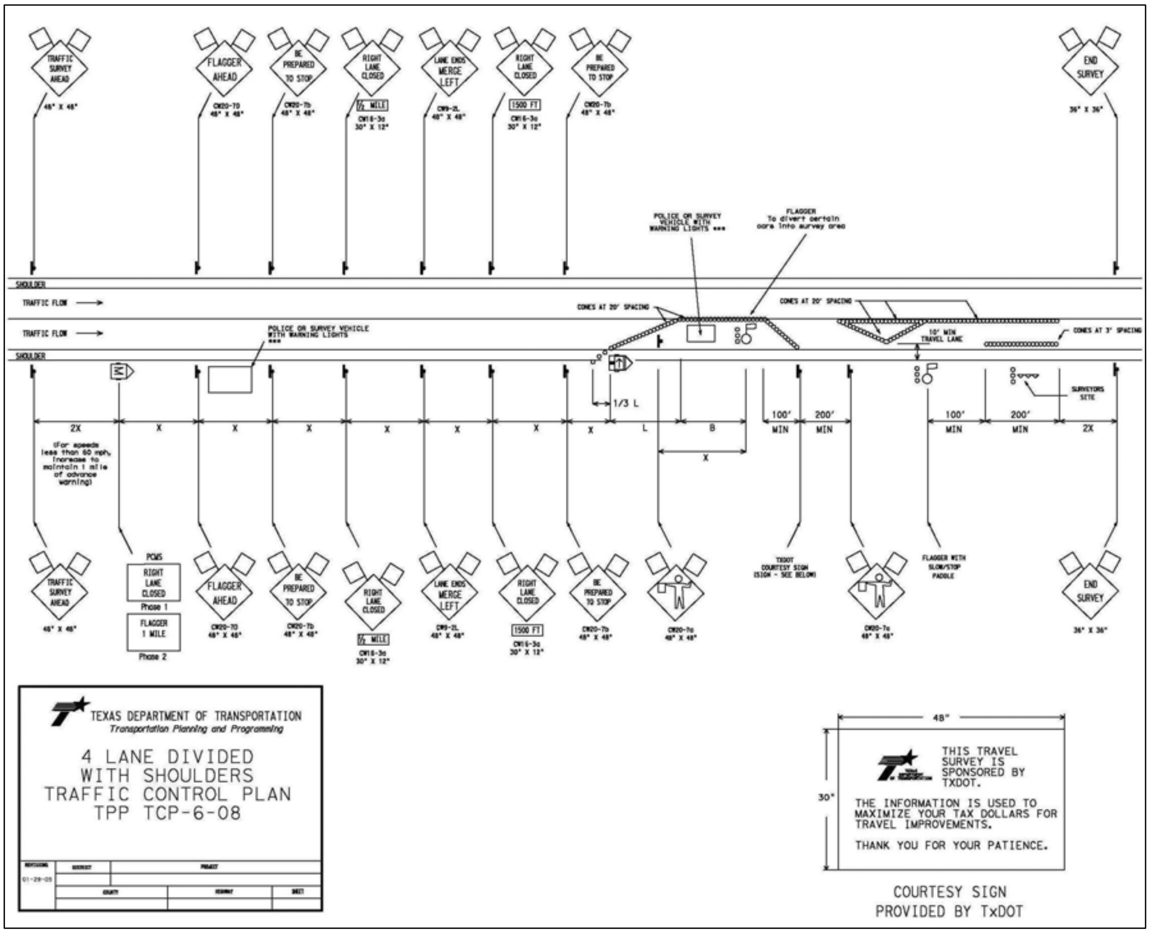

Figure 5. New TxDOT External Station TCP (with shoulders).

Figure 4 shows the new TxDOT TCP for the case of 4-lane undivided roadways without shoulders. The two-lane roadway is tapered down to one lane as the vehicles approach the station. The stop-and-ask procedure will take place on this lane and thus require all the vehicles to stop. The procedure generally continues until the survey bay is full or the queue length exceeds the acceptable threshold as specified by TxDOT.

Figure 5 illustrates the new TxDOT TCP for the case of 4-lane undivided roadways with shoulders at least 10 feet wide. The configuration is similar to the previous case but there is adequate space inside the station to set up two lanes. The flagger will divert the specified number of vehicles from the open main lane into the receiving bay. The traffic will not be stopped on the open main lane. At the receiving bay, the drivers will be stopped and asked to participate in the survey. The accepted drivers will then take a right lane to continue to where the surveyors are located. The rejected drivers can take a left lane to bypass the survey procedure and leave the survey station.

The major difference between the new TCPs with and without shoulders is the location of the receiving bay. With the presence of at least 10-foot wide shoulders, the drivers will be pulled over and stopped inside the bay and asked whether they are willing to do the survey. In this case, the outer main lane will remain open to the traffic. In the case of no shoulder or inadequate shoulder width, the stop-and-ask procedure will take place on the open main lane and thus require all of the traffic to be stopped.

ANALYTICAL PROCEDURE

In a previous TxDOT project 0-4869, TTI researchers examined the impacts of the old TCPs using the concept of queuing models and analytical discrete event simulation technique (*27, 28*). The analytical models used in the previous analysis cannot be easily extended to quantify the impacts of the new TCPs. Therefore, in this study, the researchers utilized VISSIM microscopic simulation software to replicate the operational procedure of both the old and new TCPs in order to provide more realistic and defendable results of the impacts from the new TCPs.

OVERVIEW OF MODELING TOOLS

VISSIM has become one of the industry-standard tools for microscopic simulation of transportation system. VISSIM is a microscopic, time step and behavior based simulation tool developed to model multi-modal traffic and transit operations. VISSIM has gained a significant share of users in recent years due to its capability to replicate a complex transportation system with a reasonable level of accuracy (*29*).

To replicate the operational logic and procedures of both old and new TCPs within VISSIM, researchers utilized a VISSIM proprietary add-on module known as vehicle actuated programming (VAP). VAP is designed primarily for the simulation of programmable, phase or stage based, traffic actuated signal controls. VAP provides a solution for the users to develop unconventional traffic control logics or test new strategies that have not been implemented in the conventional traffic signal controllers. In this study, it is also particularly suitable for the design of the external survey procedure where the drivers will be stopped and asked to make a decision. In such cases, VAP control logic can be programmed such that it keeps track of the bay status and determines when to start and stop diverting the vehicles into the survey bay using appropriate signalized control.

The control logic can be coded in a text file using VAP programming script. During the VISSIM simulation runs, VAP interprets the control logic commands and creates the signal control commands for the VISSIM network (*30*). VAP implementation in VISSIM is modular and highly customizable. Examples of VAP applications include ramp metering applications, advanced freeway management strategies, and transit signal priority (*31, 32*). Figure 6 shows the example of basic VAP scripts for controlling routing decisions. The example shows the diversion control where the percentage of drivers accepting the survey (referred to as a route choice in VAP) can be specified using the 'set_route' function of the VAP. This routing logic, for instance, can be part of larger subroutine that determines when to start and stop diverting the traffic.

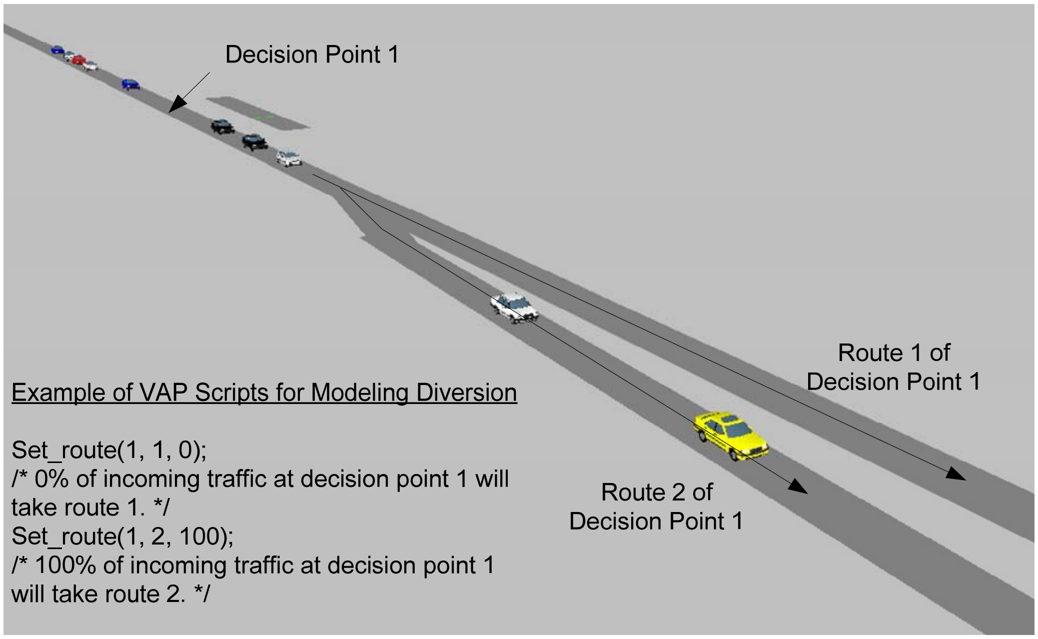

Figure 6. Example of VAP Scripts for Controlling Routing Decisions.

OPERATIONAL ASSUMPTIONS

To analyze the TCP procedures using VISSIM, some operational assumptions are required for the simulation. This section describes how the external survey procedures were replicated in the simulation.

Old TCP for 4-Lane Undivided Roadways

Under the old TCPs, the vehicles approaching the survey station were slowed down but not stopped. The two-lane roadway was tapered down into a one-lane segment. A flagger diverted all the approaching vehicles into the survey bay until the bay was full. In other words, the maximum number of vehicles allowed into the bay was equal to the number of crew members in the survey station. When the bay was full, no more vehicles were allowed to enter the survey bay. The survey procedure took place inside the survey bay and the length of survey was assumed to follow the normal distribution. The surveyed drivers could not leave the station until all the drivers in the bay had finished the survey. Once all of the drivers completed the survey, they were allowed to leave the station all at once. The bay remained shut down for an additional 15 seconds (buffer time) to ensure that all of the vehicles had completely departed the station. Once the buffer time lapsed, the bay reopened for new vehicles and the entire procedure was repeated. There was no queue condition under the old TCPs, and therefore no queue policy was required.

New TCP for 4-Lane Undivided Road without Shoulder

Under the new TCP for the case of 4-lane undivided without shoulder or with less than a 10-ft shoulder, the differences between the new and the old TCPs are:

- All drivers must be stopped and asked whether they would be willing to participate in the survey.
- Participation is optional and therefore the participation rate is a critical factor for the analysis. In contrast, the participation rate under the old TCPs is 100 percent.
- The surveyed drivers will be able to leave the station as soon as they complete the survey.
- The flagger will continuously stop and ask the vehicles to participate if there are available slots in the bay and the queue length is below the acceptable limit.
- The bay will be closed when it is full or the queue length is beyond the acceptable threshold.
- The queue policy is required. The current TxDOT threshold is 10 vehicles or approximately 250 feet (passenger cars only). Other thresholds were also examined as part of the analysis conducted in this study.

New TCP for 4-Lane Undivided Road with Shoulder

In the previous case, the flagger will stop the traffic and ask the drivers to participate in the survey on the open main lane (one lane). Under the new TCP for the four-lane undivided roadways with at least 10-ft shoulder, there is sufficient space for two lanes inside the survey bay and thus the stop-and-ask process can take place at the receiving lane inside the bay rather than on the open main lane. The following operational assumptions are used for the analysis.

- Vehicles are diverted into the bay as a platoon with the size being equal to the number of survey crews in the bay.
- The traffic will be slowed but no longer stopped on the open main lane and thus no queue condition would be formed on the open main lane.
- The queue will occur at the receiving bay inside the station but it will be limited to the size of the vehicle platoon flagged into the bay.
- The stop-and-ask process takes place inside the bay.
- The bay will shut down when the platoon size is equal to the number of survey crews.
- If the number of vehicles agreed to participate in the survey is less than the number of crews minus one, the bay will reopen and the flagger will continue to flag the vehicles into the bay in platoons of size equal to the number of survey crews. When the number of vehicles participating in the survey is equal to or greater than the number of crew members minus one, the bay will shut down. For example, for four crew members, the bay will remain open if there are less than three vehicles in the bay and will close when there are three or more vehicles in the bay.
- The vehicles that reject participating in the survey will be diverted to the left lane inside the bay to bypass the survey station.
- All the following conditions must be satisfied to reopen the bay: (a) the queue that forms inside the bay from the stop-and-ask process has been cleared for at least 20 seconds (no platoon queue remains), and (b) a period of at least 20 seconds has lapsed since the last closure.

- The surveyed vehicles can leave the station as soon as they complete the survey.

SIMULATION SETUP OF TCPS IN VISSIM

Various signal control features in VISSIM are used to replicate the operations at the external survey station.

- The stop sign control was used to simulate the stop-and-ask procedure for the new TCPs. The length of process can be varied by changing the dwell time at the stop sign. The dwell time was assumed to be normally distributed with specified mean and standard deviation values.
- VAP routing function in the VAP scripts was used to control the survey acceptance rate.
- VAP scripts and various detection features in VISSIM were used to monitor the number of vehicles inside the bay and the queue conditions that would in turn determine whether the bay should be closed.
- A timer function is VAP is used to keep track of how long the bay has been closed in order to maintain sufficient separation time between each round of surveys.
- A parking lot feature in VISSIM is used to replicate the location of the survey crews and the time it takes to complete the survey. The parking dwell time is treated as the survey completion time and it is assumed to follow normal distribution with specified mean and standard deviation values. The standard deviation is fixed at 10 percent of the mean values.
- Traffic composition is assumed to be 100 percent passenger cars. The analysis using 100 percent passenger-car mix provides the results for the best case scenario. The presence of trucks will generally degrade both the mobility of the road users as well as the efficiency of the external station.

Figure 7 shows an overview of the entire roadway segment simulated in VISSIM with speed settings and dimensions based on TxDOT TCPs. To ensure a valid comparison of the results, the dimensions used in the simulated testbed are unchanged for all the TCPs analyzed in this study.

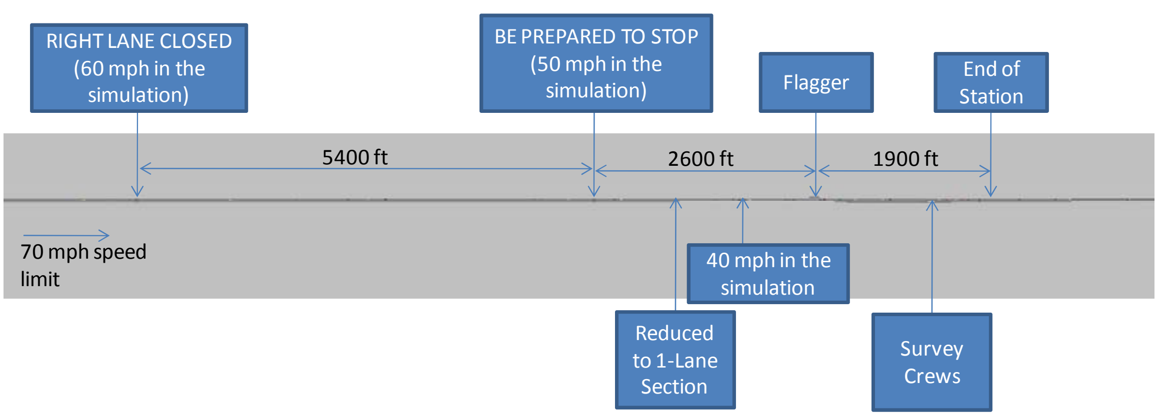

Figure 7. Overview of the Simulated Segment.

Figure 8 provides the detailed diagram of the simulated survey station. The speeds were reduced inside the station for both surveyed and non-surveyed drivers, but the non-surveyed drivers were allowed to travel at a slightly higher speed than the surveyed ones. Also of note is the location of the flagger for the new TCPs. With the presence of adequate shoulder width, the stop-and-ask procedure will take place at the receiving bay inside the survey station instead of on the outer main lane.

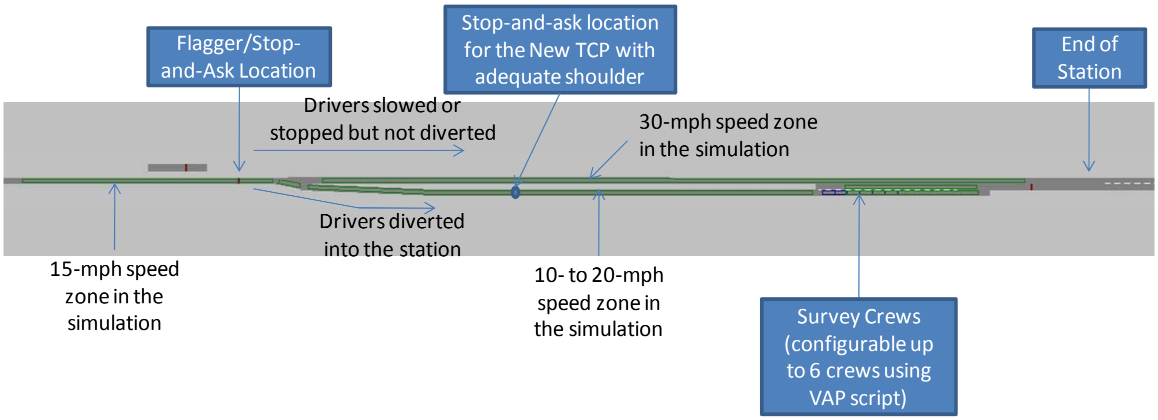

Figure 8. Detailed Diagram of the Survey Station in the Simulation.

Figure 9 displays the operation of the old TxDOT TCPs where the flagger diverts a platoon of vehicles into the survey bay. Both diverted and non-diverted vehicles are generally slowed down but not stopped as they approach the flagger. The diverted vehicles are generally surveyed and there is no stop-and-ask procedure involved. The platoon size is typically equal to the number of survey crews available. The surveyed drivers will be allowed to leave the station once everyone has completed the survey.

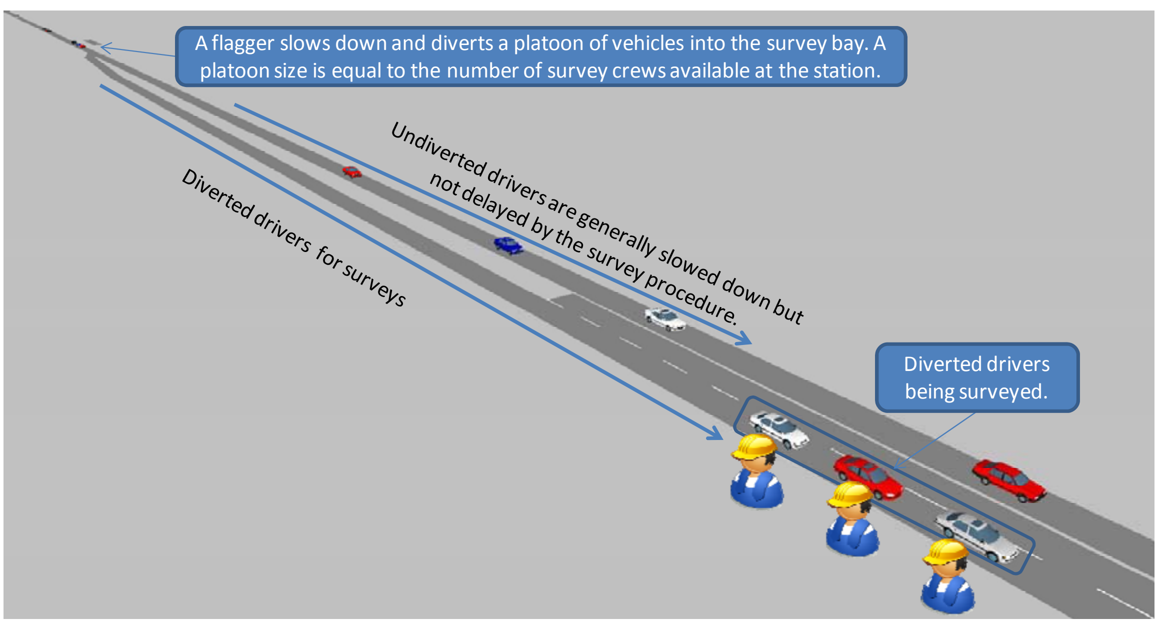

Figure 9. Simulated Operation of Old TxDOT TCPs.

Figure 10 and Figure 11 illustrate the simulated operation of the new TCPs without and with adequate shoulders, respectively. In the case of new TCPs without adequate shoulders, all the vehicles are stopped and asked on the main lane and only accepted drivers are diverted into the survey station. On the other hand, with adequate shoulders, a platoon of vehicles will be diverted into the receiving bay where they will be stopped and asked to participate in the survey. Drivers not wanting to participate will be allowed to take a bypass lane and leave the survey station. In this case, not all drivers are being stopped and delayed by the stop-and-ask procedure. In either case, the amount of delay incurred by the introduction of the stop-and-ask procedure in the new TCPs depends heavily on the participation rate. The researchers quantified the impacts of this factor in the discussions of the simulation results.

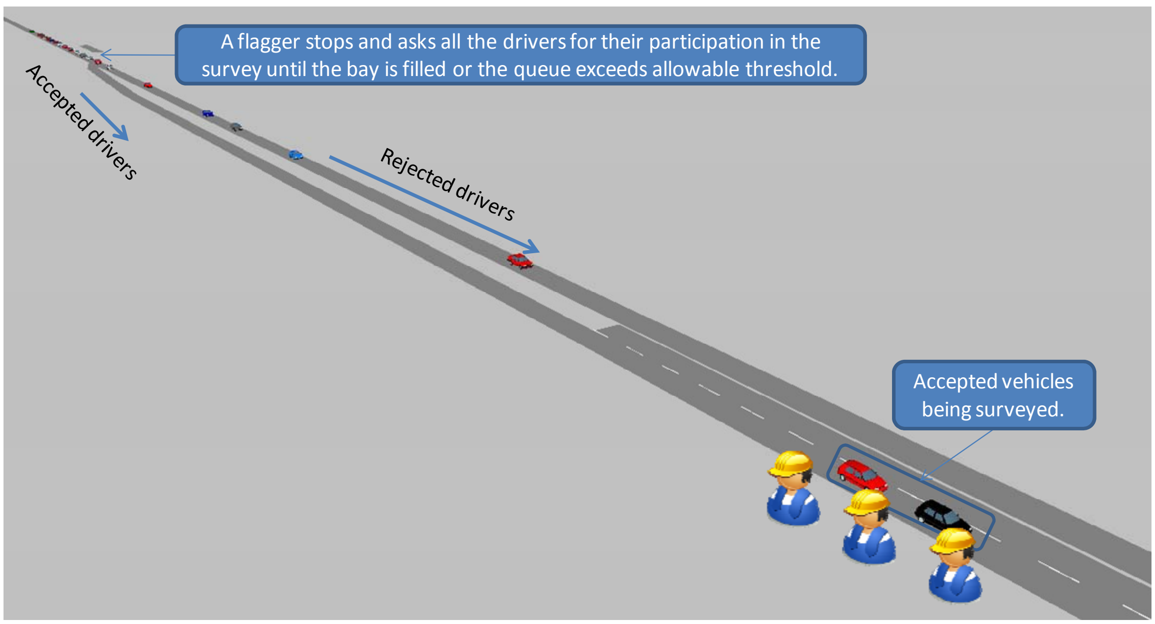

Figure 10. Simulated Operation of New TCP on Road without Shoulders.

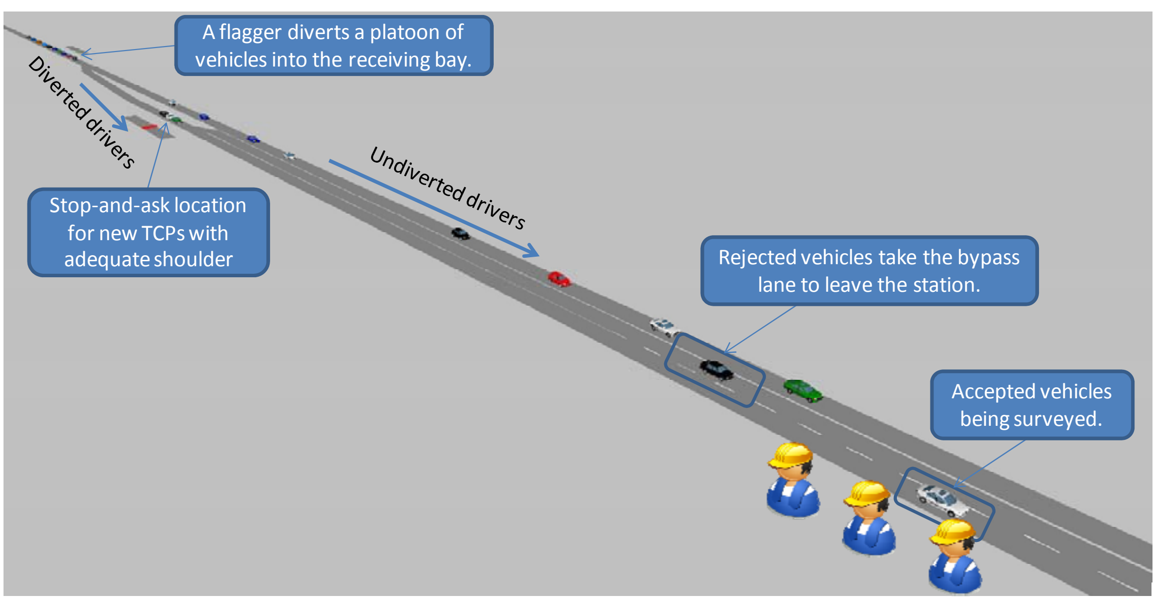

Figure 11. Simulated Operation of New TCP on Road with Shoulders.

