An Agent-based Freight Transportation Modeling Framework

BY

ZAHRA POURABDOLLAHI
B.S., Sharif University of Technology, 2002
M.S., Sharif University of Technology, 2010

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Civil and Materials Engineering
in the Graduate College of the
University of Illinois at Chicago, 2015

Chicago, Illinois

Defense Committee:

Dr. Abolfazl (Kouros) Mohammadian, Chair and Advisor
Dr. Kazuya Kawamura, Urban Planning and Policy
Dr. Jie (Jane) Lin
Dr. Sybil Derrible
Dr. Anthony Pagano, Managerial Studies

To Behzad

ACKNOWLEDGEMENTS

I am very grateful to my advisor, Dr. Abolfazl (Kouros) Mohammadian, for giving me this wonderful opportunity to pursue my research at UIC. I am very thankful to him for all his support, guidance, and patience throughout my PhD journey. His encouragement and confidence in me drove me on my way to the accomplishment of this study. I owe my deepest gratitude to my co-advisor Dr. Kazuya Kawamura for all his support and guidance during this study. I would also like to thank my thesis committee members Dr. Jie (Jane) Lin, Dr. Sybil Derrible, Dr. Anthony Pagano, and Dr. Bo Zou for their valuable advice and assistance.

I appreciate the National Center for Freight, Infrastructure, Research and Education (CFIRE) at the University of Wisconsin-Madison, and Illinois Department of Transportation for funding this study. Furthermore, I would like to thank Transportation Research Board for granting the permission to reproduce my papers, “Joint Model of Freight Mode and Shipment Size Choice,” and “Shipping Chain Choices in Long-Distance Supply Chains: Descriptive Analysis and Decision Tree Model”, published in the Transportation Research Record: Journal of The Transportation Research Board, in this dissertation.

With all of my heart, I would like to thank my parents, Sedigheh Iranmanesh and Mansour Pourabdollahi, who have been an incredible source of encouragement and inspiration throughout my life. I am forever grateful to them for their unconditional love and support throughout my life and for their patience and help during the last weeks of this project.

I am deeply thankful to my sisters Maryam, Mitra, and Fatemeh and my brother Mohammadamir who have always supported me and whatever I am doing. Their phenomenal emotional support made it much easier to pass this milestone.

Last, a very special thank you to Behzad Karimi for his emotional and practical support during the past five years. I have no suitable word to fully express my love and appreciation to him except that without him, this dissertation would never have been accomplished.

ZP

TABLE OF CONTENTS

CHAPTER **PAGE**

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The remarkable increase in freight movements and their significant impacts on transportation system, regional wellbeing, and economic growth provide sufficient motivation to develop reliable analysis tools to estimate commodity flows between zones and forecast the future demand and trends of goods movements among regions. While the need to develop freight demand model to better facilitate infrastructure planning and policy development has been clearly recognized for some time, the current state of knowledge and understanding regarding the movement of freight and behaviors of shippers and carriers lags behind those of passenger travel by a considerable margin.

Review of literature showed that only a handful of agent-based freight transportation models have been proposed and developed in the past studies, and as a result the role of key decision making agents in the freight system have been ignored in current freight models. It also emphasizes on the need to develop agent-based models that incorporate supply chain relationships and logistics components in their framework to better capture the decision-making procedures. Agent-based microsimulation logistics models can more precisely capture the complex interactions among various agents and markets in the freight system by simulating the behavior of these decision making agents. In addition, they can be integrated with micro-level models for passenger transportation to better estimate and forecast traffic volumes on networks. They would also provide a platform to test how various policy measures and major infrastructure investments can alleviate the negative impacts of freight transportation and how implementing efficient policies would make the freight transportation system more sustainable.

This study outlines a behavioral freight transportation modeling framework that will address some of critical technical and conceptual hurdles that have challenged past efforts by applying agent-based framework in which firm-level decision making processes, including supply chain formation and selecting logistics choices, are simulated. The study demonstrates the use of disaggregate, behavioral-based modeling approaches for forecasting freight movements at disaggregate firm-level and evaluating freight policy impacts at the national/regional scale.

The proposed agent-based modeling approach is unique in the focus on multiple aspects of individual firm behavior which leads to disaggregate commodity movements and ultimately to freight vehicle flows. The approach is analogous to the activity-based modeling approach to evaluating passenger travel demand, in that the freight goods and vehicle movements are modeled as derived demand arising from the needs and behaviors of individual firms. The proposed technology will be a state-of-the-art modeling system implementing advanced behavioral-based freight modeling concepts for looking at freight demand at the disaggregate firm level.

The cutting-edge supply chain and logistics choice models including a behavioral supplier selection model, a joint model of mode choice and shipment size and a decision tree clustering model for intermediate handling facility usage, are developed and incorporated in the proposed framework to enhance the precision of the model in forecasting individual shipments and their attributes. In addition, disaggregate three-dimensional crosswalks for identifying industry use-and-make shares of different commodity classes are developed to improve the accuracy of deriving industry-to-industry freight flows. Finally, the simulated truck freight flows are assigned over traffic network.

The developed agent-based microsimulation freight model provides a satisfactory analysis tool that can be used to better capture the complex interactions among decision making agents and markets in the freight system. It can be used to more precisely estimate current freight movements and forecast future commodity flows and alleviate the adverse impacts of these movements on socio-economic systems.

1. INTRODUCTION

Efficient freight movements are critical foundation of any nation’s economic strength. The United States’ 117 million household units, 7.4 million business establishments and 89,500 governmental sectors form a gigantic economy and production-consumption market that demands the effective movements of a wide variety of products (FHWA, 2011). Over the past three decades, the volume of transported goods within the United States has increased significantly and almost doubled the rate of population growth (NCHRP 606, 2008). According to the Federal Highway Administration estimates, more than 18.6 billion tons of commodities, worth more than 16 trillion dollars, were transported in the United States in 2007 (FHWA, 2011). This translates into 52 million tons of goods, worth 46 billion dollars, moved daily within U.S. transportation system.

Freight transportation has increasingly grown over time due to the population increase and expansion of economic activities in the U.S. According to the Census Bureau estimates, the U.S. population increased by 35.5 percent between 1980 and 2009 and reached to 308.7 million in April of 2010 (U.S. Census Bureau, 2011). Over the same period, the Gross Domestic Product (GDP) doubled in real terms (FHWA, 2011) and average household income grew by 30 percent based on 2012 dollar value (U.S. Census Bureau, 2011) that indicate the expansion in economic activities and economic growth in the U.S. Although the U.S. economy has been disturbed by an economic recession recently, Freight Analysis Framework (FAF) (FHWA, 2007) estimates show that commodity flows in the U.S. increased in 2010 after two consecutive declines in 2008 and

2009. It is expected that the economy will continue to grow and result in greater demand for freight transportation in long-term.

Additionally, international trade continued to evolve faster than overall economy, reaching the quintuple amount in real value between 1980 and 2009. This has resulted in significant international interconnectivity. In response to the significant level of global trade and demand growth, several projects are initiated to improve some of the critical freight infrastructure bottlenecks. The Panama Canal expansion project is an example of these improvements that will double the capacity of the canal by 2015, so more and larger ships are allowed to transit. Such improvements definitely influence the overall freight movements in the U.S.

On the other hand, freight movements have significant impacts on transportation system, environment and regional wellbeing. Many researchers and experts have widely accepted that freight transportation is a significant source of congestion, maintenance, security, and environmental costs that should be considered by transportation and urban planners. The facts and figures on freight movements and their potential effects on the quality of life at regional and national scale signify the need to incorporate freight movements in the transportation policy and planning framework. Moreover, it indicates the importance of providing decision makers with satisfactory analysis tools to assist them in taking a better picture of the current and future trends of freight movements in the region and addressing their adverse impacts. In response to this need, state departments of transportation and metropolitan planning organizations are required by federal regulations to consider freight movements in their transportation planning and travel demand studies (NCHRP 606, 2008).

Although the importance of freight transportation was recognized, developing comprehensive and reliable analysis and modeling tools for policy-making process has faced technical challenges including lack of common satisfactory modeling framework and freight data scarcity. Additionally, modern manufacturing techniques, accelerated changes in the supply chain management practices and technological advances have made logistics decision-making process more complicated than before. Therefore, sophisticated modeling tools are required for forecasting freight flows, policy assessment, and alleviating negative impacts of freight movements. It has been suggested by several researchers that developing a reliable model to forecast freight transportation flows is more complicated than development of passenger travel demand models (NCHRP 388, 1997). Is has been argued that this is mainly due to the complex nature of freight decision making process and the numerous actors and markets involved in the decision making procedure of shipping various commodity types by several modes available (Horowitz and Farmer, 1999). Thus, the development of freight transportation models, has lagged far behind the development of passenger travel demand models and unlike passenger transportation subject, freight transportation is rather a less-researched topic in terms of advanced modeling applications (Horowitz and Farmer, 1999).

This deficiency, however, has been recognized by researchers around the world and the field is developing in many directions, including advanced data collection methods, integrated modeling frameworks, and reliable operational strategies. This has led to many improvements in the development of freight transportation models used for forecasting. However, many of the models currently in use are unable to represent the implementation of newly developed and more advanced transportation mitigation strategies since they act only on aggregated flows and do not contain a behavioral component to simulate individual decision-making units' behaviors in freight transportation sector and their responses to policy changes. In addition, there has been a big loss of precision due to the aggregate nature of available practical models (Windisch et al., 2010). Moreover, many available freight transport models suffer from a lack of explicit logistics elements such as, configuration of exact supply chain, considering the use of intermediate handling facilities, determining optimum shipment size and frequency of shipments in their operational modeling framework. The current understanding of the fundamentals and nature of behavioral decision making process in freight system can be improved significantly. Therefore, there is a need for models with explicit logistics elements which represent the underlying dynamic logistics decision making process and their effects on the transportation system and communities.

This study outlines the development of an agent-based freight transportation model which incorporates the logistics decision-making process by explicitly treating each logistics choice made by individual decision making units in its framework. The study first reviews existing literature on freight transportation modeling and logistics models in Chapter 2. Next, it discusses some of the research gaps in current freight transportation studies that this work attempts to address in Chapter 3. The main objectives and contributions of this research are also explained in this chapter. In Chapter 4, the proposed framework for the model development and various steps within the framework are presented. After that, the data used for the model development are discussed and the online freight establishment survey conducted for this purpose and its results are presented in Chapter 5. Chapter 6 discusses the developed models and components of the framework in details. Results of simulation and validation are presented in Chapter 7. Finally, the thesis concludes by presenting a summary of the work, the major contribution and future directions of the study.

2. LITERATURE REVIEW

2.1. Statewide and Regional Freight Transportation Models

Current statewide freight models are reviewed and their components are discussed in several studies (NCHRP 606, 2008; NCHRP 358, 2006). Different classifications have been proposed for the current statewide freight transportation models. National cooperative highway research program (NCHRP 606, 2008) categorized statewide freight modeling approaches into five different classes ranging from the straightforward methods such as flow factoring, O-D factoring, and truck model, to the more advanced methods such as the four-step commodity-based model, and the economic activity model. This toolkit also discusses the components of each model and their differences. Finally, it provides a review of case studies in several states and developed freight models in each of them. Another paper, by FHWA, discusses freight transportation system and identifies influential factors on freight demand. It also provides information on available freight-related data and techniques used for developing freight forecast models. The study classifies freight transportation analysis methods, used in statewide studies, into four major groups including simple growth factor methods, incorporating freight into "four-step" travel forecasting, commodity models and hybrid approaches (FHWA, 2007).

However, in this section, the classification, proposed in two different studies for the statewide freight transportation models (Horowitz, 2008; Cambridge Systematics Inc, 2010), is used as the reference. These studies categorize forecast analysis methods of statewide freight transportation models into five classes (Horowitz, 2008; Cambridge Systematics Inc, 2010). Based on the proposed classification, freight models can be incorporated in the transportation plan of states using five analysis categories as follows.

- *No freight model:* In states where modeling may be restricted to MPO's and no model is in operation by the state DOT that include all of the MPO's in the state; or in states where the operational statewide model used by DOT only covers passenger trips or report total highway vehicle without distinguishing between commercial and passenger vehicles.
- *Truck model:* A travel demand model is used by state which can separately forecast the passenger vehicles and trucks demand and performance on the highway network. These types of model only can forecast road freight modes. Usually are used to analyze congestion effects of road freight movements on highways.
- *Direct commodity table freight model:* A travel demand model is operated by DOT which includes explicit modules that calculate a freight vehicle trip table. The model converts a multimodal commodity table, acquired directly, into a table for freight truck trips. Results of this conversion are assigned to a highway network either independently or jointly with other vehicle types.
- *Commodity-based four-step freight model:* This type of freight models follow the same steps as passenger four-step models, including trip generation, trip distribution, mode choice modules and network assignment. However, in the trip generation step weight of commodities by groups of commodities are produced instead of individual trips. The result of this model is a freight truck trip table that includes road and non-road modes and will be assigned to the traffic networks.
- *Economic activity freight model:* This type of model traces the commodity flows between economic sectors and zones. The results of an economic activity model are used to

calculate production and consumption of freight. The economic model difference from the four-step model is that economic indicators such as employment, which are used to calculate freight production and consumption, are not acquired exogenously. They are instead determined iteratively within the model framework.

This list can be altered to include other classifications. For example, the Oregon statewide freight model, which is classified in the economic activity category, is in fact a hybrid microsimulation model which takes the commodity flows resulted from an economic model and convert them into vehicle tours that might have intermediate stops at warehouse or distribution facilities (Cambridge Systematics Inc, 2010). Also, the Ohio statewide freight model is classified as "four-step freight model with an integrated economic model" since the freight flows are calculated from the economic value of shipments between different zones in a distinguished module, rather than over-generalizing from data scarcity (Cambridge Systematics Inc, 2010).

All the operational statewide freight models by the year 2010 are classified using the above classification (Cambridge Systematics Inc, 2010). Table 1 presents the result of this classification. Since no information was obtained on the operational freight models for the sates of Colorado, Idaho, Illinois, Kansas and South Carolina, it is assumed that these states have no operational model (Cambridge Systematics Inc, 2010). As it can be seen in the table 19 (38%) of states, are categorized in the group of "no freight model". These states either use no operational model due to modeling restrictions or the statewide travel demand model used by them do not include the freight transportation forecast.

Table 1 Classification of Current Operational Statewide Freight Models (source: Cambridge Systematics Inc, 2010)

State Name	*No model*	*Truck model*	*Direct commodity table model*	*Commodity-based four-step model*	*Economic activity model*	State Name	*No model*	*Truck model*	*Direct commodity table model*	*Commodity-based four-step model*	*Economic activity model*
Alabama		✓				Montana			✓		
Alaska	✓					Nebraska		✓			
Arizona		✓				Nevada	✓				
Arkansas	✓					New Hampshire		✓			
California	✓					New Jersey		✓			
Colorado	✓					New Mexico			✓		
Connecticut		✓				New York	✓				
Delaware		✓				North Carolina	✓				
Florida				✓		North Dakota	✓				
Georgia			✓			Ohio				✓	
Hawaii	✓					Oklahoma	✓				
Idaho	✓					Oregon					✓
Illinois	✓					Pennsylvania				✓	
Indiana				✓		Rhode Island		✓			
Iowa		✓				South Carolina	✓				
Kansas	✓					South Dakota	✓				
Kentucky			✓			Tennessee			✓		
Louisiana			✓			Texas				✓	
Maine		✓				Utah		✓			
Maryland		✓				Vermont		✓			
Massachusetts		✓				Virginia			✓		
Michigan				✓		Washington	✓				
Minnesota	✓					West Virginia	✓				
Mississippi			✓			Wisconsin				✓	
Missouri		✓				Wyoming	✓				

Fifteen states use operational truck freight models to forecast road freight transportation in the state. The truck models are commonly used in small dense urban states, such as Connecticut, Delaware, Maryland, Massachusetts, New Hampshire, and New Jersey have similar structure to the urban freight transportation models. The truck freight models are considered as typical extension of MPO urban travel demand model due to the close geographical proximity of MPO travel demand models. However, the truck models only focus on the freight trucks in the subject state and possibly a buffer of surrounding zones in proximate states and do not distinguish between short haul and long haul truck movements. Moreover, these models do not include the explicit treatment of using an intermediate handling facility as a part of long-distance supply chains and the type of commodities transported by freight trucks cannot be identified in simple truck models (Cambridge Systematics Inc, 2010).

Eight states operate direct commodity table freight models to convert multimodal commodity tables to freight vehicle trips. An alternative way to syntactically generate truck trip table using trip generation and distribution methods is to use a direct survey of commodity flows. Several statewide freight models are based on this assumption that an annual commodity flow survey can be converted to a daily freight trip table for an appropriate geographical area and used instead of synthetic trip table to be assigned to transportation network. However, two issues should be considered in development of this type of models. First, since commodity flow surveys normally report annual flows, the data need to be converted to daily flows. In addition, the large geographic zones used commonly in commodity flow surveys need to be disaggregated to a smaller geographic zone system appropriate for the statewide travel demand models and the commodity flows should be sub-allocated to smaller zones which was done in the Virginia, Mississippi and Louisiana statewide freight models.

Seven statewide models are categorized in the four-step freight model class. In the commodity-based four-step models, usually results of commodity flow survey are used to develop and calibrate trip generation, trip distribution, and modal choice steps which are then used in the assignment step. The commodity types are treated similar to travel purposes in a passenger transportation model. Therefore, it is assumed that commodities in each commodity class have similar trip generation, distribution, and mode choice behavior. The Ohio statewide freight model can be considered as the most well-developed four-step model that also incorporates economic components for the estimation of commodity production and consumption and commodity flows. Finally, only the Oregon state freight model is considered as an economic model. Economic freight models generally deal with the estimation and forecast of economic activities' extent and locations in a target area. As a result of forecasting economic activities, the production and consumption of different industry sectors will be determined. Thus, the commodity flows between suppliers and customers are identified.

2.2. Freight Modeling Frameworks

Recently, there have been some advances in developing logistics models with a primary focus on road transportation. Different freight modeling frameworks have been proposed in previous studies which vary enormously based on the methodology, criteria, scale and disaggregation level used in each of them. Several studies proposed different categories for freight modelling frameworks such as the traditional vehicle-based versus commodity-based classification or aggregate versus disaggregate modeling approaches. De Jong et al. (2004) give an overview of the recent freight transportation models that are developed around the world. The paper discusses different model components including, production, attraction and distribution models, also mode choice and network assignment models. The authors present ideas for further development of freight demand modeling practices and propose a preferred modeling structure for freight transportation. Tavasszy (2008) summarizes the cutting-edge freight models that were developed mainly in Europe and discusses major issues in freight policy that illustrate the need for freight demand models. The author outlines a conceptual framework and describes innovative fields in freight modeling which are determined by the European transport policy and are significant for The US freight policy.

Chow et al. (2010) also presents an overview on current freight forecasting models and progresses made with respect to data requirements and model structure. They categorize current freight models into seven classes and discuss each class in detail. They explored the methodology and components of each modeling approach and discuss weakness and strengths of the models in each category. They also introduced two cutting edge freight models that are ignored in NCHRP 606 (2008). These two classes are logistics models and urban vehicle touring models. These models incorporate logistics choices and supply chain configuration into their framework and replicate the behavioral aspects of different decision-making units in the supply chain. In this section, a review of current logistics and urban vehicle touring models proposed by previous studies is provided.

The study by Tavasszy et al. (1998) can be considered as the pioneer in developing a logistics modeling framework with logistics decisions. They developed an aggregate model, called the Strategic Model for Integrated Logistic Evaluation (SMILE), that includes logistics choices for the Netherlands. The theoretical bases of SMILE were founded at the beginning of 90's and the model was first used in 1997. SMILE's framework is based on a three-level chain modeling approach that includes main freight activities: production, inventory, and transportation. Logistics choices are simulated by repeating activities in the chain. Taking into account the warehouse usage and determination of the optimum locations of the warehouses were the main innovations of SMILE at that time. In 2004, a program was implemented by the Dutch Ministry of Transport, Waterways, and Public Works to improve the SMILE's model structure, which resulted in proposing a new version of the model, called SMILE+. Bovenkerk et al. (2005) explains the improvements and calibrations resulted in SMILE+ in detail.

In another effort, Boerkamps et al. (2000) developed GoodTRIP which is a disaggregate commodity based freight transportation model. The GoodTRIP model was used to forecast freight movements in Groningen city in the Netherlands. This urban freight model forecasts supply chains and truck tours in urban areas. The model framework consists of different actors and markets that function in the freight system while representing those actors' interaction through different markets in the system. Therefore, it provides insights on the logistics decision making process and their effects on freight movements in urban areas.

GoodTRIP model structure was also used by Wisetjindawat and Sano (2003), to develop an urban freight microsimulation model. This model includes three steps of the four steps in the traditional four-step modeling approach and incorporates behavior of decision makers. The model has been used to forecast freight movements in Tokyo Metropolitan Area and it has been validated with actual data. They used five percent of the operating establishments in Tokyo Metropolitan Area for the simulation purpose and provided the truck origin-destination matrices and approximate vehicle kilometer traveled in the study area by truck class as output data (Wisetjindawat et al. 2007).

Liedtke and Schepperle (2004) developed an activity-based freight transportation model. The proposed model has similar structure as the SMILE and focuses on behavioral aspects of freight by incorporating passenger activity-based transportation modeling framework. They argued that developing an inclusive freight activity-based logistics model is impossible due to the numerous influential factors and actors involved in decision-making process and insufficient data available for modeling. Therefore, they developed their framework by combining two classification methods. Total annual production in tons were obtained using employment information from the Classification of Products by Activity (RAMON 2008) and 1.7 million trips from the Standard Goods Classification for Transport Statistics were utilized to define the tour type distribution (Chow et al., 2010). Finally, a gravity model was applied to convert annual productions to tour types.

Hunt et al. (2006) developed an agent-based microsimulation model to forecast commercial truck movements in Calgary. An extensive dataset on 37,000 truck tours and 185,000 truck trips (Stefan et al., 2005) was used for the model development. The study provides valuable information on touring patterns of commercial trucks, such as route choice, in urban areas. The results of this study were used in other Canadian cities such as Edmonton and in the Ohio State in the U.S. (Yang et al., 2009).

The Statewide Integrated Model (SWIM) (Donnelly et al., 2007) was developed by the Oregon Department of Transportation as the freight component of an integrated land-use and transportation model. SWIM is an integrated economic, land use, and transport model, in which road freight movements are integrated with the passenger transportation model results to more precisely replicate micro-level truck trips (Hunt et al., 2001). Various freight data sources with a diverse range of temporal and spatial attributes were used for the development of the SWIM model.

In another study, De Jong and Ben-Akiva (2007) developed a disaggregate logistics modeling framework which ideally works at the firm-to-firm level. They considered a logistics cost function as their objective function in the framework. The logistics decisions such as shipment size and mode choice are determined by minimizing the objective logistics cost function. The paper discussed the model framework and the required data for the model set-up for national scale freight movements in Norway and Sweden. Later Ben-Akiva and de Jong (2008) reintroduced their freight model framework as the Aggregate-Disaggregate-Aggregate (ADA) freight model system. The ADA model system is a freight transportation model that can be used at different scales including national and regional level. The ADA model system includes three distinct layers. The first layer is an aggregate model that predicts production to consumption flows. The next layer is a disaggregate logistics model that forecasts logistics decisions such as shipment size and mode choice at the scale of firm-to-firm. The last layer is the network model that assigns aggregate commodity flows to the traffic network.

Liedtke (2009) developed an agent-based microsimulation model that included several logistics components. The major logistics elements that were incorporated in the framework include firm generation, supply chain replication, shipment size and carrier choice. The model also replicate tour patterns of commercial truck. However, the model only covered road freight transportation in urban area and ignored other modes such as rail for freight transportation.

Roorda et al. (2011) proposed an inclusive conceptual agent-based freight microsimulation framework. They discussed the diversity of agents in freight system and explained the interactions between those actors in different markets. The framework emphasizes on more complex supply chains and discusses new aspects of freight demand modelling such as outsourcing of logistics services to a third party logistics (3PL). However, as it is mentioned in their study, it is a controversial task and very data intensive to make the proposed framework operational.

Samimi et al. (2010) developed a freight activity-based modeling system, called FAME, with a modular structure. FAME is a microsimulation model of freight movement that considers firms as the decision-maker unit to better forecast the supply chain formation and logistics choices. This model takes as inputs total annual commodity flow at zonal level and generates disaggregate commodity flows at firm level as the outputs. Current operational framework of FAME does not include all the logistics decisions such as shipment size choice, warehousing, use of distribution and/or consolidation centers, and intermodal facilities. The model treats logistics choices separately in different tasks and only simulates truck and rail shipments. In addition, the model uses aggregate zonal-level freight flows and allocates them between firms using a straightforward process which is insensitive to behavioral aspects of decision making and can result in loss of precision.

2.3. Supplier Selection Models

Supplier evaluation and selection problems are among the most crucial logistics decisions that have been addressed extensively in supply chain management. For many business establishments, the costs related to procurement of raw material from outside suppliers constitute a significant amount of firms' total operating costs. This logistics decision is also important from the freight transportation perspective since it can affect other logistics choices, related to freight transportation, such as mode and shipment size choices. Results of supplier evaluation and selection decision identify trade relationships between business establishments and determine commodity flows between production and consumption points in a transportation network. These commodity flows are then used as input to freight transportation models to determine freight movements and their characteristics including mode choice and shipment size.

Different methods have been proposed and utilized to explore the supplier selection problem in former studies. Traditionally, potential suppliers were evaluated using only price/cost as the single influential criterion. However, since the 1960's it has been argued that selecting suppliers that offer the lowest price is not "efficient sourcing" and does not necessarily result in the least total logistics cost (Lung Ng, 2008; Dulmin and Mininno, 2003). Therefore, in the modern supply chain management, multiple factors are considered in evaluating and selecting suppliers.

This has made the supplier evaluation and selection models more complicated than before since for different decision–makers different criteria might be important in selecting a supplier and same criteria might have dissimilar importance in supplier evaluation. These differences in importance ranking of criteria result from many factors such as business establishments' logistics and purchasing policy and decision-makers' characteristics.

In an early study by Weber et al. (1991), over 74 studies about supplier selection problem were reviewed and analyzed to classify the most important factors in supplier evaluation and selection while taking into account the significant changes in logistics and supply chain management process. They argued that supplier selection process has changed significantly due to the revolutions in logistics and supply chain management methods such as improved computer communications, technical advances and growing interest toward just-in-time (JIT) strategies. They also annotated former studies that used quantitative and analytical methods to select the most efficient suppliers for a supply chain. However, the focus of the study is on the most important criteria in supplier selection process.

Review of literature showed that numerous quantitative approaches incorporating multi-criteria in supplier selection problem have been proposed in former studies. Multi-objective optimization (MOP) (Dahel, 2003; Weber, 1998; Weber, 2000; Narasimhan, 2006; Wadhwa and Ravindran, 2008; Haleh and Hamidi, 2011; Yu et al., 2012), analytic hierarchy process (AHP) (Bhutta and

Huq, 2002; Akarte et al, 2001; Chan, 2003; Levary, 2008; Mafakheri et al., 2011), data envelopment analysis (DEA) (Liu et al., 2000; Talluri and Baker, 2002; Seydel, 2006; Saen, 2007; Wu and Blackhurst, 2009; Falagario et al., 2012; Saen, 2010) and simple multi-attribute rating technique (SMART) (Seydel, 2005; Huang and Keska, 2007; Chou and Chang, 2008) are among the most common multi-criteria methods that were employed to evaluate and select suppliers in a supply chain.

De Boer et al. (2001) reviewed different approached that have been used in previous studies to solve supplier selection problem. They proposed a multi-step procedure for selecting the most efficient suppliers. The proposed procedure's steps include defining the problem, formulating criteria, evaluating potential suppliers, and choosing the most efficient suppliers. They also discussed decision methods that can be used in each step of the proposed supplier selection procedure. They also discussed thoroughly the quantitative multiple criteria models and techniques used in the two last steps of the framework, evaluation and final choice phases, to identify potential and select the most qualified ones and classified the existing models into different classes.

In another study, Ho et al (2010) extended and updated former literature reviews regarding supplier selection models. They presented a comprehensive review of academic literature on multi-criteria decision-making models for supplier evaluation and selection. They classified all models applied in former studies into two main approaches; individual approaches and integrated approaches. Individual approaches include DEA, mathematical programming, AHP, case-based reasoning, analytic network process (ANP), fuzzy set theory, SMART and genetic algorithm

(GA). Integrated approaches including integrated AHP and fuzzy models have been applied in former studies.

They annotated former studies that applied any of modeling approaches in detail and identified the most prevalent evaluating criteria used in these studies. Finally, they explore possible inadequacy of existing approaches and further developments. They review showed that AHP approach, with 17.95% share of all proposed models, is the most prevalent individual method in selecting supplier in a supply chain followed by mathematical programming models with 11.54% share. On the other hand, the integrated model of AHP and goal programming (GP) approach is the most frequently used integrated method. In addition, they found that price or cost is not the most commonly used criterion in suppliers evaluating and selecting models. Instead, quality is the most frequently used factor followed by delivery. Cost/price takes the third place among the most prevalently used criteria in selecting suppliers.

In a more recent study, Chai et al. (2013) presented a literature review on studies that were published from 2008 to 2012 and explored proposed decision models for supplier selection. In total, they reviewed 123 journal articles and identified 26 decision-making techniques used in these reviewed studies. Based on different problem solving perspectives used in these techniques, they classified them into three categories including multi-criteria decision-making approach, mathematical programming model, and artificial intelligence method.

They analyzed models in each class and explored the integration of different techniques. Based on their review, multi-criteria decision-making techniques including AHP, ANP, SMART are the most prevalent approaches used for supplier selection problem in the reviewed articles. This is followed by and Techniques for Order Performance by Similarity to Ideal Solution (TOPSIS), mathematical programming techniques including DEA, linear and non-linear programming, multi-objective programming and GP and artificial intelligence techniques including GA, neural network, Bayesian networks, decision tree etc. They discussed proposed models in each category in former studies and provided recommendations for future research regarding the employment of decision-making techniques in supplier selection problem.

All these proposed modeling methods have their advantages and shortcomings. For example, as discussed by Lung NG (2008) both multi-objective optimization and SMART models require that the exact weights of individual criterion to be determined exogenously by decision maker and used as input in the selection model. Criteria weight determination is a challenging process and requires detailed information on how each criterion affects supplier selection and how important it is to exactly determine weight values (Lung NG, 2008).

On the other hand, AHP models provide the ability of specifying weights of criteria for users (decision makers) by providing an interactive comparison environment for them. In these models, users have to have a good knowledge of relative importance and degree of relativity of each pair of criteria to be able to perform pair-wise comparison between criteria. Therefore, the implementation of these models is too demanding for users and the results of these models are highly dependent on decision makers' subjective judgments (Lung NG, 2008). Nevertheless, Lung NG (2008) classifies DEA models the most straightforward for practical implementation since weights of criteria are determined endogenously in these models using the performance scores of suppliers. However, DEA models provide no ability for decision makers to control or alter importance ranking of criteria which is an unrealistic situation.

2.4. Mode Choice and Shipment size Choice Models

Freight mode choice is among the most critical logistics decisions and has been explored exclusively in many studies using different methodologies. Gray (1982) provides a review on methodologies and approaches used in empirical studies to model freight mode choice of 70's and 80's. It examines the unit of analysis in each model (such as, commodity, firm, individual person) and assumptions from which models are driven. It also classifies behavioral approaches in freight mode choice modeling, used in past studies, into three major groups. The study by Roberts (1977) is a leading advocate of the adoption of logistics approach in freight mode choice modeling which can take into account a wide range of logistics costs when a decision maker select a mode of transport. In this study a disaggregate freight mode choice model is estimated as a multinomial logit function in which four types of explanatory variables are used in utility functions including, transportation attributes, commodity attributes, market attributes and receiver attributes (Roberts, 1977).

A general freight transportation demand model is introduced by Oum (1979) where an aggregate model is developed to present the link-specific unit transportation cost function for transporting a particular commodity group. The transportation cost function includes freight rate, level of service for different modes and distance of the link. In a study by Lewis and Widup (1982) a dynamic mode choice model for truck versus rail is proposed using the same transportation cost function presented by Oum (1979). A set of mode choice models by commodity type was developed by Nam (1997) using logit modeling structure. The models were used to explore the effects of commodity type on freight modal split which showed that commodity characteristics affects the transport time parameter in the estimated models.

Jiang et al. (1999) a disaggregate freight mode choice model using nested logit framework. The model was used to estimate modal split at the national scale for France. A multinomial logit formulation is used to develop a mode choice model for domestic shipments in Italy by Catalani (2001). Norojono and Young (2003) developed a freight mode choice model for Java, Indonesia using the results of a stated preference survey. Quality and flexibility of service were the most significant explanatory variables in the estimated models. The same data set was used in the study by Arunotayanun and Polak (2007) to develop a mixed-multinomial logit model. The model was used to investigate the effects of commodity type on modal selection. The differences in carrier and modal selection between Third Party Logistics (3PL) and other shippers was examined by Patterson et al. (2007) and Patterson et al. (2008) using a mixed-logit model. A weighted logit model was developed by Rich et al. (2009) for examining freight mode choice and crossing in the Oresund region.

A binary probit model for choice of mode (truck vs. rail) was developed by Samimi et al. (2012) for domestic shipments in the U.S. the model was used as a component of a freight activity microsimulation model, FAME (Samimi et al., 2010) and was validated using publicly available data sets. Later Pourabdollahi et al. (2013) used the results of an establishment survey (Sturm et al., 2013) to develop a logit model for freight mode choice considering four modes of transportation including truck, rail, air and courier (such as US postal service). They implemented the model into the FAME microsimulation framework and validated the model using the FAF data (FHWA, 2007). The model works at disaggregate level of firm-to-firm and the choice of explanatory variables is mainly based on publicly available freight data.

Shipment size is also an important logistics choice in freight transportation. Theoretical models of optimal shipment size have been developed and used for many decades. Generally, the early models are based on the inventory-theoretic notion, such as Economic Order Quantity (EOQ) model. Although the EOQ model was originally developed to determine the optimal shipment size for a production chain, it was also employed in freight transportation studies (Combes, 2012). In a study by Combes (2012), an EOQ model is utilized to estimate the optimal shipment size with taking into account the selected mode of transportation. The model is evaluated and validated at national scale over a heterogeneous population of shipments. The paper also highlights the inter-relationship between mode and shipment size choice. Review of literature showed that shipment size is mostly used as an explanatory covariate in the freight mode choice models (Cunningham, 1982; Gray, 1982; Roberts, 1977; Nam, 1997) rather than being considered as an independent logistics choice.

While freight mode choice is explored exclusively in many studies using different methodologies and optimal shipment size models, as a part of inventory-theoretic models, have been around for many decades, only few studies discussed joint mode-shipment size choice decisions. The study by Baumol and Vinod (1970) was probably one of the earliest studies that combined mode of transportation, shipment size, and safety stock level at destination into one disaggregate model and introduced a theoretical total transportation and inventory cost equation. In 1985, Hall

(1985), presented a simplified transportation and inventory cost function in which the inter-relationship between mode and shipment size choice is examined when they are selected simultaneously to minimize the cost function. The idea of joint model of discrete and continuous choice in freight transportation was also explored in a study by McFadden and Winston (1981). Later, they developed a joint discrete-continuous model of mode-shipment size choice (McFadden et al., 1985) for truck vs. rail freight transportation based on inventory-theoretic notion.

A joint mode-shipment size choice model is developed by Abdelwahab and Sargious (1992) using switching simultaneous equations system in which mode of transportation is modeled using a binary probit formula and the choice of shipment size is modeled using linear regression equations. The proposed model is used to examine the dependence between mode (truck vs. rail) and shipment size choice. The result suggested that two choices are highly linked and independent modeling would result in biased outcomes. The same model was later used to estimate elasticity of demand and mode choice probability in the freight transportation system (Abdelwahab, 1998). More recently, Holguin-Veras (2002) presented a joint discrete-continuous choice model of vehicle type and shipment size. The results of the model supported the interdependency of vehicle type and shipment size choice.

Despite the aforementioned studies where joint models with discrete mode and continuous shipments size is employed, there are few other studies (e.g. Chiang et al., 1980; De Jong, 2007) that have categorized shipment size into discrete groups and investigated joint decision-making problem using discrete-discrete choice models. De Jong and Johnson (2009) developed an independent discrete mode choice, a joint discrete-continuous model of mode-shipment size choice, and a joint model of discrete-discrete mode-shipment size choice and examined the results of each model and compared transportation time and cost elasticity. In an effort to include transportation cost in inventory models, Langley (1980) examined the effects of freight demand rates on mode and shipment size choices by incorporating transportation costs into the inventory-theoretic model formulation and examining the effects of simultaneous decision making.

A freight transportation model is also proposed by De Jong and Ben Akiva (2007) in that a specific notion of mode choice and shipment size is defined. The model is a comprehensive logistics cost function which is minimized to determine optimum logistics choices. In a recent study by Lloret-Batlle and Combes (2013), an Economic Order Quantity (EOQ) model is utilized to estimate the mode and shipment size choice for freight transportation jointly. The model considers the interconnection between mode and shipment size choice by including transport costs and the logistic costs of shippers in the utility specification of each transport mode. Total commodity flow, the distance and the value density of the commodity are used as the main explanatory variables in the model. The results of the proposed methodology highlight the potential improvements in freight mode choice modeling using inventory theory concept and incorporating shipment size in the modeling specification.

2.5. Shipping Chain Models

Although the field of developing advanced logistics models is rapidly expanding, many existing freight transportation models suffer from a lack of incorporating explicit logistics choices in their framework. *Shipping chain* choice is among the most important logistics choices that has been completely ignored in freight transportation models or received little attention. A *shipping chain* is defined as the physical connection between supplier and buyer of goods or origin and destination of the shipment (Pourabdollahi et al., 2012). In the freight transportation literature, the terms *distribution channel* and *transport chain* is used with the similar definition (Boerkamps et al., 2000; Wisetjindawat and Sano, 2007; Windisch et al., 2010). The *shipping chain* is defined in detail in the following sections.

The study by Tavasszy et al. (1998) can be considered as the earliest in developing a freight modeling framework with logistics decisions and treating *shipping chain* choice explicitly as an independent logistics choice in the proposed framework. They developed an aggregate model that includes logistics choices, called the Strategic Model for Integrated Logistic Evaluation (SMILE), in the Netherlands. The SMILE model takes into account the warehouse usage and determines the optimum locations of the warehouses which were the main innovations of the model at that time. SMILE considers different distribution structures for shipments and determines the spatial distribution of commodity flows by comparing cost differences and geographical and organizational resistance differences. Therefore, SMILE can be considered the first practical freight model that explicitly treated the shipping chain configuration problem in its framework. The SMILE model’s framework was improved by Bovenkerk et al. (2005). The SMILE+ model, a new version of the SMILE model that included the treatment of shipping chain choice, was proposed in 2004.

The GoodTRIP freight model, developed by Boerkamps et al. (2000), deals with the supply chain configuration and replicates urban truck tour patterns. The model framework consists of different actors and markets that function in the freight system and make logistics decisions that finally affect urban truck traffic. The proposed framework determines spatial patterns of cargo flows in the transport service market. The result of transport service market configures the shipping chain choice of each shipment and determines the supply chain patterns.

Wisetjindawat and Sano (2003) developed an urban freight microsimulation model which includes three steps of the four steps in the traditional four-step modeling approach and incorporates behavior of decision makers. They treated the shipping chain choice in an independent module in their proposed modeling structure called “commodity distribution model”. In this module the distribution channel (*shipping chain*) of shipments are determined considering different possible paths and using a system of probability equations. The model has been used to forecast freight movements in Tokyo Metropolitan Area and it has been validated with actual data. They improved the modeling structure in a later study (Wisetjindawat and Sano, 2007).

In a more sophisticated study, De Jong and Ben-Akiva (2007) developed a disaggregate logistics modeling framework which determines shipping chain configuration in a module called “transport chain choice”. The “transport chain choice” module includes three logistics choices including the number of links in the transportation chain (shipping chain), use and location of intermediate handling facility and transportation mode used for each link. They considered a logistics cost function as the objective function in the framework and determined logistics choices by minimizing the objective logistics cost function. The paper discussed the model framework and the required data for the model set-up for national freight movements in Norway and Sweden. Later Ben-Akiva and de Jong (2008) reintroduced their freight model framework as the Aggregate-Disaggregate-Aggregate (ADA) freight model system. The ADA model system includes three distinct layers. The first layer is an aggregate model that predicts production to consumption flows. The next layer is a disaggregate logistics model that forecasts logistics decisions including transport chain. The last layer is the network model that assigns aggregate commodity flows to the traffic network.

3. RESEARCH GAPS AND OBJECTIVES

The remarkable increase in freight movements and their important influence on socio-economic systems such as transportation system provide sufficient motivation to develop reliable tools to estimate commodity flows between zones and forecast the future trends of goods movements among regions. However, as the review of operational statewide and regional freight models revealed, the amount of attention dedicated to freight demand forecasting is insufficient and many states and local agencies are still operating aggregate simple models for the estimation and forecasting freight movements. Yang et al. (2009) recounted three main reasons that have made the state of practice in freight transportation modeling evolve much slower than that in passenger travel demand modeling. The stated reasons include the freight data scarcity, insufficient practical experience regarding freight forecasting and limited available input data for the current freight related models.

The NCFRP 8 (2010) argues that even the cutting edge of the state of the practice freight models do not completely fulfill needs of freight planners, modelers and decision-makers. The report recounts several key issues related to this insufficiency, including limited ties between freight planning and economic development, data limitations and limitations of existing freight demand models. Limited attention toward multimodal networks in modeling structure, neglecting the behavioral aspects of freight decision-making process in the modeling framework, ignoring freight routing and rout diversion in the network assignment step, unresponsiveness to economic changes and insensitivity to temporal changes are among the major limitations of existing freight demand modeling and analysis tools (NCFRP 8, 2010). In addition, the NCHRP 606 (2008) discusses several analytical and policy needs that are not addressed by any of the existing freight modeling classes which emphasizes on the insufficiency of current practical freight modeling frameworks.

As it is stated in the literature review of current freight demand models by Chow et al. (2010), several studies such as Hesse and Rodrigue, (2004) and Friesz and Holguin-Veras (2005) have examined the insufficiencies of current practical freight demand models. Hensher and Figliozzi (2007) discuss the drawbacks of the four-step freight demand models in capturing the effects of customer-driven economy in the 21st century on the commodity flows (Chow et al., 2010). They emphasize on the need to incorporate supply chain and logistics components in the freight transportation models. Chow et al (2010) argue that available freight demand models generally have aggregate nature which makes them insensitive to the changes in logistics behavior of decision-making units such as firms. They also discuss the major drawbacks of the common vehicle-based and commodity-based freight models and emphasize on the importance of development of more disaggregate model that incorporates behavioral aspects of decision-making process in supply chains.

Moreover, while large scale freight transportation models are effective tools to predict freight transportation demand and to better address the impacts of freight movements, many of existing freight demand models both at national or regional scale do not explicitly treat logistics choices (de Jong and Ben Akiva 2007). Current practical freight transportation models do not take into account many crucial logistics choices such as shipment size choice or configuration of shipping chain choice including the use of intermediate handling facility. Therefore, they cannot precisely replicate the supply chains and logistics decisions made by individual decision-making units in the freight system.

The main purpose of this dissertation is to develop a freight transportation model that is capable of addressing some of the outstanding issues discussed earlier and found in the current freight transportation models in a fully operational regional microsimulation framework. The proposed model works at a very disaggregate level and simulates commodity flows at firm-to-firm level. It considers firms as the decision-making units that determine trade relationships and form supply chains in the freight system. Therefore, it takes into account the behavioral aspects of decision-making process in freight system by incorporating behavioral characteristics decision makers (firms) into logistics choice models. Several advanced logistics choice models are incorporated into the microsimulation framework to explicitly treat the logistics choices that are ignored or simplified in previous practical models. A joint model of mode choice and shipment size and a shipping chain configuration model are among the advanced incorporated logistics choice models.

A procedure is also developed in this study to estimate freight generation and production rates at firm level. This procedure is used in the economic activity component of the framework to replicate commodity production-consumption rates by individual firms instead of using the aggregate publicly available commodity production and consumption at zone level. In addition, an innovative behavioral two-step optimization model is developed for supplier evaluation. The model takes into account the opinion of buyer firms in evaluating their potential suppliers and selects the most suitable suppliers by minimizing transportation costs. This model can more precisely identify trade relationships and supply chains between firms. Finally, a network analysis model completes the microsimulation framework. This step assigns commodity flows to the traffic network and provides measures validation and further analysis.

4. DEVELOPMENT OF THE FREIGHT TRANSPORTATION FRAMEWORK

4.1. Proposed Model Framework

This part of the study presents the developed framework for the agent-based freight transportation model. The proposed model considers firms or business establishments as individual decision-making units in the freight transportation system. It assumes that firms are the key actors in planning and execution of logistics decisions in supply chains. These logistics decisions which form supply chain activities include supplier selection, shipment size, mode choice and shipping chain configuration. A step-wise modeling system is used for the development of framework. This step-wise algorithm treats logistics decisions made by individual decision-makers (firms) as discrete events within the simulation and replicates logistics activities planned by decision-makers within the supply chains for each individual shipment. The fundamental underlying concept of this framework is the extension of role of individual decision-makers in performing economic and logistics activities which results in production and consumption of goods and commodity flows between producer and consumers.

The proposed modeling system is as a three-layered framework in which each layer consists of several models that perform different tasks. Figure 1 presents an outline of the proposed multilayered modeling framework and the sub-models in each layer. The objective of the proposed model is to put forward a disaggregate framework for freight transportation that includes major logistics choices, while keeping the dimensionality and complexity of the model manageable and the need for the survey and private input data the least.

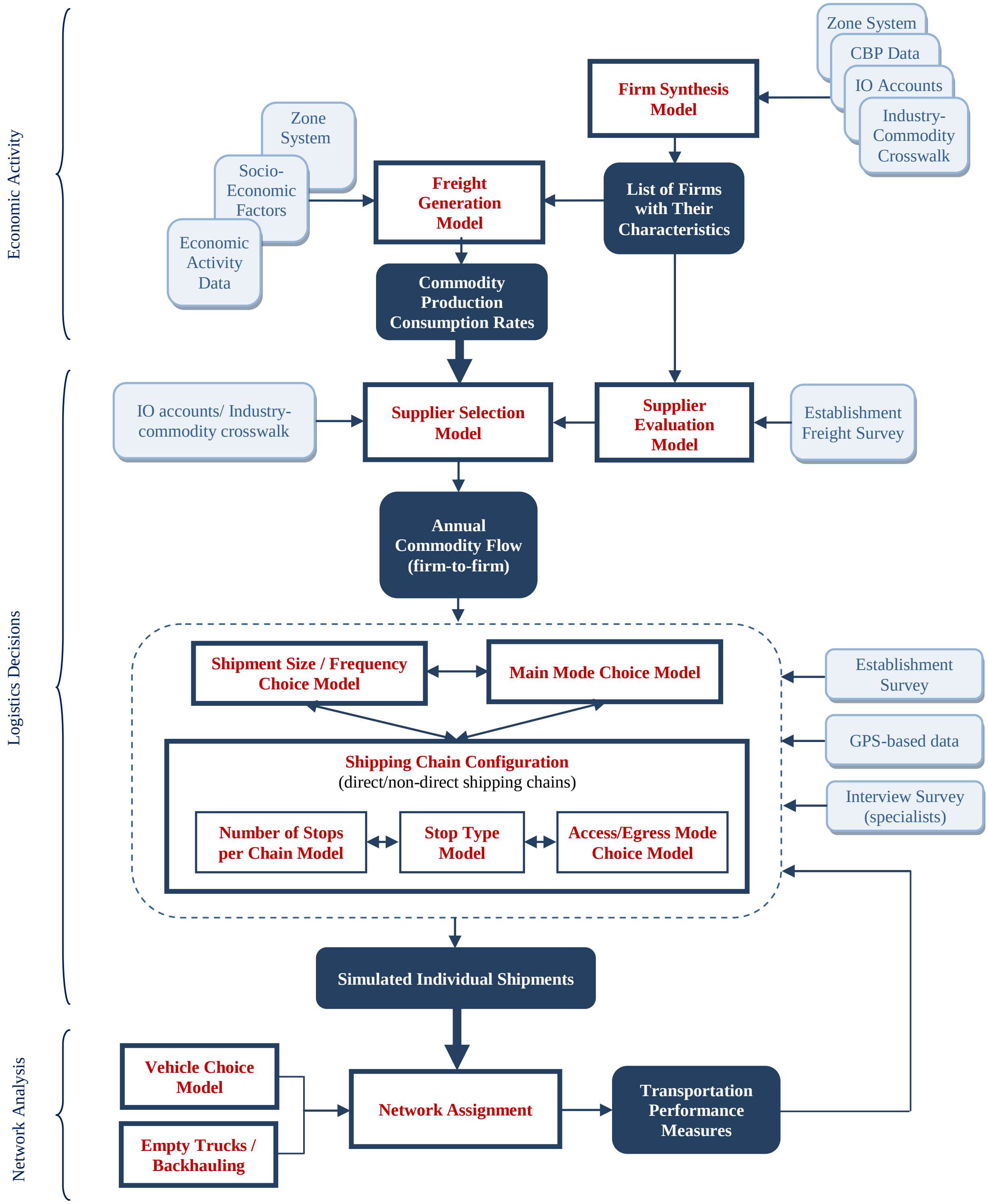


Figure 1 Outline of the proposed framework

In the first layer, “Economic Activity”, the agents (firms) in the study area are generated and their characteristics are determined. Then, economic factors, a considerable set of data sources and a complex procedure is used to determine the production and consumption values of different commodity types for these firms. The second layer is the “Logistics Decisions” in which the logistics components of supply chains are determined in a step-wise process. First, in this layer the trade relationships between firms are formed and supplier-buyer pairs are identified. A probability of partnership is estimated for each supplier-buyer pair and commodity flows between them are determined using these probabilities. Next, the logistics choices including shipment size, mode choice, shipping chain choice are determined for disaggregate flows. The final layer, “Network Analysis”, deals with the assignment of commodity flows to the transportation networks which allows further analysis and model validation. The model’s components are briefly discussed in this section and the development of its models is explained in details in Chapter 6.

4.1.1. Base Year Selection

For the development of the framework, several freight and other relevant datasets are used which have different base years and dates of publication. For example, the Freight Analysis Framework (FAF) (FHWA, 2007) is an integrated freight data source that provides estimates for freight movements among states and metropolitan areas in the U.S. and is published every 5 years. County Business Patterns (CBP) (U.S. Census Bureau, 2009) which provides national and sub-national economic data by industry is published annually. Therefore, in order to keep the consistency in the development of the framework while using different publicly available datasets, a base year has to be selected in which all datasets are obtainable. Since the year 2007 is the latest year in which most of these publicly available datasets are available, it is selected as the base year for the model development.

4.1.2. Transported Goods, Geographical Scale and Zone System

The model takes into account all commodity types that have identified value. It uses the two-digit Standard Classification of Transported Goods (SCTG) system, developed by the Bureau of Transportation Statistics (Bureau of Transportation Statistics, 2007), to categorize commodities. However, to manage computational complexity of the model some of the commodity classes that have similar nature are combined together and treated as a single commodity type. For example, four commodity types of Basic chemicals (SCTG code 20), Pharmaceutical products (SCTG code 21), Fertilizers (SCTG code 22) and Chemical products and preparations, n.e.c (SCTG code 23) are categorized as a single commodity class of "Chemical and Pharmaceutical Products".

Using this methodology, 43 classes of 2-digits SCTG commodity types are categorized into 15 classes of commodities that are used in this simulation model. However, commodity class 15, "Mixed and Unknown Freight", is excluded from the simulation since the producer and consumer agents (firms) for this commodity type can be hardly identified. Table 2 presents the commodity classification used in this study.

Since the proposed commodity classification is derived from SCTG, the commodity flow estimates provided by FAF data (FHWA, 2007) and Commodity Flow Survey (CFS) (U.S. Census Bureau, 2014) that use SCTG code for commodity classification can be easily transformed to the proposed classification system for further application in this study. The proposed model only focuses on the domestic good movements. This also includes transportation of imported goods from ports, airports or other ports of entry to other regions of the United States.

Table 2 Commodity Classes and Definitions

Commodity Class	Definition	Related SCTG
1	Agriculture and Forestry Products	1-9
2	Products of Mining	10-15
3	Petroleum Products	16-19
4	Chemical and Pharmaceutical Products	20-23
5	Wood Products	25, 26
6	Paper Products	27-29
7	Nonmetallic mineral products	31
8	Metal and Machinery Products	32-34
9	Electronic, Electrical and Precision Equipments	35, 38
10	Motorized and Transportation Vehicles and Equipments	36, 37
11	Household and Office Furniture	39
12	Plastic, Rubber and Miscellaneous Manufactured Products	24, 40
13	Textiles and Leather Products	30
14	Waste and scrap	41
15	Mixed and Unknown Freight	43, 99

The framework has a very flexible structure regarding the geographical scale. It can perform analysis at nationwide scale or can be tailored to be used as a regional model. While using the framework at nationwide scale requires some aggregation in order to deal with computational complexities, performing analysis at regional scale can focus on more detailed characteristics of freight movements. However, it may require more detailed input data as well.

The proposed framework in this study covers domestic commodity flows in the country; however, it mainly focuses on the freight movements in the Chicago region. Therefore, a variable zone system is used for this purpose. The proposed zone system comprises township level zones in the Chicago area, counties in the rest of Illinois and FAF zones (FHWA, 2007) elsewhere. Figure 2 displays the zone system developed for this study. The Chicago region that is divided into townships level zones consists of 7 counties of Cook, Dupage, Kane, Kendall, Lake, McHenry and Will. Using this variable zone system, the country is divided into 333 zones, including 120 FAF zone, 95 counties in Illinois and 118 townships in Chicago.

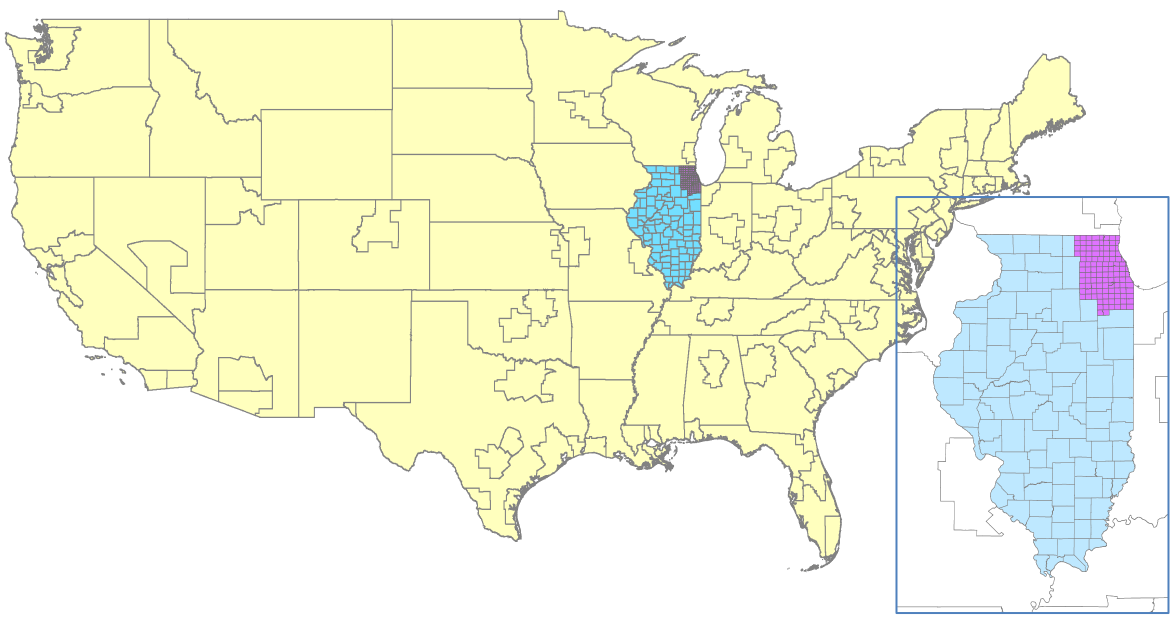


Figure 2 Zone system of the proposed framework.

4.1.3. Economic Activity Model

As it can be seen in figure 1, the Economic Activity model comprises two sub-models, the Firm Synthesis and the Freight Generation model which are described briefly as flows.

Firm Synthesis Model

The Firm Synthesis Model is the first model of the framework in which individual agents in the study area are generated. The model considers individual firms as the decision-making agents in the framework. These firms or businesses establishments can be producer or receiver of goods who form supplier-buyer pairs and specify critical logistics choices of supply chains. However, since considering all existing firms in the study area results in computational complexity in the simulation process, an aggregation method is used to address this problem. Firms with similar characteristics are categorized as a group of agents, called firm-type. A firm-type is defined as a group of firms with the same industry type, employee size and geographic location in the zone system. It is assumed that all firms in a firm-type group have similar behavior in freight decision-making process. Therefore, they can be considered as one firm-type agent in the framework. The concept of firm-type used in the proposed model is the same as what was used in the FAME model (Samimi et al. 2010).