EXPERIMENTAL DESIGN FOR THE SIMULATION ANALYSIS

There are two major simulation variables introduced as part of the new TCPs—stop dwell time and survey acceptance rate. The stop dwell time represents the time that it takes to stop and ask the drivers if they would be willing to participate in the survey. Two levels of stop dwell time were considered, which are 10 and 20 seconds. The values were assumed to be normally distributed with the standard deviations of about 5 percent of the mean values. The survey acceptance rate under the old TCPs was implicitly assumed to be 100 percent as the drivers were not asked to participate prior to being diverted into the survey station. However, this factor becomes a critical one for the new TCPs as the drivers will now have an option of whether to accept or reject participating in the survey. Consequently, it will take much longer to fill the survey station, particularly if the acceptance rate is low. Three levels of acceptance rate at 10 percent, 20 percent, and 30 percent were analyzed in this study.

The buffer time is the wait time between each round of the survey. This is implemented in the simulation mainly to prevent the overestimation of the survey efficiency. Without this buffer, the surveys can take place continuously without any breaks in between. The efficiency of the survey procedure can be overly optimistic as a consequence.

The number of survey crew members is required for the analysis of both old and new TCPs. However, under the new TCPs, the number of survey crew members can have much more impact as it can lengthen the amount of time required to fill the survey bay. When the acceptance rate is extremely low, the bay may never fill and thus require the flagger to continuously stop and ask the incoming traffic to participate. In the case of the new TCPs without adequate shoulder width, the queue length may exceed the allowable threshold before the bay is filled.

Table 3 summarizes the factors that influence the operation of the TCPs and the range of input values investigated in this study. Table 4 summarizes the factors that were varied in the simulation for the comparison of three TCP alternatives in this study.

Table 3. Factors Influencing the Operations of the TCPs.

Simulation Variables	Values
Stop dwell time	• 10 seconds or N(5, 0.5) • 20 seconds or N(15, 1.0)
Survey acceptance rate	• 10% • 20% • 30%
Number of survey crews	• 3 • 4 • 5
Survey completion time (Normal distribution with SD of 10% of the mean)	• 120 seconds • 180 seconds • 240 seconds
Buffer time (Wait time to begin the next round of diversion once the bay is cleared)	• 15 seconds
Traffic volume (vph) / ADT (K=5%) Per direction (2 lanes)	• 125 /2500 • 250/5000 • 500/10,000 • 750/15,000 • 1000/20,000
Queue Policy (unacceptable queue length; applicable to new TCPs only)	• 0 = No queue policy • 250 ft (approx. 10 vehicles) • 500 ft (approx. 20 vehicles)
% Truck	• 0% (All passenger cars) • Varied according traffic volume levels (not evaluated)

Table 4. Varied Factors for TCP Simulations.

Simulation Variables	Old TCPs	New TCPs without Shoulder	New TCPs with Shoulder
Stop dwell time		x	x
Survey acceptance rate		x	x
Number of survey crews	x	x	x
Survey completion time	x	x	x
Traffic volume (vph)	x	x	x
Queue Policy		x	

The survey completion time is assumed to follow normal distribution with three levels of mean values (2, 3, and 4 minutes). The standard deviation is fixed at 10 percent of the mean values. Additionally, four traffic volume levels were examined in this study. The volume inputs in the simulation were ranging from 125 vehicles per hour (vph) to 1,000 vph for one direction of a 4-lane roadway. Assuming the K factor of about 5 percent, the equivalent average daily traffic ranges are from 2,500 to 20,000. The traffic composition used in the simulation is 100 percent passenger cars for all TCP scenarios.

The queue policy is needed for the operation of new TCPs without adequate shoulder width. When the queue exceeds the allowable threshold, the flagger will stop asking drivers to participate and will release all of the vehicles in the queue. The flagger will resume the stop-and-ask procedure once the queue length drops below the limit. The researchers examined three options of queue policy in this study: (a) no limit, (b) 250 feet or approximately 10 passenger cars, and (c) 500 feet or approximately 20 passenger cars.

Design of Simulation Scenarios

Given the number of variables utilized in the analysis, a significant number of simulations were performed. Table 5 summarizes the design of experimental scenarios for this study. The old TCPs are considered as the base case with a total of 45 base scenarios. The researchers first analyzed the case of new TCPs on roads without adequate shoulder using full factorial experimental design, which resulted in a total of 810 scenarios. After the data were analyzed from this full-scale run, the researchers limited the number of levels considered in the case of new TCPs on roads with adequate shoulder to cover only key variables. The selected experimental runs for the second case resulted in a total of 30 scenarios. Each scenario was run with five replications. In the end, a total of 4425 simulation runs were conducted in this study.

Table 5. Experimental Setup.

Simulation Variables	Old TCPs	New TCPs without Shoulder	New TCPs with Shoulder
Stop dwell time	N/A	2 levels	1 level
Survey acceptance rate	N/A	3 levels	2 levels
Number of survey crews	3 levels	3 levels	1 level
Survey completion time	3 levels	3 levels	3 levels
Traffic volume	5 levels	5 levels	5 levels
Queue Policy	N/A	3 levels	N/A
Total Scenarios	**45**	**810**	**30**
Replications per scenario	5	5	5
Total Runs	**225**	**4050**	**150**

Data Collection and Measures of Effectiveness

A number of metrics were derived to quantify the differences between the TCP configurations. Table 6 provides a list of the measures of effectiveness (MOEs) that were observed from the simulation. Note that some MOEs are irrelevant to some TCP scenarios. For instance, the queue length is defined in the simulation as when the vehicles' travel speed is near zero. Therefore, the queues were never observed in the case of old TCPs where the vehicles are only slowed down by the flagger but not completely stopped.

Table 6. Simulation Performance Measures.

Performance Measures	Description
Delay (sec/veh)	• Measured only for non-surveyed vehicles. • *Delay_NonSurv*: Measured difference between actual travel time and desired travel time (based on desired travel speed).
Travel Time (sec)	• Average travel time measured from the beginning to the end of the simulated segment. • *TT_All*: Average travel time for all vehicles. • *TT_NonSurv*: Average travel time for non-surveyed vehicles. • *TT_Surv*: Average travel time for surveyed vehicles.
Throughput (veh/hr)	• Rate at which the vehicles are passing through the survey station. • *Veh_All*: Throughput for all vehicles. • *Veh_NonSurv*: Throughput for non-surveyed vehicles. • *Veh_Surv*: Throughput for surveyed vehicles. This is equivalent the number of surveys completed per hour.
Queue Length (feet)	• Measured at the location of the flagger. • *Q_Avg*: Average queue length at 10-second intervals throughout the observation period. • *Q_95thMax*: 95th percentile of maximum queue length observed at 10-second intervals for the entire observation period.
Surveyor Utilization (%)	• *Surv_Utz*: Percentage of time the surveyors are occupied/busy conducting the surveys.

The collected MOEs can be classified into two categories: (a) operational impacts and (b) survey efficiency. The operational impacts observed from the simulation are:

- delay – measured by the difference between the actual travel time and the desired-speed travel time (the time that a vehicle will use to travel through the segment if it were to travel at the desired speeds). This measure is applicable to only non-surveyed vehicles;
- queue length – average and 95th percentile queue length;
- throughput of non-surveyed vehicles; and
- travel time – measured separately between surveyed and non-surveyed vehicles.

The survey efficiency is measured by:

- survey output (veh/hr) – equivalent to the throughput of surveyed vehicles or the number of surveys completed per hour;
- survey output rate (veh/hr/crew) – the number of surveys completed per hour divided by the number of survey crews; and
- surveyor utilization rate (%) – this is the percent of time the surveyors are busy conducting the survey. Special data collection points are strategically placed in the simulation at the location of the survey crews to collect this data.

Surveyor utilization rate (% time busy conducting the surveys) is calculated from the occupancy percentage observed through VISSIM data collection points and is provided in Equation 1 below.

$$\% \text{ Surveyor Utilization} = \frac{\sum \text{Occupancy Rate}}{\text{Number of Surveyors}} \qquad \text{(Eq. 1)}$$

All of the simulation runs are for 70 minutes. The first 10-minute period was a warm-up period for the simulation run. The simulation data were collected for 60 minutes after the warm-up period. All the performance measures collected for each scenario were averaged from five replications. Researchers developed a small program to help with the post-processing of simulation outputs in this study.

RESULTS

This section describes the results obtained from the simulation experiment. Researchers utilized regression analyses to quantify the impacts of simulation variables investigated in this study. Only statistically significant variables (p-value < 0.05) were retained in the final model specifications presented herein.

Regression models were calibrated from full factorial simulation outputs for the cases of old TCPs and new TCPs without adequate shoulder. There are significant differences between the operations of these two TCPs, which justify the efforts to develop two separate sets of models. However, the operations of old TCPs and new TCPs with adequate shoulder share several similar characteristics. Therefore, the regression models for the case of new TCPs on roadways with adequate shoulder were not separately calibrated. Many results in this case can be inferred from the old TCPs as described subsequently in this section.

Impacts of Old TCPs

The delay incurred to the non-surveyed vehicles under the old TCPs depends almost exclusively on the volume levels and can be estimated using the Equation 2 below. The volume input is based on a directional hourly volume of a 2-lane roadway (4-lane undivided).

$$\underset{(\text{sec/veh})}{\text{Delay}} = 1.3483 + 0.0144 \times \underset{(\text{vph})}{\text{Volume}} \qquad \text{(Eq. 2)}$$

Figure 12 depicts a linear relationship between volume level and vehicular delay under the old TCPs regardless of the number of survey crews or survey completion time. The coefficient of determination (R^2) indicates an almost perfect relationship. There is also no observable queue under the old TCPs. Note that the vehicles are only slowed down and diverted into the survey bay in platoons. Therefore, no queue models were developed in this case.

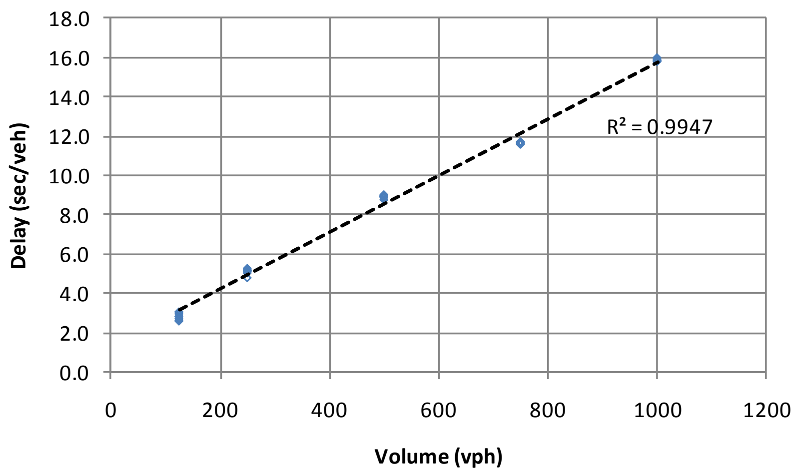

Figure 12. Impacts of Volume on Delay under Old TCP.

Under the old TCPs, the survey output rate, survey output, and the surveyor utilization can be estimated using Equations 3 through 5.

$$\underset{(\text{veh/hr/crew})}{\text{Survey Output Rate}} = 19.1036 - 0.0409 \times \underset{(\text{sec})}{\text{Completion}} + 0.0028 \times \underset{(\text{vph})}{\text{Volume}} - 0.4916 \times \underset{(\text{persons})}{\text{Crews}} \quad \text{(Eq. 3)}$$

$$\underset{(\text{veh/hr})}{\text{Survey Output}} = 30.6830 - 0.1626 \times \underset{(\text{sec})}{\text{Completion}} + 0.0117 \times \underset{(\text{vph})}{\text{Volume}} + 9.2933 \times \underset{(\text{persons})}{\text{Crews}} \quad \text{(Eq. 4)}$$

$$\underset{(\%)}{\text{Surveyor Utilization}} = 33.6789 + 0.1201 \times \underset{(\text{sec})}{\text{Completion}} + 0.0054 \times \underset{(\text{vph})}{\text{Volume}} - 2.6067 \times \underset{(\text{persons})}{\text{Crews}} \quad \text{(Eq. 5)}$$

From Equations 3 and 4, it was found that the survey output and survey output rate increase with the increase in traffic volume and the number of survey crews but reduce with the longer survey completion time. This is intuitive as more traffic and crew members should produce more surveys, but longer survey time will reduce the rate at which they can be completed.

Surveyor utilization under the old TCPs increases with longer survey and higher traffic volume. In contrast, the utilization rate decreases with more survey crews as this generally increases the likelihood of idling crew members.

Impact of New TCPS on Roads without Adequate Shoulders

In the case of new TCPs on roadways without adequate shoulder, the delay incurred from the survey is a direct result from the stop-and-ask procedure conducted on the main lane. The models were calibrated separately for the scenarios with and without a queue policy. There is no statistically significant difference on the impacts between applying 250-ft versus 500-ft queue limit policy.

The delay incurred to non-surveyed vehicles under the new TCPs on roads without adequate shoulder with the active queue policy (either 250-ft or 500-ft) can be estimated using Equation 6:

$$\underset{(\text{sec/veh})}{\text{Delay}} = 0.0694\times\underset{(\text{vph})}{\text{Volume}} - 0.1200\times\underset{(\text{sec})}{\text{Completion}} - 1.7096\times\underset{(\%)}{\text{Acceptance}} + 5.1619\times\underset{(\text{sec})}{\text{Dwell}} \qquad \text{(Eq. 6)}$$

Without queue policy, the delay impact is far greater and can be estimated using Equation 7:

$$\underset{(\text{sec/veh})}{\text{Delay}} = -482.7373 + 0.9498\times\underset{(\text{vph})}{\text{Volume}} - 1.3308\times\underset{(\text{sec})}{\text{Completion}} - 22.2459\times\underset{(\%)}{\text{Acceptance}} + 40.1124\times\underset{(\text{sec})}{\text{Dwell}} + 114.6700\times\underset{(\text{persons})}{\text{Crews}} \qquad \text{(Eq. 7)}$$

To compare the impacts from the new TCPs versus the old ones, the estimated delay equations were used to determine the combination of simulation parameters that would provide the best-case and worst-case delay scenarios as shown in Table 7. Under the old TCPs, only volume is the key factor. However, under the new TCPs, the worst-case delay scenario happens when (a) the survey completion time is short, (b) survey acceptance rate is low, (c) traffic volume is high, and (d) stop dwell time is large. The number of survey crews generally has no impacts on the delay when the queue limit policy is in effect. It is interesting to note that short survey completion time will result in a faster turn-around time for the surveyors, meaning that the flagger will have to stop and ask the traffic more frequently in order to fill the bay. This process causes the survey station to be operating in the stop-and-ask mode more often and thus resulting in greater delay to the traffic.

Table 7. Selected Scenarios for Delay Comparison.

Parameters	New TCP		Old TCP	
	Best Case	**Worst Case**	**Best Case**	**Worst Case**
Survey Completion Time (sec)	240	120	No Effect	No Effect
Survey Acceptance Rate (%)	30%	10%	N/A	N/A
Volume (vph)	125	1,000	125	1,000
Stop Dwell Time (sec)	10	20	N/A	N/A
Number of Survey Crews	No Effect with Queue Policy	No Effect with Queue Policy	No Effect	No Effect
Number of Survey Crews	3 (without Queue Policy)	5 (without Queue Policy)	No Effect	No Effect

Figure 13 illustrates the impacts of the new TCPs (inadequate shoulder) for the best-case delay scenario with a 250-foot queue policy. It was found that:

- Under the best case, the delay is under 40 seconds per vehicle for all volume range.
- There is no apparent breakpoint in volume in contrast to the worst-case scenario.

- The average queue length is maintained within 250 ft for all volume ranges, but the 95th percentile queue exceeds 250 ft when volume is greater than 600 vph.

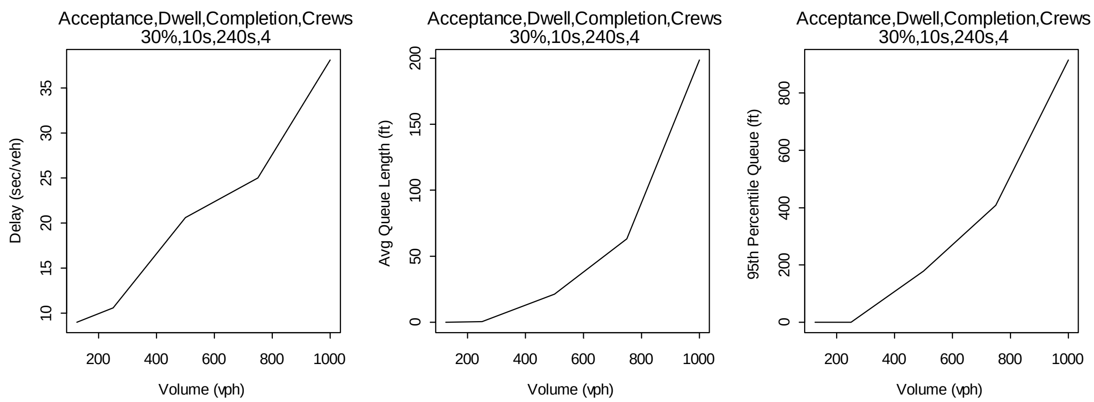

Figure 13. Best Case Scenario for New TCP on Roads without Shoulders.

Figure 14 shows the impacts of new TCPs (inadequate shoulder) under the worst-case delay scenario with a 250-foot queue policy.

- The breakpoint at which the delay experienced by the motorists increases significantly is approximately at 750 vph. The delay can exceed 100 sec/veh even at low volume conditions.
- The average queue exceeds 250 ft at 500 vph or higher.
- The 95th percentile queue exceeds 250 ft at 250 vph or higher.

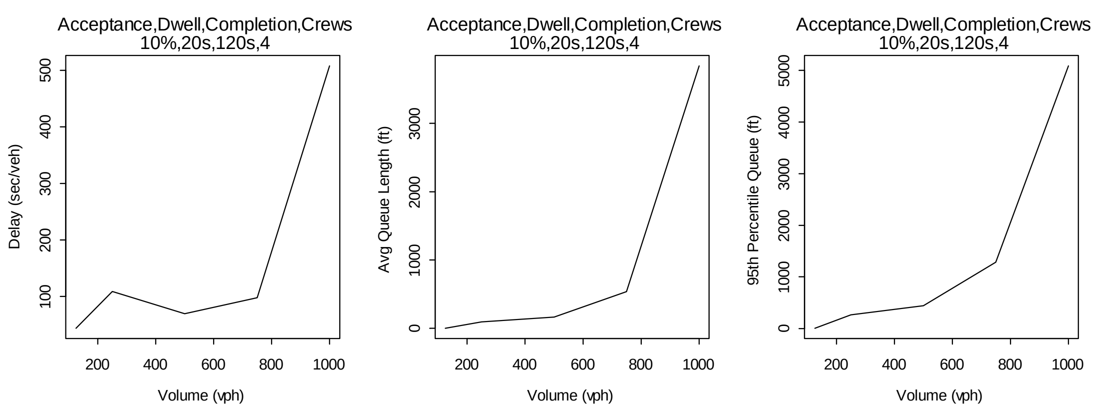

Figure 14. Worst Case Scenario for New TCP on Roads without Shoulders.

From the survey efficiency aspect, the new TCPs for roadways with inadequate shoulder are less efficient that the old TCPs in all aspects. The survey output rate for the new TCPs without adequate shoulder can be estimated using Equation 8. Under the new TCPs, higher volume significantly decreases the survey output rate especially when the queue policy is in use. When the queue exceeds the threshold, the flagger can no longer stop and ask the vehicles in order to mitigate the queue length. The queue can grow beyond the limit rapidly with the high volume and thus keeping the surveyors in the idling conditions most of the time.

$$\underset{(\text{veh/hr/crew})}{\text{Survey Output Rate}} = 12.6735 - 0.0133 \times \underset{(\text{sec})}{\text{Completion}} - 0.0014 \times \underset{(\text{vph})}{\text{Volume}}$$
$$-0.7624 \times \underset{(\text{persons})}{\text{Crews}} + 0.1762 \times \underset{(\%)}{\text{Acceptance}} - 0.2003 \times \underset{(\text{sec})}{\text{Dwell}} - 1.7470 \times \underset{(1 \text{ if limit queue})}{\text{Queue}} \quad \text{(Eq. 8)}$$

The survey output for the new TCPs on roadways without adequate shoulders can be estimated using Equation 9.

$$\underset{(\text{veh/hr})}{\text{Survey Output}} = 26.4657 - 0.0515 \times \underset{(\text{sec})}{\text{Completion}} - 0.0058 \times \underset{(\text{vph})}{\text{Volume}}$$
$$+2.8004 \times \underset{(\text{persons})}{\text{Crews}} + 0.7011 \times \underset{(\%)}{\text{Acceptance}} - 0.7894 \times \underset{(\text{sec})}{\text{Dwell}} - 7.1078 \times \underset{(1 \text{ if limit queue})}{\text{Queue}} \quad \text{(Eq. 9)}$$

Some important findings observed from the equations are:

- Under the new TCPs without adequate shoulders, higher volume also decreases the survey output especially when the queue policy is in use.
- Under the new TCP without adequate shoulders, the survey is most efficient under best-case scenario with approximately 50 percent utilized at optimal volume range (400–550 vph).
- With high acceptance rate, more crew members will increase the survey output or the number of surveys completed. However, the efficiency measured by survey output rate may decrease slightly on a per-crew basis with higher number of crews.

The surveyor utilization rate for the new TCPs on roads without adequate shoulder can be estimated using Equation 10:

$$\underset{(\%)}{\text{Surveyor Utilization}} = 32.4117 + 0.0843 \times \underset{(\text{sec})}{\text{Completion}} - 0.0054 \times \underset{(\text{vph})}{\text{Volume}}$$
$$-4.0019 \times \underset{(\text{persons})}{\text{Crews}} + 0.7935 \times \underset{(\%)}{\text{Acceptance}} - 0.8969 \times \underset{(\text{sec})}{\text{Dwell}} - 7.0463 \times \underset{(1 \text{ if limit queue})}{\text{Queue}} \quad \text{(Eq. 10)}$$

The low utilization rate indicates that the bay is empty or not filled most of the time. This rate would be at the lowest with the combinations of high traffic volume, poor acceptance rate, and active queue policy.

Impact of New TCPS on Roads with Adequate Shoulders

The operation of the new TCPs on roadways with adequate shoulder share some similar characteristics with the old TCPs in that the traffic is slowed down, but not stopped, on the main lane. The stop-and-ask procedure required as part of the new TCPs will take place in the receiving bay instead of on the open main lane and thus minimizes the impacts on the non-surveyed vehicles. A queue would not form on the main lane and the queue length at the receiving bay would be limited to the size of platoon flagged into the station. It is assumed that the flagger will limit the size of the platoon flagged at a time to the number of crew members available. The detailed operational procedures were described in earlier sections.

The delay impact to non-surveyed vehicles would be similar to those observed in the old TCPs. Therefore, the delay equation developed for the old TCPs would be applicable for the new TCPs on roadways with adequate shoulder as well. However, the efficiency impact would be different mainly as a result of the introduction of acceptance rate.

Probability of Exceeding Bay Capacity

Since the acceptance rate (10 percent or 20 percent) will result in a low number of vehicles agreeing to participate in the survey, the researchers also examined the effects of larger platoon size (including the maximum of 10-vehicle queue limit). As the survey bay capacity is unlikely to be exceeded, the researchers also examined the probability that the bay will be filled from a series of platoons. This would be more realistic for the contractor in the field as they would soon realize that, with such a low acceptance rate, the chance that the survey bay will be full from one platoon will be very slim even with the platoon size of 10.

First, the effect of one platoon was examined. Then, the exact probability that the survey bay capacity will be exceeded was calculated. The platoon size was denoted as n, the number of survey crews as k, and the probability of accepting the survey as p. The probability of x vehicles agreeing to the survey follows the binomial distribution, which is shown in Equation 11:

$$\Pr(x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad \text{(Eq. 11)}$$

Therefore, the probability that the bay capacity will be exceeded is equal to $\Pr(x > k)$. That is shown in Equation 12 and is equal to:

$$\Pr(x > k) = 1 - \Pr(x \le k) = 1 - \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} \qquad \text{(Eq. 12)}$$

For the platoon of size n = 10, the number of crew members k = 4, and the acceptance probability p = 0.1, the probability of exceeding the bay capacity is provided in Equation 13:

$$\Pr(x > 4) = 1 - \Pr(x \le 4) = 1 - \sum_{i=0}^{4} \binom{10}{i} 0.1^i (0.9)^{10-i} = 0.001635 \qquad \text{(Eq. 13)}$$

Table 8 summarizes the probabilities that the survey bay capacity will be exceeded for different number of survey crews, platoon size, and acceptance rates.

Table 8. Probability of Exceeding Survey Bay Capacity for One Platoon.