The key input data in the Firm Synthesis Model are the publicly available datasets that provide economic data including County Business Patterns (CBP) and Zip code Business Patterns (ZBP) (U.S. Census Bureau, 2009). These annual data series provide economic data by industry type for different geographic zone level. The data provide information on the number of business establishments by industry type and by employee size, employment information and payroll information. The CBP dataset provides this information at the county level and the ZBP dataset provides the information at the zip code level. The output of the Firm Synthesis Model is a complete list of all business establishments in the study area with their characteristics including industry type, employee size.

Freight Generation Model

In the Freight Generation Model, a considerable set of data sources and a complex procedure is used to determine the production and consumption values of different commodity types for the synthesized firm-types. Typically, existing freight demand models either use the publicly available freight data, which provide aggregate commodity production-consumption tables, to estimate the commodity flows or apply aggregate economic models that generate commodity production-consumption tables at the aggregate zone level. Both approaches provide production-consumption values at aggregate zone level. Therefore, they only take into account the effects of changes in the economic activities at large scale and are not capable of capturing changes in the freight decision-making process and production-consumption rates at the disaggregate firm level.

In the proposed agent-based freight model, where logistics activities are modeled for the freight agents at firm-type level, the production and consumption rates should be estimated for the same firm-type agents as well. An appropriate procedure and a huge set of databases are utilized to estimate the disaggregate commodity production-consumption (input-output) tables for the synthesized firm-types.

First, the SCTG commodity definition and grouping provided by the Bureau of Transportation Statistics (Bureau of Transportation Statistics, 2007), FAF commodity-industry crosswalk (FHWA, 2011) and the commodity-industry mapping provided by the Industry Input-Output Accounts (Bureau of Economic Analysis, 2013) are used to relate commodity groups in table 1 to 3-digits NAICS industry sectors. A disaggregate and more comprehensive three-dimensional commodity-industry crosswalk for produced and consumed commodities by industry is developed by mapping between commodity groups and industry types. The developed crosswalks provide the share of 3-digits NAICS industry sectors in production and consumption of different commodity types at national scale. It is assumed that these production and consumption shares are also valid at smaller regional and zonal scale. In addition, anther crosswalk is developed to determine the production and consumption rates by employee size using the results of an establishment survey conducted at University of Illinois at Chicago (Sturm et al., 2013).

Next, the aggregate SCTG commodity flows from FAF data (FHWA, 2007) are translated to production and consumption values at FAF zone level and transformed to the commodity classification system presented in table 2. The developed commodity-industry crosswalks are then utilized to estimate the disaggregate commodity input-output tables at firm level. The final output of the Economic Activity model is a disaggregate table of commodity production-consumption values for all synthesized firm-types. This disaggregate input-output table are used as input in the next module of the framework where the suppliers and buyers are paired and commodity flows are estimated.

4.1.4. Logistics Decisions Model

The second layer of the framework consists of several logistics sub-models that simulate logistics choices made by firm-types in a step-wise process. Development of these disaggregate logistics choice models requires detailed and comprehensive information on individual shipments' attributes, decision-makers' behavior and spatial information on the location of logistics facilities in the study area. The main input data for the development of this layer's models is the collected data in the UIC establishment survey (Sturm et al., 2013) which is discussed in details in Chapter 5. The output of this model is individual shipments at the disaggregate level of firm-to-firm with determined logistics attributes.

Supplier Evaluation and Selection Models

The first sub-model of this layer is a two-step behavioral supplier evaluation and selection model. As it was discussed in the literature review, multiple factors affect the choice of supplier in a supply chain. Different decision–makers might have different sets of important criteria for evaluating and selecting suppliers. In addition, for different decision-makers same criteria might have different importance rank. These differences in importance ranking of criteria result from many factors such as business establishments' logistics and purchasing policy and decision-makers' characteristics. Incorporating these behavioral aspects in the model structure can make supplier evaluation and selection models more complicated than before.

In this study, a two-step modeling framework is developed for supplier evaluation and selection problem which tries to take into account behavioral aspects of decision making process, transportation costs, capacity constraints, and other important factors in selecting suppliers. The first step deals with the evaluation of potential suppliers for each synthesized firm using a set of criteria. A behavioral modeling approach is used in the first step to include the effects of decision-makers' characteristics in evaluating their potential suppliers. An ordered logit structure is proposed to determine the importance rates (weights) of the criteria in evaluating suppliers for different decision-makers. A suitability score is calculated for each potential supplier based on the estimated importance rates (weights) of these criteria. In the second step, a multi-objective multi-criteria optimization model is used to select the most efficient suppliers for each firm. The optimization model in the second step uses the suitability scores from the first step to select the best suppliers with highest scores in a way to minimize the total transportation costs with consideration of production capacity of suppliers.

The benefit of using this two-step supplier selection model is that the first step considers the behavioral aspects of the problem by including the characteristics of decision-makers in the modeling structure and capturing their effects on rating important criteria. The second step takes into account the decision makers' opinion on evaluating suppliers from the first step and select best suppliers by means of minimizing the total transportation costs. It also takes into account the production capacity of suppliers and allocates the demand between suppliers with this consideration.

The results of the UIC establishment survey 2010-2011 (Sturm et al., 2013) are used to develop the models. Descriptive statistics of the gathered data used for supplier selection model development are presented in section 5. Detailed information on the model development is discussed in Chapter 6.

Mode Choice and Shipment Size Model

Mode and shipment size choice are among the most critical logistics decisions that are mostly studied separately in freight demand studies. Review of literature shows that these two choices are closely related logistics decisions and should be studied simultaneously. For example, non-road transportation such as, rail is usually used for transporting large shipments, while small shipments are transported mainly using road transportation (De Jong and Johnson, 2009).

Moreover, from the inventory-theoretic point of view, a trade-off exists between mode and shipment size choice. A freight decision-maker, for example the buyer firm, might decide to increase the frequency of shipments (decrease the shipment size) in order to decrease its inventory costs. As the frequency of shipments increases the transportation costs will rise. Therefore, the decision-maker might decide to switch to a cheaper mode of transportation to deal with transportation costs. The same interpretation exists for the reverse situation when the decision-maker decides to switch to a cheaper but slower mode of transportation. Longer transit time can cause trouble when there is an unpredictable rise in the demand for goods at the destination. Therefore, receiver firm might order larger shipments to maintain a safety stock against such incidents which may increase inventory and warehousing costs at the destination (Baumol and Vinod, 1970).

This context reveals the importance of inter-relationship between mode choice and shipment size decisions and suggests the use of joint modeling structures to better account for this simultaneous decision making process.

Different approaches are suggested and utilized for joint modeling including the popular simultaneous-equation model systems in which error terms are mostly considered normally distributed or transformed into normally distributed variables. In another approach, for each choice a suitable econometric model is developed (for example, a multinomial logit model for the mode choice and a hazard-based model for the continuous shipment size) and then by transforming random component of each model into the standard normal distribution a joint probability function is built for joint decision making problem.

The third approach was introduced by Bhat and Eluru (2009) in that a copula function is introduced for joint modeling. Copula with variety of classes allows examining a number of different dependency structures between random variables. Therefore, in contrast to previous approaches with restrictive assumption of dependency structure, it provides more flexibility in examining different dependency structures and better captures the potential effects of unobserved common factors.

The dominant approaches used to jointly model mode and shipment size impose a restrictive normality assumption on the error terms that can be incorrect and misleading. For example, Abdelwahab and Sargious (1992) utilized the switching simultaneous equations system to jointly model the logistics choices. In their model they assumed a trivariate normal distribution on the error terms of each alternative and the joint equation to derive the final form of model. McFadden et al., (1985) employed a different approach to estimate the discrete and continuous joint model of mode and shipment size choice. In their approach, they derived a marginal probability equation for the shipment size, and a conditional probability equation for mode choice. They assumed error terms of both equations to be correlated and normally distributed.

In this study, mode choice and shipment size are modeled jointly using the copula-based joint modeling approach due to its flexibility and appealing qualities. It is assumed that the joint model can capture the effects of common causal factors in choosing the interconnected alternatives. Therefore, it will provide more precise results. The UIC establishment survey results are used for the model development. The details of the model development and its results are discussed in Chapter 6.

Shipping Chain Configuration

The Shipping chain configuration is one of the key logistics components that have been ignored or treated insufficiently in current freight transportation models. The term *shipping chain* refers to the physical connection between supplier and buyer of goods or origin and destination of the shipment (Pourabdollahi et al., 2012). In the freight transportation literature, the terms *distribution channel* and *transport chain* are used with the similar definition (Boerkamps et al., 2000; Wisetjindawat and Sano, 2007; Windisch et al., 2010). A *shipping chain* can be a combination of one or more *links* depending on the number of stops per shipping chain. A *link* is the connection between two consecutive stops inside *shipping chain*. For example, a *link* can connect the origin (supplier) to the distribution center, or it can connect the supplier to the buyer directly (Pourabdollahi et al., 2012). Figure 3 illustrates an example of three possible *shipping chains* between a pair of supplier-buyer. The proposed shipping chain choice model focuses on the configuration of physical shipping chain and modeling of its relevant attributes including number of intermediate stops per chain and stop types.

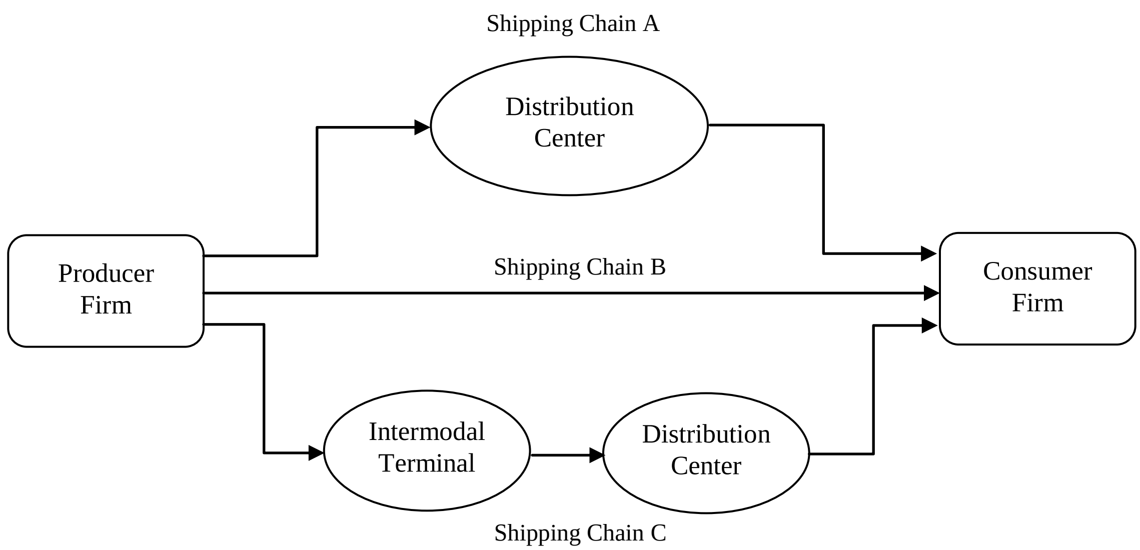


Figure 3 An industrial supply chain with 3 types of shipping chains.

A system of decision tree models is developed to determine the shipping chain configuration of freight movements at a very disaggregate level. Decision trees are uncomplicated, but powerful and effective tools for multiple variable analysis which are used to explain, describe, predict, or classify an outcome (dependent variable). Decision trees are among the most popular data mining methods which have been increasingly utilized by researchers in diverse academic fields for decision analysis and prediction purposes due to their appealing capabilities and analysis strength.

The rule-based decision tree method is used to model shipping chain configuration for individual shipments in this study. The proposed decision tree clustering model employs exhaustive CHAID algorithm as the growing method to simulate the logistics choices (dependent variables).

The simulated logistics choices are the number of stops per shipping chain and the type of each stop inside the shipping chain.

4.1.5. Network Analysis Model

The final layer of the proposed framework is the Network Analysis model which takes the disaggregate shipments with specified characteristics derived from previous model as input and assign them to the relevant networks. In this layer, first, commodity flows should be converted to vehicle loads and the vehicle choice should be determined for each single shipment. Variety of vehicle sizes in each transportation mode class, and the fact that a significant portion of vehicles might be empty or partially loaded, make this task extremely challenging.

Another challenging task that should be performed in this model is to develop detailed networks for all transportation modes considered in the study. In addition, logistics nodes that represent logistics handling facilities (such as, distribution centers, intermodal terminals, consolidation centers, etc.) have to be determined for each network. The assignment of shipments to the networks provides valuable inputs for further analyses. Particularly, it provides transportation performance measures that can be used for policy evaluation and impact analysis. The results of this model such as travel time can also be used in previous sub-models to more realistically determine logistics choices. Therefore, in the ideal modeling framework, a recursive structure should be used to feed back the generalized transportation costs into the logistics choice modeling system, and iterated until a stabilized set of commodity flows and costs are obtained.

5. DATA SOURCES

The development of an agent-based microsimulation freight model requires a significant amount of input data. Collecting the required input data is usually a challenging task. However, it is more difficult to obtain the reliable data for development of freight transportation models since the target population in freight market is usually reluctant to share what might be considered sensitive company information. In this study, it has been tried to use publicly available data for model development, when it is possible, to minimize the costs of data collection. As it is discussed in Chapter 4, CBP and ZBP data (U.S. Census Bureau, 2009), Industry Input-Output Accounts (Bureau of Economic Analysis, 2013) and FAF data (FHWA, 2007) are the major data sources used for the development of the Firm Synthesis and Freight Generation model. However, since the publicly available freight data are aggregate they cannot be used for the development of the disaggregate logistics models in the second layer of the framework. The disaggregate logistics choice models in the framework are developed using the collected data from the UIC freight Establishment survey (Sturm et al., 2013). In this chapter first the publicly available freight data that were used in this study and then UIC freight survey are discussed in the following sub-sections.

5.1. Publicly Available Freight Data

5.1.1. Economic Data on Business Establishment

The key input data in the Firm Synthesis Model are the publicly available datasets that provide economic data on business establishments. The County Business Patterns (CBP) and ZIP Code Business Patterns (ZBP), which are annually published by U.S. Census Bureau (U.S. Census

Bureau, 2009), provide such information. CBP provides information on the number of business establishments by industry type by employee size classes at county level. It also reports employment and payroll information for all establishments. ZBP data also provides information on the number of establishments by industry type by employment size classes at ZIP Code level.

Both datasets use NAICS codes to classify industry types of business establishments. While CBP industry classification are available from 2-digits aggregate NAICS to 6-digits detailed level NAICS, ZBP only provides industry classification information at 2-digits NAICS. The datasets cover most of industry sectors. However, they provide no information on the businesses that classify in the “Crop Production” NAICS 111 and “Animal Production” NAICS 112 group. Therefore these industries are excluded from this study and no firm-type of these industry classes are generated in the simulation process. These datasets provide annual economic activity information based on individual business establishments for a diverse range of geographic scale from state to ZIP code level. Therefore, they are also useful for micro-level analysis of economic activities and studying temporal changes in economic activities among other applications such as marketing and planning studies.

5.1.2. Commodity-Industry Crosswalks

A key input for the development of the Freight Generation Model at the firm level is a reliable commodity-industry crosswalk that provides mapping between SCTG commodity groups and NAICS industry classes. Freight Analysis Framework (FHWA, 2011) uses a straightforward commodity-industry crosswalk that relates SCTG codes to NAICS classes. This crosswalk is only available for the made commodities and relates most industries to only one class of SCTG commodities. Hence, it does not provide a comprehensive mapping between commodity and industry classes and cannot relate industry classes to used commodities. However, this crosswalk is used as a guideline for relating industry classes to commodity groups.

On the other hand, the Industry Input-Output Accounts (Bureau of Economic Analysis, 2008) provides a very detailed commodity-industry crosswalk that is used as a key input data for the development of the Freight Generation Model. The data is used to develop a three-dimensional commodity-industry crosswalk for produced and consumed commodities by different industry classes. The 2007 Industry Input-Output Accounts Data provides detailed information about the type and annual value of commodities that are made or used by different industry sectors in the U.S. It also provides information on the type and value of commodities required to produce a unit value of industry output for different industry types.

The commodity-industry crosswalk provided by Industry Input-Output Accounts cannot be directly used in this study for several reasons. First, the crosswalk presents the monetary relationships between commodity groups and industry classes and does not provide any information about the weight of commodities that are produced or used by industries. In other words, the reported figures are estimated based on the monetary transactions between firms not the commodity flow. Second, the crosswalk uses a special system to classify industry groups which is different from the NAICS system that is used in FAF, CFS and this study. Therefore, some essential adjustments are required before using the crosswalk, which are explained in details in Chapter 6.

Moreover, this crosswalk provides no explicit bridge between industry classes and commodity groups since it uses the same classification code for industry and commodity classes. Therefore, a reliable linkage between industry classes and commodity groups has to be developed. Other than commode-industry bridge provided in FAF data (FHWA, 2011), the SCTG commodity definition and grouping provided by the Bureau of Transportation Statistics (Bureau of Transportation Statistics, 2007) is a good source of information that is used for developing the linkage between industry and commodity classes in this study. This dataset provides detailed information on the commodity grouping in SCTG system. The information about the SCTG codes and commodities included in each SCTG code are provided at 2-digits level with 43 commodity classes to 5-digits detailed level with more than 300 commodity classes.

5.1.3. Aggregate Commodity Flow Data

Freight Analysis Framework (FAF) (FHWA, 2007) which provides freight movements data is the major dataset used in this study. FAF is an integrated freight data that is published by Federal Highway Administration (FHWA) every five years and provides estimates and forecasts of commodity flows. The last version of the FAF data (version 3), which provides estimates for the year 2007 and forecast for 2040, is used for the development of Freight Generation Model and model validation in this study. Federal Highway Administration uses a diverse range of data sources such as CFS (U.S. Census Bureau, 2014) and Transborder Freight Transportation Data (Bureau of Transportation Statistics, 2009) to develop FAF data. The FAF data provide detailed information about domestic, import, export, and in-transit freight movements in the U.S. This information includes annual weight and values of commodity flows between FAF zones. For each commodity flow record between FAF zone, commodity type, mode of transportation and type of trade.

As it is discussed in Chapter 4, FAF data is used for the development of Freight Generation Model. FAF uses SCTG system for commodity classification which makes it compatible with the commodity classification used in this study. The aggregated commodity flows in FAF data are translated into the production and consumption values at FAF zone level. These aggregate made and used values of different commodity types are then transformed to disaggregated production and consumption values at firm level using the Freight Generation Model.

Along with the publication of FAF data, FHWA provides useful publications about FAF model and data including major assumptions, descriptive statistics, methodology and model structure and network assignment and analysis. Some of this information has been used in the development of different components of the framework. As it is discussed before, the commodity-industry crosswalk provided by FAF is used as a guideline to relate commodity and industry classes. Also, the FAF zone system is used to develop the zone system in this framework which is discussed in Chapter 4. Version3 of FAF model consists of 123 domestic regions called FAF zone. The GIS specifications of 120 FAF zones are used to build the zone system in this study excluding the 3 FAF zones that comprise Illinois State.

In addition, FHWA provides transportation network specification (FHWA, 2007) for assignment of FAF flows to the traffic network and details of freight traffic analysis for FAF data (Oak Ridge National Laboratory, 2011a) which are used in the development of the proposed framework. The updates on highway network data and FAF methodology for estimating empty trucks and converting commodity flows into freight vehicle trips are among the information used in this study. These data are used in the Network Analysis layer to estimates freight truck flows and assign these flows to the traffic network.

Finally, the FAF data is used to calibrate and validate the developed logistics choice models including the supplier selection and mode choice model. Since To validate the performance of the proposed model, simulation results are aggregated and compared with aggregate FAF estimates.

5.2. Freight Transportation Surveys

5.2.1. Background

A comprehensive and reliable input data is indispensable part of developing freight transportation modeling framework. If available data sources are not sufficient for the model development, appropriate surveys must be designed to collect required data. There are three major types of freight surveys, including roadside survey, vehicle owner survey, and establishment survey (McCabe, 2007). Each type of these surveys has its own advantages and disadvantages and is conducted to accomplish a specific purpose. This section focuses on presenting a background of past efforts on freight establishment surveys and the methods used in those surveys.

Web-based surveys are newer but commonly used methods of collecting survey data in freight transportation. An author may use an existing list of e-mail leads or purchase a list of contacts from one of any reputable data marketing companies. Another possibility is the rental of e-mail addresses from one of these marketing companies, which would deliver these e-mails en-masse themselves, without revealing the names and contact information to the renters. This type of campaign is known as an e-mail blast campaign. Online surveys have the capacity to reach a wide audience, but generally suffer from poor response rates. This may be attributable to the lack of person-to-person interaction during contact, though it should be known that personalizing the address line benefits the results of the survey (Schaefer and Dillman, 1998). Nevertheless, to their benefit, the cost per response and time needed to complete is generally shorter.

UIC conducted such a study in 2009 in which an e-mail blast campaign directed participants to an online survey. In addition to the primary introductory message, three reminder e-mails were distributed 2, 7, and 14 days after the initial e-mail in order to increase the final total of responses (Samimi et al., 2010). More than 30,000 e-mails were sent which resulted in 25,997 successful contacts and over 4,000 failed e-mails. Of this number, 9.3% clicked on the survey link, and fewer completed the survey. Such a method suits those with large targetable populations, as the survey conductors will have to make do with a lower response rate than that of most other methodologies.

Telephone surveys establish the more personalized and friendly contact missing in exclusively e-mail driven studies, but suffer from other burdens. For this type of survey, low-cost employees, possibly students, are hired for the purpose of obtaining contact or survey data from potential participants. The contact data may be used for mailing or e-mailing surveys. If the caller is gathering information for the survey, it should be noted that with the time it takes to read a survey out loud and for a participant to gather information, the questionnaire must be short enough to fit into the schedule of the participant. The Oregon Department of Transportation (ODOT) conducted such a study in 2002 (Lawson and Strathman, 2002). After evaluating both telephone and mail contact methods, they concluded that the telephone method worked better for them due to a higher response rate. They also mentioned repeated person to person contact as being beneficial. An automated version of the telephone contact approach is known as the Computer Aided Telephone Interview (CATI) method. Such an interview method was conducted by a firm hired by the University of California at Irvine (Golob and Regan, 2002). In it, logistics or operations managers were directed to answer questions in an 18 minute interview. Its final result was a response rate of 22.4 out of 5,258 selected firms.

The use of parcel mail is a well established method of making contact with the target audience. Recruitment postcards are a common method of first reaching out, in order to explain the purpose of the survey and to provide the recipient with time to verify the legitimacy of the sender. After that, and possibly after verification of interest by the recipient, the survey may be sent in packet form to be filled out. A postage paid envelope should be included for return service. With this and the other methods of contact, authors may mix and match elements of these methods as suits the needs of the study. An example of this is the 2007 Region of Peel survey of shippers in the Greater Golden Horseshoe Area of Ontario, Canada (Roorda wt al., 2007). In that study, nine methods were attempted using a mixture of mail and telephone methods. Two of these methods were phone and mail hybrids, which resulted in response rates of

19% and 33% out of two samples of 100. These fared better than each of the five mail only survey methods did. Another study, conducted in Edmonton, Alberta, utilized telephone records to recruit participants, parcel mail to deliver the survey, and live individuals to collect it (ISL and Banister Research & Consulting, Inc., 2005).

Lastly, the personal interview method is a viable approach to collecting survey data. It is obviously better suited to a local survey, on no larger than a metropolitan area sized target population. While interviews could be set up with any of the above contact methods, the possibility also exists to conduct a roadside survey of passing vehicles. Such an attempt could be made at rest areas by employees of the study or, if given enough backing, the assistance of local authorities. The Alberta Ministry of Transportation (AMT) conducted one such study in 2001 with the help of its Inspection Services Branch (ISB), an organization tasked with inspection of large freight vehicles (Ishani and Meheboob, 2003). ISB vehicles established safe zones for pulling vehicles over on key routes and escorted freight vehicles to the shoulder. Reportedly truckers were largely relieved when told they were not being stopped for an inspection but for a survey. This is backed up by the response rate, which by internal definitions, resulted in only one refusal and two complaints out of 6,505 completed interviews and 6,771 observed trucks.

5.2.2. UIC Freight Establishment Survey

A previous UIC freight survey conducted in the spring of 2009 in the United States focused primarily on acquiring the data required to put together the basic components of the related freight microsimulation model, FAME (Samimi et al., 2010). Mode choice model was the main component of the framework that was drawn from this data, but the needs of the entire framework’s input were not met. For conduction of the 2009 survey, mail questionnaire, telephone interview and web-based methods were evaluated initially. lastly, the web-based method was selected due to the higher response rates, lower costs and more convenience factors. In 2010-2011, two new waves of surveys were conducted by the research team at UIC. The main purpose of the new waves of surveys was to fill in gaps and collect more detailed information on firms’ supply chain formation and logistics choices to provide the required input data for the development of the new framework’s models and components described in this proposal.

The task of selecting the appropriate data collection method is highly effective on the efficiency of the survey and quality of the data gathered (Brög and Meybur 1983). The selected method is usually resulted from a trade-off between the survey’s purposes and the available resources (Richardson et al. 1995). For conducting the new waves of UIC survey (2010-2011), a variety of techniques were employed to establish contact with potential freight survey participants and encourage them to complete an online survey record. These include telephone introductions, e-mail blast campaigns, and web crawling. Web-based survey method are generally more cost-effective than telephone interview methods but result in lower response rates compared to telephone interview surveys conducted by well-trained interviewers. Moreover, web-based surveys can include a variety of audio and visual components which can enhance the quality of survey questions (Couper et al., 2001). Furthermore, web-based surveys can be completed at any time of the day by the target participants logistics or shipping managers; therefore, they give the flexibility to the participants to complete the survey in their free time. The web-based surveys allows the automatically data collection in digital format which makes clean up and analysis easier. The web-based method was selected as the survey conduction method due to its benefits discussed above and the previous experience and success with that medium in UIC 2009 survey.

Initially a group of trained undergraduate students were recruited to call the potential participants, as it would be more memorable and persuasive, and instill a more pleasant attitude toward the survey. Approximately 27,600 telephone records were purchased from a data marketing company. Up to ten students at a time were employed in making contact with the survey prospects via Voice over Internet Protocol (VoIP) type tools. These calls were made between 10 AM and 4 PM, Monday through Thursday to try to maximize chances for contact. The callers would introduce themselves and the project, and then attempt to procure survey prospects e-mail addresses for further contact. Using an online survey management company, an online survey was created and the agreeable freight company contacts were delivered links to the survey via e-mail. These participants were delivered one introduction e-mail and up to two weekly reminder e-mails after that. This method was utilized for seven months from June to December of 2010 before it was decided that the approach was not working fast or cheap enough to meet the demands of the project. Also, as part of a brief experiment, an attempt at using web crawling methods to obtain contact information was run. It was discontinued due to lack of response.

Finally, the e-mail blast campaign which had proven successful in the previous UIC survey was applied. The new approach was faster than the web crawling approach and much cheaper than the expensive phone method of contact. An e-mail blast campaign was set up with the same data marketing company which targeted over 100,000 potential participants nationwide. In the campaign, much like before, e-mails were sent to the representatives of freight handling companies, one introductory e-mail, and two reminders. Unlike before, the e-mail addresses were simply rented from the data marketing company, and were sent by them without us being allowed to see the names of the recipients. In terms of final outcomes, this procedure was much more successful both fiscally and quantitatively. The approach brought us much nearer to our goals of sample size, but required an additional rental of 100,000 e-mail addresses to bring us to our goal. The two e-mail blast campaigns were conducted between February and April of 2011.

5.2.3. Survey Design and Results

The 2010-2011 online UIC establishment survey aimed for logistics or shipping managers of firms or someone with acceptable knowledge about shipping process of the firm as the potential survey participants. The survey included three major parts. First part asked about the relevant characteristics of the business establishment including location, employee size, value of total annual shipments, number of weekly inbound and outbound shipments, major suppliers and supply chains, etc. The second section of the survey asked for information on five most recent shipments and their attributes from participants. The information inquired very detailed and information about the attributes of individual shipments including origin, destination, mode of transportation, commodity type, value and weight of the shipment, etc. Most of these questions were similar to the previous UIC survey; however, some of the questions were modified and some new questions were added to the survey to obtain required data for the development of the new logistics components of the framework.

To estimate the behavioral disaggregate model of supplier evaluation and selection in the proposed framework, a detailed dataset on the characteristics of decision-makers and other logistics components of supply chain is required. Therefore, in the UIC establishment survey several questions were added to inquire this information. Also, to develop the disaggregate model components of the shipping chain configuration model, the survey explored the choice of shipping chain choice and its relevant logistics attributes such as the number of stops, stop type and waiting time at each stop. The results of these inquiries are explored in the next sub-section.

Approximately 219,000 contacts were attempted nationwide using all the various contact techniques including telephone introductions, web crawling, and e-mail blast campaigns. In total, 657 establishment surveys were collected which resulted in 970 useable shipment survey forms. The requirement of such a high figure to obtain 657 establishment surveys shows the difficulties involved in convincing a significant number of potential respondents to share what might be considered sensitive company information. The large numbers involved also indicate at what ease survey responses might suffer from non-response bias. Figure 4 and 5 present cumulative number of responses that were gathered using the email blast with reminder emails contact method. As it can be seen from the figures, each reminder of the e-mail blast driven waves of the survey served as a significant improvement to the total response count, represented by the sharp rises in the graphs.

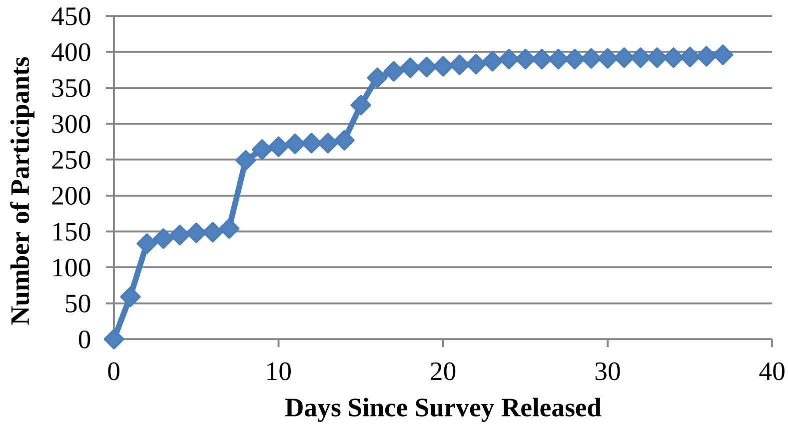


Figure 4 Cumulative number of responses of the first e-mail blast

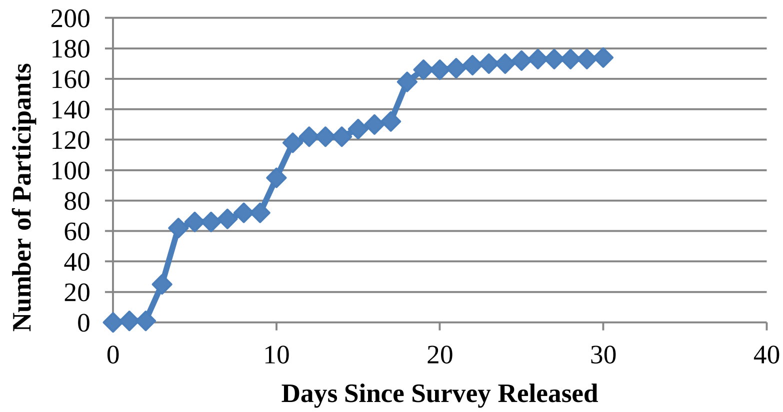


Figure 5 Cumulative number of responses of the second e-mail blast

5.2.4. Descriptive Analysis

Results were obtained from participants in 48 states and the District of Columbia. New Hampshire and Wyoming were the only two unrepresented states, though potential respondents in both states were targeted in the survey. Illinois was by far the most represented state, with 111 results from the various methods of contact. It was followed by Ohio with 53 results and

California with 47. Analysis of the geographic spread on a results per contact attempt basis was made on all three waves of the survey. Using this criterion, Missouri featured the highest response rate, at 5.02 records per thousand attempts. Illinois placed second with 4.63 and Indiana third with 4.11. The strength of the response level in Illinois and neighboring states is likely due to recognition of the university as being local and the survey therefore more relatable.

Respondents from a diverse range of industry types participated in the survey. However, a majority of respondents categorized their industry as manufacturing codes 31-33 of North America Industry Classification System (NAICS). This totaled 394, or 61.8% of the respondents. This was followed by Wholesale Trade industry (NAICS code 42) and Retail Trade industry (NAICS codes 44-45). The Mining, Quarrying, and Oil and Gas Extraction (NAICS 21) was the least represented value, with only 1% of total respondents. The large gap between types is not especially strange, as the survey specifically targeted the manufacturing industry above all others.

The employee size of companies polled leans towards smaller establishments, though with a wide range represented. 31.7% of survey takers reported belonging to companies with 20 or fewer employees, making up a plurality. 25.1% of survey takers reported belonging to companies with between 21 and 50 employees. These percentages continue to shrink as the employee size group increases in size. A much smaller figure of 5.4% of survey takers represents more sizable companies with greater than 1,000 employees.

Respondents were asked to report the number of major suppliers that their firm was in business with. Out of a total of 610 useable responses, a plurality of 33.4% identified the smallest range of suppliers, 0 to 5. This is followed by 23.1% with 6 to 10 suppliers and continues to decrease as the number of suppliers increase. Though the majority of respondents identified their number of major suppliers to be less than ten, a sizable group (about 27%) placed themselves in the larger range of up to 35 major suppliers.

A key goal of the survey was to understand the decision-making process of supplier selection and identifying the most important criteria affecting the choice of suppliers in a supply chain. Therefore, a list of eight potential criteria has been included in the e-mail blast driven wave of the survey using results of an extensive literature review. The potential variables include cost, credit and finance, delivery, distance and convenience, loyalty, management and service, manufacturing capacity and reliability, and quality and technology. Respondents were asked to rank every criterion on a scale from one to five. One represented low importance; three, medium; and five, high. Instead of simply asserting which characteristic was more valued, respondents had the ability to demonstrate how much more valued a characteristic is. The obtained average ranks are presented in figure 6.

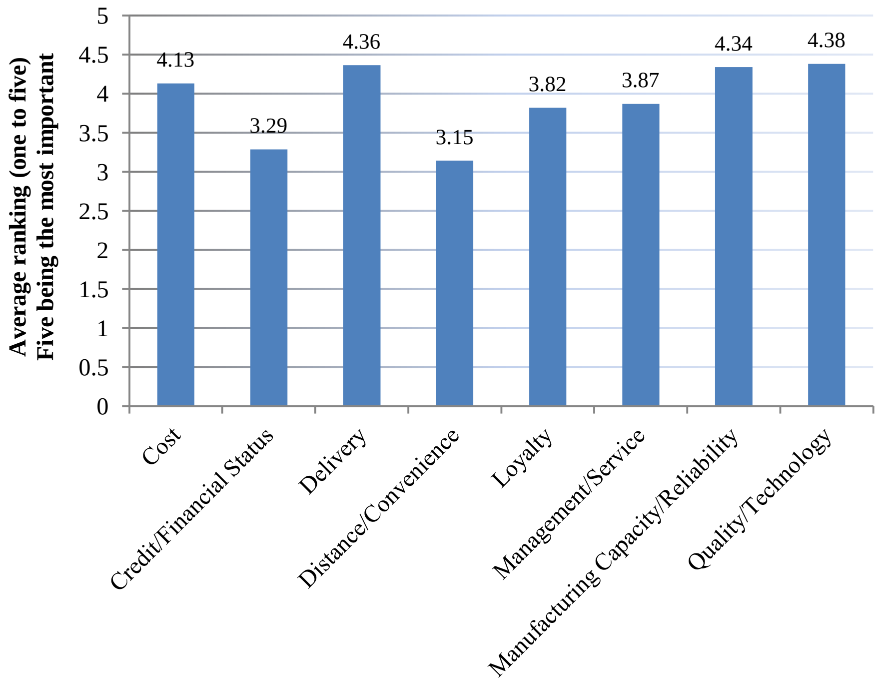


Figure 6 Average scores of supplier selection criteria.

As it is shown in Figure 6, the results depict a significantly different picture of how suppliers are chosen. All characteristics have been assigned values that average above the middle value of "medium," indicating that none are generally considered low importance. Quality and technology, delivery, manufacturing capacity, and cost are given high values with the average value of higher than four, though cost is ranked as lower than the other three factors. In addition, Figure 7 illustrates how respondents have scored each criterion and how the average values in Figure 6 are obtained.

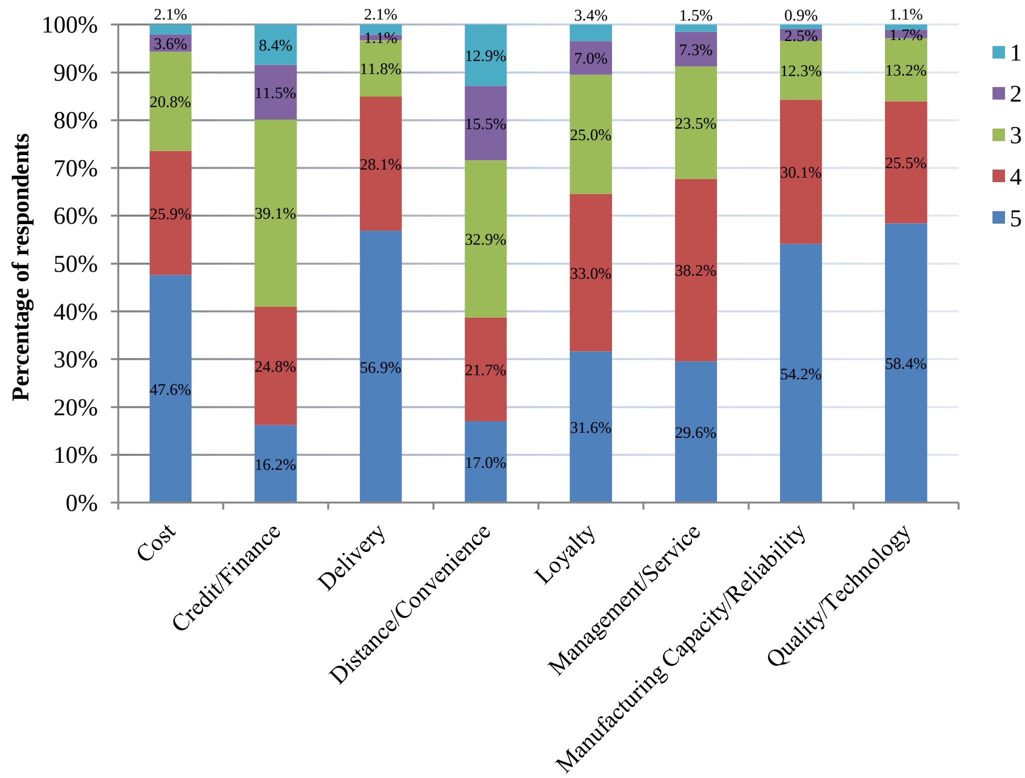


Figure 7 Rankings of supplier selection criteria.

To better understand the transportation logistic decisions made by firms, respondents were asked to provide detail information about attributes of five recent shipments and their logistics choices of those shipments. These include questions about the characteristics of shipments including commodity type, weight, value, volume, etc. and questions about the logistics choices associated with each shipment such as mode of transportation, use of intermediate handling facilities, type of activity at each intermediate handling facility, etc. Figure 8 presents surveyed shipments' characteristics including fragility, being perishable, time sensitivity, whether or not expedited, hazardousness, and liquid or dry bulk goods. The last three of these attributes are newly implemented compared to the previously run UIC freight establishment survey. Many respondents answered for some but not all of these categories for each shipment, so the number of responses ranges from 677 to 918 based on the categorical identifier.

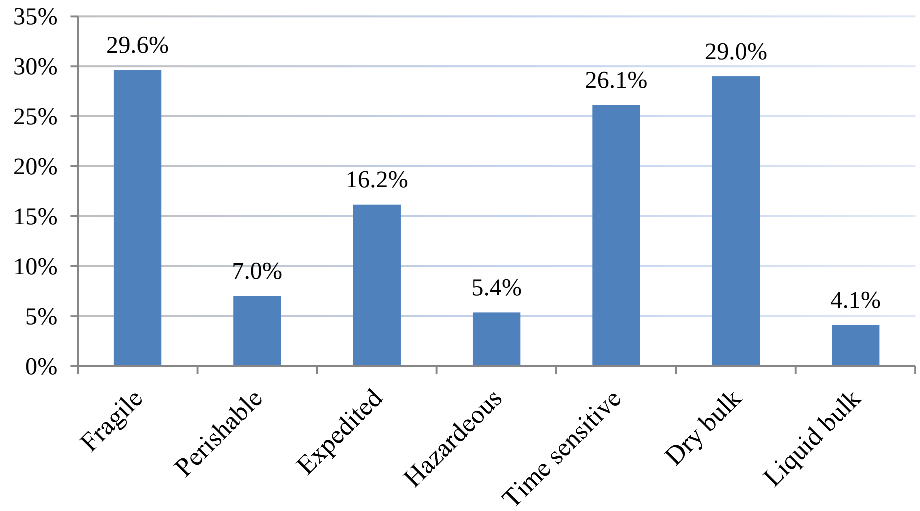


Figure 8 Shipments' characteristics

Since different industry groups were invited to participate in the survey, information of a diverse range of commodities was obtained. As illustrated in Figure 8, "Machinery and Metal products" has the highest share of 29%, while "Coal and minerals" have a share of only 1.0%. A considerable share of shipments are stated in the "Other" group of commodities by the participants, which includes commodities such as plastics, concrete products, etc. With the data coverage over a wide variety of commodity types, the demand model will be able to account for commodity heterogeneity, which is an essential issue specially for a behavioral model. Figure 9 presents the results.

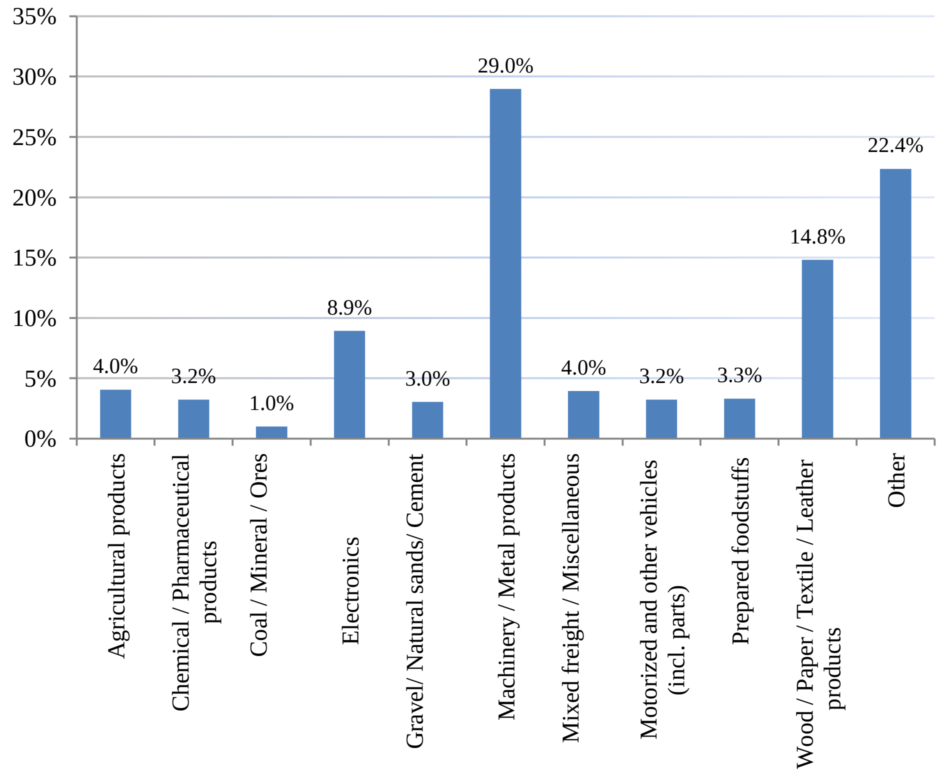


Figure 9 Commodity types in the survey.

Four different transportation modes are considered as possible choice for each shipment in the survey including truck, rail, air, and courier (including parcel, U.S. postal service or couriers such as UPS). In the survey we have gathered information on both direct and non-direct (with at least one intermediate stop) shipping chains. Therefore, some shipping chains with more than one transportation mode exist in the data set (for example, truck-rail-truck). For each shipment in the survey, one mode is selected as the main mode and reported in the following analysis. The main mode is identified as the one used for the longest distance in the shipping chain, from origin to destination. Figure 10 presents the considered main modes and the modal split in the data set based on this classification.

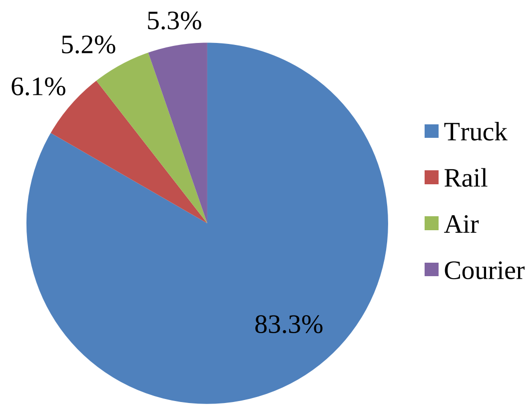


Figure 10 Modal split in data set.

The continuous shipment weight in the data set is also classified into 5 discrete shipment size groups ranging from very small to large shipments. Table 3 shows distribution of shipment size classes and average value of each class across the entire sample.

Table 3 Shipment Size Groups and Their Share in Data

Shipment size group	**weight range (lbs)**	**percentage**	**average value (lbs)**
1	Up to 200	20.7%	44.22
2	201-1,000	15.1%	563.76
3	1,001-4,000	13.4%	2191.65
4	4,001-30,000	24.8%	15409.47
5	Above 30,000	26.0%	50799.21

In addition, the split of shipment size over each transportation mode and modal split in each shipment size group are shown in Table 4 and Table 5. Figures in these tables suggest the general trend and interconnection between mode and shipment size choice to some extent. For example, it is clear that truck is dominant mode of transportation in all shipment size groups with small changes over different shipment sizes. Rail shipments mainly contain large shipment sizes in groups 4 and 5. On the other hand, only small shipments in groups 1 and 2 are transported with air.

Table 4 Distribution of Shipment Size (lbs) Over Transportation Modes

	Up to 200	**201-1,000**	**1,001-4,000**	**4,001-30,000**	**Above 30,000**
Truck	14.3%	16.8%	15.2%	25.6%	28.1%
Rail	0.0%	3.8%	8.8%	45.0%	42.5%
Air	82.4%	13.2%	1.5%	2.9%	0.0%
Courier	84.1%	4.3%	1.4%	10.1%	0.0%

Table 5 Modal Split in Shipment Size (lbs) Groups

	Truck	**Rail**	**Air**	**Courier**
Up to 200	57.6%	0	20.8%	21.6%
201-1,000	92.4%	1.5%	4.6%	1.5%
1,001-4,000	94.8%	4.0%	0.6%	0.6%
4,001-30,000	86.1%	11.1%	0.6%	2.2%
Above 30,000	90.0%	10.0%	0	0

In e-mail blast driven waves of survey, respondents were asked to determine the logistics choices relevant to the *shipping chain* choice for each shipment. These logistics choices include the number of stops, the type of stop and the mode used per each link between two consecutive stop. Moreover, some other information were gathered about the shipping cost and time for each link of the shipping chain and wait time at each stop. Based on the number and type of stops used for each shipment, different shipping chain categories can be defined.

From the 570 completed surveys in the two e-mail blast driven waves of survey, 504 useable shipment forms were obtained which include information about individual shipments. Out of 504 shipments, a plurality of 47.2% has direct shipping chains with no intermediate stop between origin and final destination. The frequency of shipping chains decreases as the number of stops increases per shipping chain. Figure 11 presents distribution of shipping chain type based on number of intermediate stops in the gathered data through survey. Since shares of three and four-stop shipping chains are very small in comparison with other shipping chain types, these two categories of shipping chain are combined with two-stop chains and considered as a single category (2+ stop) in the analysis.

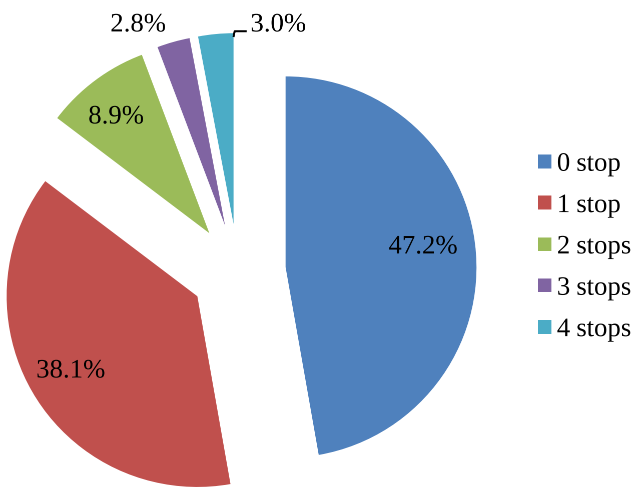


Figure 11 Distribution of different shipping chain types in the survey.

For each shipping chain, traveled distance is calculated as the Great Circle Distance (GCD) using the origin and destination zip code. The average GCD of shipping chains increases as the number stops increases per shipping chain. Figure 12 shows the average GCD for direct shipping chains with no intermediate stop, one-stop shipping chains and shipping chains with two and more (2+) stops. Table 6 presents distribution of shipment sizes (weight) over shipping chains with different number of intermediate stops. The total numbers in the last row and column of table presents actual frequencies in each group.

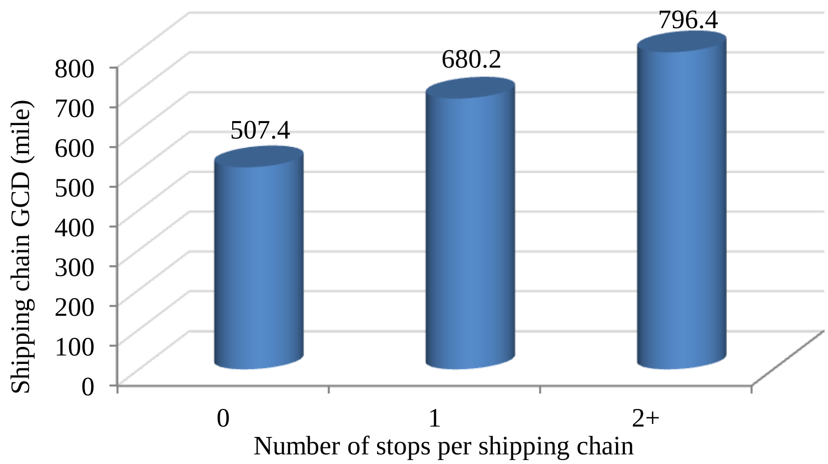


Figure 12 Average GCD for different types of shipping chains in survey.

Table 6 Distribution of Shipment Sizes Over Shipping Chains

Number of stops per shipping chain	**Shipment size (LBS)**					total
	0-200	**201-1000**	**1001-4000**	**4001-30000**	**>30000**	
0	41%	39%	55%	49%	55%	238
1	36%	51%	31%	43%	30%	192
2+	23%	10%	14%	7%	15%	74
total	129	97	77	95	106	504

Table 7 presents frequency of different main modes in shipping chains with different number of stops. As the table shows, air and rail are used mostly in shipping chains with one or more stops. Table 8 also presents distribution of different commodity types over shipping chains with different number of stops.

Table 7 Distribution of Different Main Modes in Shipping Chain

number of stop per shipping chain	**Air**	**Courier**	**Rail**	**Truck**	total
0	15%	79%	14%	48%	238
1	30%	15%	43%	40%	192
2+	55%	6%	43%	12%	74
total	20	34	21	429	504

Table 8 Distribution of Commodity Types Over Shipping Chains

number of stops per shipping chain	**Agricultural products**	**Chemical / Pharmaceutical products**	**Coal / Mineral / Ores**	**Electronics**	**Prepared foodstuffs**	**Gravel/ Natural sands/ Cements**	**Machinery / Metal products**	**Mixed freight / Miscellaneous**	**Motorized and other vehicles (incl. parts)**	**Wood / Paper / Textile / Leather products**	**Other**	total
0	50%	47%	33%	38%	17%	71%	47%	20%	71%	55%	52%	238
1	22%	36%	0%	33%	78%	21%	38%	76%	14%	34%	36%	192
2+	28%	17%	67%	29%	4%	7%	15%	4%	14%	11%	12%	74
total	18	36	3	48	23	14	133	25	14	91	99	504

Seven types of intermediate stops were considered in the survey including consolidation center, distribution center, warehouse, airport, port, truck/rail intermodal facility and an intermediate customer when shipments with different destinations are consolidated. In total, these seven types of intermediate stops were used 390 times in the shipping chains of all shipments with at least one stop. the most used intermediate stop in shipping chains belong to warehouses with more than 25% share of all 390 stops and the least type of stop showed in the shipping chains is an intermediate customer. Figure 13 presents distribution of each type of stop in the shipping chains.

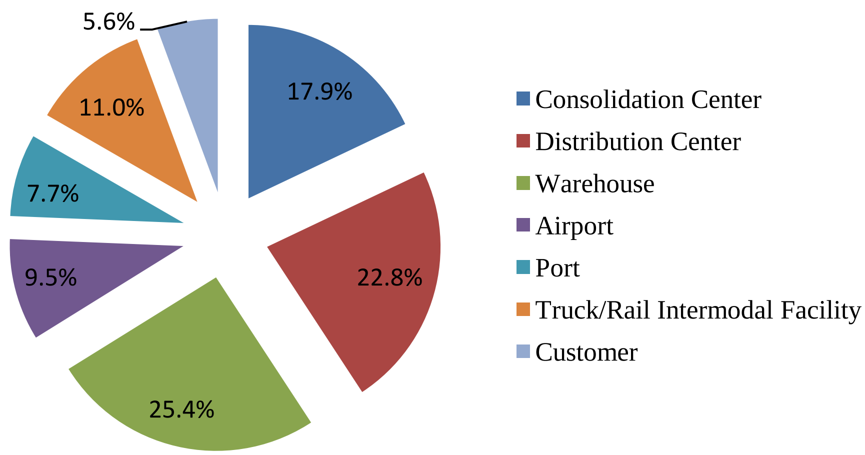
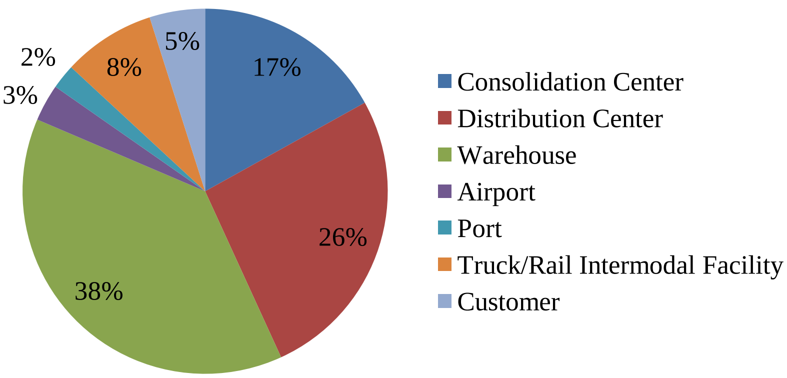
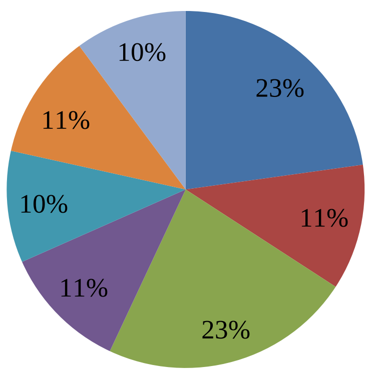


Figure 13 Distribution of different types of intermediate stops in survey.

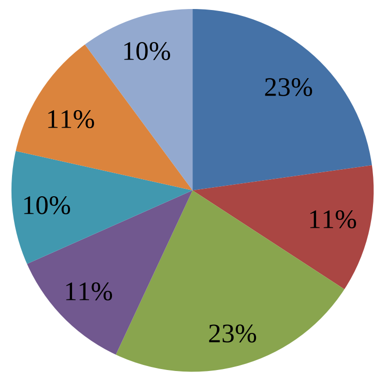
For more clarification on the configuration of shipping chains in the survey data, number of stops per shipping chain and type of stop were studied together. Figure 14 shows distribution of different type of intermediate stops used in shipping chains with different number of stops. It also presents the type of sequential stops in chains with more than one stop. Figure 14 (a) presents distribution of type of stops in shipping chains with only one stop. Figure 14 (b) shows distribution of stop type for the first stop in shipping chains with two or more stops and Figure 14 (c) presents distribution of stop type for the second, third and forth stops in chains with two or more stops.



(a)



(b)



(c)

Figure 14 Distribution of type of stops in shipping chains (a) distribution of type of stops in one-stop shipping chains (b) distribution of stop type for the 1st stop in 2+ stops shipping chains (c) distribution of stop type for the 2nd, 3rd and 4th forth stops in 2+ stops chains.