Number of Survey Crews	Platoon n = 4		Platoon n = 7		Platoon n = 10	
	p = 0.1	p = 0.2	p = 0.1	p = 0.2	p = 0.1	p = 0.2
1	5.23%	18.08%	14.97%	42.33%	26.39%	62.42%
2	0.37%	2.72%	2.57%	14.80%	7.02%	32.22%
3	0.01%	0.16%	0.27%	3.33%	1.28%	12.09%
4	0.00%	0.00%	0.02%	0.47%	0.16%	3.28%
5	N/A	N/A	0.00%	0.04%	0.01%	0.64%
6	N/A	N/A	0.00%	0.00%	0.00%	0.09%

If vehicles continue to be flagged in a series of platoons of the same size, the chance that the bay will exceed the capacity will increase with the number of platoons. That is, assuming none of the vehicles that participated in the survey have left the survey station. Hence, the number of vehicles that has to be asked in order to fill up the bay follows the negative binomial distribution (the number of Bernoulli trials needed to get k successes).

Therefore, with the acceptance probability of p, the average number of vehicles that need to be asked to fill up the bay with k crew members is k/p. For example, with 4 crew members and 10 percent acceptance rate, it takes on average 4.0/0.1 or 40 vehicles (or equivalently 10 platoons of 4 vehicles each) to fill up the bay. With a conservative assumption that it takes one minute to ask a platoon of 4 vehicles, it is very likely that by the time 10 platoons have been asked (over 10 minutes) some of drivers will have completed the survey and left the station already. As a result, it is very unlikely that the bay capacity will ever be exceeded given a low acceptance probability and a series of small platoons.

Survey Efficiency

The following range of parameter values were examined in the case of new TCPs on roadways with adequate shoulders. Only efficiency measures were observed and discussed in this section.

- survey completion time: 120, 180, and 240 seconds;
- dwell time: 20 seconds (inside the bay);
- survey acceptance rate: 10 percent and 20 percent;
- survey crews: 4 crews;
- platoon size: 4 vehicles (set to be equal to the size of survey crews); and
- volume level: 125, 250, 500, 750, and 1000 vph.

Five replications were run for each scenario. The simulation results were averaged to obtain the corresponding MOEs. The results are presented in form of look-up tables in this case since the full factorial experimental runs were not conducted.

Table 9 and Table 10 summarize the survey efficiency measures observed from the simulation of the new TCPs on roadways with adequate shoulder. Operational impacts on motorists were not reported here as there was no queue observed on the open main lane.

The queue inside the bay is limited to the size of the platoon. The delay to non-surveyed vehicles can be estimated using Equation 2.

Survey acceptance rates play the most important role in the efficiency of the new TCPs on roads with adequate shoulder. When the acceptance rate is increased from 10 percent to 20 percent, the number of survey outputs and surveyor utilization percentage also increase by at least 50 percent to almost 100 percent.

Table 9. Completed Surveys* – New TCP on Roads with Adequate Shoulders.

Volume	*Survey Completion Time = 120 s*		*Survey Completion Time = 180 s*		*Survey Completion Time = 240 s*	
(veh/hr)	10% Acceptance	20% Acceptance	10% Acceptance	20% Acceptance	10% Acceptance	20% Acceptance
125	9.0	16.2	9.4	16.6	9.0	15.6
250	10.2	18.4	13.0	21.8	10.0	17.2
500	12.0	21.6	16.2	27.4	11.4	18.6
750	12.4	22.2	17.2	29.4	11.2	20.0
1000	12.4	22.6	18.6	31.6	12.2	20.2

*The numbers in the table are the number of completed surveys per hour averaged from five replications.

Table 10. Surveyor Utilization – New TCPs on Roads with Adequate Shoulder.

Volume	*Survey Completion Time = 120 s*		*Survey Completion Time = 180 s*		*Survey Completion Time = 240 s*	
(veh/hr)	10% Acceptance	20% Acceptance	10% Acceptance	20% Acceptance	10% Acceptance	20% Acceptance
125	7.4%	13.5%	11.6%	19.8%	14.1%	23.0%
250	8.7%	15.0%	15.0%	25.6%	15.5%	25.8%
500	10.0%	16.5%	19.2%	30.2%	17.5%	26.7%
750	10.3%	16.7%	19.5%	30.5%	17.8%	29.2%
1000	10.7%	17.7%	21.1%	33.7%	18.0%	28.3%

SUMMARY

The researchers quantified and compared the impacts from the old and new TxDOT TCPs using VISSIM microscopic traffic simulation. The new TCPs introduce the stop-and-ask procedure that did not exist under the old TCPs. The analysis of the new TCPs was separated into two cases: (a) 4-lane undivided roadway without adequate shoulder (less than 10 feet) and (b) 4-lane undivided road with adequate shoulder (10 feet or greater). The major difference between these two is the location where the stop-and-ask procedure takes place. With adequate shoulder, the stop-and-ask procedure takes place at the receiving bay inside the station instead of on the outer main lane. Simulation experiments were set up with varying parameters in order to capture their effects on the impacts, and over 4000 simulation runs were conducted. The impacts measured from the simulation were categorized into two groups: operational impacts and efficiency impacts. The results from the simulation were also used to calibrate regression models to quantify the impact from each design factor. It was found that the delay incurred by the new TCPs without adequate shoulder would be unacceptable without the queue limit policy. However, the queue policy when in use severely limits the ability to meet the minimum survey requirements per day particularly at 10 percent to 20 percent survey participation rate. The operational impacts were found to be much more manageable for the new TCPs on roads with adequate shoulders. Nevertheless, the survey efficiency in terms of the survey outputs and site utilization rate would be significantly reduced under the new TCPs. In addition, the cost per individual survey would likely be more than double when comparing with that of the old TCPs.

Some important findings observed from the analysis of the impacts of the proposed new TCPs are the following:

- New TCPs without adequate shoulder cannot be used in practice without the queue limit policy.
- Queuing conditions observed under the new TCPs without adequate shoulder can be managed with the queue policy but the survey site will still suffer from highly underutilized condition.
- Survey bay will be rarely filled. Unrealistic assumptions such as 30 percent acceptance rate would be needed for the survey site to be functional.
- Survey completion times actually will be higher with interview on passenger side.
- The analysis currently assumes 100 percent non-commercial vehicles. The actual performance will further degrade with the increasing mix of commercial vehicles.
- The number of surveys completed in an hour is generally cut by more than half with the new TCPs.
- Under the new TCPs, the survey crews can be idling at least 80 percent of the time under general traffic and site conditions.
- Cost per individual survey would be more than double.

Table 11 provides a summarized comparison of motorist impacts and survey efficiency observed from this study for both old and new TCPs.

Table 11. Comparison of Impacts from Old versus New TCPs.

Impacts	Old TCPs	New TCPs without Shoulder	New TCPs with Shoulder
Motorist Impacts			
Delay to Non-Surveyed Motorists (seconds/vehicle)	• Delay is primarily a function of traffic volume. • Average delay per vehicle is less than 20 seconds even when the volume is as high as 1000 vph.	• In addition to volume, delay depends on acceptance rate, survey completion time, and stop dwell time. • Under the best-case scenario, the delay is less than 40 sec/veh. • Under the worst-case scenario, the delay is less than 2 minutes/veh until the traffic level reaches 750 vph. Then, the delay rises significantly to 5 or more minutes per vehicle.	• Delay to non-surveyed vehicles is similar to the case of old TCPs.
Queue (ft)	• No queue up to 1200 vph.	• The queue is apparent even at low-volume conditions. • Only under the best-case scenario, the average queue can be maintained below 250 ft. • Under the worst-case scenario, the average queue exceeds 250 ft at 500 vph.	• No queue on the open main lane. • The queue formed at the receiving bay is limited to the size of platoon diverted into the bay by the flagger.
Survey Efficiency			
Survey output (veh/hr)	• 30–70 veh/hr depending on the volume, survey completion time, and the number of crews.	• 20–35 veh/hr under the best-case scenario. • 3-15 veh/hr under the worst-case scenario.	• 9–32 veh/hr depending on the volume, survey completion time, and acceptance rate.
Survey output rate (veh/hr/crew)	• 8–15 veh/hr/crew depending on the same factors.	• 5–9 veh/hr/crew under the best-case scenario. • 1–4 veh/hr/crew under the worst-case scenario	• 2–8 veh/hr/crew.
Surveyor Utilization (% of time occupied)	• 35%–60%. High utilization with higher volume, longer completion time, and fewer crews.	• 35–50% under the best-case scenario. • 3–12% under the worst-case scenario.	• 10%–35%. Higher utilization with higher volume.

CHAPTER 4. ALTERNATIVE METHODS – USE OF OTHER TRAVEL SURVEY DATA

TxDOT currently supports a robust travel survey program that includes household, work place, and commercial vehicle surveys as well as external station surveys. Each type of survey is intended to capture travel information and characteristics of specific elements of travel within an urban area. While the external station surveys are intended to capture data and information on travel in and out and travel through the urban area, some data on external travel are collected in the household, work place, and commercial vehicle surveys. In the absence of an external station survey, the question raised was whether these data could be used to estimate the external local movements in an urban area. The objective of this sub-task is to examine this question and using actual survey data, develop estimates of external local travel, and compare those estimates to the results of the external survey.

The principal data elements of an external survey used in travel demand models are the estimates of external local travel, the external through trip table, and the average trip lengths for the external local travel. These estimates are typically prepared for non-commercial and commercial vehicles. Data from the external survey for non-commercial vehicles are typically stratified into two categories, resident travel and visitor travel. Travel by residents are those trips made by persons that live within the urban area while travel by visitors are those trips made by persons that do not live within the area.

HOUSEHOLD SURVEYS

The household survey captures information on a randomly selected sample of households within the urban area. That information includes data on the trips made by each person in the household during a 24-hour period. Included are the origin and destination of each trip. If individuals in a sampled household made trips outside the urban area, these would be included in the survey. These trips when weighted to represent travel for all households in the study area would be expected to be comparable to the resident trips observed in the external station survey. Data from the household surveys in the Rio Grande Valley, Austin, and San Antonio were analyzed and weighted to produce estimates of the total amount of external travel (vehicle trips) being made on a typical week day. These estimates were then compared to the estimates for resident travel obtained from the external surveys in those study areas. Table 12 presents the results of these comparisons for each study area.

Table 12. Resident External Travel Estimates.

Study Area	Number of Resident External Local Vehicle Trips		
	Household Survey	External Station Survey	Percent Difference
Rio Grande Valley	29,921	61,426	- 51.3 %
Austin	77,639	93,020	- 16.5 %
San Antonio	66,258	85,572	- 22.6 %

WORK PLACE SURVEYS

The work place survey captures information on the destination end of travel in an urban area. Work places are randomly selected based on a stratified sampling plan where the establishments are stratified by employment type and area type. Employee and visitors surveys are conducted at each establishment to capture information on the origin and destination of travel. Person or vehicle counts are conducted at the establishment to weight the survey data. These data are used to develop attraction rates (i.e., number of trips attracted to establishments per employee) by trip purpose. Since data are captured on the address of each person surveyed, the number of trips by non-residents may be identified and external attractions estimated based on the origin and destinations of those trips. External trip rates were developed for work places stratified by employment type and area type. These rates were then applied to the total employment in each category to estimate the total number of external trips to and from work places in the study area. These estimates are considered comparable to the visitor trips identified at the external station surveys in the areas. Table 13 presents these estimates and the percent difference with the data from the external station surveys.

Table 13. Visitor External Travel Estimates.

Study Area	Number of Visitor External Local Vehicle Trips		
	Work Place Survey	External Station Survey	Percent Difference
Rio Grande Valley	34,091	56,938	- 40.1 %
Austin	310,403	105,595	+ 194.0 %
San Antonio	154,870	126,088	+ 22.8 %

COMMERCIAL VEHICLE SURVEYS

Commercial vehicle surveys are conducted to obtain information on travel by commercial vehicles. Commercial vehicle trips are estimated and modeled directly in the travel demand models for urban areas. These surveys provide data that are used to develop estimates of the total number of commercial vehicle trips and their average trip lengths for use in the travel demand models. Work establishments are randomly selected in an urban area and if they have vehicles used for commercial purposes, recruited to participate in the commercial vehicle survey. Those agreeing to participate are asked to complete a 24-hour diary detailing each trip made by the commercial vehicle, the origin and destination, purpose, type of cargo, etc. As part of this survey, some of the trips recorded are external local in nature and may be used to estimate the number of commercial vehicle external local trips in the study area. The commercial vehicle survey data for the Rio Grande Valley, Austin, and San Antonio were analyzed to identify the external local trips. The data were expanded using the same expansion factors as computed for the internal commercial vehicle trips. The resulting estimates are shown in Table 14 with the number of commercial vehicle trips found in the external station surveys.

Table 14. Commercial Vehicle External Local Travel Estimates.

Study Area	Number of Commercial Vehicle External Local Vehicle Trips		
	Commercial Vehicle Survey	External Station Survey	Percent Difference
Rio Grande Valley	4,390	17,834	- 75.4 %
Austin	15,624	34,156	- 54.3 %
San Antonio	15,753	43,573	- 64.0 %

AVERAGE TRIP LENGTH ESTIMATES

One data element from the external station surveys used in the travel demand models is the average trip length for non-commercial and commercial external local trips. This estimate is used in the distribution of the trips between the external stations and the internal traffic analysis zones. This distribution step produces the estimate of internal vehicle miles of travel that may be attributed to external travel in the urban area. The identification of the external trips from the household, work place, and commercial vehicle surveys is straightforward. The household surveys coded the facility the trip was on when they exited the study area in most cases but did not code the zone number assigned to the external station associated with those facilities. This had to be done manually in order to use the network skims data to find the travel time and distance between the internal zone and the external station. This step was done for each of the three household surveys. The coding of the trips in the work place and commercial vehicle surveys included the external station zone number and simplified the computation of average trip length for the external local trips. Table 15 presents the comparison of the average trip lengths of non-commercial vehicles for external local trips based on the network skims for each study area, while

Table 16 provides the same information for commercial vehicles. The number of trips in the surveys used in computing the average trip lengths are also shown. This information provides the analyst with a sense of the reliability of the estimates based on data from the household, work place, and commercial vehicle surveys. The trips from the household and work place surveys were combined in computing the average trip lengths to maintain comparability with the data from the external station surveys.

Table 15. Non-Commercial Vehicle External Local Trip Length Comparisons.

Study Area	Household/Work Place Survey			External Station Survey			Percent Differences	
	Number Trips	Avg Trip Length		Number Trips	Avg Trip Length			
		Miles	Minutes		Miles	Minutes	Miles	Minutes
Valley	474	20.4	23.8	5,751	15.8	20.6	+29.1 %	+15.5 %
Austin	382	33.7	46.5	6,842	51.1	72.3	-34.1 %	-35.7 %
San Antonio	266	23.9	31.9	7,791	23.0	31.3	+3.9 %	+1.9 %

Table 16. Commercial Vehicle External Local Trip Length Comparisons.

Study Area	Commercial Vehicle Survey			External Station Survey			Percent Differences	
	Number Trips	Avg Trip Length		Number Trips	Avg Trip Length			
		Miles	Minutes		Miles	Minutes	Miles	Minutes
Valley	158	33.4	37.4	1,012	20.6	25.6	+ 62.1 %	+ 46.1 %
Austin	159	48.6	68.1	888	45.9	65.2	+ 5.9 %	+ 4.4 %
San Antonio	148	30.7	40.7	1,090	29.3	39.1	+ 4.8 %	+ 4.1 %

SUMMARY

As shown, it is possible to estimate external local travel (number of vehicle trips and average trip length) for both non-commercial and commercial vehicles using data from non-external station surveys (i.e., household, work place, and commercial vehicle surveys). This task developed those estimates for three urban areas and compared the results with the external station surveys done in those areas. The resulting estimates from non-external station surveys ranged from poor to good. The total non-commercial vehicle trips estimated from the non-external surveys was less than the estimate from the external survey by 46 percent. For Austin, the estimate was 95 percent greater and for San Antonio, the estimate was 4 percent greater. The estimates of commercial vehicle external local trips based on the commercial vehicle surveys for the three areas were all significantly less than those from the external surveys, ranging from 54 percent to 75percent.

The estimates of average trip length showed similar results. For non-commercial vehicles, the estimates ranged from a negative 34 percent to a plus 29 percent. The estimates for San Antonio were very close. For commercial vehicle trips, the average trip lengths for the Austin and San Antonio area were very similar to those from the external surveys while the estimates for the Valley were plus 46 and 62 percent, minutes and miles, respectively.

It should be recognized that the household, work place, and commercial vehicle surveys were not designed to produce reasonable estimates of external travel. External travel, while captured during the course of these surveys, is a fairly rare event and with all of the surveys being small samples, the estimates of external travel should not be expected to be very accurate. This task has demonstrated that these estimates may be obtained from these surveys. For a reasonable expectation of accurate estimates, it will be necessary to increase the sample sizes in all of these surveys by two to three times what is currently obtained. The survey designs will also need to be modified to more accurately capture the amount of external travel and the external stations involved in that travel. It should also be recognized that these surveys will not reflect any information on the amount of external through travel for an urban area. These estimates will have to be obtained through limited external surveys or by synthetic models.

CHAPTER 5. ALTERNATIVE METHODS – STATEWIDE ANALYSIS MODEL

This chapter examines the viability of applying the Statewide Analysis Model as an alternative method for ascertaining urban external travel movements. Specifically, the application of SAM would be used to derive external local and external through apportions for urban area external stations and as the basis for developing urban area external through trip matrices.

The primary purpose for such an endeavor was to analyze and evaluate the potential for developing synthetic external travel patterns derived from SAM results in lieu of conducting external surveys. The principle objective is to determine how well and to what extent the SAM is capable of replicating or being consistent with known external travel patterns. Because the proposed methodology's effectiveness is dependent on emulating observed travel patterns, this chapter focuses on the comparative assessment between SAM results and observed travel survey data.

The remainder of this chapter summarizes the efforts undertaken to determine how well the SAM model currently replicates external local and through trip patterns within Texas urban areas in comparison to observed travel survey data based on two test cases, Austin and Corpus Christi. Following is a discussion and summary of three topics:

- statewide analysis model application,
- development of the Austin and Corpus Christi external matrices, and
- statewide analysis model results.

STATEWIDE ANALYSIS MODEL APPLICATION

The application of the SAM to derive external local and external through apportions for urban area external stations is a fairly straightforward process. The SAM includes a user interface with functions to analyze travel patterns. One of the functions is the External Urban Links Analysis (EULA) application. This function allows the user to create an external trip table for any of the 25 urban areas in Texas. The basis for extracting an urban area specific external trip table from the entire statewide trip matrix is the delineation of pre-defined urban area boundaries (UABs) for all 25 Texas urban areas as shown in Figure 15.

The network links that traverse a UAB are used by the EULA application to define the location of the external stations for a given urban area. The SAM assigned volumes on the individual links that cross the UAB will subsequently represent the total external volumes at each external station. Thus, the external trip table that is extracted by the EULA application includes both external through trips and external local trips with the total trips at each external station corresponding to the link assigned volume.

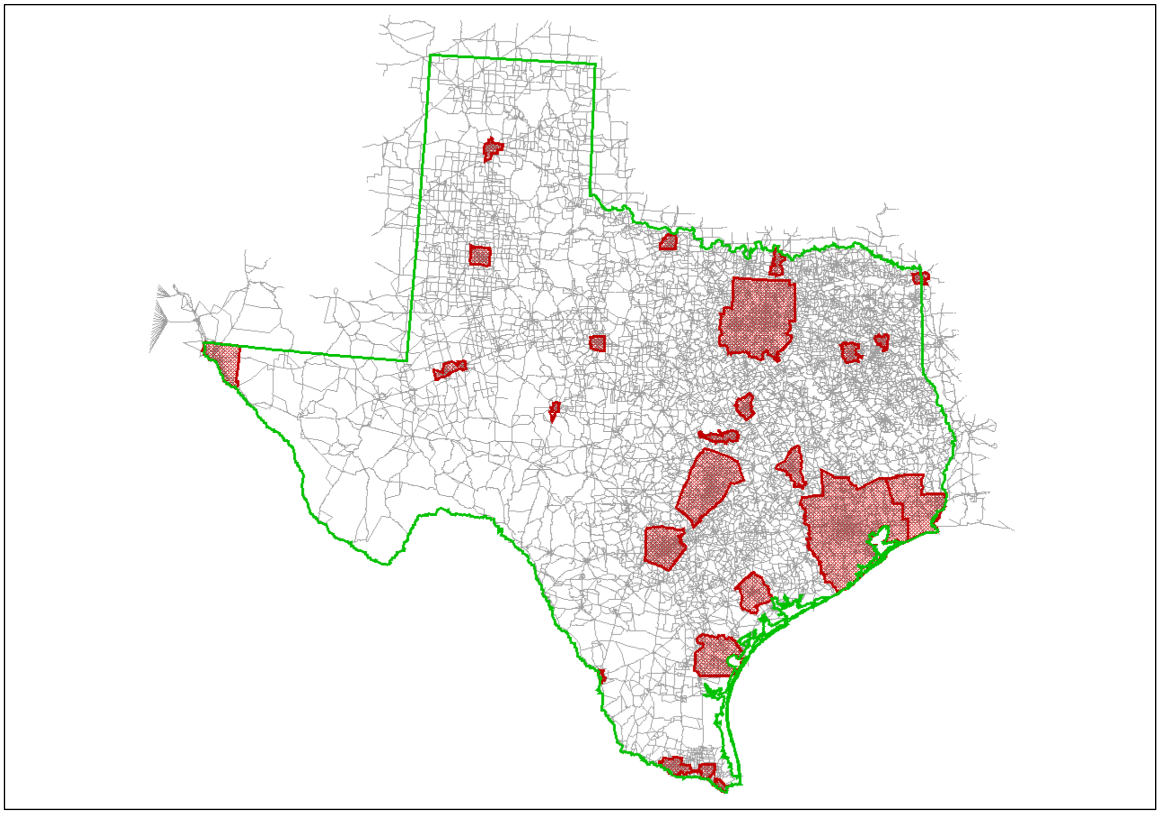

Figure 15. SAM Urban Area Boundaries.

The external trip table produced by the External Urban Links Analysis application provides a good starting point for estimating through trips. It can be used to estimate both percentage of external through trips for each external station, and the distribution of external through trips between external stations. The current version of SAM has base and forecast years of 1998 and 2025, respectively; consequently, any trip matrix extraction is limited to those two years. Estimating trips for intermediate years would require either new demographic data, which may be difficult to obtain, or interpolation, which may not produce reasonable results. Since the intent is to compare SAM data results with actual observed data, the process is then limited to surveys that occurred around 1998. Due to this fact, two study areas that had travel surveys performed around 1998 were chosen as test cases for the analysis. The areas chosen for analysis purposes were Austin and Corpus Christi; Austin had an external survey conducted in 1998 and Corpus Christi's external survey occurred in 1996.

DEVELOPMENT OF THE AUSTIN AND CORPUS CHRISTI EXTERNAL MATRICES

For both areas that were chosen as test cases (Austin and Corpus Christi), the number of local and through trips by external station were developed using the EULA application.

Comparison of UABs to External Survey Locations

Since the UAB is used to determine the location of each external station, the SAM UABs for Austin and Corpus Christi were compared to the designated locations where external surveys were conducted for both urban areas. The process entailed examining each link that the UAB traversed and comparing that location to the actual location where the survey was conducted to ensure that both locations were comparable. If the UAB did not traverse a roadway at the correct location the UAB was subsequently repositioned to the actual location where the external survey was conducted. For both Austin and Corpus Christi the UAB was relocated for a number of external stations to ensure geographic consistency between the survey data and the extracted SAM external matrices.

Development of External Local and External Through Trip Matrices

Once the UABs for Austin and Corpus Christi were revised to coincide with the known external survey locations the EULA application was applied twice to extract a trip matrix for both urban areas. Both resulting trip matrices were comprised of three types of trip movements:

- Internal-Internal Trips,
- Internal-External/External-Internal Trips (i.e., external local trips), and
- External-External Trips (i.e., external through trips).

For analysis purposes the Internal-Internal trips were of no interest and were subsequently subtracted out from the Austin and Corpus Christi trip matrices. This yielded two trip matrices comprised entirely of external trip movements. In both instances the external local trips were separated from the external through trips to provide two distinct trip matrices (external local and external through) for both urban areas. These final matrices were then used to compare external local and external through percentage splits against original external survey data to assess how well and to what extent the SAM was capable of replicating observed external travel patterns.

STATEWIDE ANALYSIS MODEL RESULTS

The EULA-extracted SAM external trip matrix results were compared to existing survey data. Due to the site by site variation in the number of trips, it was determined that the initial test of effectiveness would be to review the results in aggregate format. Table 17 provides the results of the initial test. The table shows the number of local, through, and total trips for Austin and Corpus Christi that were reported during the survey as well as developed using the EULA application of the SAM. The table also provides the percent of trips that are local and through for both the external survey and the SAM.

While the EULA application provided a reasonable estimation of external local and external through splits in terms of percentages, the total number of trips was underestimated by a significant margin. In Corpus Christi, the total number of trips estimated by SAM was nearly 50 percent less than what was developed during the external survey. In Austin, SAM underestimated the total number of trips identified in the external survey by more than 40 percent.

Table 17. Summary of Survey and SAM Results.

Study Area	Survey Results			SAM Results		
	Through Trips	Local Trips	Total Trips	Through Trips	Local Trips	Total Trips
Corpus Christi (1996)	6,218 (8%)	67,558 (92%)	73,776	3,009 (8%)	36,227 (92%)	39,236
Austin (1998)	2,289 (2%)	91,623 (98%)	93,912	2,194 (4%)	51,371 (96%)	53,565

A more detailed comparison is provided in Table 18 through Table 21:

- Table 18 compares unexpanded Austin survey data to SAM model results.
- Table 19 compares expanded Austin survey data to SAM model results.
- Table 20 compares unexpanded Corpus Christi survey data to SAM model results.
- Table 21 compares expanded Corpus Christi survey data to SAM model results.