6. DEVELOPMENT OF MODEL'S COMPONENTS

6.1. Firm Synthesis Model

The Firm Synthesis Model is the first model of the framework in which individual agents in the study area are generated. The model considers individual firms as the decision-making agents in the framework. These firms or businesses establishments can be producer or receiver of goods who form supplier-buyer pairs and specify critical logistics choices of supply chains. However, since considering all existing firms in the study area results in computational complexity in the simulation process, an aggregation method is used to address this problem. Firms with similar characteristics are categorized as a group of agents, called firm-type. A firm-type is defined as a group of firms with the same industry type, employee size and geographic location in the zone system. It is assumed that all firms in a firm-type group have similar behavior in freight decision-making process. Therefore, they can be considered as one firm-type agent in the framework. The concept of firm-type used in the proposed model is the same as what was used in the FAME model (Samimi et al. 2010).

The key input data in the Firm Synthesis Model are the publicly available datasets that provide economic data including County Business Patterns (CBP) and Zip code Business Patterns (ZBP) (U.S. Census Bureau, 2009). These annual data series provide economic data by industry type for different geographic zone level. The data provide information on the number of business establishments by industry type and by employee size, employment information during the week of March 12, annual payroll and first quarter payroll. The CBP dataset provides this information at the county level and the ZBP dataset provides the information at the zip code level.

As it is shown in Figure 2, the zone system in this study consists of 118 township-level zones in the Chicago area, 95 counties in the rest of Illinois State and 120 FAF zones (FHWA, 2007) elsewhere. Since this zone system is not compatible with the geographic zone system used in the CBP and ZBP datasets, a procedure is applied to transform the CBP and ZBP information on business establishments from counties and zip code areas to the proposed zone system in this study.

For the 120 FAF zones that exist in the zone system of this study, the CBP data is used. Each FAF zone consists of several whole counties. Therefore, county-level information can be easily aggregated to obtain the corresponding information in the FAF zones. For the 95 counties in Illinois State, the exact same information provided by CBP is used without any conversion or aggregation. However, for the 118 township zones in the Chicago region, a GIS-based process is used to obtain the required information from the ZBP data. All the zip code areas that have their centroids in one township are combined and considered as the zip code areas that comprise the corresponding township. The ZBP information on business establishments for the zip code areas are then aggregated to obtain the required information for the corresponding townships. Only 15 townships in the outer part of the Chicago region were much smaller than the zip code areas that no zip code centroid was located in them. Therefore, no information on the business establishments was obtained for these townships.

The North American Classification System (NAICS) is used in the CBP and ZPB data to determine the industry type of business establishments. The economic data provided in CBP are available at aggregate 2-digits NAICS to 6-digits detailed level NAICS. However, the ZBP data only provides this information for the 2-gidits NAICS. Since the number of establishments decreases in the zip code areas, the information provided by ZBP are reported at a more aggregated level to keep the confidentially standards.

Analogous to the CBP and ZBP data, in this model, the NAICS codes are used to classify industry type of firm-types. 2-digits level NAICS data is too aggregate and inclusive to provide detailed information on the type of industry and produced/consumed commodities for the firms in the study. Therefore, 2-digits NAICS industry type classification is of little use as it provides limited information about the specification of the firms. On the other hand, 6-digits level NAICS data is too disaggregate that suffers from missing values and unreleased information due to confidentiality issue. In addition, selecting the 6-digit for industry type classification may results in computational complexity in the simulation process in the framework. Thus, the 3-digits level NAICS is selected to classify firms into firm-types and determine the industry type of the synthesized agents in the study.

The CBP data is used to extract the number of 3-digit NAICS business establishments by employee size in FAF zones and Illinois counties of the proposed zone system in this study. However, since the ZBP data only provides economic information for 2-digits NAICS industries, for the townships in the Chicago area, a more complicated procedure has to be used to extract the same information for 3-digits NAICS. an Iterative Proportional Fitting (IPF) method is used to generate the number of 3-digits NAICS establishments by employee size at zip code level. For each of the 7 counties in Chicago region, it is assumed that the distribution of 3-digits NAICS firms in zip code areas is the same as the distribution of 2-digits NAICS firms provided in ZBP.

Using this assumption the share of 3-digits NAICS establishments in zip code areas are obtained for each county. From the CBP data the total number of 3-digits NAICS firms is known at county level. The total numbers of 3-digits NAICS firms in each county are multiplied by the estimated share of 3-digit NAICS firms at zip code level to generate number of 3-digit NAICS firms in zip code areas of the Chicago region. Finally, to validate the results the estimated numbers of 3-digits NAICS establishments are aggregated to 2-digits NAICS and summed up for each zip code area. Then the sum value is compared to the reported number of 2-digits NAICS establishment in ZBP.

In total, 87 industry classes of 3-digits NAICS are considered in this study. Since the datasets provide no information on the businesses that are classified in the "Crop Production" NAICS 111 and "Animal Production" NAICS 112 group, these industries are excluded and no firm of these industry classes are generated in the Firm Synthesis Model. In addition, 7 classes of employee size are used for the classification of firms and defining the employee size of firm-types. These classes of employee size are the same employee sizes used in CBP and ZBP data. Table 9 shows the employee size classification used to categorize establishments and synthesize firm-types.

Table 9 Employee Size Classification in Firm Synthesis Model

Employee Size Class	Definition
1	Establishments with 1 to 19 employees
2	Establishments with 20 to 49 employees
3	Establishments with 50 to 99 employees
4	Establishments with 100 to 249 employees
5	Establishments with 250 to 499 employees
6	Establishments with 500 to 999 employees
7	Establishments with 1,000 employees or more

There are more than 7.6 million business establishments in the U.S. in 2007 (U.S. Census Bureau, 2009). The firm Synthesis model uses 333 zone, 87 industry classes and 7 employee size groups to categorize these business establishments. Using the firm-type concept, these firms are categorized into 70,116 firm-type groups which are considered as the agents in the simulation process. Each firm-type has a unique id that shows the geographic zone, industry type and employee size of the actual firms that comprise that firm-type group. Also, the number of actual firms in each firm-type group is kept in the id and used in the next model of the framework to estimates production and consumption values for firm-types.

6.2. Freight Generation Model

The Freight Generation Model uses economic data and a procedure to estimate the production and consumption values of different commodity types for the synthesized firms. The required data for the development of this model are discussed in Chapter 5. The methodology used for the estimation of production and a consumption values at firm level can be divided into two steps. In the first step a three-dimensional commodity-industry crosswalk is developed for the made and used commodities by 3-digits NAICS industry and in the second step, the developed crosswalk is used to apportion aggregate production and consumption values between firms.

As it is discussed in Chapter 5, the commodity-industry crosswalk provided by Industry Input-Output Accounts is the base input for development of the proposed crosswalk. However, the crosswalk cannot be directly used as input in this study for several reasons. First, the crosswalk figures are estimated based on the monetary transaction between industries not commodity flow. Second, the crosswalk uses a special system to classify industry groups which is not compatible with the NAICS system that is used in this study. Moreover, this crosswalk provides no explicit bridge between industry classes and commodity groups since it uses the same classification code for industry and commodity classes. Hence, a reliable linkage between industry classes and commodity groups has to be developed. Therefore, some essential adjustments are required before using the crosswalk which are presented in a flow diagram in Figure 15.

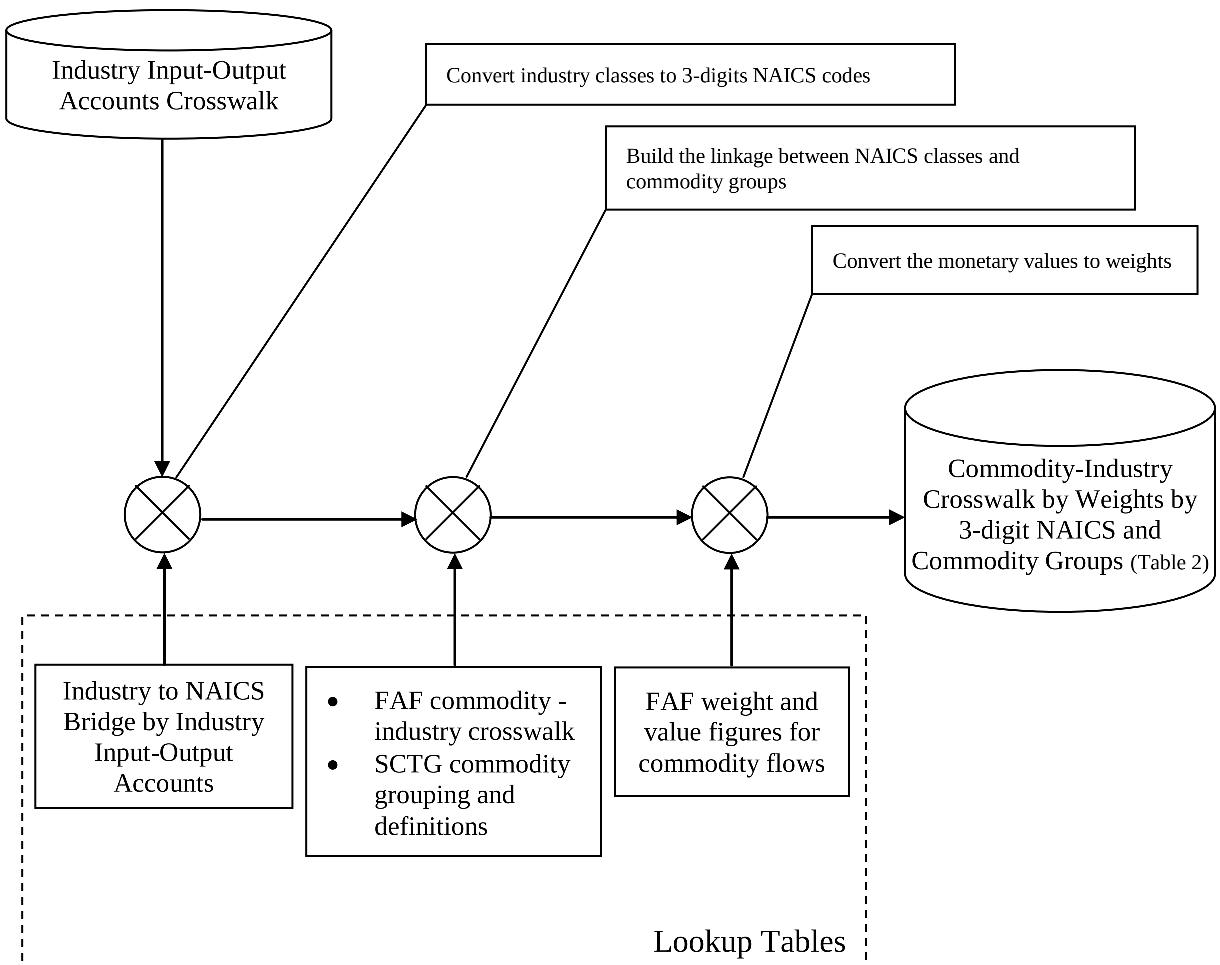


Figure 15 Development of commodity-industry crosswalk flow diagram.

As figure 15 shows, the industry to NAICS bridge provided by Industry Input-Output Accounts (Bureau of Economic Analysis, 2008) is used for converting industry classes to NAICS codes. The mapping between NAICS industries and SCTG commodities provided in FAF data (FHWA, 2011) and SCTG commodity grouping and definitions (Bureau of Transportation Statistics, 2007) are used to relate industry classes to commodity groups. Finally, the average unit value of commodity groups are estimated using the FAF data (FHWA, 2007). The FAF data provides weight and value of commodity flows between regions. For each commodity group, the weights and values of commodities are summed over regions and the average unit value ($ per lbs) of each group is estimated. The estimated average unit values of commodity groups, presented in Table 10, are used to convert the dollar values to commodity weights in the crosswalk.

Table 10 Average Unit Value of Commodity Groups

Commodity Class	**Definition**	**Unit Value ($/lbs)**
1	Agriculture and Forestry Products	0.317085
2	Products of Mining	0.012764
3	Petroleum Products	0.298583
4	Chemical and Pharmaceutical Products	1.031893
5	Wood Products	0.127216
6	Paper Products	0.794224
7	Nonmetallic mineral products	0.081471
8	Metal and Machinery Products	1.633708
9	Electronic, Electrical and Precision Equipments	10.02304
10	Motorized and Transportation Vehicles and Equipments	3.464202
11	Household and Office Furniture	2.587798
12	Plastic, Rubber and Miscellaneous Manufactured Products	1.736014
13	Textiles and Leather Products	5.136977
14	Waste and scrap	0.043154

The developed crosswalk using algorithm in Figure 15 is employed to estimate share of 3-digits NAICS industries in production and consumption of different commodity types. As it is mentioned before, commodity class 15, “Mixed and Unknown Freight”, is excluded from the model and simulation since the producer and consumer industries for this commodity type can be hardly identified. Table 11 and 12 presents these shares for produced and consumed commodities by industry class respectively.

Table 11 Percentage of Produced Commodity Groups By Industry Classes

NAICS \ Commodity Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14
111	14.7	0.0	0.0	0.0	0.6	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
112	14.3	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
113	0.0	0.0	0.0	0.0	16.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
114	0.7	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
115	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.7	0.0	0.0	0.0	0.0	0.0	0.0
211	0.0	0.2	31.5	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
212	0.0	98.8	0.0	0.3	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
213	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.4	0.0	7.9	0.0	0.0	0.0	0.0
321	0.0	0.0	0.0	0.0	81.5	0.0	0.0	0.0	0.0	0.0	0.4	0.1	0.0	0.7
327	0.0	0.6	0.0	0.7	0.0	0.0	92.7	0.0	0.0	0.0	0.0	0.1	0.3	0.4
331	0.0	0.0	0.0	0.1	0.0	0.0	0.1	22.5	0.0	0.0	0.0	0.0	0.0	19.0
332	0.0	0.0	0.0	0.1	0.3	0.0	0.1	29.2	0.0	0.1	0.5	0.5	0.0	24.5
333	0.0	0.0	0.0	0.1	0.0	0.0	0.1	29.1	0.2	0.2	0.1	0.2	0.0	5.4
334	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.1	24.6	0.4	0.1	2.4	0.0	0.0
335	0.0	0.0	0.0	0.0	0.0	0.0	0.1	1.5	7.1	0.0	0.0	0.0	0.0	0.0
336	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9.2	1.4	91.3	0.1	0.1	0.1	19.6
337	0.0	0.0	0.0	0.0	0.4	0.0	0.1	0.0	0.0	0.0	88.0	0.1	0.1	0.0
339	0.0	0.0	0.0	0.8	0.0	0.1	0.0	0.1	4.6	0.0	10.2	15.7	0.2	0.0
311	56.3	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0
312	13.8	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
313	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	36.7	0.2
314	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	31.1	0.0
315	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	23.6	0.0
316	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.5	0.0
322	0.0	0.0	0.0	0.1	0.1	51.8	0.0	0.0	0.0	0.0	0.0	0.3	0.2	5.5
323	0.0	0.0	0.0	0.0	0.0	21.1	0.0	0.6	0.0	0.0	0.2	0.2	0.2	3.8
324	0.0	0.3	68.4	5.1	0.0	0.0	6.3	0.1	0.0	0.0	0.0	0.0	0.0	0.0
325	0.2	0.0	0.1	92.1	0.0	0.1	0.0	0.1	0.1	0.0	0.0	28.0	0.0	0.0
326	0.0	0.0	0.0	0.0	0.3	0.4	0.1	0.2	0.0	0.1	0.3	51.2	0.8	1.7
511	0.0	0.0	0.0	0.0	0.0	23.5	0.0	0.0	9.1	0.0	0.0	0.0	0.0	0.0
512	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9.0	0.0	0.0	0.0	0.0	0.0
515	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.6	0.0	0.0	0.0	0.0	0.0
517	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	33.7	0.0	0.0	0.0	0.0	0.0
518	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.8	0.0	0.0	0.0	0.0	0.0
519	0.0	0.0	0.0	0.0	0.0	2.6	0.0	0.0	0.8	0.0	0.0	0.0	0.0	0.0
562	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	17.5
722	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	1.6
	100%	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

Table 12 Percentage of Consumed Commodity Groups By Industry Classes

NAICS \ Commodity Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14
111	2.5	1.3	1.1	3.7	0.2	0.2	0.0	2.1	0.1	0.1	0.0	0.5	0.4	0.0
112	10.8	1.0	0.6	0.4	0.1	0.1	0.0	0.7	0.1	0.0	0.0	0.1	0.1	0.0
113	0.0	0.0	0.1	0.0	2.9	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0
114	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0
115	0.2	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0
211	0.0	0.1	5.0	1.3	0.0	0.0	0.2	1.1	0.1	0.0	0.0	0.0	0.0	0.0
212	0.0	10.8	0.8	0.4	0.1	0.1	0.2	0.8	0.0	0.2	0.0	0.4	0.0	0.0
213	0.0	0.1	0.3	0.2	0.0	0.0	0.2	0.7	0.1	0.1	0.0	0.1	0.0	0.0
221	0.0	8.9	11.6	0.7	0.0	0.1	0.2	0.2	0.3	0.0	0.0	0.0	0.0	0.0
236	0.1	5.3	2.0	0.9	9.7	0.5	17.9	4.0	1.9	1.1	16.4	3.0	1.0	0.0
237	0.1	5.3	2.0	0.9	9.7	0.5	17.9	4.0	1.9	1.1	16.4	3.0	1.0	0.0
238	0.1	5.3	2.0	0.9	9.7	0.5	17.9	4.0	1.9	1.1	16.4	3.0	1.0	0.0
311	58.5	1.8	0.3	1.1	0.2	8.4	0.8	1.0	0.4	0.4	0.0	3.6	0.2	0.0
312	4.0	0.2	0.1	0.1	0.2	1.8	2.8	1.3	0.2	0.2	0.0	2.3	0.1	0.0
313	0.5	0.1	0.0	0.3	0.0	0.2	0.1	0.0	0.1	0.0	0.0	2.5	9.5	0.0
314	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.1	0.0	0.0	1.8	12.9	0.1
315	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.1	10.2	0.0
316	0.1	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	1.7	0.0
321	0.0	0.0	0.1	0.4	25.9	0.4	0.8	0.5	0.3	0.4	2.3	0.6	1.4	0.0
322	0.1	2.4	0.3	2.1	5.6	19.5	0.2	0.6	0.4	0.2	0.0	1.2	4.3	9.6
323	0.2	0.0	0.3	1.3	0.0	6.1	0.0	0.6	0.3	0.1	0.0	0.2	2.7	0.0
324	0.0	1.4	49.1	2.3	0.0	0.1	0.9	0.1	0.1	0.1	0.0	0.1	0.1	0.0
325	1.5	5.0	3.5	48.9	0.5	3.2	1.1	1.2	0.8	0.3	0.0	8.2	0.0	0.1
326	0.0	0.2	0.1	2.8	2.5	2.3	1.1	1.0	0.5	0.2	0.9	21.7	4.4	0.2
327	0.0	11.7	0.2	1.3	0.3	0.9	11.7	0.5	0.3	0.2	0.0	0.5	0.6	0.6
331	0.0	27.3	0.2	0.4	0.3	0.7	1.7	9.4	0.5	0.3	0.0	0.4	0.0	88.3
332	0.0	0.7	0.1	1.9	0.2	1.1	0.3	13.6	1.1	0.5	0.0	1.1	0.0	1.0
333	0.0	0.2	0.2	1.0	0.5	1.2	0.8	11.4	3.0	2.8	1.0	3.9	1.7	0.0
334	0.0	0.1	0.1	1.5	0.5	0.9	0.5	2.8	11.7	0.5	1.7	1.9	0.0	0.0
335	0.0	0.3	0.2	0.3	0.2	0.7	1.1	3.6	1.7	0.1	0.0	1.7	0.0	0.0
336	0.0	0.3	0.1	1.4	2.2	1.6	4.0	20.3	5.2	74.0	1.1	9.9	10.1	0.0
337	0.0	0.0	0.0	0.4	5.8	0.7	0.1	0.9	0.3	0.1	9.3	1.9	6.8	0.0
339	0.1	0.3	0.1	0.6	0.9	1.1	0.3	1.6	1.2	0.2	1.9	4.5	4.6	0.0
423	0.0	0.0	0.2	0.1	0.4	1.3	0.0	0.1	0.8	0.3	0.5	0.7	1.0	0.0
424	0.0	0.0	0.2	0.1	0.4	1.3	0.0	0.1	0.8	0.3	0.5	0.7	1.0	0.0
425	0.0	0.0	0.2	0.1	0.4	1.3	0.0	0.1	0.8	0.3	0.5	0.7	1.0	0.0
441	0.0	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.3	0.2	0.3	0.2	0.1	0.0
445	0.1	0.0	0.1	0.0	0.3	0.5	0.0	0.0	0.2	0.3	0.0	0.5	0.2	0.0

Table 12 Percentage of Consumed Commodity Groups By Industry Classes (Continue)

NAICS \ Commodity Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14
442	0.0	0.0	0.0	0.0	0.1	0.3	0.0	0.0	0.2	0.1	0.2	0.2	0.4	0.0
443	0.0	0.0	0.0	0.0	0.1	0.3	0.0	0.0	0.2	0.1	0.2	0.2	0.4	0.0
446	0.0	0.0	0.0	0.0	0.1	0.3	0.0	0.0	0.2	0.1	0.2	0.2	0.4	0.0
447	0.0	0.0	0.0	0.0	0.1	0.3	0.0	0.0	0.2	0.1	0.2	0.2	0.4	0.0
448	0.0	0.0	0.0	0.0	0.1	0.3	0.0	0.0	0.2	0.1	0.2	0.2	0.4	0.0
451	0.0	0.0	0.0	0.0	0.1	0.3	0.0	0.0	0.2	0.1	0.2	0.2	0.4	0.0
453	0.0	0.0	0.0	0.0	0.1	0.3	0.0	0.0	0.2	0.1	0.2	0.2	0.4	0.0
454	0.0	0.0	0.0	0.0	0.1	0.3	0.0	0.0	0.2	0.1	0.2	0.2	0.4	0.0
452	0.0	0.0	0.1	0.0	0.1	0.1	0.0	0.0	0.2	0.0	0.0	0.1	4.6	0.0
481	0.0	0.0	4.3	0.0	0.0	0.0	0.0	0.1	0.2	0.1	0.0	0.0	0.0	0.0
482	0.0	0.0	1.1	0.0	1.1	0.0	0.0	0.2	0.0	0.8	0.0	0.0	0.0	0.0
483	0.0	0.0	1.1	0.0	0.0	0.0	0.0	0.2	0.1	0.5	0.0	0.0	0.1	0.0
484	0.0	0.0	5.3	0.1	0.1	0.3	0.0	0.4	0.6	1.8	0.0	0.5	0.0	0.0
485	0.0	0.0	0.4	0.0	0.0	0.1	0.0	0.1	0.1	0.2	0.0	0.0	0.0	0.0
486	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0
487	0.0	0.0	0.1	0.1	0.0	0.1	0.0	0.2	0.0	0.1	0.0	0.0	0.0	0.0
488	0.0	0.0	0.1	0.1	0.0	0.1	0.0	0.2	0.0	0.1	0.0	0.0	0.0	0.0
491	0.0	0.0	0.2	0.0	0.0	0.1	0.0	0.1	0.0	0.7	0.0	0.0	0.1	0.0
492	0.0	0.0	0.9	0.0	0.1	0.5	0.0	0.1	0.1	0.3	0.0	0.8	0.1	0.0
493	0.0	0.0	0.1	0.0	0.0	0.1	0.0	0.1	0.2	0.1	0.0	0.1	0.0	0.0
511	0.0	0.1	0.0	0.1	0.1	7.8	0.0	0.4	2.3	0.2	0.0	0.2	0.0	0.0
512	0.0	0.0	0.0	0.0	0.1	0.6	0.0	0.0	3.4	0.0	0.0	0.0	0.4	0.0
515	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.2	3.1	0.1	0.0	0.1	0.0	0.0
517	0.0	0.2	0.0	0.2	0.2	0.3	0.5	0.6	16.9	0.1	0.2	0.4	0.3	0.0
518	0.0	0.0	0.0	0.0	0.6	0.2	0.5	0.1	0.8	0.0	0.9	0.1	0.0	0.0
519	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0
521	0.0	0.0	0.1	0.0	0.0	0.5	0.0	0.0	0.5	0.0	0.1	0.0	0.0	0.0
522	0.0	0.0	0.1	0.0	0.0	0.4	0.0	0.0	1.2	0.1	0.1	0.0	0.0	0.0
523	0.0	0.0	0.0	0.0	0.0	2.7	0.0	0.0	4.3	0.1	0.1	0.0	0.0	0.0
524	0.0	0.0	0.0	0.0	0.0	2.4	0.0	0.0	0.7	0.0	0.0	0.2	0.0	0.0
525	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
531	0.1	5.5	0.2	1.2	13.2	0.4	6.7	2.0	1.6	0.0	22.3	2.0	0.7	0.0
532	0.0	0.5	0.2	0.1	0.4	0.1	0.1	0.2	0.2	0.2	0.2	0.2	0.2	0.0
533	0.0	0.0	0.0	0.0	0.0	0.7	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0
541	0.2	0.9	0.2	2.2	0.4	5.9	2.3	1.2	6.2	1.2	0.6	2.0	0.5	0.0
551	0.0	0.1	0.3	0.6	0.0	0.8	0.1	0.2	2.7	0.2	0.0	0.1	0.0	0.0
561	0.1	0.3	0.5	1.2	0.4	2.7	0.6	0.6	3.0	1.2	0.0	0.8	0.7	0.0
562	0.0	0.1	0.2	0.2	0.0	0.1	0.1	0.7	0.4	0.4	0.0	0.1	0.4	0.0

Table 12 Percentage of Consumed Commodity Groups By Industry Classes (Continue)

NAICS \ Commodity Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14
611	0.9	0.3	0.1	0.2	0.0	0.7	0.1	0.2	1.1	0.1	0.0	0.4	0.1	0.0
621	0.0	0.0	0.1	7.1	0.1	2.0	1.5	0.1	2.8	0.1	0.3	1.9	0.1	0.0
622	1.4	0.1	0.4	5.9	0.1	2.4	0.6	0.1	2.2	0.0	0.0	0.9	1.1	0.0
623	0.9	0.0	0.2	0.5	0.0	0.6	0.1	0.0	0.4	0.0	0.0	0.4	1.4	0.0
624	0.7	0.0	0.1	0.2	0.4	1.1	0.1	0.1	0.4	0.3	0.2	0.6	0.9	0.0
711	0.1	0.2	0.0	0.0	0.0	0.2	0.0	0.0	0.2	0.0	1.4	0.0	0.2	0.0
712	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
713	0.4	0.6	0.1	0.2	0.1	0.2	0.1	0.1	0.2	0.0	0.7	0.1	0.9	0.0
721	0.4	0.1	0.2	0.0	0.3	1.4	0.1	0.1	0.3	0.0	0.0	0.1	1.0	0.0
722	15.2	0.3	0.5	0.2	1.7	2.0	2.7	0.8	1.0	0.4	1.9	2.4	0.8	0.0
811	0.0	0.1	0.1	0.4	0.0	0.3	0.2	1.6	1.8	3.7	0.3	1.8	0.5	0.0
812	0.0	0.4	0.1	0.7	0.0	0.7	0.4	0.1	0.4	0.1	0.1	0.7	1.4	0.0
813	0.1	0.1	0.1	0.0	0.1	0.6	0.0	0.0	0.4	0.0	0.0	0.1	1.9	0.0
814	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	0.0
	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

Although the crosswalks are estimated based on the values of produced and consumed commodities at national scale, it is assumed that the estimated shares are valid for smaller regions. Therefore, the crosswalks can be used to estimate commodity production and consumption rates by industry at FAF zone, county and township level. In other word, it is assumed that the production and consumption rates do not change for firms in different regions when the regional commodity production and consumption values are breaking down between them.

Employee size is another important factor in firm-types' characteristics and it should be considered when estimating production/consumption amounts at firm level. Since the production and consumption values change by the size of establishment, another crosswalk is developed to capture the commodity production and consumption share of establishments by employee size. The UIC establishment survey is used for the development of this crosswalk. In the survey, the establishments' annual commodity production and consumption data were collected. These values are used to estimate average commodity production and consumption share of establishments by employee size. Table 13 presents the estimated crosswalk.

Table 13 Commodity Production and Consumption Share By Employee Size

Employee Size Class	Number of Employees	Production Share	Consumption Share
1	1 to 19	0.7%	0.3%
2	20 to 49	1.5%	2.9%
3	50 to 99	3.6%	3.9%
4	100 to 249	10.2%	8.6%
5	250 to 499	22.8%	12.9%
6	500 to 999	27.9%	19.6%
7	1,000 or more	33.4%	51.9%

Using the developed crosswalks in table 11, 12 and 13, zone level commodity production and consumption values provided from FAF data are apportioned between synthesized firms in the study area. The zone level production and consumption values are obtained by aggregating domestic FAF commodity flows (FHWA, 2007) over destinations and origins respectively. The model searches for producers and consumers of commodities in each zone and allocate the zone level production and consumption values between firms based on their production and consumption shares obtained from crosswalks in table 11, 12 and 13. In addition, the total number of actual firms in each firm-type group is considered in the allocation process. The production and consumption shares are weighted by the number of actual firms in each firm-type group. Thus, firm-types with more actual firms produce and consume more amounts of commodities than firm-types with less number of actual firms. The result of this model is firm level production and consumption values which are used as input data in the Logistics Decisions layer where producer and consumer are paired as supplier and buyer and annual commodity flows between them is estimated.