Unexpanded survey data are the actual number of useable surveys that were collected at each external station whereas expanded survey data has been factored to the total traffic count at each individual external station.

Table 18. Summary of Unexpanded Austin Survey Data and SAM Results.

External Station	Total Trips	Survey Local	Survey Through	Survey Percent Local	Survey Percent Through	Total Trips	SAM Local	SAM Through	SAM Percent Local	SAM Percent Through
IH 35 North	No Survey Conducted									
FM 2115	120	101	19	84.2%	15.8%	160	0	160	0.0%	100.0%
SH 95 North	521	432	89	82.9%	17.1%	2,497	2,401	96	96.2%	3.8%
FM 487	133	120	13	90.2%	9.8%	970	807	163	83.2%	16.8%
FM 1331	122	110	12	90.2%	9.8%	41	40	1	97.6%	2.4%
US 79 East	382	347	35	90.8%	9.2%	3,147	2,932	215	93.2%	6.8%
FM 619	75	72	3	96.0%	4.0%	220	84	136	38.2%	61.8%
SH 95 South	375	313	62	83.5%	16.5%	3,199	2,983	216	93.2%	6.8%
US 290 East	No Survey Conducted									
FM 969	324	315	9	97.2%	2.8%	1,199	1,199	0	100.0%	0.0%
SH 71 East	No Survey Conducted									
FM 812	404	399	5	98.8%	1.2%	2,761	2,743	18	99.3%	0.7%
US 183 South	423	401	22	94.8%	5.2%	8,151	7,972	179	97.8%	2.2%
SH 21	423	234	189	55.3%	44.7%	1,163	878	285	75.5%	24.5%
SH 80	668	588	80	88.0%	12.0%	5,857	5,613	244	95.8%	4.2%
FM 621	404	387	17	95.8%	4.2%	998	997	1	99.9%	0.1%
SH 123	613	546	67	89.1%	10.9%	7,229	7,214	15	99.8%	0.2%
IH 35 South	No Survey Conducted									
FM 2439	455	418	37	91.9%	8.1%	1,067	1,067	0	100.0%	0.0%
FM 32	535	475	60	88.8%	11.2%	1,597	1,565	32	98.0%	2.0%
FM 165	279	255	24	91.4%	8.6%	15	15	0	100.0%	0.0%
US 290 West	431	379	52	87.9%	12.1%	1,081	1,036	45	95.8%	4.2%
SH 71 West	504	472	32	93.7%	6.3%	2,894	2,887	7	99.8%	0.2%
FM 1869	74	69	5	93.2%	6.8%	613	597	16	97.4%	2.6%
SH 29	535	454	81	84.9%	15.1%	2,464	2,411	53	97.8%	2.2%
US 183 North	466	419	47	89.9%	10.1%	2,623	2,418	205	92.2%	7.8%
SH 195	553	516	37	93.3%	6.7%	3,619	3,512	107	97.0%	3.0%

Since Table 18 only provides unexpanded survey data a comparison of actual trips is not appropriate as the survey data are merely a sample of the total volume at each external station and thus would never approximate SAM assigned volumes. On the other hand, percentages of external local and external through trips can be compared to evaluate how well SAM output correlates to observed data. In comparing percent of external through trips in Table 18, a wide range of differences exist between the survey data and the SAM results. The differences extend from approximately 84 percent over (FM 2115) to roughly 20 percent under (SH 21). In between that range however, the SAM results for several facilities match the survey data quite closely; these include: FM 812, FM 969, US 79 East, US 183 North, and US 183 South. Unfortunately, there is no clear pattern or rationale as to why SAM replicates observed survey data well for some facilities and not others. Another problematic observation is that for four facilities the SAM allocated all the trips to either external local or external through (FM 2115, FM 969, FM 2439 and FM 165).

Table 19. Summary of Expanded Austin Survey Data and SAM Results.

External Station	Total Trips	Survey Local	Survey Through	Survey Percent Local	Survey Percent Through	Total Trips	SAM Local	SAM Through	SAM Percent Local	SAM Percent Through
IH 35 North	No Survey Conducted									
FM 2115	439	435	4	99.1%	0.9%	160	0	160	0.0%	100.0%
SH 95 North	4,359	4,077	282	93.5%	6.5%	2,497	2,401	96	96.2%	3.8%
FM 487	0	0	0	0.0%	0.0%	970	807	163	83.2%	16.8%
FM 1331	565	565	0	100.0%	0.0%	41	40	1	97.6%	2.4%
US 79 East	5,802	5,620	182	96.9%	3.1%	3,147	2,932	215	93.2%	6.8%
FM 619	306	306	0	100.0%	0.0%	220	84	136	38.2%	61.8%
SH 95 South	3,629	3,292	337	90.7%	9.3%	3,199	2,983	216	93.2%	6.8%
US 290 East	No Survey Conducted									
FM 969	2,211	2,207	4	99.8%	0.2%	1,199	1,199	0	100.0%	0.0%
SH 71 East	No Survey Conducted									
FM 812	6,047	6,040	7	99.9%	0.1%	2,761	2,743	18	99.3%	0.7%
US 183 South	9,637	9,367	270	97.2%	2.8%	8,151	7,972	179	97.8%	2.2%
SH 21	4,169	3,957	212	94.9%	5.1%	1,163	878	285	75.5%	24.5%
SH 80	11,814	11,667	147	98.8%	1.2%	5,857	5,613	244	95.8%	4.2%
FM 621	4,073	4,067	6	99.9%	0.1%	998	997	1	99.9%	0.1%
SH 123	8,807	8,705	102	98.8%	1.2%	7,229	7,214	15	99.8%	0.2%
IH 35 South	No Survey Conducted									
FM 2439	2,204	2,018	186	91.6%	8.4%	1,067	1,067	0	100.0%	0.0%
FM 32	2,116	2,000	116	94.5%	5.5%	1,597	1,565	32	98.0%	2.0%
FM 165	1,079	1,074	5	99.5%	0.5%	15	15	0	100.0%	0.0%
US 290 West	3,622	3,496	126	96.5%	3.5%	1,081	1,036	45	95.8%	4.2%
SH 71 West	5,215	5,194	21	99.6%	0.4%	2,894	2,887	7	99.8%	0.2%
FM 1869	312	312	0	100.0%	0.0%	613	597	16	97.4%	2.6%
SH 29	7,326	7,181	145	98.0%	2.0%	2,464	2,411	53	97.8%	2.2%
US 183 North	4,092	3,965	127	96.9%	3.1%	2,623	2,418	205	92.2%	7.8%
SH 195	6,088	6,078	10	99.8%	0.2%	3,619	3,512	107	97.0%	3.0%

In Table 19, which compares expanded Austin survey data to SAM results, the differences in percent external though trips is comparable to Table 18 though the range of differences is slightly less, 99 percent over (FM 2115) to 8 percent under (FM 2439). Table 19 also indicates that the SAM estimate of total external trips for the Austin area is nearly always less than the actual count. In only two instances (FM 487 and FM 1869) did SAM estimate a higher volume that the observed count, for all other external stations the SAM volume was lower than the counted volume.

Table 20. Summary of Unexpanded Corpus Christi Survey Data and SAM Results.

External Station	Total Trips	Survey Local	Survey Through	Survey Percent Local	Survey Percent Through	Total Trips	SAM Local	SAM Through	SAM Percent Local	SAM Percent Through
US 77 N.	354	243	111	68.6%	31.4%	1,962	1,601	361	81.6%	18.4%
FM 136	436	404	32	92.7%	7.3%	313	243	70	77.6%	22.4%
SH 188	338	95	243	28.1%	71.9%	333	237	96	71.2%	28.8%
FM 1069	406	397	9	97.8%	2.2%	312	264	48	84.6%	15.4%
SH 35	441	421	20	95.5%	4.5%	5,885	5,224	661	88.8%	11.2%
Park Rd 22	245	211	34	86.1%	13.9%	9,218	9,167	51	99.4%	0.6%
US 77 S.	510	408	102	80.0%	20.0%	9,024	8,549	475	94.7%	5.3%
Bus. US 77	515	492	23	95.5%	4.5%	0	0	0		
FM 665	457	274	183	60.0%	40.0%	691	607	84	87.8%	12.2%
SH 44	425	408	17	96.0%	4.0%	3,162	3,148	14	99.6%	0.4%
Co. Rd. 352	Facility not included in SAM network									
FM 624	482	464	18	96.3%	3.7%	1,069	1,031	38	96.4%	3.6%
FM 70	145	118	27	81.4%	18.6%	0	0	0		
SH 359 N.	374	278	96	74.3%	25.7%	1,180	772	408	65.4%	34.6%
FM 3024	231	206	25	89.2%	10.8%	0	0	0		
IH 37	504	436	68	86.5%	13.5%	2,726	2,502	224	91.8%	8.2%
SH 359 S.	447	367	80	82.1%	17.9%	2,230	1,829	401	82.0%	18.0%
US 181	415	376	39	90.6%	9.4%	1,131	1,053	78	93.1%	6.9%

Table 21. Summary of Expanded Corpus Christi Survey Data and SAM Results.

External Station	Total Trips	Survey Local	Survey Through	Survey Percent Local	Survey Percent Through	Total Trips	SAM Local	SAM Through	SAM Percent Local	SAM Percent Through
US 77 N.	8,094	6,691	1,403	82.7%	17.3%	1,962	1,601	361	81.6%	18.4%
FM 136	3,030	2,724	306	89.9%	10.1%	313	243	70	77.6%	22.4%
SH 188	784	593	191	75.6%	24.4%	333	237	96	71.2%	28.8%
FM 1069	2,052	2,052	0	100.0%	0.0%	312	264	48	84.6%	15.4%
SH 35	8,576	8,424	152	98.2%	1.8%	5,885	5,224	661	88.8%	11.2%
Park Rd 22	717	645	72	90.0%	10.0%	9,218	9,167	51	99.4%	0.6%
US 77 S.	12,761	10,361	2,400	81.2%	18.8%	9,024	8,549	475	94.7%	5.3%
Bus. US 77	6,991	6,958	33	99.5%	0.5%	0	0	0		
FM 665	2,229	2,196	33	98.5%	1.5%	691	607	84	87.8%	12.2%
SH 44	8,058	7,867	191	97.6%	2.4%	3,162	3,148	14	99.6%	0.4%
Co. Rd. 352	Facility not included in SAM network									
FM 624	2,982	2,957	25	99.2%	0.8%	1,069	1,031	38	96.4%	3.6%
FM 70	422	338	84	80.1%	19.9%	0	0	0		
SH 359 N.	3,367	3,133	234	93.1%	6.9%	1,180	772	408	65.4%	34.6%
FM 3024	663	657	6	99.1%	0.9%	0	0	0		
IH 37	7,509	6,909	600	92.0%	8.0%	2,726	2,502	224	91.8%	8.2%
SH 359 S.	2,653	2,406	247	90.7%	9.3%	2,230	1,829	401	82.0%	18.0%
US 181	2,888	2,647	241	91.7%	8.3%	1,131	1,053	78	93.1%	6.9%

Table 20 and Table 21 summarize the differences between Corpus Christi survey data and SAM results. Comparable to Table 18, Table 20 shows a range of differences between the survey data and the SAM results in comparing percent external through trips. The differences extend from approximately 15 percent over (FM 136) to more than 43 percent under (SH 188). The comparison of total trips in Table 21 is also comparable to Table 19 results in that SAM underestimated total external trips for all the Corpus Christi external stations except one (Park Road 22) even though the SAM is a 1998 model and the Corpus Christi survey data were collected in 1996.

Table 21 also indicates some unique issues in attempting to use SAM results in lieu of collecting survey data. For three facilities (Business US 77, FM 70 and FM 3024) SAM did not assign any traffic to those roadways at the point where the UAB traversed the facility and thus an estimate of percent external local and external through trips is not available from the SAM. In addition the assigned volume on Park Road 22 is inordinately different (9,218 versus a count of 717) from the traffic count though that appears to be a unique anomaly.

SUMMARY

In addition to underestimating the number of trips by external station, the two test cases also indicate that a limitation of the SAM is the inability to reasonably replicate the percentage of external local and external through trips for a majority of the external stations. A further limitation is that it does not have the capability to estimate non-resident trip information needed by modelers.

The lack of data for numerous years is also a challenge in that the SAM model would not encompass all the years for which TxDOT-TPP develops base year models. Another consideration is whether or not funding for SAM updates will be continued. While this may not be an issue at this time, if funding is discontinued and no new data are collected to update the model, then the utility of this planning tool would be greatly diminished.

Lastly, due to the different zone structures that exist between urban area models and the SAM, a method for estimating the average trip length for external local trips would still need to be developed for use in urban area models. The summary of the two test cases noted in this chapter indicate that in the absence of survey data the SAM results do not approximate external local and external through trips well enough to be considered as an alternative approach to conducting external surveys. Consequently, at this time, it would appear that the EULA application is not a feasible method for developing reasonable estimates of external related travel.

CHAPTER 6. ALTERNATIVE METHODS – LOGIT MODELS

OVERVIEW OF PREVIOUS RESEARCH

An assessment of general patterns in previous research approaches to estimating through trip patterns helped create a starting point for this research. First, most of the models developed in previous research were some type of statistical regression model. Five of the models used linear regression. Only the Horowitz model and the three spatial economic-models tested by Anderson are non-statistical models. Second, of the five regression models, four were actually a system of two models, where one model (the stage one model) predicted the proportion of through trips at an external station, and the second model (the stage two model) distributed the through trips between external stations. The other researchers used a combined model, where one model predicted both the percent through trips, and the distribution of through trips between external stations. A final pattern is that the earlier models did not include variables to account for the regional context of the study area, and the later models sought to improve predictive power by including such variables. These patterns served as a guide when forming the research approach.

One other research project not discussed yet also influenced the research approach. This was the work of Martchouk and Fricker (*33*), who recently proposed modeling through trips using logistic regression rather than linear regression, as all previous regression-based models had done. They developed a logistic model that uses a set a variables similar to those of the Modlin and Pigman models, but that is a one-stage model rather than a two-stage model.

Table 22 through Table 24 summarize the results of previous research. Table 22 lists the variables that were considered for at least one stage one model. A check mark indicates that the variable was included in the final model. Table 23 and Table 24 present the same information for combined models and stage two models, respectively. These results served as a starting point for choosing the candidate predictor variables for this research.

Table 22. Variables for Previous Stage One Models.

	Modlin 1974	Pigman 1979	Modlin 1982	Han 2008
Characteristics of the survey station				
Roadway characteristics				
Functional classification	✓			✓
Number of lanes				
Traffic characteristics				
Average daily traffic	✓	✓	✓	✓
Percent heavy trucks	✓	✓	✓	✓
Percent light trucks (pickups and panels)	✓			
Regional context				
Zipf's probability factor				
Huff's probability factor				
Characteristics of the study area				
Demographic characteristics				
Population	✓	✓	✓	✓
Employment				✓
Income				
Geographic characteristic				
Area				✓
Miscellaneous				
Marginal highway route				✓

Table 23. Variables Considered for Previous Combined Models.

	Anderson 2006	Martchouk 2009
Characteristics of the survey station		
Traffic characteristics		
Average daily traffic		
Percent trucks		
Roadway characteristics		
Number of lanes		
Regional context		
Nearby major city	✓	
Characteristics of the entry station		
Traffic characteristic		
Average daily traffic	✓	✓
Percent trucks		
ADT as a portion of total ADT for the study area		✓
Roadway characteristics		
Number of lanes		
Functional classification		
Regional Context		
Nearby major city		
Characteristics of study area		
Demographic characteristics		
Population		
Employment		
Miscellaneous		
Route continuity	✓	✓
Internal-external factor	✓	✓

Table 24. Variables for Previous Stage Two Models.

	Modlin 1974	Pigman 1979	Modlin 1982	Han 2008
Characteristics of the survey external station				
Roadway characteristics				
Functional classification	✓	✓	✓	
Number of lanes				
Traffic characteristics				
Average daily traffic				
ADT as a portion of total ADT for study area				
Regional context				
Zipf's probability factor				
Huff's probability factor				
Miscellaneous				
Percent through trip ends from stage one	✓			
Marginal highway route				
Characteristics of the entry external station				
Roadway characteristics				
functional Classification	✓			
Number of lanes				✓
Traffic characteristics				
Average daily traffic	✓	✓	✓	
Percent heavy trucks	✓	✓	✓	
Percent light trucks (pickups and panels)				
ADT as a portion of total ADT for study area		✓	✓	✓
Regional Context				
Zipf's probability factor				✓
Huff's probability factor				
Miscellaneous				
Percent through trip ends from stage one	✓	✓	✓	✓
marginal highway route				
Characteristics of the study area				
Demographic characteristics				
Population	✓			✓
Employment				
Income				
Geographic characteristics				
area				
Miscellaneous				
Route continuity			✓	✓
Angle between survey and choice station				
Horowitz's weight				✓

MODEL OVERVIEW

A logit model is a special type of regression model that can be used to estimate percentages. Logit models, in general, predict proportions while ensuring that the sum of the proportions is equal to one. The model developed as part of this research consists of two logit models. The first model (Model I) estimates the percentage of external-through and external-local traffic at each external station in a study area. A second model (Model II) then estimates the percentage of through traffic that is going to or coming from each of the other external stations within the study area. Figure 16 illustrates how the two models work together. Once the percentages are calculated for each of the external stations within the study area, a through trip table can be developed using traffic counts obtained at the various stations.

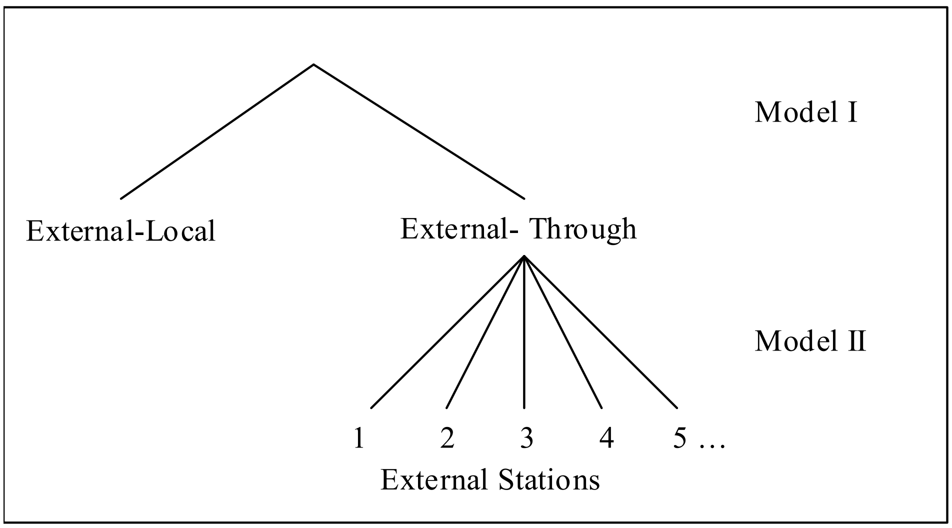

Figure 16. Interaction of Two Logit Models.

Table 25 and Table 26 provide the equations and variables for Model I and model II. Model I has two equations: one for commercial vehicles and one for non-commercial vehicles. Each of these two equations is applied once to each external station. Model II has one equation for both vehicle types. It is applied once in each direction for every possible pair of external stations in the urban area.

The variables *PINTTH*, *INTTL1*, and $PINT1_{ij}$ are functions of the interaction score, which results from a simple gravity model developed specifically for this research. The simple gravity model generates the interaction score as an estimate of the relative amount of travel between urban areas in and around the study area. The interaction score is based on the populations of the urban areas and the distances between them. The variable $ROUTE_{ij}$ signifies whether the least time route between external stations i and j is valid. A route is considered valid of it passes through the study area and if it does not pass through any other external stations.

Table 25. Logit Model I.

$$p_{thru,com,j} = \frac{\exp(V_{thru,com,j})}{1+\exp(V_{thru,com,j})}$$

$$p_{thru,non,j} = \frac{\exp(V_{thru,non,j})}{1+\exp(V_{thru,non,j})}$$

$$V_{thru,com,j} = -1.7816 + 2.3762 \times PINTTH_j - 0.2349 \times INTTL1_j$$

$$V_{thru,non,j} = -2.9375 + 2.3762 \times PINTTH_j - 0.2349 \times INTTL1_j$$

$p_{thru,com}$ = the proportion through trips for commercial vehicles at external station j

$p_{thru,non}$ = the proportion through trips for non-commercial vehicles at external station j

$PINTTH_j$ = the proportion through trips at external station j as estimated by the interaction score

$INTTL1_j$ = 1 if the total interaction score at external station j is greater than zero, and is equal to 0 otherwise.

Table 26. Logit Model II.

$$p_{ij} = \frac{\exp(V_{ij})}{\sum_{k \neq j} \exp(V_{kj})}$$

$$V_{ij} = 2.2 \times PINT1_{ij} - 1.1 \times \sqrt{TURNS_{ij}} + 0.57 \times \ln(PADT_{ij}) + 0.81 \times ROUTE_{ij}$$

i	=	an index variable referring to external stations where through trips enter the study area
j	=	an index variable referring to external stations where through trips exit the study area
p_{ij}	=	the proportion of through trips exiting the study area at external station j that entered the study area at external station i
k	=	the index variable for the summation
$PINT1_{ij}$	=	1 if the proportion through trips entering at i as estimated by the interaction score is greater than one, and is equal to 0 otherwise
$TURNS_{ij}$	=	the number of turns on the least time route between external stations i and j
$PADT_{ij}$	=	the ADT at external station i as a proportion of the total ADT across all external stations, excluding external station j
$ROUTE_{ij}$	=	1 if the least time route between external station i and j is valid

The results from Models I and II can be used to develop commercial and non-commercial through trip tables using the equations in Table 27.

Table 27. Equations for Developing Through Trip Tables.

$$t_{ij,com} = t_{ji,com} = (1/4)\left(p_{thru,com,j} \times p_{ij} \times ADTLV_j + p_{thru,com,i} \times p_{ji} \times ADTLV_i\right)$$

$$t_{ij,non} = t_{ji,non} = (1/4)\left(p_{thru,non,j} \times p_{ij} \times ADTSV_j + p_{thru,non,i} \times p_{ji} \times ADTSV_i\right)$$

where

$t_{ij,com}$ =	number of commercial vehicle through trips entering the urban area at external station i and exiting the urban area at external station j
$t_{ji,com}$ =	number of commercial vehicle through trips entering the urban area at external station j and exiting the urban area at external station i
$t_{ij,non}$ =	number of non-commercial vehicle through trips entering the urban area at external station i and exiting the urban area at external station j
$t_{ji,non}$ =	number of non-commercial vehicle through trips entering the urban area at external station j and exiting the urban area at external station i
$ADTLV_j$ =	the average daily traffic (ADT) at external station j for large vehicles
$ADTLV_i$ =	the average daily traffic (ADT) at external station i for large vehicles
$ADTSV_j$ =	the average daily traffic (ADT) at external station j for small vehicles
$ADTSV_i$ =	the average daily traffic (ADT) at external station i for small vehicles

MODEL DATA SOURCES

The research approach is largely controlled by the data that are available. Over the last 10 years, TxDOT has conducted external surveys in numerous cities around the state. The data from these surveys provide a means to develop robust models as well as an opportunity to assess the ability of the logit model to estimate through trips. For the purpose of this research, a total of 13 study areas were included in the analyses. Table 28 provides the study area and year of the external survey, as well as the number of external stations in the study area and the number of external stations that were surveyed. Additionally, the table provides the total number of commercial and non-commercial vehicle surveys that were completed as well as the total number of through trips that were reported by commercial and non-commercial vehicles during the conduct of the survey.

Table 28. Summary of External Surveys in Texas.

Study Area	Year of Survey	Total External Stations	External Stations Surveyed	Total Surveys	Through Surveys
Abilene	2005	21	11	3,329	164
Amarillo	2005	24	12	4,234	360
Austin	2005	42	22	8,298	618
Dallas-Fort Worth	2005	79	32	12,642	628
Longview	2004	60	30	8,426	1,712
Lubbock	2005	23	17	3,988	239
Midland-Odessa	2002	19	13	4,023	339
San Angelo	2004	23	11	4,031	334
San Antonio	2005	42	22	9,892	1,244
Sherman-Denison	2005	20	10	3,975	535
Tyler	2004	32	18	5,124	549
Waco	2006	24	15	4,557	583
Wichita Falls	2005	19	11	3,093	177
Total		428	224	75,612	7,482

MODEL APPROACH

To achieve the research objective, this research developed two models. The first model (Model I) estimates the proportion of *all* trips at an external station that are through trips (external-external, or E-E). The remaining proportion is the proportion of *all* trips that are local (external-internal, or E-I/I-E) trips. The second model (Model II) estimates the proportion of *through* trips at an external station that entered the study area at each of the other external stations. Multiplying the result for each entry external station from Model II by the result from Model I estimates the proportion of *all* trips that are through trips that entered at each of the other external stations.

Models I and II are both logit models. Logit models are appropriate when the response is one of a finite number of outcomes. Model I is a binary logit model, where the response has two possible outcomes (local or through). Model II is a conditional logit model, where the response has three or more possible outcomes (each of the external stations). For Model I, the through trip response is Y = 1, and the local trip response is Y = 0. Model I has the form shown in Equation 14:

$$g(\mathbf{x}) = \ln\left[\frac{P(Y=1)}{1-P(Y=1)}\right] \quad \text{(Eq. 14a)}$$

where:

$$g(\mathbf{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p \qquad \text{(Eq. 14b)}$$

The probability that $Y = 1$ is:

$$P(Y=1) = \frac{\exp[g(\mathbf{x})]}{1+\exp[g(\mathbf{x})]} \qquad \text{(Eq. 14c)}$$

For Model II, one of the responses is the baseline response, coded as Y = 0. The other responses are Y = 1, 2, …, J, where J + 1 is the number of possible responses. Model II has the form provided in Equation 15:

$$\begin{aligned} g(\mathbf{x}_0) &= 0 \\ g(\mathbf{x}_1) &= \ln\left[\frac{P(Y=1)}{P(Y=0)}\right] \\ g(\mathbf{x}_2) &= \ln\left[\frac{P(Y=2)}{P(Y=0)}\right] \\ &\vdots \\ g(\mathbf{x}_J) &= \ln\left[\frac{P(Y=J)}{P(Y=0)}\right] \end{aligned} \qquad \text{(Eq. 15a)}$$

where:

$$g(\mathbf{x}_j) = \beta_1 x_{1j} + \beta_2 x_{2j} + \cdots + \beta_p x_{pj} \qquad \text{(Eq. 15b)}$$

The probability that $Y = j$ is

$$P(Y=j) = \frac{\exp[\ g(\mathbf{x}_j)\ ]}{\sum_{k=0}^{J} \exp[g(\mathbf{x}_k)]} \qquad \text{(Eq. 15c)}$$

In both models the estimates for the parameters $\beta_0, \beta_1, \beta_2, \cdots + \beta_p$ result from maximizing the likelihood of the observed responses.

Most previous research has used linear models instead of logit models to predict through trip proportions. This research uses logit models because they have statistical and practical advantages over linear models. Previous research has fit linear models to the through trip proportions estimated from external surveys. Using the through trip proportions, rather than the number of responses for each possible outcome, results in a loss of all information about sample size. In addition, the linear model can result in proportion predictions that are more than one or less than zero and can result in estimates that do not sum to one. The logit models retain information about sample size, and the estimated proportions always sum to one as they should, with no estimates greater than one or less than zero.

Another possible model form is the nested logit model, where a single model could replace Model I and Model II. However, preliminary analysis showed erratic and poor results for the nested logit model, probably resulting from the fact that attributes of two very different kinds describe the E-I/I-E outcome and the E-E outcomes. In addition, even if a good nested model exists for this problem, the model would probably be hard to understand and interpret because of the different kinds of attributes for the outcomes.

MODEL I DEVELOPMENT AND EVALUATION

The Model I development process started with choosing candidate predictor variables. Then a subset of the candidate predictor variables was chosen to form a preliminary model, and the fit of the preliminary model was evaluated using model diagnostics. Then a final variable selection was made, this time from the variables in the preliminary model. The possibility of refining the model composed of the variables from the final selection using transformation and interactions was then investigated. The final model was then evaluated to determine its goodness of fit and practical applicability.

Candidate Predictor Variables

The model development process started by selecting candidate predictor variables, which are variables that have good potential for predicting through trips and merit further analysis. After defining each of the candidate predictor variables, this section discusses some of the more complicated variables, and variables that are new to predicting through trips. This section also explains why some of the variables that were used in previous research are not considered in this research. Finally, this section describes the data source for each of the predictor variables.

Variable Definitions

Table 29 defines each of the candidate predictor variables for Model I and divides them into groups by type of variable. The groups include traffic characteristics, roadway characteristics, study area characteristics, interaction score, and route validity. Several of the variables depend on the interaction score, which is defined in Table 30. The subscript j refers to the external station for which the through trip estimation is made, also called the survey station.

Table 29. Candidate Variable Definitions for Model I.

Traffic Characteristics	
$ADTALL_j$	The average daily traffic for external station j for all vehicle types, where ADT is the average non-holiday weekday 24-hour two-way count of vehicles passing through the external station.
$ADTLV_j$	The ADT for external station j for large vehicles, where large vehicles are vehicles belonging to classes 4 through 13 of the Federal Highway Administration (FHWA) vehicle classification system.

$PROPLV_j$	$$\frac{ADTLV_j}{ADTALL_j}$$
$PADTSV_j$	$$\frac{ADTSV_j}{\sum_{q \in E} ADTSV_q}$$ Where E is the set of all external stations in the study area, $ADTSV_j$ and $ADTSV_q$ are the ADTs for small vehicles for external stations j and q respectively, and small vehicles are vehicles belonging to classes 1 through 3 of the FHWA vehicle classification system.
$PADTLV_j$	$$\frac{ADTLV_j}{\sum_{q \in E} ADTLV_q}$$ Where $ADTLV_j$ and $ADTLV_q$ are the ADTs for large vehicles for external stations j and q, respectively.
$PADTAL_j$	$$\frac{ADTALL_j}{\sum_{q \in E} ADTALL_q}$$ Where $ADTALL_j$ and $ADTALL_q$ are the ADTs for all vehicles for external stations i and q, respectively.
Roadway Characteristics	
$LANES_j$	Total number of lanes in both directions at external station j. For example, the value of $LANES_j$ for an external station with two lanes in each direction would be 4. The lane count only includes main through lanes. Any turning lanes, median left turn lanes, climbing lanes or passing lanes are not counted.
$DVIDED_j$	A binary variable which is 1 when, in the area of external station j: (1) the two directions of traffic are separated by either a non-traversable barrier, such as a wall or railing, or by a non-paved area which is not intended for traffic, such as a grassy median; and (2) opportunities for left turns across the barrier or non-paved area at an intersection are less frequent than is typical for an urban arterial. The variable is 0 otherwise.

$LIMTED_j$	A binary variable, which is 1 when the roadway in the area of external station j is a limited-access facility, which means that access to the roadway is only provided by ramps. For areas where the roadway transitions from limited access to non-limited access the variable is 1. The variable is 0 otherwise.
Interaction Score Variables	
$INTTHR_j$	$\sum_{\{q \in E, q \neq j\}} INT_{qj}$ Where INT_{qj} is the through interaction score for entry external station q and survey external station j, as defined in Table 30.
$INTTTL_j$	$INTTHR_j + INTLCL_j$ Where $INTLCL_j$ is the local interaction score for survey station j, as defined in Table 30.
$PINTTH_j$	$\frac{INTTHR_j}{INTTTL_j}$ if $INTTTL_j = 0$ then $PINTTH_j = 0$
$INTTL1_j$	A binary variable, which is 1 if $INTTTL_j$ is greater than 0, and is 0 otherwise.
$PINTTH1_j$	A binary variable, which is 1 if $PINTTH_j$ is greater than 0, and is 0 otherwise.
Route Validity	
$RTELCL_j$	A binary variable, which is 1 if the non-congested least time route from the centroid of at least one U.S. Census urban area or urban cluster whose centroid is with the study area to external station j is valid, and is 0 otherwise. A route is valid if (1) it passes through the study area and (2) it crosses the study area boundary only at external station j.
Characteristics of the Study Area	
POP	Population of the study area.
EMP	Employment in the study area.
INC	Average household income of residents of the study area.
$AREA$	Surface area of the study area in square miles.

ADTALL	The ADT for all vehicles summed across all external stations.
ADTLV	The ADT for large vehicles summed across all external stations.
Is Commercial Vehicle	
ISCV	A binary variable, which is 1 if the vehicle is a commercial vehicle, and is 0 otherwise. Here a commercial vehicle is any vehicle used for a commercial purpose, regardless of size or type of vehicle.
Average Turns	
$AVGTRN_j$	$\sum_{\{q \in E, q \neq j\}} \left(\frac{DIST_{qj} \cdot ROUTE_{qj} \cdot TURNS_{qj}}{\sum_{\{r \in E, r \neq j\}} DIST_{rj} \cdot ROUTE_{rj}} \right)$ Where $DIST_{qj}$ and $DIST_{rj}$ are the results from Model II, and $ROUTE_{qj}$, $ROUTE_{rj}$, and $TURNS_{qj}$ are defined in Table 42; j is the survey external station; and q and r are entry external stations.

Table 30. Interaction Score Definition.

INT_{ij}	$\sum_{\{v\in U, v\neq w\}}\left(\sum_{w\in U}\left(\frac{P_v\times P_w}{10^6\times D_{vw}^2}\times f_{vwij}\right)\right)$
$INTLCL$	$\sum_{\{v\in U, v\neq w\}}\left(\sum_{w\in U}\left(\frac{P_v\times P_w}{10^6\times D_{vw}^2}\times g_{vwij}\right)\right)$
	where:
INT_{ij}	is the through interaction score for entry external station i and survey external station j;
$INTLCL$	is the local interaction score for survey external station j;
U	is the set of each U.S. Census Bureau urban area and urban cluster which has its centroid within the study area; or has its centroid within 50 miles of the study area boundary; or has a population of at least 50,000 people and has its centroid within 250 miles of the urban area boundary;
P_v, P_w	are the populations of v and w;
D_{vw}	is the non-congested least time route distance in miles from the centroid of v to the centroid of w;
f_{vwij}	is a binary variable, which is 1 if the non-congested least time route from the centroid of v to the centroid of w passes through external stations i and j, and if the route segment between i and j is valid. The route segment is considered valid if (1) it passes through i before j; and (2) it passes through the study area; and (3) it crosses the study area boundary only i and j. Otherwise, the variable is 0.
g_{vwj}	is a binary variable, which is 1 if the non-congested least time route from the centroid of v to the centroid of w passes through external station j; and if the centroid of v is inside the study area; and if the route segment between v and j is valid. The route segment is considered valid if it crosses the study area boundary only at j. Otherwise, the variable is 0.

For most of the candidate predictor variables, the meaning and importance of the variable is obvious from the variable definition. However, the interaction score variables, the route validity variable, and the average turns variable are complicated and warrant further discussion to make their meaning and importance more obvious. The "is commercial vehicle" variable is also discussed here, since it is an important new variable that has not been included in previous research.

Interaction Score Variables

The interaction score generates and distributes relative amounts of trips using a simple gravity model, assigns the trips to the roadway network, then checks to see if the trips pass through the study area. The gravity model assumes that all trips originate and terminate at the centroid of urban areas, and that the relative amount of travel between two urban areas is proportional to the product of the urban areas' populations and inversely proportional to the square of the distance on the least time route between the urban areas centroids.

The results from the gravity model are assigned to the least-time route between urban area centroids. The routes are then checked to determine if they enter or exit the study area, or both, and then are checked to determine which external stations they use. At this point the gravity model results are assigned to the appropriate external stations or external station pairs.

The interaction score, and the variables that are based on it, take into account the geographical distribution of land uses that generate and attract significant numbers of trips, and the configuration of the roadway network that connects the land uses. These predictors are the basis of most travel demand models and are two of the most important predictors of travel demand.

Route Validity Variable

For some external stations, the least-time route from one or more of the urban areas centroids inside the study area to the external station passes through another external station. In this case the route is not valid. If none of the routes from internal urban area centroids are valid, then it is less likely that trips passing through the external station are E-I/I-E trips, and more likely that they are through trips. The variable $RTELCL_j$ reflects this observation. It is 1 when at least one of the routes is valid and is 0 when no routes are valid.

A simple example best explains the route variable. Figure 17 shows a study area boundary (the dashed line), roads (the solid lines), four external stations (the small squares), and an internal urban area (the small circle). The least time routes from the urban area to external stations *w* and *x* are valid. The least time routes from the urban area to external stations *y* and *z* are not valid, since both routes pass through external station *x*. Therefore, *y* and *z* are less likely to have E-I/I-E trips and more likely to have through trips.

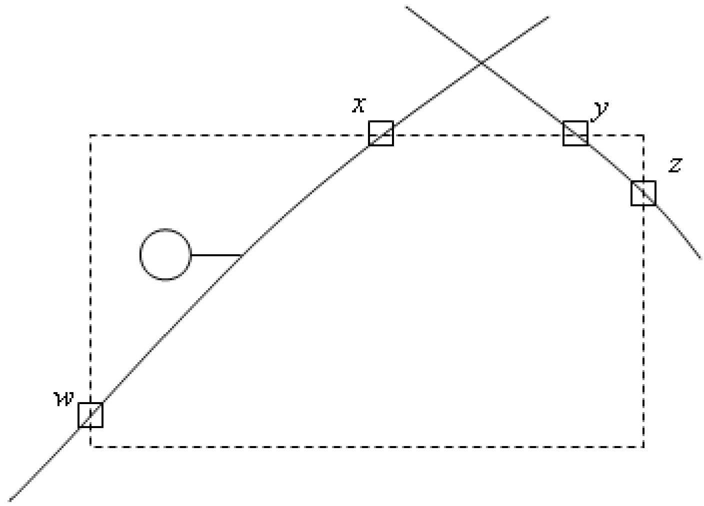

Figure 17. Local Route Validity Example.

Average Turns Variable

The variable $AVGTRN_j$ is a weighted average of the number of turns on the routes from each of the other external stations to external station *j*. The average is weighted by the through trip distribution results from Model II. The purpose of $AVGTRN_j$ is to measure the directness for through trips exiting the study area at external station *j*. A high value of the variable means that, on average, through trips exiting the study area at external station *j* would have made a high number of turns while inside the study area. A low value of the variable means that, on average, through trips would have made a low number of turns inside the study area. In the first case, the proportion through trips is probably very low, since the directness offered to through trips is low. In the second case, the number of through trips is probably high, since the directness offered to through trips is high.

"Is Commercial Vehicle" Variable

The variable *ISCVj* is new in this research. None of the previous research made separate predictions for commercial vehicles and non-commercial vehicles. However, such a separation is important, since commercial vehicles likely have different through trip patterns. On average, commercial vehicles have longer trips than non-commercial vehicles, so a greater proportion of commercial vehicle trips may be through trips than of non-commercial vehicle trips. In addition, since most commercial vehicles are also large vehicles, they place different pavement, traffic flow, and air quality demands on the transportation system than do non-commercial vehicles.

Variables Not Selected as Candidate Predictor Variables

Previous research on developing synthetic external models utilized a variety of variables. The set of candidate variables for this research includes many of those variables but not all of them. The following section explains why some of those variables utilized previously are not included in the development of this model.

The variables functional classification, percent pickups and vans, and marginal highway route were used in the final version of at least one previous model but are not considered as candidate variables for this research. The variable functional classification is not included for a number of reasons. First, functional classification tends to be somewhat subjective, especially for areas that transition from urban to rural. Since most external stations are in these types of areas, a functional classification variable may lead to inconsistencies in model development and application.

Second, several of the candidate predictor variables provide information that is very similar to that which would be provided by a functional classification variable. The variables roadway $LANES_j$, $DVIDED_j$, and $LIMITED_j$ along with the traffic variables appear to provide more than enough information to make up for the absence of the functional classification variable. Finally, previous research has not proven conclusively that functional classification is always a good predictor of through trips. The Modlin 1974 stage one model included functional classification, but it is unclear whether this model was compared with a model that did not include functional classification. Even Modlin himself did not include functional classification in his 1982 stage one model. The Han 2008 stage one model includes functional classification but only as a single dummy variable indicating whether or not the road is a collector or local road.

The percent pickups and vans was not included as a candidate predictor variable because previous research did not show that it was always a good predictor. It was included in the Modlin 1974 model, but Modlin did not include it in his 1982 model. In addition, these data are not available for some of the study areas, because pickup trucks and vans were aggregated with smaller cars and motorcycles in the vehicle classification schemes.

Marginal highway route is a variable in the Han 2008 model which indicates whether an external station is on a highway route that cuts through the corner of a study area or almost parallels the study area boundary to create two external stations very close together on the same highway. Marginal highway route is not included as a candidate predictor variable for this research because several of the other candidate variables provided the same information in a less subjective way, such as the interaction score variables, $RTELCL_j$, and $AVGTRN_j$.

Additionally, the variables "nearby major city," "route continuity," and "internal-external factor" were included in at least one previous model but are not candidate variables for this research. "Internal-external factor" and "route continuity" are not included because they are only necessary for combined models, and this research created a two-stage model. The variable "nearby major city" is not included because the interaction score variables provide the same information in a more comprehensive and less subjective way.

Data Sources for Predictor Variables

Data for the traffic characteristics variables were obtained from pneumatic tube vehicle classification counts conducted with the external surveys. Roadway data came from Google Earth, which provides satellite images from several different years, so that the year of the image used and the year of the survey never differ by more than a few years. The study area characteristics were obtained from data provided by the U.S. Census Bureau and the U.S. Bureau of Labor Statistics. The source for data for the average turns variable is explained later in this chapter.

The interaction score variables depend on the location and population of urban areas, and on the least time routes between urban areas. As stated in the interaction score definition, the data for urban area locations and populations are provided by the U.S. Census Bureau, which publishes population estimates for each Census urban area and urban cluster, as well as provides a GIS file with polygons for all urban areas and clusters throughout the United States.

Least time routes between urban area centroids were extracted from the Bing Maps web service using an MS Visual Basic 2008 utility. After extracting the routes, the utility analyzed them to determine which external stations the route passes through (if any) and if the route segments are valid, as described in the interaction score definition. The same utility was also used to extract routes for the route validity variable.

Preliminary Variable Selection

The selection of candidate variables was based on the work of previous researchers, and on new theories about what variables have good potential for predicting through trips. From this set of candidate predictor variables, a new selection of variables was made based on forward selection, which was a more rigorous variable selection technique.

Forward selection begins with a model containing only a constant. Then, the one variable that has the lowest p-value in a likelihood ratio test when added to the constant only model is added to create a new model with one variable and the constant. Then, the one variable that has the lowest p-value when added to the one-variable model is added to create a new model with two variables and the constant. The forward selection process continues in this manner, with variables added one at a time according to the results from a likelihood ratio test (*34*).

Usually, the forward selection process continues until no variable can be added with a p-value smaller than some pre-specified value, such as 0.05 or 0.01. However, preliminary analysis showed that following such a rule would result in selecting most of the candidate variables, and such a large model is not desirable for multiple reasons. First, selecting too many variables can result in a model that over-fits the data, meaning that the model fits noise in the data rather than the true pattern. Second, a large model would be more difficult to understand and interpret than a smaller model. Third, a large model would have higher data collection costs than a smaller model.

To help limit the number of variables selected, the model development process used the Akaike information criterion (*AIC*), the Bayesian information criterion (*BIC*), and adjusted rho-square ($\overline{\rho}_C^2$) as defined in Equations 16, 17, and 18 (*35, 36*). Each of these criteria is a measure of the log-likelihood of the model, penalized for the number of variables in the model. Lower values of *AIC* and *BIC*, and higher values of $\overline{\rho}_C^2$ indicate a better model.

$$AIC = \frac{-2(LL(\hat{\beta}) - K)}{N} \quad \text{(Eq. 16)}$$

where:

AIC = Akaike information criterion,

$LL(\hat{\beta})$ = the log-likelihood for the estimated model,

K = the number of parameters in the estimated model, and

N = the number of covariate patterns (survey external stations) in the sample.

$$BIC = \frac{-2(LL(\hat{\beta}) - K \times \log(K))}{N} \quad \text{(Eq. 17)}$$

where:

BIC = Bayesian information criterion.

$$\bar{\rho}_C^2 = 1 - \frac{LL(\hat{\beta}) - K}{LL(C) - K_{MS}} \quad \text{(Eq. 18)}$$

where:

$\bar{\rho}_C^2$ = adjusted rho squared with respect to the constants only model,

$LL(C)$ = the log-likelihood for the constant only model, and

K_{MS} = the number of parameters in the constants only model (here equal to 1).

Normally, the model from forward selection with the best value of a criterion would be chosen. However, preliminary analysis showed that following this rule would also choose most of the predictor variables. Rather than the absolute value of each criterion, the rate of change of each criterion is used as a guide for forward selection. With this rule, a significant decrease in the rate of improvement of the criteria would suggest ending forward selection.

In addition to AIC, BIC and $\bar{\rho}_C^2$, the root mean square error ($RMSE$) as defined in Equation 19 is used as a guide for forward selection, where a lower value indicates a better model. Although $RMSE$ is not as statistical valid as the other three criteria, it does give a practical and intuitive sense of how well a model fits.

$$RMSE = \left(\frac{\sum_{n=1,\cdots,N} (\hat{\pi}_n - p_n)^2}{N} \right)^{1/2} \quad \text{(Eq. 19)}$$

where:

$RMSE$ = the root mean square error,

n = an index for each covariate pattern (survey external station) in the sample,

p_n = the sample proportion through trips for survey external station n,

$\hat{\pi}_n$ = the estimated proportion through trips for survey external station n.

Table 31 lists the results of the forward selection process, with the variables appearing in the order that they were added. Each row gives the values of the criteria for the model including the variables on that row and all previous rows. Each row also gives the p-value for a likelihood ratio test for the model with the variable on that row and all previous rows, compared to a model with only the variables on the previous rows.

Table 31. Forward Selection Results for Model I.

Variable	AIC	BIC	$\bar{\rho}_C^2$	*RMSE*	P
Constant	110.2	110.2		16.4	
$PINTTH_j$	102.7	102.7	0.068	13.8	$< 10^{-100}$
$ISCV_j$	99.7	99.7	0.095	11.9	$< 10^{-100}$
$LANES_j$	99.3	99.4	0.099	11.9	10^{-38}
$PROPLV_j$	98.8	98.9	0.103	11.6	10^{-52}
EMP_j	98.7	98.7	0.105	11.7	10^{-15}
$ADTALL$	98.3	98.4	0.108	11.7	10^{-34}
$RTELCL_j$	98.2	98.3	0.109	11.3	10^{-16}

Figure 18 graphically presents the values of the criteria as a function of the number of variables in the model. The plots show that each criterion improved quickly up to the second variable, where the rate of improvement of the criteria slowed significantly, suggesting that the model with two variables is the best model. However, to allow for the possibility that additional variables would be important to the model, variable selection continued.

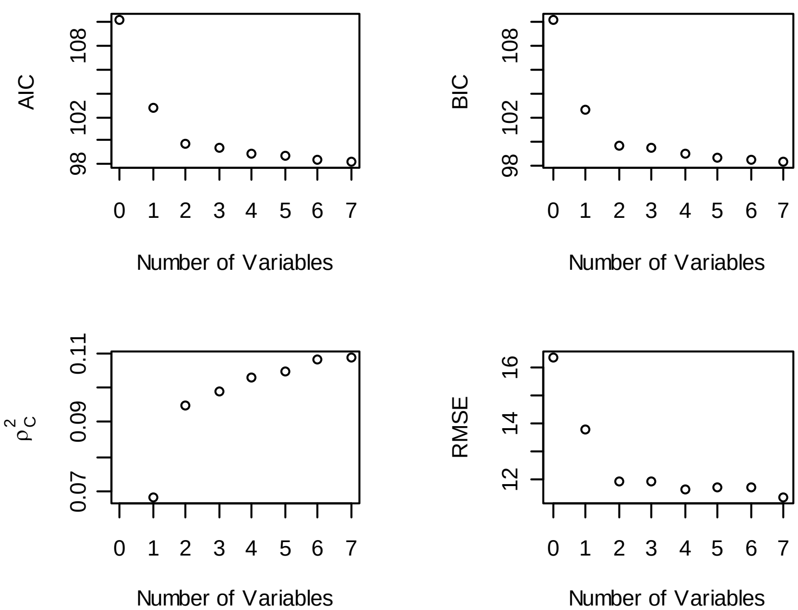

Figure 18. Forward Selection Results for Model I.

Variable selection stopped at the seventh variable, because the rate of improvement of the criteria continued to slow, and because all of the important variables had already been selected. Initial analysis showed that the last variable, $RTELCL_j$, performed poorly, so it was dropped and the first six variables formed the preliminary model, called Model I-a. Table 32 presents Model I-a.

Table 32. Model I-a.

Variable	Coeff.	Std. Err.	z	P	Mean of X
Constant	-3.10	0.0609	-50.9	<0.0001	
$PINTTH_j$	1.92	0.0417	46.2	<0.0001	0.216
$ISCV_j$	1.15	0.0304	38.0	<0.0001	0.125
$LANES_j$	-0.189	0.0144	-13.1	<0.0001	2.68
$PROPLV_j$	1.81	0.142	12.7	<0.0001	0.155
EMP_j	-4.84×10^{-7}	-3.39×10^{-8}	-14.3	<0.0001	6.82×10^{5}
$ADTALL$	2.92×10^{-6}	2.42×10^{-7}	12.1	<0.0001	2.11×10^{5}

Diagnostics

Before continuing to the final variable selection, the model development process checks each observation using three model diagnostics: ΔX^2_n, ΔD_n, and $\Delta\beta_n$. The first two diagnostics measure the effect of the observations with covariate pattern n on the model Pearson chi-square statistic and the model deviance, which are two summary measures of goodness of fit. The third diagnostic, $\Delta\beta_n$, detects covariate patterns whose observations have a large effect on the parameter estimates. Especially poor (high) values of these three diagnostics are useful in detecting covariate patterns that have data errors or whose observations are not fit well by the model.

Figure 19 presents ΔX_n^2 and ΔD_n versus $\hat{\pi}_n$. Three points are especially high compared to the other points in each figure. These three points correspond to external stations 1213 in San Antonio, GR20 in Sherman-Denison, and 711 in Tyler, where in each case the model under-predicted the proportion through trips. An investigation revealed no data errors (many errors have been identified and corrected in exploratory analysis) and showed that the observations are plausible. Attempts to find a variable that would improve the fit of the model for these points while not over-fitting the model were unsuccessful.