6.3. Supplier Evaluation and Selection Model

6.3.1. Multi-criteria Supplier Evaluation Methodology

Providing rating is among the common questions asked from participant in surveys. The question, asked in the UIC survey about the importance rate of different criteria in selecting suppliers, is an example of this type questions. As it is discussed by Train (2003) the key characteristic of rating questions, from a modeling point of view, is that the potential responses are ordered. Due to the ordered nature of alternatives (rates), some sort of dependency exists among them. There is more similarity between closer alternatives and more dissimilarity between alternatives that are further away. Therefore, in contrary to standard logit structure, the assumption of independent error terms for each alternative (IIA) does not apply here and a standard logit specification cannot handle the ordered nature of this problem. Some alternative modeling specifications include nested logit, mixed logit, or probit model that takes into account the similarity and variation patterns among the alternatives. However, a more accepted specification is the ordered logit which uses the logistic distribution for ordered alternatives. This modeling approach is used in this study to rate the important criteria in supplier selection problem.

The decision-making process in the ordered logit model is based on this assumption that respondents have some level of utility or opinion associated with the objective of question and use this associated utility to make their choice (Train, 2003). For example, in the UIC survey, it is assumed that the respondent has an opinion on how important each criterion is in selecting suppliers. Variable U is assigned to this opinion where higher levels of U mean that the respondent thinks the criterion is more important and lower levels mean in respondent's opinion the questioned criterion is less important in selecting suppliers. In the survey, respondents were asked to express their opinion on the importance of each criterion using a scale from one to five. One represented the lowest importance; three, medium; and five, the highest. Although U can take many different values representing various levels of importance of the criterion in selecting suppliers, the question allows only five responses. Respondent's decision making process, presented in Table 14, is based on the level of his U in which respondent considers some cutoff to choose his/her answer

Table 14 Decision Making Process in The Survey

Selected importance score	**Condition**
5	If $U > k_4$
4	If $k_3 < U < k_4$
3	If $k_2 < U < k_3$
2	If $k_1 < U < k_2$
1	If $U < k_1$

Many observed and unobserved factors affect decision maker's (respondent's) opinion on the importance of different criteria in supplier selection problem. Observed factors include characteristics of decision maker and buyer business establishment including employee size, industry type, number and amount of inbound shipments, etc. Assume that U, the random utility level accrued by decision maker for each criterion, include observed and unobserved components and can be written as follows.

$$U = \beta' X + \varepsilon \tag{6-1}$$

where β is a vector of coefficients that corresponds to the vector of independent variables X (observed factors) and ε represents the unobserved factors (error terms). Distribution of error terms determines the propensity of the five potential responses. Figure 16 depicts the distribution of U around $\beta' X$ in which the shape of distribution depends on the distribution of ε. The variables $k_1, \dots, k_4$ represent cutoff points for the potential responses. As the figure illustrates, the probability of choosing rate 1 for a criterion is equivalent to the probability of U being less than k_1 which is the area in the left tail of the distribution. Similarly the probability of choosing other alternatives can be obtained.

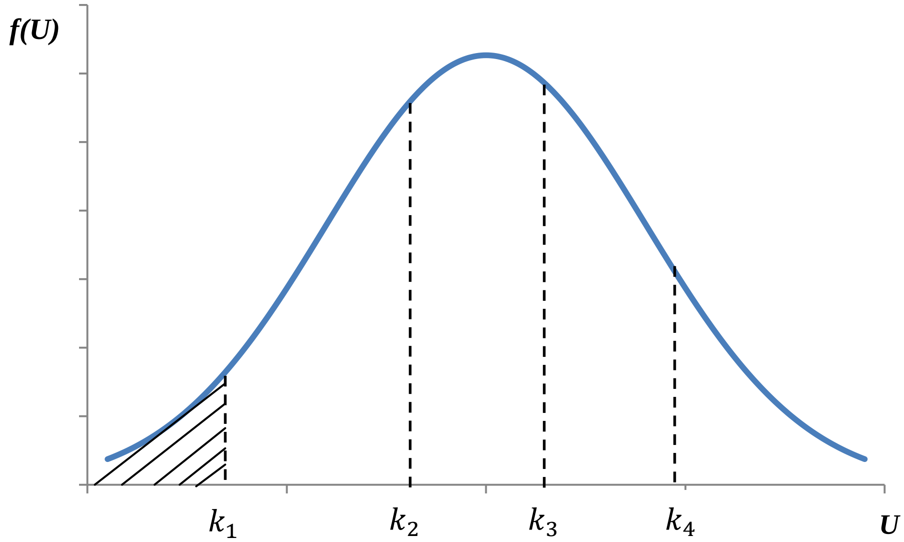


Figure 16 Distribution of U and cutoff points.

To obtain the propensity of choosing a potential rank, only the distribution of ε is needed. In this study, it is assumed that error terms have logistics distribution with a cumulative distribution function given by *F(.)* which is defined as follows.

$$F(\varepsilon) = \frac{e^{\varepsilon}}{1+e^{\varepsilon}} \tag{6-2}$$

For example, the probability of choosing 5 as an answer to the rating question can be calculated as follows.

$$\begin{aligned} P(5) &= P(U > k_4) = 1 - P(U < k_4) \\ &= 1 - P(\beta' X + \varepsilon < k_4) \\ &= 1 - P(\varepsilon < k_4 - \beta' X) \\ &= 1 - \frac{e^{k_4 - \beta' X}}{1 + e^{k_4 - \beta' X}} \end{aligned} \tag{6-3}$$

The probabilities enter the log-likelihood function and parameters are estimated by maximizing the log-likelihood using an econometrics software. The parameters to be estimated include coefficient vectors β, which gives the effect of explanatory variables on the importance of criteria and the cutoff points $k_1, \dots, k_4$.

Once the model parameters are estimated, the model is used to evaluate suppliers by translating measures under multiple criteria into a unique score using the obtained probabilities of importance rating for all criteria. Assume that decision maker n has a subset of I_n $(i = 1, 2, 3, \dots, I)$ for potential suppliers and there are J criteria for evaluating these potential suppliers. Let r_{nj} be the obtained importance rate of criterion j $(j = 1, 2, \dots, J)$ for decision maker n which is estimated using the ordered logit model. r_{nj} is transformed to the normalized weight of the criterion j for decision maker n, w_{nj}, using the simple transformation formula in the following.

$$w_{nj} = \frac{r_{nj}}{\sum_{j=1}^{8} r_{nj}} \tag{6-4}$$

Assume that x_{ij} $(j = 1,2, \dots 8)$ presents the measure for supplier i under criterion j. It should be noted that x_{ij} $(\forall\, j; j = 1,2, \dots 8)$ can be either positively or negatively related to the score of supplier i. Therefore, the negatively related measures are multiplied by minus one to account for their negative impact on the supplier score. In addition, to obtain more consistent results, all measures of supplier i under all criteria (x_{ij} $(\forall\, j; j = 1,2, \dots 8)$) are normalized into a common 0-1 scale (-1 to 0 scale for negative measures) using a linear transformation method. Therefore, no criterion takes domination over others due to the large scale of measure (Lung Ng, 2008). The score of supplier i under all criteria for decision maker n, S_{ni} can be obtained as follows.

$$S_{ni} = \sum_{j=1}^{8} w_{nj} x_{ij} \tag{6-5}$$

Using the obtained scores for all potential suppliers, decision makers can evaluate suppliers and identify those with highest scores.

6.3.2. Supplier Evaluation Model Estimation Results

For each of 8 criteria presented in the UIC survey and in Figure 6, an ordered logit models is developed using SAS econometrics software (SAS Institute Inc., 2008). For each model, the Score Test for the Proportional Odds Assumption is performed by the software which is the Chi-Square Score Test for the Proportional Odds Assumption. This test investigates if estimating one equation model for all levels (rates) of the dependent variable is valid or not. For all models, we failed to reject the null hypothesis based on the result of this test. Therefore, it can be concluded that assuming equal coefficients across all levels of the dependent variable and estimating only one equation model is a valid assumption.

The models' fit is assessed by the Akaike Information Criterion (AIC) and Schwarz Criterion (SC) which are based on the likelihood function at convergence. The calculated AIC and SC criteria are used to evaluate the estimated models' fit to the data. The following equation is used to calculate AIC.

$$AIC = -2ln(L) + 2((k-1) + s) \tag{6-6}$$

where L is the maximum value of likelihood function at convergence, k is the number of levels (rates) of the dependent variable and s is the number of exogenous variables that are estimated in the model. SC is defined using the following equation.

$$SC = -2ln(L) + ((k-1) + s) \times \ln\left(\sum f_i\right) \quad (6\text{-}7)$$

where $fi's$ are the frequency of the i^{th} observation, and k and s are defined as above. As it can be seen in equations (6-6) and (6-7). AIC and SC are used to compare the estimated models. Ultimately, the model with the lowest AIC and SC is selected as the preferred model with the best fit.

In addition, the global null hypothesis is tested using three asymptotically equivalent Chi-Square tests: Likelihood Ratio, Score and Wald test. This hypothesis implies that all of the estimated parameters for exogenous variables in the model are equal to zero. The Likelihood Ratio Chi-Square test, which is calculated using the following equation, examines that at least one of the estimated parameters for exogenous variables is not equal to zero in the estimated model.

$$LR = -2\,ln(L_{null}) - 2ln\,(L_{fitted}) \quad (6\text{-}8)$$

where LR is the Likelihood Ratio Chi-Square test value, L_{null} is the maximum value of likelihood function of the model with no predictors and intercept only, and L_{fitted}, refers to the value of likelihood function of the model with intercept and all independent variables. The Score and Wald Chi-Square tests also examine if at least one of the estimated parameters for exogenous variables is not equal to zero in the estimated model.

Table 15 presents results of model estimation for rating of all 8 criteria. Exogenous variables that are considered in the models estimation are decision-makers' (i.e., firms) characteristics, including industry type, production/consumption rates, and number of employees. P-values for estimated coefficients of predictors are presented in parentheses in the table. For the model estimation the ordered value of the dependant variable was set in descending order in SAS software (SAS Institute Inc., 2008). Thus, a positive estimated parameter is positively related to the importance rate of criterion which means that an increase in the value of the corresponding explanatory variable results in higher rate of the criterion. Similarly, a negative estimated parameter is negatively related to the importance rate of criterion and an increase in the value of respective explanatory variable results in lower rate estimate for criterion. Also, it should be noted that the model constant for each importance rate (1, 2, …, 5) is specified with the corresponding number. For example, “Intercept 1” refers to the intercept of the rate 1 with lowest importance value.

Once the required information and measures under the criteria for potential suppliers are known (e.g. cost or price of order), the suitability scores of suppliers under multiple criteria can be calculated using the equations (6-4) and (6-5). Using the obtained scores, potential suppliers can be ranked based on their scores and decision-makers can identify the most suitable candidates that meet their criteria.

Table 15 Ordered Logit Model Estimate For Importance Rating of Supplier Evaluation Criteria

Variables \ Criterion	Cost	Credit/ Financial Status	Delivery	Distance/ Convenience	Loyalty	Management/ Service	Capacity/ Reliability	Quality/ Technology
Model Constant								
Intercept 1	-0.3039 (0.17)	-1.6509 (0.00)	0.4454 (0.02)	-2.0280 (0.00)	-0.9023 (0.00)	-1.8005 (0.00)	-0.4290 (0.15)	0.5241 (0.00)
Intercept 2	0.8297 (0.00)	-0.2977 (0.27)	1.9999 (0.00)	-0.8842 (0.00)	0.5495 (0.00)	-0.0851 (0.76)	1.1621 (0.00)	1.8603 (0.00)
Intercept 3	2.6240 (0.00)	1.5475 (0.00)	3.8827 (0.00)	0.5300 (0.05)	2.1817 (0.00)	1.5957 (0.00)	2.7760 (0.00)	3.7563 (0.00)
Intercept 4	3.6627 (0.00)	2.5639 (0.00)	4.3449 (0.00)	1.5138 (0.00)	3.5671 (0.00)	3.5179 (0.00)	4.0389 (0.00)	4.6205 (0.00)
Industry type								
Agriculture, Forestry, Fishing and Hunting (NAICS 11)	-	-	-	-	-	-1.5721 (0.01)	-1.0328 (0.08)	-
Construction (NAICS 23)	1.7117 (0.11)	-	-	-	-	-	-1.7381 (0.02)	-0.8759 (0.20)
Manufacturing (NAICS 31-33)	0.3604 (0.06)	0.3079 (0.22)	-	-	-	-	-	-
Wholesale Trade (NAICS 42)	-	0.3697 (0.26)	-	-0.3404 (0.14)	0.7795 (0.00)	-	0.3502 (0.21)	-
Retail Trade (NAICS 44-45)	-	0.2776 (0.43)	-	-	0.4192 (0.16)	0.5675 (0.05)	-	-
Transportation and Warehousing (NAICS 48-49)	-	-	-1.0692 (0.13)	0.6466 (0.28)	-	-	-	-0.6467 (0.30)
Information (NAICS 51)	-	-	-	1.0559 (0.13)	-	1.7459 (0.02)	-2.2909 (0.02)	-
Production/Consumption								
Average annual value of inbound shipments	2.168E-9 (0.24)	-	-901E-12 (0.24)	-	-	-	-	-
Average annual value of outbound shipments	-1.15E-9 (0.19)	-	-	-	-	-	-264E-12 (014)	-
Number of Employee								
1-50	-0.2877 (0.15)	-0.7581 (0.00)	-0.3842 (0.07)	-	-	-	-	-0.2919 (013)
>250	-	-	-	-	-0.3367 (0.21)	-	-	-
Log-Likelihood	-528.6	-720.0	-428.2	-797.7	-630.7	-610.3	-478.8	-525.1

It should be noted that some of the evaluation criteria including Delivery, Loyalty, Management/Service, and Quality/Technology are unmeasured factors and cannot be easily quantified. Therefore, only those criteria that could be quantified including Cost/Price, Distance/Convenience, Capacity/Reliability, and Credit/Financial Status are considered in this study for evaluating supplier firms. The unit value of produced (supplied) commodity is used as measure under Cost/Price, production limit of supplier firms is used as measure under Capacity/Reliability, and annual value of produced commodities is used as proxy measure for Credit/Financial Status of suppliers.

6.3.3. Multi-criteria Supplier Evaluation Model Structure

The obtained suitability scores can be used by decision makers to evaluate suppliers and rank them. It should be considered that a specific supplier might be the potential supplier of different buyers and due to the capacity constraints; this supplier might be unable to cover the demand of all buyers. Moreover, another key logistics component in supplier selection is the consideration of transportation cost. For example, a potential supplier might have a high suitability score. However, it might lose its attraction when considering the transportation costs associated with shipments from this supplier. As it was mentioned before, the cost criterion considered in supplier evaluation step refers to the order cost and price of the shipped commodity and should be differentiated from the logistics costs associated with shipping and handling costs.

Using the results of the importance rate model and calculated suppliers' scores, a multi-objective mathematical optimization model is developed in this study to identify the best suppliers considering their suitability score, production capacity and transportation costs. The proposed multi-objective linear model is formulated as follows.

$$Maximize\ z_1 = \sum_{n \in N} \sum_{i \in Q_n} S_{ni}\, y_{ni} \tag{6-9}$$

$$Minimize\ z_2 = \sum_{n \in N} \sum_{i \in Q_n} C_{ni}\, d_n y_{ni} \tag{6-10}$$

Subject to

$$\sum_{n \in N} d_n y_{ni} - P_i \leq 0 \qquad (\forall\ i \in Q_n) \tag{6-11}$$

$$\sum_{i \in Q_n} y_{ni} = 1 \qquad (\forall\ n \in N) \tag{6-12}$$

$$0 \leq y_{ni} \leq 1 \qquad (\forall\ n \in N, \forall i \in Q_n) \tag{6-13}$$

where the variables used in the model are defined as follows.

N : set of all buyers (firms) looking for best suppliers to cover their demand

Q_n : the set of potential suppliers for the buyer n

S_{ni} : the score of supplier i under all criteria for decision maker n calculated using equation (6-5)

y_{ni} : coverage fraction, the probability of selecting supplier i to cover the demand of buyer n (which is also defined as the fraction of demand of buyer n covered by supplier i)

C_{ni} : total transportation cost of unit of shipment from supplier i to buyer n

d_n : total quantity of demand of buyer n

P_i : Production capacity of supplier i

Objective function z_1 maximizes the total score of selected suppliers while objective function z_2 minimizes the total transportation costs for total commodity flows. Constrain (6-11) indicates that total commodity demand of buyer firms have to be covered by the selected suppliers and makes sure that suppliers are not overloaded by demand levels that are bigger than their production capacity. Constraint (6-12) determines y_{ni}'s as the fraction of demand that will be covered by a supplier.

6.3.4. Solution Method

To solve the multi-objective model, different approaches can be used including the scalarization technique, ε-constraints method, goal programming (GP) and multi-level programming. One of the hybrid approaches that has been in the research focus recently is the fuzzy goal programming (FGP). FGP is obtained by applying fuzzy set theory in goal programming. In GP, for each objective function, specific value or bound (goal) is defined and model tries to minimize sum of deviations from defined goals. GP is an extension form of linear programming (LP) and in the case that LP is infeasible; GP can provide a close answer for the defined goals. It is very difficult especially in real world to determine the goals for objective functions. Therefore, to deal with this fuzziness and imprecision, FGP applies fuzzy set theory to GP by assigning a fuzzy membership function to each objective function. The first FGP method was introduced by Narasimhan (Narasimhan, 1980) and later Hannan (1981) proposed an equivalent LP method to Narasimhan's method. Several studies used FGP method to solve multi-objective optimization problem (Chanas and Kuchta, 2002). The FGP method is also used to solve the proposed optimization problem in this study. The General Algebraic Modeling System (GAMS) is used for modeling and solving the proposed supplier selection problem.

6.4. Mode Choice and Shipment Size Models

6.4.1. Copula-based Joint Modeling Approach

A copula-based joint modeling framework is employed to model freight mode and shipment size choices simultaneously. A copula is a multivariate distribution function that determines the interdependency between random variables and generates the joint distribution of them using the given marginal distributions for each random component.

In contrast to other joint modeling approaches, the copula approach eliminates restricting assumptions on the distribution of error terms and allows the random components to take any various type of distributions. Copula provides more flexibility in examining different distributions of random components to better determine the dependency structure between unobserved common factors. In other words, the analyst can assume any kind of distribution for the error terms of alternatives and then a suitable type of copula from a diverse set of copula models can be used to derive the joint probability of choices.

A basic assumption of such modeling system is that common causes affect both logistics choices simultaneously and both choices are determined by common observed and unobserved factors (Train, 1986). For example, commodity type can be among common observed factors that have influence on both choices. Also, there are unobserved factors that affect both choices simultaneously and generates a dependency between them. The effects of these common unobserved factors on joint decision can be captured using a sound copula-based joint model.

In the proposed joint copula model, both mode and shipment size are considered as discrete choices. A multinomial logit - multinomial logit (MNL-MNL) copula modeling framework is developed in which both mode choice and shipment size are modeled using multinomial logit structures.

Moreover, copula models can be easily estimated within the common maximum likelihood framework. Thus, the estimation of the copula model does not impose any more computational difficulty than other joint modeling approaches. However, one of the challenges in the development of a copula-based joint model is the selection of best-fitted copula function. Since there are various types of copula functions with different pre-specified assumptions and formulas, it is time-consuming to estimate different models with different sets of explanatory variables and find the best fitted one. Also, because of the limited number of studies in transportation literature, where copula-based models are used for the simultaneous decision making problems, little help could be obtained in model estimation and selection from previous experiences.

Several studies have proposed different copula-based modeling structures for joint modeling in passenger travel demand models (Bhat and Eluru, 2009; Portoghese et al., 2011). However, the proposed model is the first effort contributing to the development of a copula-based joint model for simultaneous decision-making problem in freight transportation studies. The same data, gathered through UIC establishment survey 2010-2011, is used for the model estimation. The model formulation and estimation are presented in the following section.

6.4.2. Model Structure

Assume that decision maker n has a subset of M_n as alternative set for mode i which represents the mode alternative. Let U_{in} be the random utility when decision maker n choooses transportation mode i.

$$U_{in} = \beta' X_{in} + \varepsilon_{in} \tag{6-14}$$

Where β is a vector of coefficients that corresponds to the vector of independent variables, X_{in} and ε_{in} represents error terms which are identically and independently Gumbel-distributed with a location parameter of 0 and a scale parameter of 1. Therefore,U_{in} is also Gumbel-distributed with parameters ($\beta' X_{in}$, 1).

Also, assume the following random utility function for classified shipment size q and decision maker n who has a subset of S_n as alternative set.

$$V_{qn} = \gamma' Y_{qn} + \delta_{qn} \tag{6-15}$$

Where γ is a vector of coefficients that corresponds to the vector of independent variables Y_{qn} and δ_{qn} represents error terms which are identically and independently Gumbel-distributed with parameters (0,1).

Therefore, the probability that decision maker n chooses mode i and shipment size q may be written as follows:

$$P_n(i,q) = Pr\left(\vartheta_{in} \leq \beta' X_{in} - ln \sum_{\substack{j \neq i \\ j \in M_n}} e^{(\beta' X_{jn})}, \mu_{qn} \leq \gamma' Y_{qn} - ln \sum_{\substack{j \neq q \\ j \in S_n}} e^{(\gamma' Y_{jn})}\right) \tag{6-16}$$

Where ϑ_{in} and μ_{qn} are random variables corresponding to the mode and shipment size choice and are defined using equations (6-17) to (6-20), given that ε_n^* and δ_n^* are random variables Gumbel-distributed with parameters (0,1) (Ben-Akiva and Lerman, 1985).

$$\vartheta_{in} = \varepsilon_n^* - \varepsilon_{in} \tag{6-17}$$

$$\mathrm{U}_\mathrm{n}^* = \max_{\substack{j \in M_n \\ j \neq i}}(\beta' X_{jn} + \varepsilon_{jn}) = ln \sum_{\substack{j \neq i \\ j \in M_n}} e^{(\beta' X_{jn})} + \varepsilon_n^* \tag{6-18}$$

$$\mu_{qn} = \delta_n^* - \delta_{qn} \tag{6-19}$$

$$\mathrm{V}_\mathrm{n}^* = \max_{\substack{j \in S_n \\ j \neq q}}(\gamma' Y_{jn} + \delta_{jn}) = ln \sum_{\substack{j \neq q \\ j \in S_n}} e^{(\gamma' Y_{jn})} + \delta_n^* \tag{6-20}$$

The presented joint probability in equation (6-16) is directly related to the dependence structure between random variables ϑ_{in} and μ_{qn}.

6.4.3. *Copula Function and Model Estimation*

A copula is a distribution function which is used to determine the dependence relationship between random variables and provide joint distribution given the marginal distribution of random variables (Trivedi and Zimmer, 2007). The definition of copula function and its importance in Statistics is explained in Sklar's Theorem (Sklar, 1973). Assume that X and Y are two random variables and F and G are the given marginal distribution functions for X and Y, respectively. Based on the Sklar's Theorem (Sklar, 1973), if H describes the joint distribution function for X and Y, then for all x, y in R a 2-dimentional copula with a dependence parameter of θ can be defined using the following equation (Quesada-Molina et al., 2003).

$$H(x, y) = C_\theta(F(x), G(y)) \tag{6-21}$$

Different classes of bivariate copulas $C_\theta(F(x), G(y))$ can be used to determine the joint distribution function for two random variables X and Y. Archimedean class of copulas is one of the most frequently used copula class that consists of Clayton, Gumbel, Frank, and Joe copula models. Interested readers are referred to (Nelsen, 2006) for a detailed definition and discussion of various copulas. In this study, the Archimedean class of copulas is used for modeling the dependency structure between choices. This class of copula models can be easily derived and possess appealing properties. For example, unlike Gaussian and Farlie-Gumbel-Morgenstern (FGM) copulas, Archimedean class of copulas can generate asymmetric dependence structure between random variables.

Using the copula expression, the joint probability that any mode i and shipment size q is chosen by decision-maker n from equation (6-13) may be written as:

$$P_n(i,q) = Pr\left(\vartheta_{in} \leq \beta' X_{in} - ln \sum_{j \neq i} e^{(\beta' X_{jn})}, \mu_{qn} \leq \gamma' Y_{qn} - ln \sum_{j \neq q} e^{(\gamma' Y_{jn})}\right)$$

$$= C_\theta\left[F\left(\beta' X_{in} - ln \textstyle\sum_{j \neq i} e^{(\beta' X_{jn})}\right), G\left(\gamma' Y_{qn} - ln \textstyle\sum_{j \neq q} e^{(\gamma' Y_{jn})}\right)\right] \tag{6-22}$$

Where,

$$F\left(\beta' X_{in} - ln \textstyle\sum_{j \neq i} e^{(\beta' X_{jn})}\right) = \frac{1}{1+e^{-\beta' X_{in} + ln \sum_{j \neq i} e^{(\beta' X_{jn})}}} \tag{6-23}$$

$$G\left(\gamma' Y_{qn} - ln \textstyle\sum_{j \neq q} e^{(\gamma' Y_{jn})}\right) = \frac{1}{1+e^{-\gamma' Y_{qn} + ln \sum_{j \neq q} e^{(\gamma' Y_{jn})}}} \tag{6-24}$$

The parameters to be estimated include coefficient vectors β and γ, and the dependence parameter θ. The log-likelihood function with following form is used for the copula model:

$$\log L = \sum_{n=1}^{N} \left(\sum_{i \in C_n^m} \sum_{q \in C_n^s} D_{iqn} \log \left(P_n(i,q) \right) \right) \tag{6-25}$$

where for each mode $i \in C_n^m$ and each shipment size $q \in C_n^s$, D_{iqn} is equal to 1 if decision maker n has chosen mode i and shipment size q and take the value of 0 otherwise. Parameters in the model are then estimated by writing a code and maximizing the log-likelihood using SAS software (SAS Institute Inc., 2008).

The exogenous variables considered for the model estimation can be classified into three main categories:

- commodity type and characteristics (including value per unit of weight, fragility and whether the shipment is containerized, or not)
- shipping attributes (such as, transportation cost, whether the shipping chain is a part of an international shipping chain, shipping time and shipping distance)
- decision maker characteristics (such as, industry type and employee size)

The descriptive statistics of exogenous variables used in utility functions of mode and shipment size choices is discussed in Chapter 5.

6.4.4. Model Estimation Results

The joint copula MNL-MNL framework is developed using three types of copula function from Archimedean class including Gumbel, Clayton and Frank copula. As noted before, various types of copula function provide the flexibility to test different dependency structures between choices. The estimation of copula models, using maximum likelihood estimator, results in the non-nested type of models. The traditional F and likelihood ratio test cannot be used to assess the non-nested copula models. The most accepted method, used to test the data fit of copula models and select the best fitted model is the Bayesian Information Criterion (BIC) (Trivedi and Zimmer, 2007).

Similar to AIC, the BIC is a criterion for model selection and is calculated using the likelihood function at convergence. For a given copula model the BIC criterion can be obtained as follows.

$$\mathrm{BIC} = -2 \times \ln(\mathrm{L}) + \mathrm{k} \times \ln(\mathrm{n}) \tag{6-26}$$

Where, L is the maximum value of likelihood function at convergence, k is the number of parameters that are estimated and n is the number of observations. However, if the number of estimated parameters is the same for all competing copula models, the second part of the BIC equation, $(\mathrm{k} \times \ln(\mathrm{n}))$, will be the same for all of them. In this case, the competing copula models would be compared using the log-likelihood value. In this study, the number of explanatory variables used in estimated copula models was different. Therefore, the number of estimated parameters was different and copula models are compared and evaluated based on the BIC values.

The lower the BIC value, the better the explanatory power of the model (Trivedi and Zimmer, 2007). Among the estimated copula models, the Frank copula resulted in the smallest value of BIC and selected as the preferred copula model with the best fit. The Frank copula is a symmetric copula of Archimedean class presented by the following equation.

$$\mathrm{C}_{\theta}\big(\mathrm{F}(x), \mathrm{G}(y)\big) = -\frac{1}{\theta}\ln\left(1 + \frac{\left(e^{-\theta \times \mathrm{F}(x)} - 1\right)\left(e^{-\theta \times \mathrm{G}(y)} - 1\right)}{e^{-\theta} - 1}\right) \tag{6-27}$$

Where, $F(x)$ and $G(y)$ are equal to $F\left(\beta' X_{in} - \ln \sum_{j \neq i} e^{(\beta' X_{jn})}\right)$ and $G\left(\gamma' Y_{qn} - \ln \sum_{j \neq q} e^{(\gamma' Y_{jn})}\right)$ respectively and are obtained using equations (6-23) and (6-24). The model estimation results are presented in Table 16 (p-values for estimated parameters are presented in parentheses).