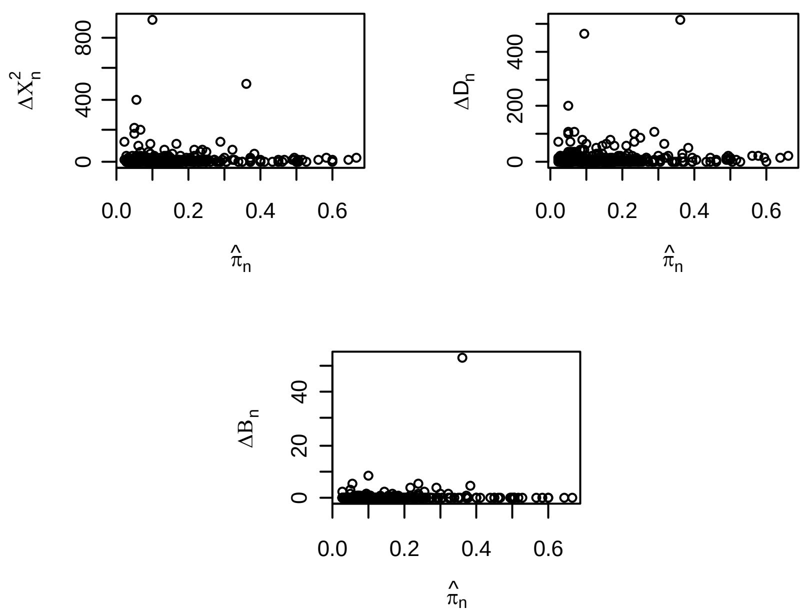

Figure 19. Model I-a Diagnostics.

Figure 19 also presents $\Delta\beta_n$ versus $\hat{\pi}_n$. One point, corresponding to external station 1213 in San Antonio, was especially high compared to the other points. To further investigate its effect on the parameter estimates, the observations from the external station were removed from the data and Model I-a was fit again.

Table 33 compares the new parameter estimates to the original parameter estimates. None of the parameter estimates changed by a large amount (no parameter estimate changed by more than 19 percent), so the external station was retained in the data.

Table 33. Parameter Estimate Change after Removing External Station 1213.

Variable	Coeff	Percent Change
Constant	-3.01	-3
$PINTTH_j$	1.73	-10
$ISCV_j$	1.19	3
$LANES_j$	-0.160	-15
$PROPLV_j$	1.46	-19
EMP_j	-4.81×10^{-7}	-1
$ADTALL$	2.53×10^{-6}	-13

Final Variable Selection

The results from the forward selection process suggested that a model with only two variables was the most appropriate model. However, to allow for the possibility that a richer model would actually be more appropriate, six variables were selected and retained throughout the model diagnostics process. Now the model development process more thoroughly investigated the hypothesis that the two-variable model was the better model.

The investigation was based on the following experiment: take a random sample of study areas from the set of all 13 study areas, then fit a model using the results from the external surveys for the sampled study areas. Repeat this process a number of times and then compare the parameter estimates from each repetition. Variables whose parameter estimates change relatively little between repetitions of the experiment are better predictors than variables whose parameter estimates change much.

The results from this approach roughly give some of the same information as the standard error in Table 32, since both measure the variability of the parameter estimates. However, this approach has the advantage that it measures parameter estimate variability by sampling whole study areas at a time, which gives confidence that the results can be extended to an entirely new study area.

To carry out the experiment, the set of 13 study areas were randomly divided into four groups, as presented in Table 34. Then each group of study areas was removed from the dataset, and a model was fit to the remaining data, to produce four sets of parameter estimates.

Table 35 presents each new parameter estimate as a relative change from the original parameter estimates. The first two variables have significantly smaller changes than do the last four. The largest change in the first two variables is 16 percent, while the last four variables change by at least 29 percent at least once, and three of the last four change by at least 51 percent at least once.

Table 34. Study Area Groups.

Group 1	Amarillo, San Antonio, and Waco
Group 2	Austin, Lubbock, and Sherman-Denison
Group 3	Dallas-Fort Worth, Longview, San Angelo, and Wichita Falls
Group 4	Abilene, Midland-Odessa, and Tyler

Table 35. Parameter Estimate Changes when Removing Study Area Groups (percent).

	Group Removed			
Variable	1	2	3	4
Constant	-7	21	-1	-9
$PINTTH_j$	-12	1	7	4
$ISCV_j$	16	-6	-9	-2
$LANES_j$	9	-51	17	19
$PROPLV_j$	-23	22	29	-16
EMP_j	4	23	-56	-14
ADTALL	-14	59	-20	-29

The results from this investigation confirmed the hypothesis that the smaller, two variable model may be the better model. One of these two variables is *PINTTHj*, which is the interaction score for through trips as a portion of the interaction score for all trips. However, some of the external stations have no interaction scores at all. For these external stations the portion is not defined, and the variable is set to be zero. Thus, the meaning of a zero value for this variable is ambiguous, because it could mean that the external station has no interactions at all, or it could mean that the external station has interactions, but none of them are interactions for through trips. To allow the model to distinguish between these two situations, the variable *INTTL1j* was added to the model. This variable is 0 when the external station has no interaction scores, and is 1 when the external station has any interaction score. Thus it serves as an adjustment to the model to distinguish between the two cases when *PINTTHj* is zero. The resulting model is Model I-b, and is presented in Table 36.

Table 36. Model I-b

Variable	Coeff.	Std. Err.	z	P	Mean of X
Constant	-2.94	0.0292	-101	<0.0001	
$PINTTH_j$	2.38	0.0466	51.0	<0.0001	0.216
$ISCV_j$	1.16	0.0301	38.4	<0.0001	0.125
$INTTL1_j$	-0.235	0.0383	-6.13	<0.0001	0.697

Model Refinement

To this point, the model development assumed that the continuous variables are linear in the logit, and the effect of each variable does not vary across the levels of any of the other variables. The model development process then tested each of these assumptions.

To test the first assumption, the model development process used the logit step test, which was performed as follows. First, the continuous variable was divided into four groups of equal intervals or divided into four groups based on the quartiles of the variable. Then the continuous variable was recoded as a categorical variable with a set of three design variables. The design variables correspond to the second, third, and fourth groups of the continuous variable, and the first group acts as the base class. Next, researchers fit a model, replacing the continuous variable with the new categorical variable. Finally, the parameter estimates were plotted for each design variable against the midpoint of the corresponding group of the continuous variable. For the first group, zero was plotted against its midpoint. If the continuous variable was linear in the logit, then the plotted points would show a linear relationship. If not, then the shape of the plot would suggest possible transformations of the continuous variable to make it linear in the logit.

Figure 20 shows the plot for the step test of $PINTTH_j$, the only continuous variable in the model. The plot does not show evidence that $PINTTH_j$ is not linear in the logit, so the variable is not transformed.

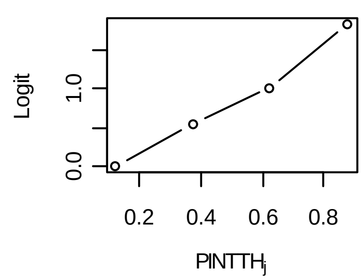

Figure 20. Step Test Plot for *PINTTH$_j$*.

To test the second assumption that the effect of each variable does not vary across the levels of any other variable, the model development process compared Model I-b to a model with added interactions. With three variables in Model I-b, three two-variable interactions were possible, but only the interactions with $ISCV_j$ were tested. The two interactions are defined in Equations 20 and 21. The two interactions were added to Model I-b to create Model I-c, which is presented in Table 37. The interactions do not contribute to the model significantly, as evidenced by the relatively high P values. Therefore, Model I-c is not retained, and Model I-b is the final model.

$$SIPINTTH = ISCV_j \times PINTTH_j \quad \text{(Eq. 20)}$$

$$SIINTTL1 = ISCV_j \times INTTL1_j \quad \text{(Eq. 21)}$$

Table 37. Model I-c.

Variable	Coeff.	Std. Err.	z	P	Mean of X
Constant	-2.97	0.0326	-91.0	<0.0001	
$PINTTH_j$	2.43	0.0529	45.9	<0.0001	0.216
$ISCV_j$	1.30	0.0668	19.5	<0.0001	0.125
$INTTL1_j$	-0.218	0.0443	-4.91	<0.0001	0.697
$SIPINTTH_j$	-0.256	0.111	-2.31	0.0212	0.0316
SIINTTL1j	0.0749	0.0883	-0.849	0.3961	0.0957

Model Evaluation

After the final model was chosen, researchers then evaluated its performance. The goal of this research is to develop a model that can be applied with reasonably accurate results to Texas study areas, including study areas that are not in the dataset for this research. The model evaluation simulates applying the model to new study areas using cross validation, which fits the model using a randomly selected segment of the available data, then tests the model on the remaining data. The model fitting and testing is repeated four times, using each of the four groups as defined in Table 34 as the test datasets, and the remaining data in each case as the model fitting dataset. The fit of the model to each of the training data sets is evaluated using cross-classification tables. Table 38 through Table 41 present the cross-classification tables, which classify each observation by its observed and predicted percent through trips.

Table 38. Cross Classification of Observations for Group 1.

Observed	Predicted (percent)						
(percent)	0–5	6–10	11–20	21–40	41–60	61–80	81–100
0–5	19	3	15	0	0	0	0
6–10	6	4	5	0	0	0	0
11–20	1	4	8	7	0	0	0
21–40	1	1	8	6	3	0	0
41–60	0	0	2	1	2	0	0
61–80	0	0	0	0	0	0	0
81–100	0	0	0	1	1	0	0

Table 39. Cross Classification of Observations for Group 2.

Observed	Predicted (percent)						
(percent)	0–5	6–10	11–20	21–40	41–60	61–80	81–100
0–5	16	7	6	0	0	0	0
6–10	7	3	6	1	0	0	0
11–20	4	3	14	2	0	0	0
21–40	2	0	11	4	0	0	0
41–60	0	0	2	2	0	0	0
61–80	0	0	1	0	1	0	0
81–100	0	0	0	0	0	0	0

Table 40. Cross Classification of Observations for Group 3.

Observed	Predicted (percent)						
(percent)	0–5	6–10	11–20	21–40	41–60	61–80	81–100
0–5	12	30	20	1	0	0	0
6–10	3	7	15	3	0	0	0
11–20	1	4	17	2	0	0	0
21–40	0	1	19	13	6	0	0
41–60	0	0	2	1	2	1	0
61–80	0	0	0	2	1	0	0
81–100	0	0	0	0	0	1	0

Table 41. Cross Classification of Observations for Group 4.

Observed	Predicted (percent)						
(percent)	0–5	6–10	11–20	21–40	41–60	61–80	81–100
0–5	5	18	12	1	0	0	0
6–10	0	6	5	2	0	0	0
11–20	0	1	11	2	1	0	0
21–40	0	1	5	3	1	0	0
41–60	0	0	0	3	0	0	0
61–80	0	1	0	0	0	0	0
81–100	0	0	0	1	1	0	0

MODEL II DEVELOPMENT AND EVALUATION

Model II was the second model in the two-model system developed by this research. The first model, Model I, predicted the portion of all trips at an external station that are through trips. Model II was developed to distribute the through trips by predicting the proportion of all through trips exiting the study area at an external station j that entered the study area at each external station i. The development of Model II followed the same general process as that of Model I. The first step was to choose candidate predictor variables. Then the preliminary model was formed by choosing variables from the set of candidate variables based on the results of forward selection. Then, a second and final variable selection was made from the variables in the preliminary model. Next, transformations of the selected variables were tested to form the final model. Finally, the performance of the final model was evaluated. This section describes each of these parts of the Model II development process.

Candidate Predictor Variables

The first step of the model development process was to select a set of candidate predictor variables. The purpose of this step was to provide a set of variables that merit further analysis. The selection of some of the candidate variables was based on the results of previous research, but other candidate variables are new. This section defines each variable, discusses the new and more complicated variables in more detail, explains why some variables were not selected as candidate predictor variables, and gives the variable data sources. Table 42 defines each of the candidate predictor variables for Model II and divides them into groups by variable type.

Table 42. Variable Definitions for Model II.

<u>Traffic Characteristics</u>	
$ADTALL_i$	The average daily traffic for external station i for all vehicle types, where ADT is average non-holiday weekday 24-hour two-way count of vehicles passing through the external station.
$ADTLV_i$	The ADT for external station i for large vehicles, where large vehicles are vehicles belonging to classes 4 through 13 of the FHWA vehicle classification system.
$PROPLV_i$	$\frac{ADTLV_i}{ADTALL_i}$
$PADTSV_{ij}$	$\frac{ADTSV_i}{\sum_{\{q \in E, q \neq j\}} ADTSV_q}$ where $ADTSV_i$ and $ADTSV_q$ are the ADTs for small vehicles for external stations i and q, respectively; j is the survey external station; E is the set of all external stations in the study area; and small vehicles are vehicles belonging to classes 1 through 3 of the FHWA vehicle classification system.
$PADTLV_{ij}$	$\frac{ADTLV_i}{\sum_{\{q \in E, q \neq j\}} ADTLV_q}$ where $ADTLV_i$ and $ADTLV_q$ are the ADTs for large vehicles for external stations i and q, respectively, and j is the survey external station.

$PADTAL_{ij}$	$\frac{ADTALL_i}{\sum_{\{q \in E, q \neq j\}} ADTALL_q}$ where $ADTALL_i$ and $ADTALL_q$ are the ADTs for all vehicles for external stations *i* and *k*, respectively, and *j* is the survey external station.

<u>Roadway Characteristics</u>

$LANES_i$	Total number of lanes in both directions at external station *i*. For example, the value of $LANES_i$ for an external station with two lanes in each direction would be 4. The lane count only includes main through lanes. Any turning lanes, median left turn lanes, climbing lanes, or passing lanes are not counted.
$4LANE_i$	A binary variable, which is 1 when $LANES_i$ is greater than or equal to 4, and is 0 otherwise.
$DVIDED_i$	A binary variable, which is 1 when, in the area of external station *i*: (1) the two directions of traffic are separated by either a non-traversable barrier, such as a wall or railing, or by a non-paved area, which is not intended for traffic, such as a grassy median; and (2) opportunities for left turns across the barrier or non-paved area at an intersection are less frequent than is typical for an urban arterial. The variable is 0 otherwise.
$LIMTED_i$	A binary variable, which is 1 when the roadway in the area of external station *i* is a limited-access facility, which means that access to the roadway is only provided by ramps. For areas where the roadway transitions from limited access to non-limited access the variable is 1. The variable is 0 otherwise.

<u>Measures of Separation between External Stations</u>

$DISTRO_{ij}$	The great circle distance in miles divided by the non-congested least time route distance in miles from external station *i* to external station *j*.
$RSPEED_{ij}$	The great circle distance in miles divided by the non-congested least time route duration in hours from external station *i* to external station *j*.
$TURNS_{ij}$	The number of turns on the non-congested least time route from external station *i* to external station *j*.

$RAMPS_{ij}$	The number of freeway ramps, including on-ramps, off-ramps, and freeway-to-freeway ramps, on the non-congested least time route from external station *i* to external station *j*.
Interaction Score Variables	
INT_{ij}	The interaction score for external stations *i* and *j* as defined in Table 30.
$PINT_{ij}$	$\dfrac{INT_{ij}}{\sum_{\{q \in E, q \neq j\}} INT_{qj}}$
$PINT1_{ij}$	A binary variable which is 1 if $PINT_{ij}$ is greater than 0, and is 0 otherwise.
Route Validity	
$ROUTE_{ij}$	A binary variable which is 1 if the least time route from external station *i* to external station *j* is valid, and is 0 otherwise. The route is valid if (1) it passes through the study area, and (2) it crosses the study area boundary only at external stations *i* and *j*.
Results from Model I	
$CVSPLT_i$	Proportion through trips at external station *i* for commercial vehicles as predicted by Model I.
$NCSPLT_i$	Proportion through trips at external station *i* for non-commercial vehicles as predicted by Model I.
$ADTTHR_i$	$CVSPLT_i \times ADTLV_i + NCSPLT_i \times ADTSV_i$

Discussion of Candidate Predictor Variables

The meaning and importance of most of the candidate predictor variables is clear from the variable definitions. However, some of the variables are more complicated or are new variables that have not been used by other researchers. These variables are discussed here in greater detail, with the exception of the variables based on the interaction score, which is discussed in previous sections.

Measures of Separation between External Stations

The purpose of the measures of separation between external station pairs was to quantify the likelihood that a trip passing through one external station would also pass through a second external station. The approach was to measure the directness of travel between external station pairs, with the assumption that external station pairs with more direct connecting routes were more likely to share trips than external station pairs with less direct connecting routes. $DISTRO_{ij}$ and $RSPEED_{ij}$ measured the distance and time separation between external stations, while normalizing for the great circle distance between external stations. High values of these two variables indicated high directness. $TURNS_{ij}$ and $RAMPS_{ij}$ measured the separation between external stations using the number of turns and ramps. High values of these variables indicated low directness.

Route Validity Variable

For some pairs of external stations, the least time route connecting the two external stations passed through one or more other external stations, or the route did not pass through the study area. In either of these cases the route was not valid. It is unlikely that two external stations that do not have a valid connecting route would exchange through trips.

The route validity variable is illustrated in Figure 21, where the dotted line represents the study area boundary, the solid lines represent roads, and the small squares represent external stations. If external station *w* is the survey external station, then the route from external station *x* is valid, but the routes from external stations *y* and *z* are not valid. Therefore, *w* probably exchanges it's through trips mostly with external station *x*.

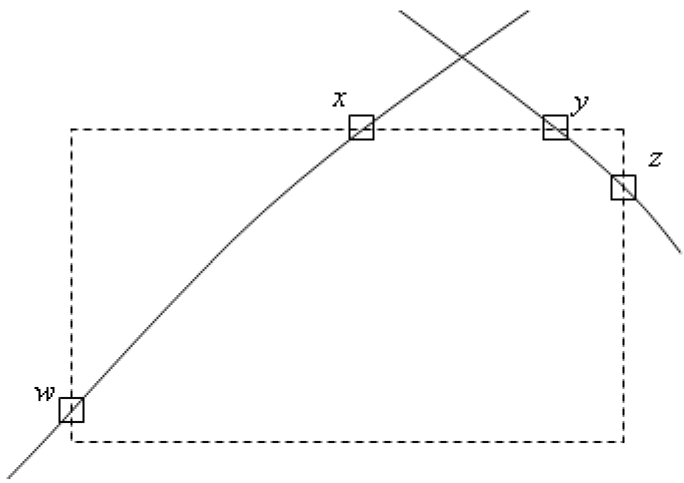

Figure 21. Illustration of Through Route Validity Variable.

Variables Not Selected as Candidate Predictor Variables

Previous research on combined and stage two models utilized numerous variables. Of these variables, most are included as candidate variables for this research. However, some are not candidate variables. This section explains why they are not included.

Functional classification was not included for a number of reasons. First, as discussed previously, functional classification tends to be subjective, and this subjectivity can lead to inconsistencies in developing and applying the model. Second, several of the candidate predictor variables provide information that is very similar to that which would be provided by a functional classification variable. The traffic and roadway variables probably provide more than enough information to make up for the absence of a functional classification variable.

Third, previous research has not proven conclusively that functional classification is a good predictor of through trip distribution. It is included in the Modlin 1974, Pigman 1979, and Modlin 1982 stage two models, but it is unclear whether they were compared to models without the functional classification variable. The variable "route continuity" is not included because the external station separation variables give the same information in a more comprehensive way. Zipf's probability factor and Horowitz's weight are replaced by the interaction score variables.

Data Sources for Predictor Variables

Data for the traffic characteristics variables was obtained from pneumatic tube vehicle classification counts conducted with the external surveys. Roadway data came from Google Earth, which provides satellite images from several different years, so that the year of the image used and the year of the survey never differ by more than a few years.

The interaction score variables depended on the location and population of urban areas, and on the least time routes between urban areas. As stated in the interaction score definition, the data for urban area locations and populations was provided by the U.S. Census Bureau, which publishes population estimates for each Census urban area and urban cluster, as well as provides a GIS file with polygons for all urban areas and clusters throughout the United States.

Least time routes between urban area centroids were extracted from the Bing Maps web service using a MS Visual Basic 2008 utility. After extracting the routes, the utility analyzed them to determine which external stations the route passed through (if any) and if the route segments were valid, as described in the interaction score definition. The same utility was also used to extract routes for the route validity variable and for the external station separation variables.

Preliminary Variable Selection

Forward selection as described in the previous chapter was used to select variables from the set of candidate predictor variables, with two changes. First, the criterion $\overline{\rho}_C^2$ was replaced by the criterion $\overline{\rho}_0^2$, which is defined in Equation 22. Both of these criteria measured the log-likelihood of the model, while penalizing larger models. The first criterion, $\overline{\rho}_C^2$, measured the log-likelihood of the model with respect to a model with only a constant. For Model II, a model with only a constant would not be appropriate, since all alternatives have the same utility function. The variable $\overline{\rho}_0^2$ measures the log-likelihood with respect to a model, which gives equal likelihood to each alternative. The second change was that the forward selection process starts with the equal-likelihood model, rather than the constant only model.

$$\overline{\rho}_0^2 = 1 - \frac{LL(\hat{\beta}) - K}{LL(0)} \qquad \text{(Eq. 22)}$$

where:

$\overline{\rho}_0^2$ = adjusted rho squared with respect to the equal likelihood model and

$LL(0)$ = the log-likelihood for the model assigning equal likelihood to all alternatives.

Table 43 presents the results from the forward selection process. The row labeled "None" presents the results for the equal-likelihood model, and each following row give the variables in the order they were added to the model. The results are presented graphically in Figure 22 where the points corresponding to 0 variables are for the equal-likelihood model. As expected, each of the criteria improved significantly from the equal-likelihood model to the model with 1 variable. For the models with one or more variables, the rate of improvement of the criteria slowed after the model with 3 variables. Forward selection stopped at seven variables, because the rate of improvement of the criteria continued to slow and because the most important variables had been selected.

Table 43. Forward Selection Results for Model II.

Variable	AIC	BIC	$\overline{\rho}_0^2$	*RMSE*	P
None	247.1	247.1		9.64	
$PINT1_{ij}$	172.1	172.1	0.304	8.60	$< 10^{-100}$
$TURNS_{ij}$	159.8	159.9	0.353	8.32	$< 10^{-100}$
$PADTAL_{ij}$	150.8	150.9	0.390	7.91	$< 10^{-100}$
$RAMPS_{ij}$	148.3	148.4	0.400	7.89	$< 10^{-100}$
$ROUTE_{ij}$	146.3	146.4	0.408	7.76	10^{-94}
$PINT_{ij}$	144.6	144.7	0.415	7.61	10^{-81}
$DISTRO_{ij}$	144.1	144.2	0.417	7.57	10^{-24}

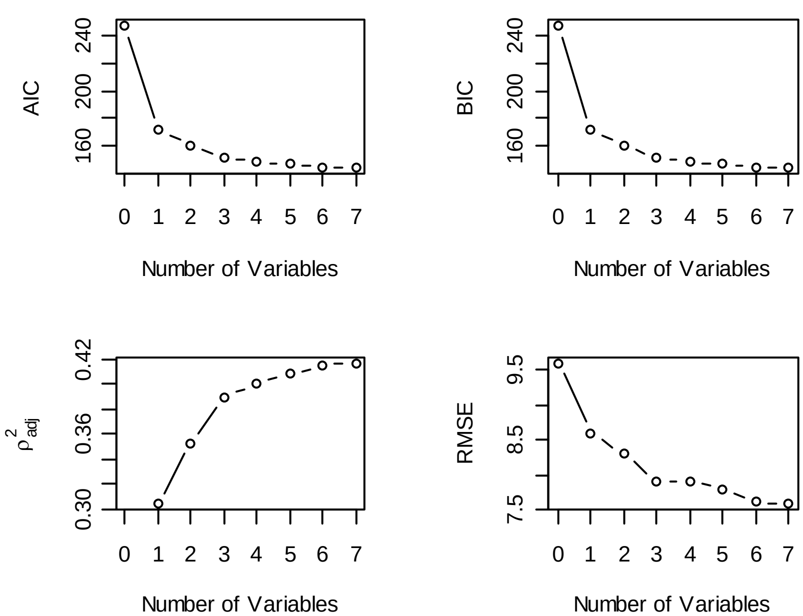

Figure 22. Forward Selection Results for Model II.

The results from the forward selection process suggest that the best model is the three-variable model. However, to allow for the possibility that the best model actually includes more variables, all seven variables were retained at this stage of the model development process. These seven variables form the preliminary model, Model II-a, which is presented in Table 44.

Table 44. Model II-a.

Variable	Coeff.	Std. Err.	z	P
$PINT1_{ij}$	1.87	0.0491	38.1	<0.0001
$TURNS_{ij}$	-0.361	0.0117	-30.8	<0.0001
$PADTAL_{ij}$	8.03	0.188	42.7	<0.0001
$RAMPS_{ij}$	-0.269	0.0139	-19.3	<0.0001
$ROUTE_{ij}$	0.884	0.0466	19.0	<0.0001
$PINT_{ij}$	0.941	0.0538	17.5	<0.0001
$DISTRO_{ij}$	1.45	0.145	9.99	<0.0001

Final Variable Selection

To further investigate each variable in the preliminary model, the model development process used the experiment described in the Model I development section with the same set of study area groups. Table 45 presents the results from this experiment.