Table 16 Frank Copula Joint Model Estimate

Variables	Mode Choice (MNL)				Shipment size (MNL)				
	Truck	Rail	Air	Courier	size 1	size 2	size 3	size 4	size 5
Model Constant	-	-14.96	-4.534	-3.825	0.725	-0.205	-	1.515	1.934
Commodity Type and Characteristics									
Commodity type									
Agricultural/Forestry	0.530 (0.35)	-	-	-	-	-	0.399 (0.25)	1.005 (0.00)	1.709 (0.00)
Chemical/Pharmaceutical	-	-	1.354 (0.01)	-2.256 (0.03)	-	-	-	-	0.893 (0.00)
Electronics	-	-	0.631 (0.23)	-	0.467 (0.04)	-	-	-	-
Machinery	-	-	-	-	-	0.338 (0.06)	-	-	0.666 (0.00)
Mining Products	-0.873 (0.14)	-	-	-	-	-	1.477 (0.04)	1.079 (0.12)	2.33 (0.00)
Motorized and Other Vehicles (incl. parts)	-	0.953 (0.19)	-	-	-	-	1.084 (0.00)	0.74 (0.04)	-
Wood/Paper/Textile	-	0.584 (0.36)	-	-	-	-	-	-	-
Other Commodity Types	-0.452 (0.12)	-	-	-	-	-	-	-	-
Commodity Characteristics									
Fragile	-	-	0.956 (0.03)	-	-	-	-	-	-
Perishable	-	-	-	-1.766 (0.37)	-	-	-	-	-
Commodity Value									
Unit Value ($ per lbs)	-	-	-	-	0.006 (0.00)	-	-0.004 (0.20)	-	-0.255 (0.00)
Square root of Unit Value	-	-	0.078 (0.00)	-	-	-	-	-	-
Shipping Characteristics									
Shipping Cost ($)									
Cost (in Ln form)	-0.391 (0.00)	-0.378 (0.00)	-0.723 (0.00)	-0.752 (0.00)	-	-	-	-	-
Shipping Distance (mile)									
Great Circle Distance	-	-	-	-	0.0004 (0.00)	0.0003 (0.04)	-	0.0002 (0.13)	-
Network Distance (in Ln form)	-	0.865 (0.02)	-	-	-	-	-	-	-
Decision-maker Characteristics									
Number of Employee									
1-100	-	-	-	-	-	1.353 (0.00)	1.24 (0.00)	-	-
101-1,000	-	-	-	-	-	1.02 (0.01)	0.932 (0.02)	-	0.293 (0.10)
>1,000	-	-	-	-	-1.06 (0.00)	-	-	-0.808 (0.00)	-
Model parameters									
Copula Parameter θ	6.165	4.63	6.566	7.469					
Kendal's tau τ	0.52	0.43	0.54	0.58					
Log-likelihood = -1857.6, Number of Observations = 1302									

The selected Frank copula resulted in a log-likelihood value of -1857.6 which is much higher than the log-likelihood value of -2294 for the independent copula model (where the dependency parameter, θ, is set to zero in the Frank copula and two choices are considered to be completely independent). In addition, the estimated Frank copula model and the independent model are compared using the nested likelihood ratio test. The test resulted in a value of 872.8, with a degree of freedom equals to one (the Frank copula model has only one more free parameter than the independent model which is the dependency parameter θ) which statistically greater than any value in the chi-squared table with one degree of freedom. The comparison of log-likelihood and ratio test clearly reject the hypothesis of independence between the mode and shipment size choice and implies that there is a radially symmetric dependence relationship between error terms ϑ_{in} and μ_{qn}.

As shown in Table 16, the estimated dependency parameter of Frank copula model for truck, rail, air and courier mode are equal to 6.16, 4.63, 6.57 and 7.47 respectively. The t-value of estimated copula parameters are 9.05, 3.046, 3.7 and 4.15 which suggest that copula parameters are significantly different from zero and implies that there is a strong dependency between mode choice and shipment size. Therefore, it can be concluded that there are unobserved factors that simultaneously affect both mode and shipment size choice. The dependency parameters can be converted into the Kendall correlation coefficient, τ, which is commonly referred to as Kendall's tau (τ). The Kendall's tau is a measure of monotonic dependence between two random variables (error terms ϑ_{in} and μ_{qn}). The measure is based on the concordance concept. The concordance concept implies "that large values of one random variable are associated with large values of another" (Trivedi and Zimmer, 2007) while discordance implies that "large values of one random variable are associated with small values of the other" (Trivedi and Zimmer, 2007). The τ value for the Frank copula with parameter θ can be obtained using equations (6-28) and (6-29).

$$\frac{[D_1(\theta)-1]}{\theta} = \frac{-1+\tau}{4} \tag{6-28}$$

Where D_1 is the Debye function of the first kind and is calculated as follows.

$$D_1(\theta) = \frac{1}{\theta}\int_0^{\theta} \frac{t}{e^t-1} dt \tag{6-29}$$

The τ takes on a value between -1 and 1. The value of 1 implies a perfect correspondence between two variables and the value of -1 shows the perfect disagreement. The value of zero means that variables are completely independent (Nelsen, 2006). The Kendall's tau is used to test the statistical dependence between two random variables. The estimated Kendall correlation coefficients for truck, rail, air and courier mode of transportation are 0.52, 0.43, 0.54, and 0.58 respectively. The τ values indicate that there is significant correlation (dependency) between error terms ϑ_{in} and μ_{qn} and accordingly between logistics choices. Therefore, there are some unobserved factors that simultaneously affect both choices and can be captured using the copula-based joint modeling formulation. The positive value of τ indicates that the factors that increase (decrease) the probability of choosing a mode i also increase (decrease) the probability of choosing a shipment of size q to be transported by that mode.

Several functional forms of variables were tested to choose the best form in the final model. For example, shipping cost was tested as a linear continues variable, a categorical dummy variable, and in the logarithmic form. In each case, the statistical fit of the model, significance level of variables, variables used in previous studies, applications of the model using public data sets, and parsimony in specification were considered to select the final model specification.

Shipping cost (total cost including transportation and handling costs) turned out to be the most significant variable in the mode choice with negative values for all modes. As it was expected, the cost parameter has the least value for truck and the most value for courier. For the longer shipping distances, the propensity of choosing rail as the shipping mode is higher. However, the effect of distance on the shipment size is negligible and it seems there would be no difference between shipment size choices in different shipping distances. Expensive commodities (with higher unit value) have a tendency to be transported with air and in small shipments. This would be the same for fragile commodities which are more likely to be transported by air. Perishable commodities are less likely to be transported with courier as they tend to have specific shipping chains with a well defined distribution network. Decision makers tend to ship agricultural products with truck and in large shipment sizes. Also, bulky commodities (such as, mining product) and agricultural/forestry products are more probable to be transported in large shipment sizes.

In general, commodity type seemed to be an influential variable on shipment size choice. Only chemical and pharmaceutical products appeared to be significant in mode choice decision, which are more likely to be transported with air and less likely with courier. This could be another reflection of the advantages of using a copula-based model where decisions are made simultaneously and since commodity type is already a factor in shipment size selection, its influence in mode choice selection has been limited. Number of employees at the decision-maker firm was a significant explanatory variable only in the shipment size model. Small and medium size firms (with employee size of 1-1,000) are more likely to select the smaller shipment sizes for shipping their commodities than the large size firms.

Reviewing literature showed that modeling results are in line with previous studies concerning the influential factors on mode and shipment size choice. Transportation attributes, such as freight rate charge, transit time, distance, wait time, reliability and commodity attributes, such as type, value, weight and shelf life are the most frequently used variables in mode choice models (Gray, 1982; Oum, 1979; Cunningham, 1982; Nam, 1997).

Although the choice of optimal shipment size has been studied for a long time using inventory-theoretic models such as Economic Order Quantity, it is usually included as an exogenous variable in the mode choice models (Cunningham, 1982; Nam, 1997). In early studies (Baumol and Vinod, 1970; Hall, 1985; Langley, 1980), that inventory-theoretic concept was employed to jointly model mode and shipment size choice, shipping and inventory costs were the main factors considered in the models' specifications and analyzed to determine optimal shipment size and mode choice. The explanatory variables used in the econometric model of joint discrete-continuous choice of mode and shipment size, proposed in (McFadden et al., 1985), include freight charges, commodity value, transit time and shipment's weight for mode choice equation and transit time, fixed and marginal rates of each mode for the shipment size equation. In (Abdelwahab, 1998) where the simultaneous equations system is utilized for joint modeling, the explanatory variables used in the model specification include shipping cost, transportation time, commodity attributes (such as value, density and total demand at destination) and reliability of transportation mode. The choice of explanatory variables in the proposed study is in accordance with the studies described above. However, some other factors also influenced the choice of explanatory variables in the final model specification, presented in Table 11, including practical issues and data availability, along with statistical issues, such as the existence of high correlation among some variables.

The findings from this model offer an advanced analytical methodology that can be used to study other simultaneous decision-making problems in freight transportation. The MNL-MNL copula-based framework can be employed to jointly model any pair of interdependent discrete alternatives.

6.5. Shipping Chain Configuration Model

As discussed earlier, one of the most important logistics choices, is the *shipping chain* choice which has been ignored or treated insufficiently in current freight transportation models. A brief descriptive statistics of different shipping chain configurations used by diverse decision-makers makers (business establishments) in the UIC survey was presented in Chapter 5. The results of the survey are used in this section to explore shipping chain choice in supply chains. This part of thesis focuses on the shipping chain configuration and modeling of its relevant logistics choices including number of stops and stop type. A system of hierarchical database models of shipping chain choice for freight transportation is presented. A system of decision tree models is developed to determine the shipping chain configuration of freight transportation at a very disaggregate level.

6.5.1. Decision Tree Analysis Method

The rule-based decision tree clustering analysis method is used to model shipping chain configuration for individual shipments by employing the Exhaustive *Chi*-squared Automatic Interaction Detector (Exhaustive CHAID) algorithm (Biggs et al., 1991) for classification. Decision trees like many other data mining approaches are mainly used to explore and analyze a group of observations that form a dataset and identify meaningful and systematic correlations among the variables in the data. The rule-based decision tree methods utilize various rules and criteria to identify the significant interdependencies between target (dependent) variable and predictor (independent/input) variables in the data and classify the data into several homogeneous clusters in which members are assumed to share similar attributes.

Decision trees provide unique capabilities that make them significantly superior to various data mining models (such as neural networks) and make them more appropriate substitute for traditional statistical forms of analysis (SAS Publication, 2013). The main strength of the decision tree models is their multiple variable analysis capability that allows us to explore and describe correlations between variables in the context of multiple influences. The multiple variable analysis results in more accurate outcomes in current problem solving and outperforms the simple single-factor analysis which can only describe one-cause, one-effect relationships between variables and may lead to misleading outcomes (SAS Publication, 2013). Several multiple variable techniques exist that can be used for analysis. However, the prediction power, accuracy, robustness with a variety of data and levels of measurement, ease of use and ease of interpretation of decision tree models make them increasingly appealing for modeling purposes and analyses in many research fields (SAS Publication, 2013). The incremental development and presentation of decision trees make them easier to interpret and understand than many complex, multiple variable models. In fact, decision trees are combination of several one-cause, one-effects correlations presented sequentially, in recursive form. This structure is more compatible with human short-term memory limitations and presents the significant correlations between target and predictor variables in a simple, but effective form.

Decision trees explore the relationships between predictor and target variables and identify those predictor values that have the significant correlations to the target value. All the detected predictor values with high correlations are then grouped in a bin that determines a branch of the decision tree. Thus, decision trees are generated by taking each input variable, determining how the values of input variable are correlated to the values of target variable, splitting the data into branch-like homogeneous segments based on the input-target variable correlation to improve the model ability to predict the value of the target variable. Several algorithms can be used to detect the correlation between the object of analysis (target variable) and the input (predictor) variables and derive the decision rules to split the data and create the branches of the tree. The most common methods used to obtain the optimal decision tree include the Classification and Regression Tree (C&RT), Quick, Unbiased and Efficient Statistical Tree (QUEST), and Chi-squared Automatic Interaction Detector (CHAID) algorithm.

The Exhaustive CHAID algorithm is used in this study to estimate the optimal decision tree. The Exhaustive CHAID was proposed by Biggs et al. (1991) as a modification to the basic CHAID algorithm which based upon adjusted significance testing (Bonferroni testing) (Kass, 1980). The Exhaustive CHAID algorithm includes three major steps: merging, splitting and stopping. A tree starts to grow by using this algorithm on the root node that represents target variable and repeat the algorithm on other nodes. Similar to the CHAID algorithm, Exhaustive CHAID allows multiple splits of a node, but uses a more advanced testing of predictor variables and performs a more comprehensive merging. Therefore, the computing time to estimate the optimal tree is more than that of the basic CHAID. An advantage of Exhaustive CHAID is that it allows multiple splitting of a node and goes beyond the generating binary trees, which makes it easier to interpret and more appealing for the researchers.

The collected data through the UIC freight survey are used to estimate the decision tree models. The target variables are the logistics choices that are related to the shipping chain configuration. These logistics choices are the number of stops per shipping chain, including zero, one and two or more stops, and the type of each stop inside the shipping chain which is selected from the seven considered intermediate handling facilities presented in Figure 13. The predictor variables used in the proposed models include shipments' characteristics such as, commodity type, shipment size and mode of transportation, shipping distance and duration and decision maker attributes such as industry type.

The survey data is divided into train and test samples. The train sample consists of eighty percent of the observations of the survey data and the test sample includes the remaining twenty percent of observations. Trees are developed using the train dataset. The test dataset is used to examine the prediction power of the models and validate them. To validate the models, the estimated distribution of the target variable values for the train data are compared with the observed distribution of target variable values in the test dataset. The next section presents the results of the model estimation for the number of stops and stop types for the first and second stop in the shipping chains. It also discusses the results of validation and comparison of the estimated distribution with the observed distribution for these target variables.

6.5.2. Model Estimation Results

Trees are developed based on three growth criteria by employing the SPSS software which is a powerful package for statistical analysis. A minimum number of 50 observations in parent nodes and 25 in child nodes are required. The maximum growth level of each tree equals to four and P-value of 0.05 is selected for the significance value for splitting nodes and merging categories. The main findings of the study are presented in this section and the best-fitted decision tree models for number of stops and stop type are depicted in Figures 17, 18 and 19.

Each inverted tree originates from a root node at the top of the tree that represents the dependent variable. The Exhaustive CHAID algorithm is then employed to split the data into branch-like segments and form the inverted decision tree. The predictor variables in the next nodes are selected based on their (Bonferroni) adjusted P-value in a way to result in the most significant split (smallest P-value). This process continues until one of the growth criteria is violated. At each node of the trees, the distribution of the values of the target variable are displayed for both train data and test. It also presents the chi-square and P-value for each predictor variable.

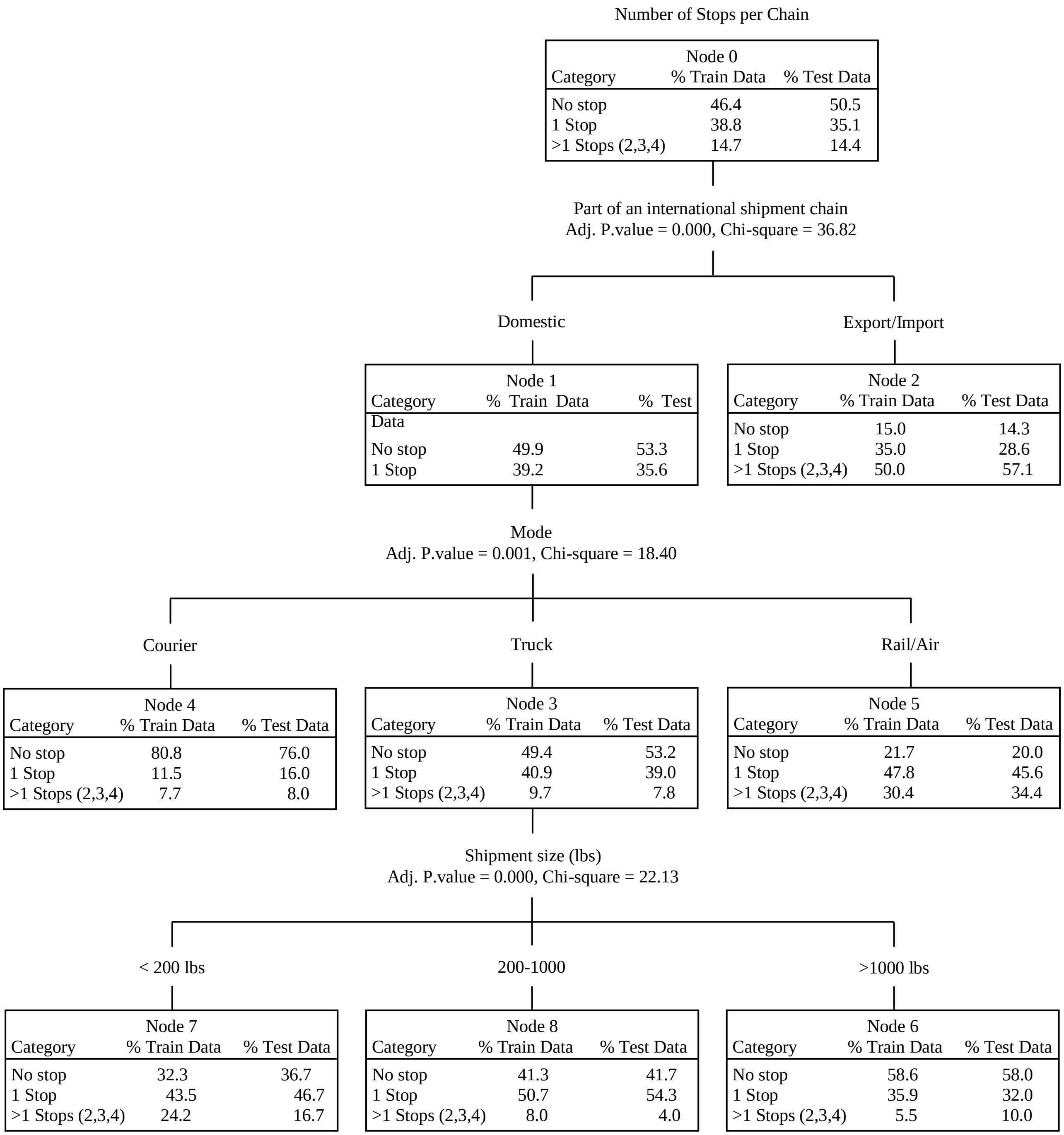


Figure 17 Decision tree cluster for number of stops per shipping Chain.

Stop Type (First Stop)

Category	% Train	% Test
Consolidation Center	18.8	18.5
Distribution Center	22.6	18.5
Warehouse	31.7	40.7
Airport	6.2	3.7
Port	5.3	1.9
Truck/Rail Terminal	9.1	9.3
Customer	6.2	7.4

Mode
Adj. P.value = 0.000, Chi-square = 94.71

Air

Category	% Train	% Test
Consolidation Center	35.7	33.3
Distribution Center	7.1	0.0
Warehouse	7.1	0.0
Airport	50.0	66.7
Port	0.0	0.0
Truck/Rail Terminal	0.0	0.0
Customer	0.0	0.0

Truck/Courier

Category	% Train	% Test
Consolidation Center	18.0	20.4
Distribution Center	25.8	18.4
Warehouse	34.3	40.9
Airport	3.4	2.0
Port	5.1	4.0
Truck/Rail Terminal	6.2	6.1
Customer	7.3	8.2

Rail

Category	% Train	% Test
Consolidation Center	12.5	0.0
Distribution Center	0.0	0.0
Warehouse	25.0	20.0
Airport	0.0	0.0
Port	12.5	20.0
Truck/Rail Terminal	50.0	57.1
Customer	0.0	0.0

Part of an international shipment chain
Adj. P.value = 0.000, Chi-square = 69.59

Domestic

Category	% Train	% Test
Consolidation Center	19.0	19.0
Distribution Center	28.5	21.4
Warehouse	36.1	45.2
Airport	1.9	0.0
Port	0.6	0.0
Truck/Rail Terminal	6.3	4.8
Customer	7.6	9.5

Export/Import

Category	% Train	% Test
Consolidation Center	10.0	22.6
Distribution Center	5.0	6.0
Warehouse	20.0	32.9
Airport	15.0	10.0
Port	40.0	24.3
Truck/Rail Terminal	5.0	4.3
Customer	5.0	0.0

Commodity Typr
Adj. P.value = 0.000, Chi-square = 69.78

Agricultural/
Motorized/Other

Category	% Train	% Test
Consolidation Center	18.9	8.3
Distribution Center	13.5	33.3
Warehouse	51.4	41.7
Airport	0.0	0.0
Port	0.0	0.0
Truck/Rail Terminal	0.0	0.0
Customer	16.2	16.7

Electrical/Gravel/
Sand/Cement

Category	% Train	% Test
Consolidation Center	27.8	35.4
Distribution Center	11.1	7.3
Warehouse	5.6	0.0
Airport	5.6	0.0
Port	5.6	0.0
Truck/Rail Terminal	33.3	50.0
Customer	11.1	7.3

Prepared Foodstuff/ Wood/Paper/
Textile/Leather/Mixed products

Category	% Train	% Test
Consolidation Center	8.6	7.1
Distribution Center	44.8	24.3
Warehouse	32.8	40.0
Airport	0.0	0.0
Port	0.0	0.0
Truck/Rail Terminal	6.9	14.3
Customer	6.9	14.3

Machinery/ Metal
products/ Chemicals

Category	% Train	% Test
Consolidation Center	28.9	38.0
Distribution Center	26.7	20.0
Warehouse	40.0	40.0
Airport	4.4	2.0
Port	0.0	0.0
Truck/Rail Terminal	0.0	0.0
Customer	0.0	0.0

Figure 18 Decision tree cluster for the first stop type.

Stop Type (Second Stop)

Category	% Train	% Test
Consolidation Center	18.8	18.5
Distribution Center	22.6	18.5
Warehouse	31.7	40.7
Airport	6.2	3.7
Port	5.3	1.9
Truck/Rail Terminal	9.1	9.3
Customer	6.2	7.4

Mode
Adj. P.value = 0.000, Chi-square = 69.79

Air/Courier

Category	% Train	% Test
Consolidation Center	31.2	0.0
Distribution Center	12.5	20.0
Warehouse	0.0	0.0
Airport	56.2	80.0
Port	0.0	0.0
Truck/Rail Terminal	0.0	0.0
Customer	0.0	0.0

Truck

Category	% Train	% Test
Consolidation Center	19.4	23.1
Distribution Center	38.8	23.1
Warehouse	11.9	15.4
Airport	7.5	7.7
Port	6.0	7.7
Truck/Rail Terminal	10.4	15.4
Customer	6.0	7.7

Rail

Category	% Train	% Test
Consolidation Center	0.0	0.0
Distribution Center	0.0	20.0
Warehouse	7.7	0.0
Airport	0.0	0.0
Port	23.1	13.3
Truck/Rail Terminal	69.2	66.7
Customer	0.0	0.0

Type of Previous Stop
Adj. P.value = 0.000, Chi-square = 59.37

Warehouse/Consolidation Center

Category	% Train	% Test
Consolidation Center	40.0	33.3
Distribution Center	36.0	33.3
Warehouse	20.0	16.7
Airport	4.0	0.0
Port	0.0	16.7
Truck/Rail Terminal	0.0	0.0
Customer	0.0	0.0

Distribution Center/Port

Category	% Train	% Test
Consolidation Center	8.3	0.0
Distribution Center	66.7	50.0
Warehouse	4.2	0.0
Airport	0.0	0.0
Port	16.7	20.0
Truck/Rail Terminal	4.2	30.0
Customer	0.0	0.0

Truck-Rail Terminal/ Airport/ Customer

Category	% Train	% Test
Consolidation Center	5.6	20.0
Distribution Center	5.6	0.0
Warehouse	11.1	20.0
Airport	22.2	20.0
Port	0.0	0.0
Truck/Rail Terminal	33.3	20.0
Customer	22.2	20.0

Figure 19 Decision tree cluster for the second stop type.

In the estimated models for number of stops and stop type, other logistics choices and shipment's attributes such as mode choice and shipment size are among the most influential variables. Whether the shipping chain is a part of a bigger import/export chain or not is also found to be significant. Commodity type was also another influential variable in predicting the type of stop for the first stop in the chains. Shipments which are parts of import and export chains have more intermediate stops in their shipping chains. This is the same trend for the shipments that are transported with air or rail. Shipments, which are included in wood, paper, textile, machinery, coal, and mineral ore commodity groups, usually are transported with more intermediate stops in the shipping chain. However, heavier shipments are transported with less intermediate stops at the intermediate logistics facilities. It is worth to note that other characteristics of shipments such as distance, shipment value and shipping cost and decision makers' characteristics including industry type and employee size were also considered in the modeling process but were excluded due to the high p-value and not being significant.

For the first stop type, 50 percent of shipments that are transported by rail have at least one stop at a truck/ rail intermediate handling facilities and around 50 percent of air shipments have a stop at the airport. For truck and courier shipments, warehouses, distribution centers and consolidation centers have the highest share with 36.6, 24.2 and 18.5 percent respectively. Commodity type is also found to be influential for determining the intermediate handling facility type for domestic shipments that are transported with truck or courier. However, for the second stop type, the most significant variables are mode of transportation and type of the previous stop if the shipment is transported with truck or courier. More than 56 percent of air and courier shipments stop at an airport in their second stop while 69.2 percent of rail shipments stop at a truck/rail intermediate handling facility. As it was explained before, the mode that was used in the longest link of the shipping chain was selected as the main mode and considered in the analysis. Therefore, it should not be expected that all air shipments have the first stop at an airport because the air mode might be used in the third or the fourth link of the chain.

As it is illustrated in the figures, the distribution of the alternatives of target variable for the estimated trees (train data) are similar to the observed distributions of these alternatives in the test data. This confirms that the estimated models can predict the distribution of the target variables for the test data with an acceptable precision and hence can be used for predicting shipping chain configuration for other datasets. However, as the figures illustrate, the distribution of the target variable values are more similar in branches with more observations. In other words, as the growth level of the tree (number of splits) increases and the number of observations in the clusters decreases, the differences between distributions of target variable in train and test data increase. Also, as it can be seen in the figure 19, the difference in the distribution of target variable values for the second stop type are larger since the tree is estimated based on a smaller data sample (only the shipments with 2 or more stops). This implies that decision tree models work more effectively for rather large samples.

The estimated decision trees determine the significant predictor variables that have strong correlation with the dependant variables. As it is displayed in the figures, the trees classify the data into homogeneous branch-like clusters in which observations (members) share similar attributes. The distributions of the target variable values are also estimated for each cluster. These distributions along with the determined predictors in branches can be used to predict the target variable values in other datasets. A superior way to use these models for prediction is to find the best-fitted distribution for all clusters in the developed trees and assumes that these distributions are valid for other datasets. Therefore, for any dataset with known predictor variables and unknown target variable the proposed trees can be used to cluster the data and simulate the target variable by using the best-fitted distributions.

7. SIMULATION RESULTS AND VALIDATION

This chapter presents results of model application in estimating domestic commodity flows in the U.S. in 2007. It should be noted that none of the models and components of the framework are calibrated to fit the existing patterns and trends, except the mode choice model that has been calibrated using CFS dataset.

7.1. Firm Synthesis

There are more than 7.6 million business establishments in the U.S. in 2007 (U.S. Census Bureau, 2009). The firm Synthesis model uses 333 zone, 87 industry classes and 7 employee size groups to categorize these business establishments. Using the Firm Synthesis Model, these business establishments are categorized into 70,116 firm-type groups which are considered as the agents in the simulation process. Each firm-type has a unique id that shows the geographic zone, industry type and employee size of the actual firms that comprise that firm-type group. Also, the number of actual firms in each firm-type group is kept in the id and used in the next model of the framework to estimates production and consumption values for firm-types.

For example, the id 130 236 1 (17) belongs to an actual synthesized firm-type. 130 shows the zone in which this firm-type is located which is Menard County in Illinois in this case. 236 represents the industry type for this firm-type which is the "Construction of Buildings". The single digit 1 shows that the number of employees of this firm-type is between 1 to 19 employees. Finally, the number in the parentheses shows the number of firms in this firm-type group. Figure 20 presents an example of generated firm-types in the study area. Distribution of

Food Manufacturing industries (NAICS 311) of 7 employee size classes are shown in the figure. The size of circles represents the actual number of firms in the corresponding firm-type.