The changes in parameter estimates for $PINT_{ij}$, $TURNS_{ij}$, $PADTAL_{ij}$, and $ROUTE_{ij}$, are equal to or less than 22 percent, whereas the other three parameters change by at least 29 percent at least once. Thus the variables $PINT1_{ij}$, $TURNS_{ij}$, $PADTAL_{ij}$, and $ROUTE_{ij}$ are retained as the better variables. These variables form Model II-b, which is presented in Table 46.

Table 45. Relative Change in Parameter Estimates after Removing Study Area Groups.

Variable	Group Removed			
	1	2	3	4
$PINT1_{ij}$	-1	9	-19	5
$TURNS_{ij}$	-11	4	22	-6
$PADTAL_{ij}$	11	-4	-6	0
$RAMPS_{ij}$	-29	20	2	6
$ROUTE_{ij}$	-10	1	16	-3
$PINT_{ij}$	21	-31	28	-10
$DISTRO_{ij}$	-31	56	-25	-6

Table 46. Model II-b.

Variable	Coeff.	Std. Err.	z	P
$PINT1_{ij}$	2.60	0.0385	67.6	<0.0001
$TURNS_{ij}$	-0.479	0.0115	-41.7	<0.0001
$PADTAL_{ij}$	7.72	0.172	44.9	<0.0001
$ROUTE_{ij}$	0.869	0.0471	18.4	<0.0001

Model Refinement

The model development process now tests the assumptions that the continuous variables are linear in the logit. This assumption is tested using the step test as described in previous sections. Figure 23 presents the results of the step test for variables $TURNS_{ij}$ and $PADTAL_{ij}$. Each graph provides evidence that the variable is not linear in the logit. Based on these graphs, two transformed variables were created to replace $TURNS_{ij}$ and $PADTAL_{ij}$, as defined in Equations 23 and 24. Figure 23 also presents the results of the step test for the new transformed variables. The new plots do not show evidence that the new transformed variables were not linear in the logit, so the new transformed variables were retained to replace the original variables. The model with the transformed variables is Model II-c, and it is the final model. Table 47 presents the final model.

$$TURN02_{ij} = \sqrt{TURNS_{ij}} \quad \text{(Eq. 23)}$$

$$PADTLG_{ij} = \ln(PADTAL_{ij}) \quad \text{(Eq. 24)}$$

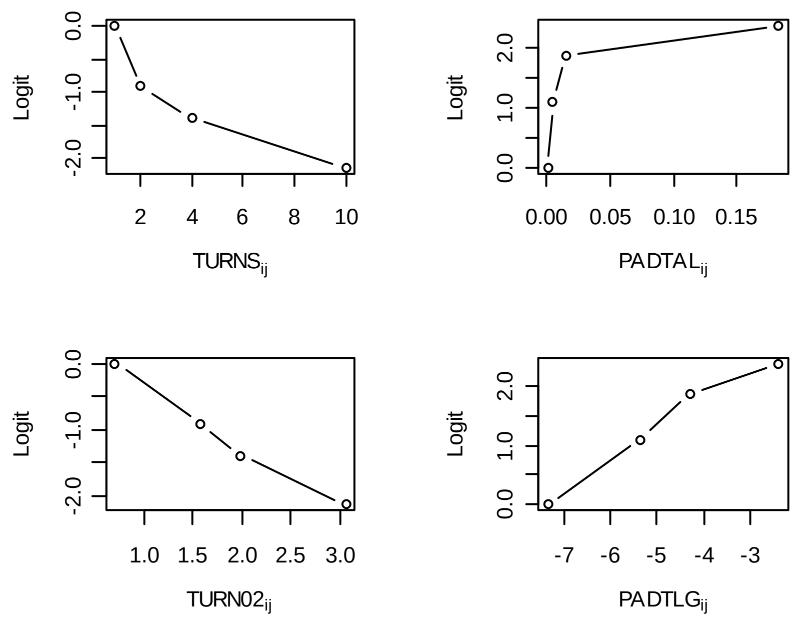

Figure 23. Step Rest Results for Model II.

Table 47. Model II-c.

Variable	Coeff.	Std. Err.	z	P
$PINT1_{ij}$	2.24	0.0392	57.2	<0.0001
$TURN02_{ij}$	-1.09	0.0234	-46.5	<0.0001
$PADTLG_{ij}$	0.572	0.0119	48.0	<0.0001
$ROUTE_{ij}$	0.814	0.0477	17.1	<0.0001

Model Evaluation

The final step of the model development process was to evaluate the model. As for Model I, the evaluation used the four study area groups defined in Table 34 as test data sets, and the remaining data in each case as the model building dataset. The evaluation is made using cross-classification tables. Table 48 through Table 51 present the cross-classification tables, which classify each observation by its observed and predicted distributions.

Table 48. Cross Classification of Observations for Group 1.

Observed	Predicted (percent)							
(percent)	0–1	1–5	6–10	11–20	21–40	41–60	61–80	81–100
0–1	950	321	47	21	13	0	0	0
1–5	24	26	7	4	6	0	0	0
6–10	10	26	14	9	6	0	0	0
11–20	10	21	9	9	5	0	0	0
21–40	4	7	15	6	9	1	1	0
41–60	3	2	4	2	4	3	1	0
61–80	0	1	2	0	5	1	1	0
81–100	0	2	0	0	1	2	1	0

Table 49. Cross Classification of Observations for Group 2.

Observed	Predicted (percent)							
(percent)	0–1	1–5	6–10	11–20	21–40	41–60	61–80	81–100
0–1	739	337	51	34	10	1	0	0
1–5	18	26	6	1	5	1	0	0
6–10	8	10	12	9	3	2	0	0
11–20	4	24	13	9	8	3	0	0
21–40	3	15	11	6	5	1	2	0
41–60	1	3	3	7	5	2	0	0
61–80	0	1	0	5	0	1	0	0
81–100	0	1	0	0	1	3	0	0

Table 50. Cross Classification of Observations for Group 3.

Observed	Predicted (percent)							
(percent)	0–1	1–5	6–10	11–20	21–40	41–60	61–80	81–100
0–1	3034	810	113	61	13	0	0	0
1–5	59	53	8	9	5	0	0	0
6–10	23	40	14	6	9	0	0	0
11–20	37	29	24	13	8	2	0	0
21–40	17	20	17	15	5	2	0	0
41–60	3	8	4	8	6	4	0	0
61–80	0	3	4	4	5	2	1	0
81–100	0	0	0	0	1	1	2	0

Table 51. Cross Classification of Observations for Group 4.

Observed	Predicted (percent)							
(percent)	0–1	1–5	6–10	11–20	21–40	41–60	61–80	81–100
0–1	440	260	53	18	7	0	0	0
1–5	10	13	3	7	2	0	0	0
6–10	3	11	3	6	2	1	0	0
11–20	4	15	11	8	6	1	0	0
21–40	0	4	13	5	4	2	0	0
41–60	0	2	4	6	6	3	0	0
61–80	0	0	0	0	3	2	0	0
81–100	0	0	0	1	1	0	4	1

MODEL RESULTS AND SAMPLE APPLICATIONS

The previous sections explained how Model I and Model II were developed. This section summarizes the final models and then demonstrates how the results from the models can be used to create commercial and non-commercial through trip tables for study areas in Texas.

Table 52 and Table 53 present the final models. Model I has two equations: one for commercial vehicles and one for non-commercial vehicles. Each of these two equations is applied once to each external station. Model II has one equation for both vehicle types. It is applied once in each direction for every possible pair of external stations in the urban area.

Table 52. Final Model I Equations.

$$\hat{\pi}_{split,com,j} = \frac{\exp(V_{thru,com,j})}{1+\exp(V_{thru,com,j})}$$

$$\hat{\pi}_{split,non,j} = \frac{\exp(V_{thru,non,j})}{1+\exp(V_{thru,non,j})}$$

$$V_{split,com,j} = -1.7816 + 2.3762 \times PINTTH_j - 0.2349 \times INTTL1_j$$

$$V_{split,non,j} = -2.9375 + 2.3762 \times PINTTH_j - 0.2349 \times INTTL1_j$$

$\hat{\pi}_{split,\,com,\,j}$	=	the proportion of all commercial vehicle trips at external station *j* that are through trips
$\hat{\pi}_{split,non,j}$	=	the proportion of all non-commercial vehicle trips at external station *j* that are through trips
$PINTTH_j$	=	the proportion through trips at external station *j* as estimated by the interaction score
$INTTL1_j$	=	1 if the total interaction score at external station *j* is greater than zero, and is equal to 0 otherwise.

The results from Models I and II can be used to develop commercial and non-commercial through trip tables using Equations 25 and 26, which are commonly used to develop through trip tables and were not developed as part of this research. Once the through trip tables are created, the total number of through trips in each direction at each external station can be calculated by summing each row and each column. The number of E-I/I-E trips can then be calculated by subtracting the total number of through trips from the total average daily traffic at the external station.

$$t_{com,ij} = (1/4)\left(p_{split,com,j} \times p_{dist,com,ij} \times ADTLV_j + p_{split,com,i} \times p_{dist,com,ji} \times ADTLV_i\right) \quad \text{(Eq. 25)}$$

$$t_{non,ij} = (1/4)\left(p_{split,non,j} \times p_{dist,non,ij} \times ADTSV_j + p_{split,non,i} \times p_{dist,non,ji} \times ADTSV_i\right) \quad \text{(Eq. 26)}$$

Table 53. Final Model II Equations.

$$\hat{\pi}_{dist,com,ij} = \hat{\pi}_{dist,non,ij} = \frac{\exp\left(V_{dist,ij}\right)}{\sum_{\{q \in E, q \neq j\}} \exp\left(V_{dist,qj}\right)}$$

$$V_{dist,ij} = 2.2 \times PINT1_{ij} - 1.1 \times \sqrt{TURNS_{ij}} + 0.57 \times \ln(PADT_{ij}) + 0.81 \times ROUTE_{ij}$$

i	=	an index variable referring to external stations where through trips enter the study area
j	=	an index variable referring to external stations where through trips exit the study area
p_{ij}	=	the proportion of through trips exiting the study area at external station j that entered the study area at external station i
q	=	the index variable for the summation
$PINT1_{ij}$	=	1 if the proportion through trips entering at i as estimated by the interaction score is greater than one, and is equal to 0 otherwise
$TURNS_{ij}$	=	the number of turns on the least time route between external stations i and j
$PADT_{ij}$	=	the ADT at external station i as a proportion of the total ADT across all external stations, excluding external station j
$ROUTE_{ij}$	=	1 if the least time route between external station i and j is valid

Selected Results

The previous sections detailed the methodology for developing external through trip tables for specific study areas. After developing external through trip estimates, it is necessary to gauge the results. A means to do that is to compare the model results with results from actual surveys that have been performed. Table 54 through Table 59 provide examples of the through trip tables that were developed using the results from Model I and Model II. These through trip tables are the commercial and non-commercial tables for the Wichita Falls, Amarillo, and Austin study areas. Based on these through trip tables, the total numbers of commercial and non-commercial through and local vehicle trips were also calculated for each study area. These results are presented in Table 60 through Table 62, alongside the results from the estimations based on the external surveys. These tables allow the results from Model I and II to be compared to the results from the external surveys.

Table 54. Wichita Falls Commercial Vehicle Through Trip Table.

	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518
500	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
501	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0
502	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
503	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
504	0	0	0	0	0	1	1	18	12	0	32	0	1	0	0	8	0	0	4
505	0	0	0	0	1	0	2	1	1	0	1	0	0	0	0	1	0	0	1
506	0	1	0	0	1	2	0	55	39	1	89	1	2	0	0	181	0	1	108
507	0	0	0	0	18	1	55	0	1	0	5	0	1	0	0	97	0	0	54
508	0	0	0	0	12	1	39	1	0	0	0	0	0	0	0	8	0	0	38
509	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	2	0	0	1
510	0	0	0	0	32	1	89	5	0	1	0	0	1	0	0	4	0	0	85
511	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	1	0	0	1
512	0	0	0	0	1	0	2	1	0	0	1	1	0	0	0	1	0	0	2
513	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
514	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
515	0	1	1	0	8	1	181	97	8	2	4	1	1	1	0	0	0	1	199
516	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
517	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0
518	0	0	1	0	4	1	108	54	38	1	85	1	2	0	0	199	0	0	0

Table 55. Wichita Falls Non-Commercial Vehicle Through Trip Table.

	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518
500	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1
501	0	0	0	0	0	0	2	1	1	0	1	0	0	0	0	1	0	0	1
502	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1
503	0	0	0	0	0	0	1	1	1	0	1	0	0	0	0	1	0	0	1
504	0	0	0	0	0	1	3	58	44	1	88	1	2	0	0	12	0	0	13
505	0	0	0	0	1	0	3	1	1	0	2	0	0	0	0	2	0	0	3
506	1	2	1	1	3	3	0	176	137	3	249	1	4	1	1	280	0	1	348
507	0	1	0	1	58	1	176	0	2	1	13	0	2	0	1	148	0	1	175
508	0	1	0	1	44	1	137	2	0	1	1	0	1	0	0	13	0	1	136
509	0	0	0	0	1	0	3	1	1	0	3	0	1	0	1	5	0	0	5
510	0	1	0	1	88	2	249	13	1	3	0	1	2	0	0	7	0	1	243
511	0	0	0	0	1	0	1	0	0	0	1	0	2	1	1	2	0	0	1
512	0	0	0	0	2	0	4	2	1	1	2	2	0	1	1	3	0	0	4
513	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	1	0	0	2
514	0	0	0	0	0	0	1	1	0	1	0	1	1	0	0	1	0	0	1
515	1	1	1	1	12	2	280	148	13	5	7	2	3	1	1	0	0	1	362
516	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
517	0	0	0	0	0	0	1	1	1	0	1	0	0	0	0	1	0	0	0
518	1	1	1	1	13	3	348	175	136	5	243	1	4	2	1	362	0	0	0

Table 56. Amarillo Commercial Vehicle Through Trip Table.

	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450
427	0	0	0	0	0	0	3	0	82	65	0	0	0	117	1	1	72	1	0	0	5	1	0	0
428	0	0	0	2	0	0	1	0	0	0	0	0	0	19	0	0	9	0	0	0	36	1	0	0
429	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
430	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
431	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
432	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
433	3	1	0	0	0	0	0	4	4	3	0	0	0	23	0	0	15	0	0	0	46	2	0	0
434	0	0	0	0	0	0	4	0	1	0	0	0	0	1	0	0	0	0	0	0	2	0	0	0
435	82	0	0	0	0	0	4	1	0	4	1	0	0	13	1	1	71	1	0	0	287	3	0	0
436	65	0	0	0	0	0	3	0	4	0	0	0	0	10	1	1	56	1	0	0	225	3	0	0
437	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0
438	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
439	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
440	117	19	0	0	0	0	23	1	13	10	1	0	0	0	1	0	1	0	0	0	50	4	0	0
441	1	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	1	0	0	0	1	0	0	0
442	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
443	72	9	0	0	0	0	15	0	71	56	0	0	0	1	1	0	0	0	0	0	2	0	0	0
444	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
445	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
446	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
447	5	36	0	1	0	0	46	2	287	225	1	1	0	50	1	0	2	0	0	0	0	4	1	0
448	1	1	0	0	0	0	2	0	3	3	0	0	0	4	0	0	0	0	0	0	4	0	0	0
449	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
450	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 57. Amarillo Non-Commercial Vehicle Through Trip Table.

	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450
427	0	1	0	0	1	1	4	0	44	30	1	0	0	97	2	1	67	1	0	0	3	1	0	0
428	1	0	0	5	0	0	2	0	0	0	1	0	0	28	1	1	20	0	0	0	33	2	0	0
429	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
430	0	5	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0
431	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	0	0
432	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
433	4	2	0	1	1	1	0	14	7	5	1	0	0	56	1	1	45	1	0	0	72	3	1	0
434	0	0	0	0	0	0	14	0	1	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0
435	44	0	0	0	0	0	7	1	0	2	1	0	0	9	2	1	56	1	0	0	148	4	1	0
436	30	0	0	0	0	0	5	0	2	0	0	0	0	7	2	1	40	1	0	0	111	3	0	0
437	1	1	0	0	0	0	1	0	1	0	0	0	0	2	0	0	1	0	0	0	2	1	0	0
438	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
439	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
440	97	28	0	1	1	0	56	1	9	7	2	0	0	0	2	0	1	0	0	0	32	5	1	0
441	2	1	0	0	0	0	1	0	2	2	0	0	0	2	0	0	3	0	0	0	1	0	0	0
442	1	1	0	0	0	0	1	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0
443	67	20	0	1	1	0	45	1	56	40	1	0	0	1	3	1	0	0	0	0	1	0	0	0
444	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
445	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
446	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
447	3	33	0	1	1	0	72	1	148	111	2	1	0	32	1	0	1	0	0	0	0	3	1	0
448	1	2	0	0	1	1	3	0	4	3	1	0	0	5	0	0	0	0	0	0	3	0	0	0
449	0	0	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0
450	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 58. Austin Commercial Vehicle Through Trip Table.

	1324	1325	1326	1327	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1346	1347	1348	1349	1350	1352	1353	1354	1355	1357	1358	1359	1360	1362
1324	0	0	4	0	0	1	0	2	0	0	1	0	1	1	0	0	2	1	0	0	7	1	2	1	1	1	0	0	1	1	9	1	0	0
1325	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0
1326	4	1	0	1	2	3	2	18	1	1	3	2	5	8	1	3	138	5	0	2	30	51	9	5	3	5	0	2	3	68	782	5	2	0
1327	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	2	0	0	0
1329	0	0	2	0	0	0	1	2	0	0	1	0	1	2	0	0	1	0	0	0	4	1	1	1	0	1	0	0	0	1	5	1	0	0
1330	1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0
1331	0	0	2	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	2	0	0	0
1332	2	0	18	0	2	0	1	0	0	0	1	0	1	2	0	0	1	0	0	0	2	1	2	1	0	1	0	0	1	10	59	1	0	0
1333	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
1334	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
1335	1	0	3	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	2	0	0	0
1336	0	0	2	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	4	1	0	0
1337	1	0	5	0	1	0	0	1	0	0	0	1	0	2	0	0	1	3	0	0	2	0	1	4	0	1	0	0	0	5	23	0	0	0
1338	1	0	8	0	2	0	1	2	0	0	1	1	2	0	1	0	2	1	0	0	6	0	4	0	0	1	0	0	0	1	24	0	0	0
1339	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0
1340	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0
1341	2	0	138	0	1	0	0	1	0	0	0	0	1	2	1	1	0	1	0	0	3	0	2	0	0	2	0	0	1	1	4	1	1	0
1342	1	0	5	0	0	0	0	0	0	0	0	0	3	1	0	0	1	0	0	0	2	0	2	0	0	0	0	0	0	0	2	0	0	0
1343	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
1344	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	2	0	0	0	0	0	0	0	1	0	0	0
1346	7	1	30	2	4	1	1	2	0	0	1	1	2	6	1	1	3	2	1	2	0	14	1308	5	3	5	1	3	2	8	24	3	6	2
1347	1	0	51	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	14	0	266	0	3	1	0	0	0	1	5	1	1	0
1348	2	0	9	1	1	1	0	2	0	0	1	1	1	4	1	1	2	2	1	2	1308	266	0	3	3	2	0	1	2	6	9	1	2	1
1349	1	0	5	0	1	0	0	1	0	0	0	0	4	0	0	0	0	0	0	0	5	0	3	0	1	1	0	0	0	1	4	1	1	0
1350	1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3	3	1	0	0	0	0	0	0	1	0	0	0
1352	1	0	5	0	1	0	0	1	0	0	0	0	1	1	0	0	2	0	0	0	5	1	2	1	0	0	0	1	1	1	6	1	1	0
1353	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
1354	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	1	0	0	1	0	0	1	0	2	0	0	0
1355	1	0	3	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	2	0	2	0	0	1	0	1	0	1	4	1	0	0
1357	1	0	68	0	1	0	0	10	0	0	0	1	5	1	0	0	1	0	0	0	8	1	6	1	0	1	0	0	1	0	6	1	1	0
1358	9	1	782	2	5	1	2	59	1	1	2	4	23	24	1	1	4	2	0	1	24	5	9	4	1	6	0	2	4	6	0	3	3	1
1359	1	0	5	0	1	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	3	1	1	1	0	1	0	0	1	1	3	0	1	1
1360	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	6	1	2	1	0	1	0	0	0	1	3	1	0	0
1362	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	1	1	0	0
1363	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
1364	1	0	48	0	1	0	0	7	0	0	0	0	9	7	0	0	14	1	0	0	9	1	3	1	0	0	0	0	0	0	3	1	0	0
1367	0	0	14	0	1	0	0	2	0	0	0	1	11	11	0	1	46	2	0	1	9	39	3	3	1	1	0	0	1	13	7	1	0	0
1368	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	5	1	1	1	0	0	0	0	0	1	4	1	0	0
1369	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	1	0	0	0
1370	0	0	60	0	1	0	7	16	0	0	1	1	1	18	0	0	18	1	0	0	11	14	3	2	1	1	0	0	1	12	81	1	0	0
1371	1	0	1	0	0	0	0	3	0	0	0	0	1	26	0	0	25	1	0	0	11	18	3	2	1	1	0	0	0	8	132	1	0	0
1372	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	2	0	0	0

Table 59. Austin Non-Commercial Vehicle Through Trip Table.

	1324	1325	1326	1327	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1346	1347	1348	1349	1350	1352	1353	1354	1355	1357	1358	1359	1360	1362
1324	0	1	12	1	1	2	1	5	0	1	2	1	3	5	1	1	5	2	0	1	13	3	4	3	2	2	0	1	4	5	22	2	1	0
1325	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	2	0	0	0
1326	12	2	0	1	3	7	5	25	1	2	7	3	6	11	1	4	202	7	1	3	43	84	14	8	4	5	1	4	11	133	1151	7	4	1
1327	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	1	1	3	0	0	0
1329	1	0	3	0	0	1	3	4	0	0	1	0	3	6	1	0	2	1	0	1	6	1	2	1	1	1	0	1	2	2	10	1	1	0
1330	2	0	7	1	1	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	2	0	0	0
1331	1	0	5	0	3	0	0	2	0	0	1	0	1	2	0	0	1	0	0	0	2	1	1	1	0	1	0	0	1	1	5	1	0	0
1332	5	0	25	0	4	1	2	0	0	0	2	0	1	3	0	0	1	0	0	0	3	1	2	1	1	1	0	1	2	21	91	1	1	0
1333	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
1334	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0
1335	2	0	7	0	1	0	1	2	0	0	0	1	1	2	0	0	1	0	0	0	1	1	2	0	0	0	0	0	1	1	4	1	0	0
1336	1	0	3	0	0	0	0	0	0	0	1	0	2	3	0	0	1	0	0	0	1	1	2	1	1	1	0	0	1	2	8	1	1	0
1337	3	0	6	0	3	1	1	1	0	0	1	2	0	3	0	0	1	5	0	1	3	0	2	7	0	1	0	0	1	10	36	1	1	0
1338	5	0	11	0	6	1	2	3	0	0	2	3	3	0	1	1	4	2	0	1	9	1	7	1	1	1	0	1	1	1	41	1	1	0
1339	1	0	1	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	2	0	1	0	0	0	0	0	0	0	1	0	0	0
1340	1	0	4	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	0	2	0	2	0	0	0	0	0	0	0	1	0	0	0
1341	5	0	202	0	2	0	1	1	0	0	1	1	1	4	1	1	0	1	0	1	5	0	4	0	0	2	0	1	2	2	7	1	2	0
1342	2	0	7	0	1	0	0	0	0	0	0	0	5	2	0	1	1	0	0	0	3	0	3	0	0	0	0	0	0	1	2	0	1	0
1343	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
1344	1	0	3	0	1	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	4	0	3	0	0	0	0	0	0	1	1	0	0	0
1346	13	2	43	2	6	1	2	3	0	0	1	1	3	9	2	2	5	3	1	4	0	22	1951	7	5	7	2	5	5	14	35	5	9	4
1347	3	0	84	0	1	0	1	1	0	0	1	1	0	1	0	0	0	0	0	0	22	0	414	1	5	2	0	1	2	2	8	1	2	1
1348	4	1	14	1	2	1	1	2	0	0	2	2	2	7	1	2	4	3	1	3	1951	414	0	5	4	2	1	2	4	10	15	2	3	1
1349	3	0	8	0	1	0	1	1	0	0	0	1	7	1	0	0	0	0	0	0	7	1	5	0	1	1	0	1	1	1	7	1	2	1
1350	2	0	4	0	1	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	5	5	4	1	0	0	0	0	1	1	2	0	1	0
1352	2	0	5	0	1	0	1	1	0	0	0	1	1	1	0	0	2	0	0	0	7	2	2	1	0	0	0	1	2	1	7	1	2	1
1353	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	1	0	1	0	0	0
1354	1	0	4	0	1	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	5	1	2	1	0	1	0	0	4	1	5	1	1	1
1355	4	0	11	1	2	0	1	2	0	0	1	1	1	1	0	0	2	0	0	0	5	2	4	1	1	2	1	4	0	5	15	2	3	1
1357	5	0	133	1	2	0	1	21	0	0	1	2	10	1	0	0	2	1	0	1	14	2	10	1	1	1	0	1	5	0	13	2	3	1
1358	22	2	1151	3	10	2	5	91	1	2	4	8	36	41	1	1	7	2	0	1	35	8	15	7	2	7	1	5	15	13	0	4	5	2
1359	2	0	7	0	1	0	1	1	0	0	1	1	1	1	0	0	1	0	0	0	5	1	2	1	0	1	0	1	2	2	4	0	3	1
1360	1	0	4	0	1	0	0	1	0	0	0	1	1	1	0	0	2	1	0	0	9	2	3	2	1	2	0	1	3	3	5	3	0	1
1362	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	1	1	1	0	1	0	1	1	1	2	1	1	0
1363	2	0	5	0	1	0	0	1	0	0	0	1	1	1	0	0	2	1	0	0	2	0	1	0	0	0	0	0	0	0	1	0	0	0
1364	3	0	110	0	2	0	1	15	0	0	1	1	15	15	0	1	41	1	0	1	14	4	4	3	0	1	0	0	1	1	5	1	1	0
1367	1	0	17	1	2	0	1	2	0	0	1	1	15	17	1	1	60	2	0	1	13	51	4	4	2	1	0	1	2	21	9	1	1	0
1368	0	0	1	0	0	0	0	1	0	0	1	1	1	2	0	0	2	1	0	0	7	2	2	2	1	1	0	0	1	2	8	1	0	0
1369	1	0	2	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	3	1	1	0	0	0	0	0	1	1	3	0	0	0
1370	1	0	96	0	2	0	22	29	0	1	2	1	2	36	0	1	36	1	0	1	16	30	5	3	2	1	0	1	3	31	141	2	1	0
1371	2	0	1	0	0	0	1	6	0	0	1	1	2	47	0	0	52	2	0	1	17	42	5	5	1	1	0	1	3	25	223	2	1	0
1372	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	4	1	1	1	0	0	0	0	1	1	4	0	0	0

Table 60. Wichita Falls Comparison of Trip Estimates.

External Station	Commercial Vehicles				Non-Commercial Vehicles			
	Through Trips		Local Trips		Through Trips		Local Trips	
	Survey	Model	Survey	Model	Survey	Model	Survey	Model
500	0	0	25	25	0	6	258	252
501	0	4	46	42	8	14	416	410
502	0	4	55	51	0	6	248	242
503	0	0	53	53	0	12	399	387
504	51	154	207	104	122	446	4,312	3,988
505	7	16	172	163	4	26	825	803
506	1,141	962	1,062	1,241	2,078	2,424	16,047	15,701
507	139	464	328	3	169	1,160	6,369	5,378
508	22	198	198	22	281	678	4,602	4,205
509	0	10	121	111	3	42	1,195	1,156
510	242	436	562	368	271	1,224	6,541	5,588
511	0	10	98	88	49	20	557	586
512	12	18	191	185	76	46	1,272	1,302
513	0	2	32	30	4	12	423	415
514	0	2	52	50	0	14	372	358
515	956	1,012	2,371	2,315	1,757	1,682	8,121	8,196
516	0	0	21	21	0	0	94	94
517	0	4	91	87	0	10	376	366
518	493	988	732	237	615	2,594	16,234	14,255
Total	3,063	4,284	6,417	5,196	5,437	10,416	68,661	63,682

Table 61. Amarillo Comparison of Trip Estimates.

External Station	Commercial Vehicles				Non-Commercial Vehicles			
	Through Trips		Local Trips		Through Trips		Local Trips	
	Survey	Model	Survey	Model	Survey	Model	Survey	Model
427	721	696	1,615	1,640	530	508	5,704	5,726
428	100	136	317	281	162	188	4,663	4,637
429	0	0	3	3	0	0	87	87
430	0	6	9	3	23	18	366	371
431	0	0	12	12	0	12	553	541
432	0	0	26	26	0	6	208	202
433	86	202	764	648	178	432	9,174	8,920
434	0	16	59	43	0	36	332	296
435	1,595	938	3,386	4,043	1,160	554	5,825	6,431
436	419	738	3,169	2,850	296	404	3,838	3,730
437	0	6	70	64	35	20	611	626
438	0	2	13	11	0	2	129	127
439	0	0	28	28	0	0	126	126
440	668	480	1,648	1,836	638	486	5,966	6,118
441	130	12	23	141	52	28	1,068	1,092
442	0	6	65	59	0	12	319	307
443	257	454	698	501	253	476	5,726	5,503
444	0	6	51	45	0	8	259	251
445	0	0	4	4	0	0	11	11
446	0	0	45	45	0	0	143	143
447	1,231	1,324	3,609	3,516	1,056	822	5,900	6,134
448	16	36	250	230	80	48	1,293	1,325
449	0	2	11	9	0	8	263	255
450	0	0	6	6	0	0	85	85
Total	5,223	5,060	15,881	16,044	4,463	4,068	52,649	53,044

Table 62. Austin Comparison of Trip Estimates.

External Station	Commercial Vehicles				Non-Commercial Vehicles			
	Through Trips		Local Trips		Through Trips		Local Trips	
	Survey	Model	Survey	Model	Survey	Model	Survey	Model
1324	50	80	611	581	223	236	6,810	6,797
1325	0	6	101	95	0	16	550	534
1326	4,542	2,582	5,478	7,438	4,764	4,012	30,269	31,021
1327	8	12	112	108	43	24	421	440
1329	46	56	421	411	125	126	3,487	3,486
1330	0	14	131	117	61	36	1,308	1,333
1331	0	30	99	69	0	110	639	529
1332	507	268	801	1,040	203	450	5,131	4,884
1333	0	4	78	74	27	4	291	314
1334	0	4	79	75	49	12	413	450
1335	7	24	83	66	5	70	544	479
1336	19	28	188	179	2	74	1,617	1,545
1337	152	148	710	714	260	252	3,159	3,167
1338	164	242	2,147	2,069	350	462	11,908	11,796
1339	0	12	130	118	6	20	645	631
1340	11	16	59	54	0	32	382	350
1341	138	536	1,220	822	106	886	6,898	6,118
1342	59	44	216	231	26	74	814	766
1343	6	4	33	35	2	6	143	139
1344	22	16	109	115	38	42	901	897
1346	3,814	3,004	2,841	3,651	8,997	4,500	10,241	14,738
1347	313	838	663	138	548	1,370	4,951	4,129
1348	2,886	3,304	3,073	2,655	9,272	4,980	8,000	12,292
1349	74	76	538	536	59	140	3,051	2,970
1350	318	36	0	282	107	72	1,543	1,578
1352	0	70	712	642	5	94	1,957	1,868
1353	0	2	46	44	0	12	415	403
1354	0	20	119	99	0	70	1,394	1,324
1355	0	40	230	190	49	170	4,685	4,564
1357	173	296	965	842	156	634	9,266	8,788
1358	4,547	2,446	7,233	9,334	5,385	3,806	43,642	45,221
1359	0	58	606	548	0	94	2,500	2,406
1360	85	38	179	226	141	104	2,277	2,314
1362	0	10	92	82	3	32	795	766
1363	16	10	131	137	15	48	1,497	1,464
1364	129	218	393	304	124	502	5,964	5,586
1367	93	342	1,864	1,615	69	478	6,146	5,737
1368	0	34	188	154	23	80	1,698	1,641
1369	0	10	70	60	0	34	604	570
1370	149	504	1,177	822	460	942	6,817	6,335
1371	385	472	145	58	118	892	4,287	3,513
1372	0	10	86	76	17	34	555	538
Total	18,713	15,964	34,157	36,906	31,838	26,032	198,615	204,421

While the previous tables provide model results from three study areas, trip estimates for a total of 13 study areas were developed as part of this research. Table 63 provides the total number of through and local trips that resulted from the modeling process as compared to the survey process for both commercial and non-commercial vehicles.

Table 63. Comparison of Trip Estimates for all Study Areas Reviewed.

Study Area	Commercial Vehicles				Non-Commercial Vehicles			
	Through Trips		Local Trips		Through Trips		Local Trips	
	Survey	Model	Survey	Model	Survey	Model	Survey	Model
Abilene	7,569	5,784	6,150	10,062	5,539	9,072	60,360	53,996
Amarillo	5,223	5,060	15,881	16,044	4,463	4,068	52,649	53,044
Austin	18,713	15,964	34,157	36,906	31,838	26,032	198,615	204,421
Dallas/Fort Worth	20,609	15,864	70,272	75,017	19,139	21,404	318,286	316,021
Longview	13,944	18,544	34,573	29,313	21,650	23,980	123,291	120,961
Lubbock	2,001	1,824	9,236	9,962	2,240	3,308	59,367	57,690
Midland/ Odessa	2,555	3,208	11,421	11,740	956	2,452	36,540	35,044
San Angelo	1,497	1,664	5,805	5,635	2,908	3,732	39,055	38,231
San Antonio	9,710	9,816	43,573	43,962	23,570	14,940	211,661	214,558
Sherman/ Denison	4,727	8,212	16,645	13,160	15,423	13,616	81,809	83,616
Tyler	10,095	13,596	27,667	24,166	29,752	20,020	109,029	118,761
Waco	11,833	14,896	21,730	18,809	35,273	26,336	104,838	114,944
Wichita Falls	3,063	4,284	6,417	5,196	5,437	10,416	68,661	63,682

The primary objective of the logit model development was to devise a methodology for estimate external through trips in lieu of the conduct of an external survey. After estimating the total number of through trips for each of the study areas, the results were compared to those obtained in the external survey. Table 64 provides the total number of modeled and surveyed through trips for each of the 13 study areas reviewed. Additionally, the table provides the percent difference between the survey and model results, for commercial and non-commercial vehicles combined.

Table 64. Percent Difference for Modeled Through Trips.

Study Area	Commercial		Non-Commercial		Total (com + non-com)		% Diff
	Survey	Model	Survey	Model	Survey	Model	
Abilene	7,569	5,784	5,539	9,072	13,108	14,856	13.34
Amarillo	5,223	5,060	4,463	4,068	9,686	9,128	-5.76
Austin	18,713	15,964	31,838	26,032	50,551	41,996	-16.92
Dallas/Fort Worth	20,609	15,864	19,139	21,404	39,748	37,268	-6.24
Longview	13,944	18,544	21,650	23,980	35,594	42,524	19.47
Lubbock	2,001	1,824	2,240	3,308	4,241	5,132	21.01
Midland/ Odessa	2,555	3,208	956	2,452	3,511	5,660	61.21
San Angelo	1,497	1,664	2,908	3,732	4,405	5,396	22.50
San Antonio	9,710	9,816	23,570	14,940	33,280	24,756	-25.61
Sherman/ Denison	4,727	8,212	15,423	13,616	20,150	21,828	8.33
Tyler	10,095	13,596	29,752	20,020	39,847	33,616	-15.64
Waco	11,833	14,896	35,273	26,336	47,106	41,232	-12.47
Wichita Falls	3,063	4,284	5,437	10,416	8,500	14,700	72.94
Total	111,539	118,716	198,188	179,376	309,727	298,092	-3.76

As shown in Table 64, eight of the 13 study areas had a percent difference between the model and survey results of less than ± 20 percent. Only two study areas, Midland/Odessa and Wichita Falls, had substantially more modeled through trips than surveyed through trips. When looking at all of the study areas in aggregate, the model underestimated the number of through trips by approximately 4 percent.

SUMMARY

This chapter has described a system of two logit models used to estimate urban area external travel in Texas. Overall, the logit models produce reasonable estimates of external travel that can be useful for travel demand modeling. That the logit model estimation method does not always closely replicate the external survey results does not necessarily indicate that the logit model estimates are not accurate. Given that both the external survey results and the logit model results are simply estimates of the true travel patterns, an unreasonably high correlation between the results from the two estimation methods should not be expected. Additionally, part of the difference between the external survey results and the logit model results may arise from the logit models actually producing better estimates then the surveys, which may be possible since the logit model is based on a much larger sample size than any one set of external surveys.

Even if the accuracy of the logit model results could be improved, the improvement may not be necessary, since the current logit model estimates may be “close enough” for most practical purposes. For example, for an estimation error to actually make a difference in a transportation investment decision, the error may have to change the estimated value by many times its true value, and by more than a few hundred or thousand trips. In this situation, the logit model results are probably close enough to produce the same investment decision as if the “true” volumes were known.

In addition to producing reasonable and useful estimates of external surveys, the model estimation method is very easy and cheap to implement. The only field data collection that is required is to count and classify the vehicles at the external stations. Otherwise, all the variables depend on data that can be obtained from maps of the study area and surrounding region, and from data provided by the U.S. Census Bureau. In addition, preparing the data and performing the calculations are fairly straightforward. The complete process of data collection, preparation, and calculation should take no more than two or three person-working-days, so the overall costs of this method are less than one hundredth of the cost of a set of external surveys.

A disadvantage of the logit model estimation method is that it does not provide all the data that would be collected from an external survey. Particularly, it does not produce information on the distribution of external-local trips to zones within the study area. This information would need to be developed using a different method. Additionally, the method does not provide information related to non-resident trips that are made within the study area.

CHAPTER 7. SUMMARY, RECOMMENDATIONS, AND CONCLUSIONS

The primary focus of the research was to evaluate various methods for developing external related survey data in lieu of conducting external travel surveys. Through the research process, three primary methods for developing external travel data were identified and evaluated.

SUMMARY OF ALTERNATIVE METHODS

The preceding chapters detailed various methods tested for their ability to replicate results produced by external surveys. Table 65 provides a summary of the three methods reviewed as compared to the current method and illustrates the model inputs that each method has the capability to produce. However, it is important to understand that although a particular method has the capability to produce a specific model input, it does not mean that the method can generate the input with a level of accuracy that would make it usable in a travel demand model. The following sections provide a summary of the methods and the model inputs that they can produce.

Table 65. Overview of Current and Alternative Methods.

Model Inputs	External Survey	Household, Workplace, and Commercial Vehicle Surveys	SAM	Logit
Through/local split	✓	✓	✓	✓
Commercial/ non-commercial split	✓	✓	✓	✓
Non-resident travel	✓	I-E/E-I only		
Average trip length	E-E and I-E/E-I	I-E/E-I	E-E and I-E/E-I	E-E
Through trip table (expanded)	✓		✓	✓
Local trip table (expanded)	✓	✓	✓	

Other Survey Types

The use of other survey types such as household, workplace, and commercial vehicle surveys appears on the surface to offer estimates of a majority of the model inputs. However, given the limited number of external-related trips that are obtained from these survey types, the statistical validity of these estimates would be suspect. Current survey sample sizes would need to be increased two to three times in order to develop reasonable estimates of external-related traffic. Additionally, the estimates derived from this method would be limited to external local trip types. While it is possible for these surveys to capture external through trip movements during the conduct of the survey, the observed frequency of this occurring is extremely rare.

EULA Application in SAM

Like the use of other survey types, the EULA application of the SAM appears to provide estimates of a majority of the model inputs. The primary input that it lacks the ability to estimate is the amount of non-resident travel within a study area. However, another significant obstacle is the base and forecast years of the model. In order to provide estimates for years in between these years, it would require analysts to extrapolate the data. This could degrade the accuracy of the estimates. Additionally, due to the differences between urban area and SAM zone structures, methods would need to be devised to develop average trip lengths for external local trips.

Logit Model

The implementation of the two logit models can produce estimates of all of the model inputs except non-resident trips and external local trip tables. Additionally, without the ability to produce external local trip tables, it is not able to produce average trip lengths for external local trips. Despite these limitations, the logit model approach appears to offer more robust estimates than either of the other two methods.

ADDITIONAL METHODS TO CONSIDER

In addition to the methods analyzed in this research, there are several other methods that researchers identified as potentially having some utility in the future. These methods are briefly reviewed in the sections below.

Internet-Based Travel Surveys

The use of internet-based travel surveys is not a new concept, and it is a method that TxDOT has implemented in the past. As the number of people in the United States that have internet access continues to grow, employing this method to obtain travel survey data is not an unrealistic expectation. However, as with most methods, there are some issues that would need to be addressed before implementing the method. A primary issue is the recruitment of eligible participants that is performed in a manner that has minimal sample bias. Additionally, the identification of persons that would be making external related trips on the survey day would be a difficult task to accomplish.

Postcard Surveys

Postcard surveys have been implemented in various parts of the United States with varying degrees of success. One primary negative associated with this method is the response rate. It is not uncommon for postcard surveys to have a 10 percent or less response rate. Developing a method to disseminate the postcards is also a significant consideration. If the postcards are distributed on the roadways, then traffic control plans must be developed in order to provide for safe conditions for motorists and survey personnel. Postcard surveys can be combined with internet-based surveys by including an internet website on the postcard, thus allowing survey recipients the opportunity to complete the survey online.

GPS Enabled Cellular Phone Data

Using GPS data from cellular phones to develop travel estimates is an area that is gaining increasing interest from planners and modelers in the United States. Given the large percentage of Americans that own cellular phones, this technology offers a lot of promise. However, privacy and legal issues related to the use of the data present a significant challenge to people interested in the data. Additionally, deciphering the raw data can be challenging for analysts. For example, the raw data do not provide information on the trip purpose, the type of place at a trip end, or the vehicle occupancy. There have been a number of successful programs in the United States that have recruited people in metropolitan areas for ongoing access to their cellular phone GPS data for the purpose of monitoring traffic congestion and travel times. In these instances, the recruitment process offers more of an incentive to the potential recruit in that there is a perceived immediate benefit.

RECOMMENDATIONS

Through the course of the research, it was revealed that there are a multitude of methods available for developing external travel estimates in urban areas in Texas. Each method has advantages and disadvantages.

CONCLUSIONS

External travel surveys provide several key components to the external distribution portion of the modeling process. Without the conduct of external surveys, there are essentially two approaches that can be taken. The first is to develop new methods to collect the same types of data. The second approach is to develop methods to synthetically derive the travel estimates.

While the methods reviewed to collect travel data using new means such as internet-based, postcard, and GPS-based surveys appear to have the capability to collect detailed travel information similar to that of the roadside survey, the primary difficulty was in the implementation of the method.

For the methods that involved synthetically deriving travel estimates, the methods are less intrusive and easier to implement, but they do not provide all of the key elements that a traditional external travel survey provides. Additionally, these types of methods also offer the advantage of being lower cost methods for developing external travel estimates.

REFERENCES

1. Griffiths, R; Richardson, AJ: Lee-Godsselin, MEH. "Travel Surveys." *Transportation in the New Millennium*. Transportation Research Board, A1D10, Committee on Travel Survey Methods, Washington, DC. 2000, pp. 1-7.

2. "Research Mode to Go By the Wayside." *Engineering News-Record*, 08919526, 8/12/2002, Vol. 249, Issue 7.

3. Hagen, L; Zhou, Huaguo; Pirinccioglu, Fatih. *Development of Revised Methodology for Collecting Origin-Destination Data*. Florida Department of Transportation Report BD 544-30. Center for Urban Transportation Research, Tampa, Florida. 2006.

4. Modlin, D.G. "Synthesized Through-Trip Table for Small Urban Areas". *Transportation Research Record 1685, Transportation Research Board*, Washington, D.C., 1982, pp. 16-21.

5. Bell, M.G.H. "The Estimation of an Origin-Destination Matrix from Traffic Counts". *Transportation Science*, 17(2), 1983, pp. 198-217.

6. Pigman, J.G. *Simulation of External-Internal and Through Trips for Small Urban Areas*. Research Report 493, Division of Research, Kentucky Bureau of Highways. 1978.

7. Arun, C. and Raja, M., "Synthetic Models for Through Trips in Small Urban Areas," Transportation Research Board Presentation. 1987.

8. Reeder, P.R. "Comparison of Synthetic Thru Trips to Recent External O&D Data." Paper delivered at the Transportation Research Board's Fourth National Conference on Transportation Planning Methods Applications, Daytona Beach, Florida. 1993.

9. Martin, W.A. and McGuckin, N.A Travel Estimation Techniques for Urban Planning. NCHRP Report 365. Transportation Research Board, Washington, D.C. 1998.

10. Horowitz, A. and Patel, M. "Through-Trip Tables for Small Urban Areas: A Methodology for Quick-Response Travel Forecasting." Transportation Research Record 1685, Transportation Research Board, Washington, D.C., 1999, pp. 57-64.

11. Madden, M. Internet Penetration and Impact. Pew Internet and American Life Project. http://www.pewinternet.org/~/media//Files/Reports/2006/PIP_Internet_Impact.pdf.. Accessed April 2009.

12. Kerr, R. US Mobile Internet Usage to Near 50% in 2013. Vator News. http://vator.tv/news/show/2009-08-26-us-mobile-internet-usage-to-near-50-in-2013. Accessed August 2009.

13. Anderson, M. Internet Based Travel Surveys. University Transportation Center for Alabama. 2008.

14. B.R. Ullman, C.L. Dudek, N.D. Trout, S.K. Schoeneman. Amber Alert, Disaster Response and Evacuation, Planned Special Events, Adverse Weather and Environmental Conditions, and Other Messages for Display on Dynamic Message Signs. Research Report 0-4023-4. Texas Transportation Institute, College Station, TX. October 2005.

15. C.L. Dudek, S.D. Schrock, B.R. Ullman. License Plate and Telephone Numbers in Changeable Message Sign Amber Alert Messages. In Transportation Research Record: Journal of the Transportation Research Board, No. 2012, Transportation Research Board of the National Academies, Washington, D.C. 2007, pp. 64-71.

16. North Caroline Department of Transportation, Transportation Planning Branch. US 64 – NC 49 Corridor Study. Postcard Survey Technical Report. May 2004.

17. Urbitran Associates. Sussex County Ten-Year Mobility Study; Chapter 3 – Origin/Destination Survey. http://www.sussex.nj.us/documents/planning/mobility/Mobility%20Study%20Chapter%203.pdf . Accessed August 2009.

18. CTIA – The Wireless Association. Wireless Quick Facts. http://www.ctia.org/media/industry_info/index.cfm/AID/10323 . Accessed August 2009.

19. University of California, Berkeley. http://traffic.berkeley.edu/theproject.html

20. Winters, P.L., Barbeau, S.J., and N.L. Georggi. Smart Phone Application to Influence Travel Behavior (TRAC-IT Phase 3). National Center for Transit Research, Center for Urban Transportation, University of South Florida. Tampa, FL. March 2008.

21. Anderson, M. D., Y. M. Abdullah, S. E. Gholston and S. L. Jones. Development of a Methodology to Predict Through-Trip Rates for Small Communities. ASCE Journal of Urban Planning and Development, Vol. 132, No. 2, 2006, pp. 112-114

22. Anderson, M. D. Evaluation of Models to Forecast External-External Trip Percentages. *ASCE Journal of Urban Planning and Development,* Vol. 125, No. 3, 1999, pp. 110–120.

23. Anderson, M. D. Spatial Economic Model for Forecasting the Percentage Splits of External Trips on Highways Approaching Small Communities. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1931*, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 68-73.

24. Han, Y., and J.R. Stone. Synthesized Through-Trip Models for Small and Medium Urban Areas. In *Transportation Research Record: Journal of the Transportation Research Board, No 2077*, Transportation Research Board of the National Academies, Washington, D.C. 2008, pp. 148 - 155.

25. Yang Han, John R. Stone, and Leta F. Hunsinger. Manual Traffic Allocation for Small-Town Networks. Transportation Research Record, No. 1997, 2007, p. 24-34.

26. Barton-Aschman Associates. Model Validation and Reasonableness Checking Manual. Federal Highway Administration, 1997. http://tmip.fhwa.dot.gov/clearinghouse/docs/mvrcm/. Accessed 5 September 2008.

27. E.N. Hard, S.P. Farnsworth, D.F. Pearson, P. Songchitruksa, D.L. Spillane, T.L. Forrest. Evaluation of External Station Survey Methodologies for High Volume Locations, Research Report 0-4869-1. Texas Transportation Institute, College Station, TX. May 2006.

28. Songchitruksa, P. and E. N. Hard. Queuing Simulation of Roadside Survey Station: Blocked Traffic Lane. *Transportation Research Part A -- Policy and Practice*, Vol. 42, No. 6, 2008, pp. 857-873.

29. PTV VISSIM 5.10 User Manual. PTV Vision, July 2008.

30. VAP 2.16 Interface User Manual, PTV AG, 2007.

31. Tian, Z., C. Messer, K. Balke and T. Urbanik. Integration of Diamond Interchange and Ramp Metering Operations. *Transportation Research Record*, No. 1925, 2005, pp. 101-111.

32. Balke, K. N., P. Songchitruksa, N. A. Chaudhary, C. Chu, S. Sunkari, P. Nelson, S. Kuchangi and V. Tyagi. Development of a Prototype Dynamic Congestion and Incident Prediction System. *Technical Report 0-4946-2*, Texas Transportation Institute, College Station, TX, March 2007.

33. Martchouk, M., and Fricker, J. D. "Through-Trip Matrices Using Discrete Choice Models: Planning Tool for Smaller Cities." Presented at the 88th Annual Meeting of the Transportation Research Board, Washington, D.C. 2009.

34. Hosmer, David W, and NetLibrary, Inc. *Applied Logistic Regression*. 2nd ed. Wiley series in probability and statistics. New York: Wiley. 2000.

35. Ben-Akiva, Moshe E. *Discrete Choice Analysis Theory and Application to Travel Demand*. MIT Press Series in Transportation Studies 9. Cambridge, Mass: MIT Press. 1985.

36. Koppelman, Frank S., and Chandra Bhat. *A Self Instructing Course in Mode Choice Modeling: Multinomial and Nested Logit Models*. January 31, 2006. http://www.civil.northwestern.edu/people/koppelman/PDFs/LM_Draft_060131Final-060630.pdf.