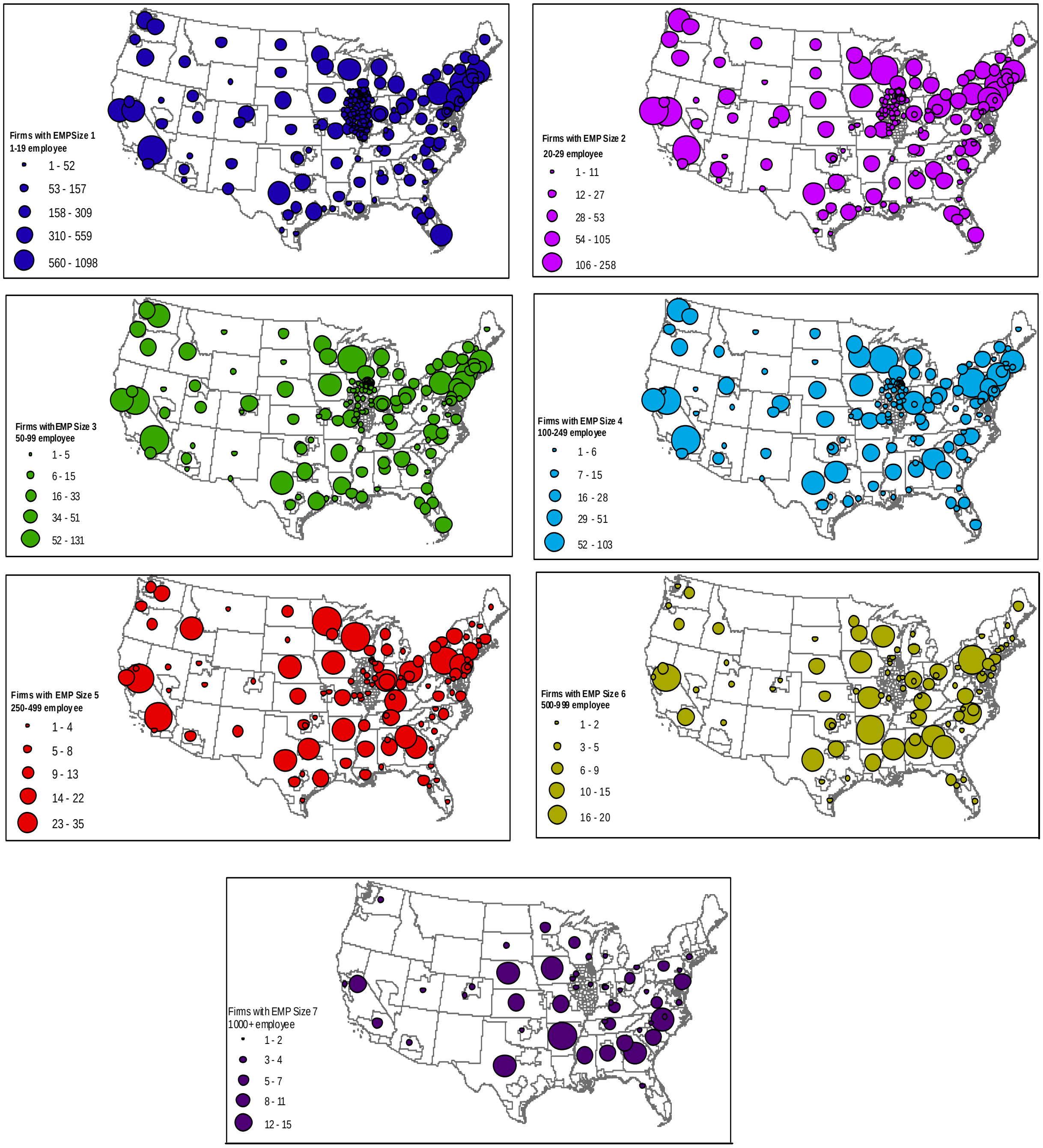


Figure 20 Distribution of food manufacturing industries with different employee size.

7.2. Freight Generation

As it is explained in Chapter 6, a multi-step procedure is used to estimate the production and consumption values of different commodities for synthesized firms. First, using the developed commodity-industry crosswalks, the percentages of produced and consumed commodities by industry types are estimated. Then, the zone level production and consumption values provided by FAF data are allocated between firms using the estimated percentages. The zone level production and consumption values are obtained by aggregating domestic FAF (FHWA, 2007) commodity flows over destinations and origins respectively. The model takes more than 14.3 billion tons of zone-level produced and consumed commodities provided by FAF as input and distribute them between synthesized firms.

The zone-level production and consumption values, used as input in the model, are less than total domestic commodity flows (18.6 billion tons) estimated in FAF data (FHWA, 2007) for two main reasons. First, these values are derived from those FAF commodity flows that are transported by only four modes of truck, rail, air, and courier. The reason for this selection is to make the model results compatible with the mode choice model in this study and provide possibility of further comparison and validation against FAF data. Also, as it is mentioned before, commodity class 15, “Mixed and Unknown Freight”, is excluded from this simulation since the producer and consumer industries for this commodity type can be hardly identified.

While the zone level values from FAF are used as the main input data to estimate firm-level production and consumption, the developed commodity-industry crosswalks and explained procedure are the key tools in the Freight Generation model to perform the allocation. Therefore, the estimated firm-level production and consumption values can be aggregated and compared with the zone level values provided by FAF to validate the proposed procedure and evaluate accuracy of the developed crosswalks.

Table 17 presents the total estimated production and consumption values for all 14 classes of commodities and compares them with the reported values in FAF data. As the table shows, the total estimated production and consumption values are slightly different from the reported values in FAF data, except for commodity class 1, “Agricultural and Forestry Products”. The significant difference between estimated and reported values for commodity class 1 is mainly due to the absence of economic data on establishments in the “Crop Production” (NAICS 111) and “Animal Production” (NAICS 112) industries in CBP and ZBP datasets. As it was mentioned before these industries are excluded from this study and no firm-type of these industry classes are generated in the simulation process.

However, based on the developed commodity-industry crosswalks, the “Crop Production” (NAICS 111) and “Animal Production” (NAICS 112) industries produce 29% and consume 13.3% of Agricultural and Forestry Products. Therefore, 29% of produced and 13.3% of consumed Agricultural and Forestry Products could not be allocated between any firms in the Freight Generation Model. In total, more than 13.4 billion tons of produced commodities worth more than 10.6 million dollars and more than 13.7 billion tons of consumed commodities worth more than 10.7 million dollars are allocated between firms in the Freight Generation Model.

Table 17 Total Estimated and FAF Production and Consumption Values (KTON)

Commodity Class	FAF Production /Consumption	Estimated Production	Percentage of Difference for Production	Estimated Consumption	Percentage of Difference for Consumption
1	2842869	2017092	29.0%	2447513	13.9%
2	4371775	4365197	0.2%	4265851	2.4%
3	1448457	1406551	2.9%	1389943	4.0%
4	670304.2	669552.4	0.1%	639476	4.6%
5	875362.8	868341.9	0.8%	865240	1.2%
6	265028.3	263942	0.4%	263275	0.7%
7	1316807	1313754	0.2%	1311939	0.4%
8	679572.4	677782.8	0.3%	657884	3.2%
9	60107.02	60099.46	0.0%	59796	0.5%
10	141681.6	141612.4	0.0%	139001	1.9%
11	35406.1	35406.02	0.0%	35400	0.0%
12	274304.66	273952.2	0.1%	271861	0.9%
13	51625	51473.99	0.3%	51202	0.8%
14	1276473	1273808	0.2%	1267943	0.7%

As the table shows, the difference between FAF and estimated production and consumption values at national scale is slim. This confirms that the developed crosswalks and proposed procedure can accurately distribute produced and consumed values between firms. The small difference between FAF and estimated values origins from data insufficiency and missing values in development of the crosswalks and allocation of produced and consumed values between firms. For example, there might be a producer or consumer industry that has been ignored in the development of crosswalks, or when the model is distributing zone level produced values between firms, it might be unable to find any relevant producer establishment in that zone. The first problem can be result of data insufficiency in development of the crosswalk and later is the result of missing values in CBP and ZBP data that have been used to synthesized firms in the study area.

Figure 21 depicts the estimated production and consumption values for commodity class 6 “Paper Products” to provide an example of the output of the Freight Generation Model. The presented values in the figure are firm level production and consumption values that have been aggregated to zone level. Since very small zone system is used for the Chicago area and Illinois State, these two areas are depicted separately in figure 22.

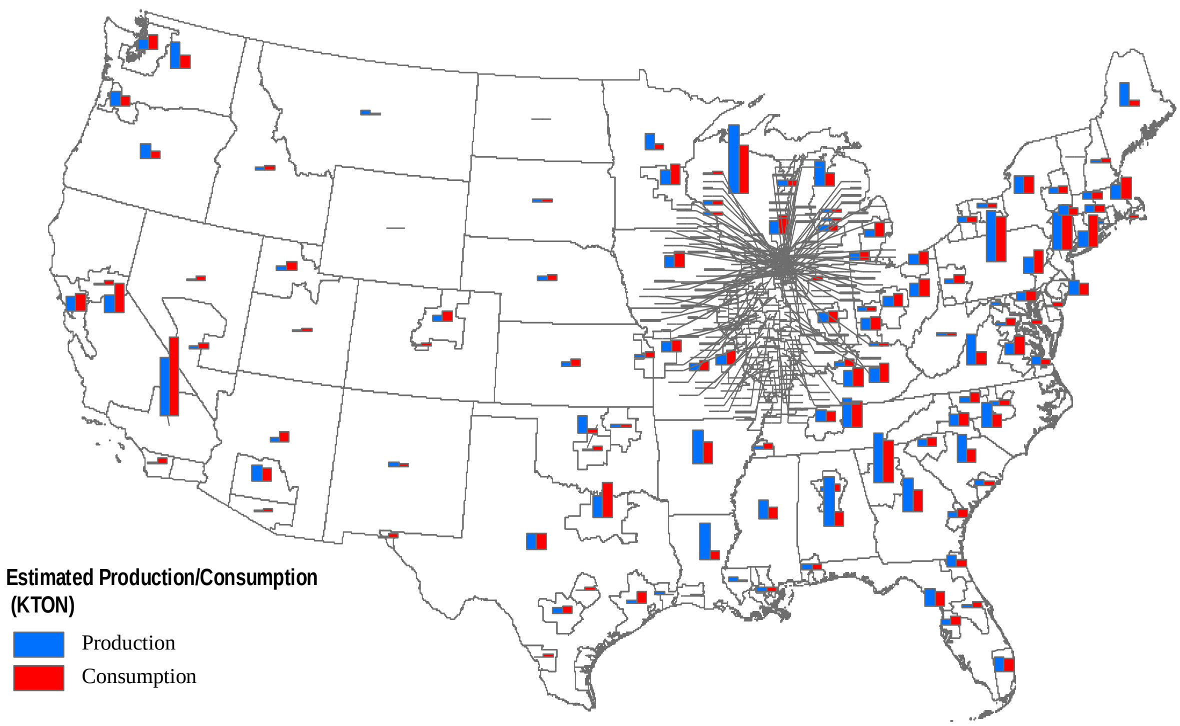


Figure 21 Aggregated production and consumption values for “Paper Products”.

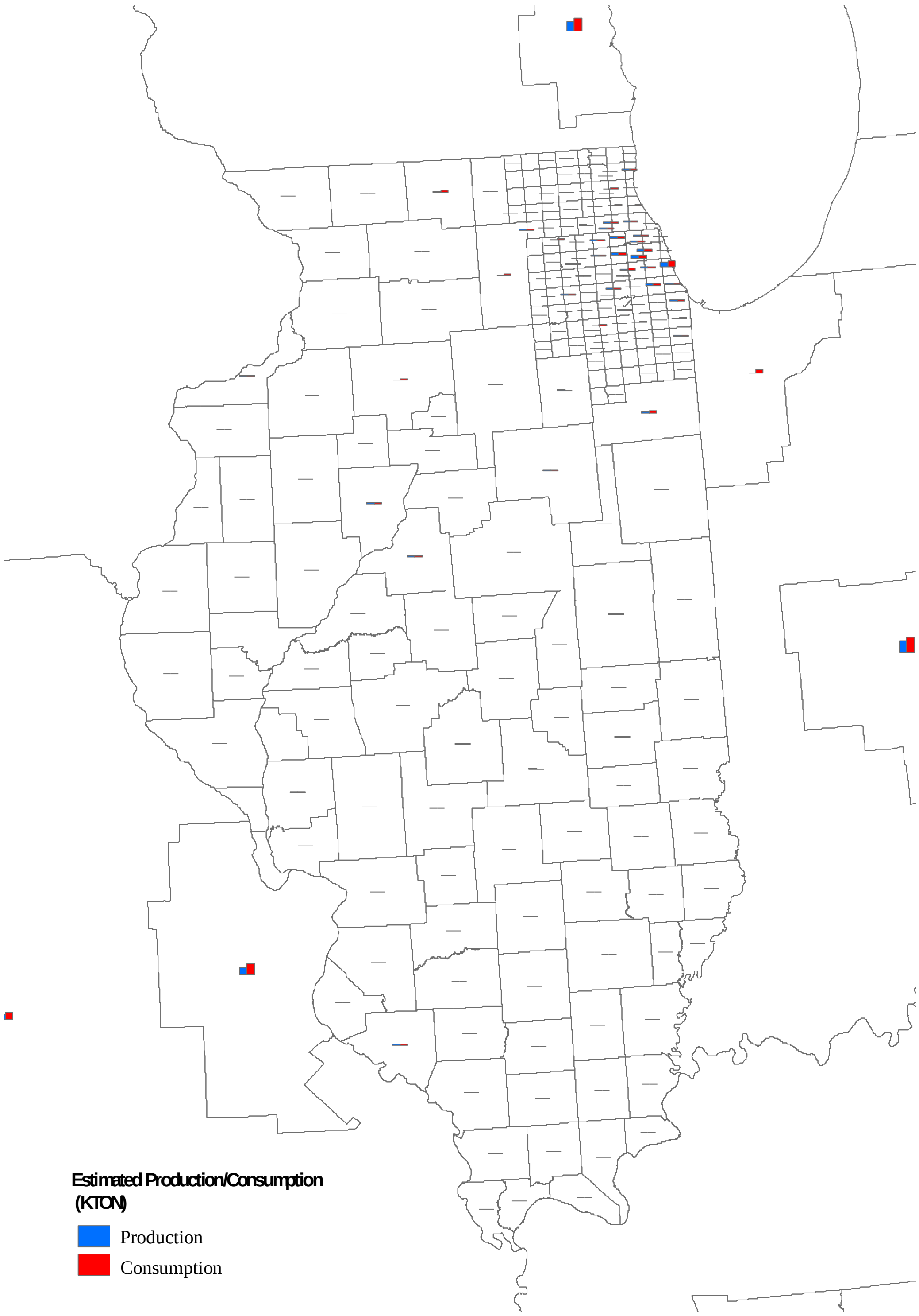


Figure 22 Aggregated production and consumption values for “Paper Products” in Illinois State.

The comparison of the estimated and FAF values and validation can be also performed for smaller regions in the zone system for all types of commodities to evaluate performance of the developed commodity-industry crosswalks and proposed procedure at more disaggregate level. Figure 23 is an example of comparing the estimated production values vs. FAF production values for “Paper Products” in regions within the zone system. The consumption values are compared in Figure 24. As the figures show, for areas other than Illinois State the estimated production and consumption values closely match the reported values in FAF data which validates the developed crosswalks and used procedure. However, since the values for Illinois State could not be clearly presented in figures the aggregate estimated values in Illinois State are compared with the reported values by FAF in Table 18 and 19.

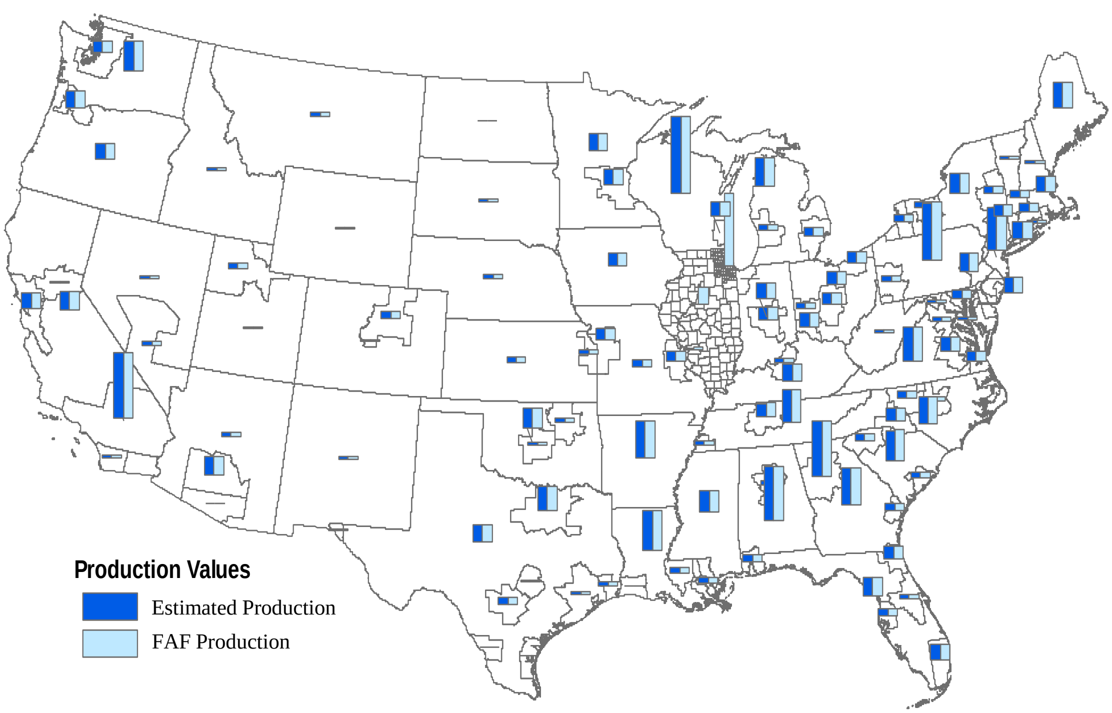


Figure 23 Comparison of estimated and FAF production values for “Paper Products”.

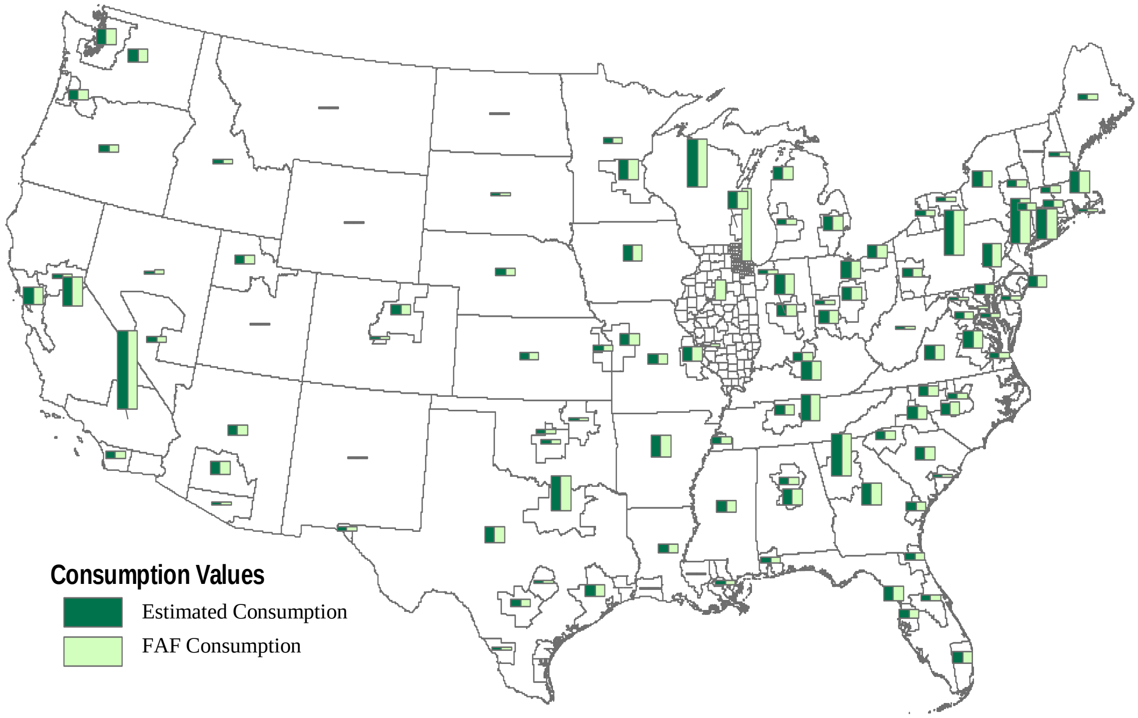


Figure 24 Comparison of estimated and FAF consumption values for “Paper Products”.

Table 18 Comparison of Estimated And FAF Production of “Paper Products” in Illinois

FAF Zone	Definition	FAF Production (KTON)	Estimated Production (KTON)
171	Chicago, IL-IN-WI CSA (IL Part)	12085.94	12037.6
172	St. Louis, MO-IL CSA (IL Part)	305.0249	303.8
179	Remainder of Illinois	2861.234	2849.79

Table 19 Comparison of Estimated And FAF Consumption of “Paper Products” in Illinois

FAF Zone	Definition	FAF Production (KTON)	Estimated Production (KTON)
171	Chicago, IL-IN-WI CSA (IL Part)	13807.08	13751.86
172	St. Louis, MO-IL CSA (IL Part)	343.54	340.44
179	Remainder of Illinois	3850.733	3816.078

As it can be seen from Table 18 and 19, three FAF zones constitute Illinois State. The estimated production and consumption values for firms that are locate in these areas are aggregated and compared with the FAF estimates. As the tables show, there is a close match between estimated values and FAF data. This again confirms that the developed crosswalks and proposed procedure can accurately distribute produced and consumed values between firms at smaller zone scale. The small difference between FAF and estimated values results from the same issues data insufficiency and missing values that are explained before. This analysis was performed for other 13 commodity classes which resulted in similar outcome. Therefore, it can be concluded that the developed crosswalks and proposed procedure work with an acceptable precision in the Freight Generation Model. Hence, the model results can be used in the next modules of the framework without any concerns regarding the procedure's validation.

7.3. Supply Chain Formation

Supplier-buyer pairs are identified and supply chains are replicated using the behavioral two-step supplier evaluation and selection. The proposed behavioral model connect supplier firms to buyer firms and form supply chains for all 14 types of commodity at very disaggregate level. The model provides total annual commodity flows at firm-to-firm level. Based on the results of Freight Generation Model, 621,325 consumption points (records) have to be connected to 99,986 potential production points (records) of different commodities. Running the supplier selection optimization model and simulating all the disaggregate commodity flows for these production and consumption points is impossible due to the computational complexity and insufficiency of available software and hardware tools. Therefore, two techniques are used to counteract this simulation problem and reduce the problem size. First, the input production and consumption records are trimmed out and some of the records are excluded from the input data. The excluded records are the smallest records that constitute 10% to 25% of total production and consumption values for each commodity type. These small production and consumption records are excluded until the rest of records could be simulated in the framework. Secondly, the simulation is performed for one commodity class at a time. Table 20 presents the total estimated commodity flows for all classes of commodities and compare the statistics of the estimated flows with FAF data. It should be noted that the estimated flows are aggregated to the FAF zone level, so they can be compared with FAF flows.

Table 20 Total FAF and Estimated Commodity Flows For All Classes of Commodities

Commodity Class	FAF Flows (KTON)				Estimated Flows (KTON)			
	Total Flow	Minimum Flow	Maximum Flow	Average Flow	Total Flow	Minimum Flow	Maximum Flow	Average Flow
1	2842869	0.001	165058.9	245.16	2017092	1.07	98033	3564.9
2	4371775	0.001	96643.5	551.3	3788499	1.04	106852	6051
3	1448457	0.001	69444.71	217.42	1229671	0.052	22832	2231
4	668304	0.001	76502.55	56.53	498310	0.024	13217	592
5	875363	0.001	76502.55	105.35	775975	0.002	16761	1363
6	265028	0.001	7514.97	24.21	187974	0.773	1986	205
7	1316807	0.001	55543.18	136.87	1151567	9.06	23697	1771
8	679572	0.001	15697.1	52.25	577992	0.14	6385	395
9	60107	0.001	2294.32	4.1	43054	0.064	1218	26
10	141682	0.001	5792.73	14.93	122828	0	2198	139
11	35406	0.001	1140.37	2.52	31352	0	222	23
12	274305	0.001	9915.57	20.83	210286	0.011	2961	164
13	51625	0.001	4340.88	5.18	45562	0	951	37
14	1276473	0.001	59798.58	145.25	1118231	2.19	16658	2803

Due to reduction of problem size by trimming out the production and consumption values the total estimated commodity flows are less than FAF flows. As it can be obtained from the table, the model simulates 85.1% (around 12.2 million tons) of FAF domestic commodity flows that are transported by truck, rail, air and courier modes. As the table shows, although the maximum value of estimated flows are much smaller than maximum flow in FAF data, the average value of estimated flows is much bigger than the FAF data. Considering the huge intervals between minimum and maximum values of FAF commodity flows, it can be concluded that the distribution FAF flows is skewed toward smaller commodity flows while the proposed model simulates more evenly distributed commodity flows.

Another factor that can be explored and compared with FAF data is the Ton-Mile variable. At this step of the simulation, the volume of commodities (kilo tons) that have been transported between zones can be calculated and multiplied by the transport distance (GCD) to estimate Ton-Mile for the simulated commodity flows. These values are estimated and compared with the corresponding values obtained from FAF data in Table 21.

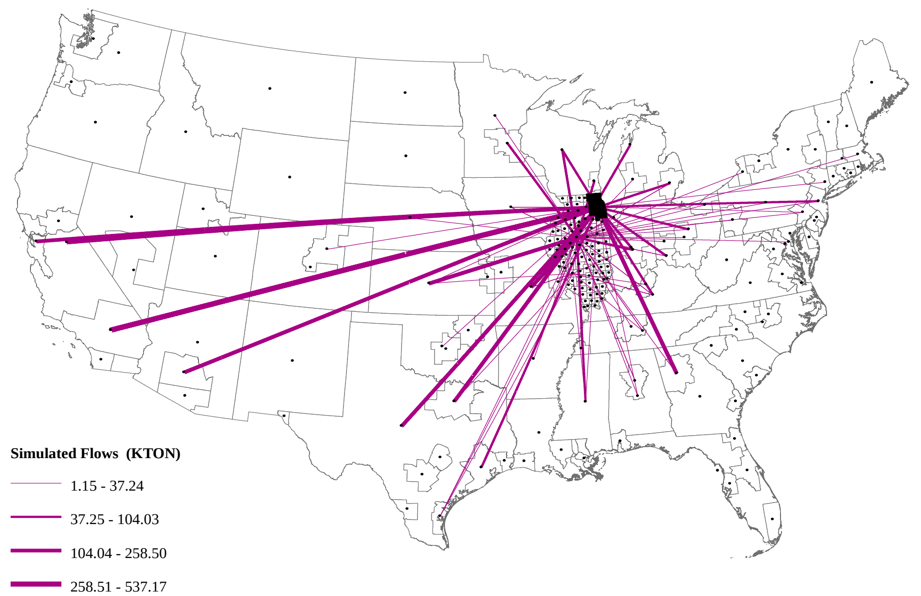
Table 21 Total KTon-Mile for FAF and Estimated Commodity Flows by Commodity Types

Commodity Class	FAF Flows (KTON-MILE)				Estimated Flows (KTON-MILE)			
	Total KTon-Mile	Average	Min	Max	Total KTon-Mile	Average	Min	Max
1	458,379,628	31,918	-	8,584,452	208,656,616	311,428	-	9,703,291
2	732,537,913	50,618	-	48,546,960	593,362,264	947,863	-	9,290,331
3	116,666,477	8,360	-	3,122,218	205,902,801	373,689	-	5,025,646
4	198,479,602	13,720	-	3,920,781	154,458,415	183,660	-	3,824,229
5	78,368,598	5,609	-	1,351,493	172,841,116	303,763	-	2,897,447
6	94,699,410	6,860	-	629,947	52,853,717	57,638	-	829,640
7	79,512,185	5,495	-	1,750,566	206,226,767	317,272	-	4,467,046
8	156,058,505	10,606	-	1,593,482	195,572,250	133,954	-	2,627,501
9	28,364,731	1,886	-	436,993	19,488,673	11,705	-	920,573
10	45,227,806	3,136	-	712,157	48,809,975	55,403	-	1,663,714
11	13,685,199	923	-	407,229	16,028,493	11,820	-	366,114
12	103,634,595	7,095	-	902,535	81,773,825	63,886	-	2,758,691
13	22,777,822	1,575	-	1,260,704	22,305,483	18,061	-	449,472
14	77,860,318	5,357	-	1,859,783	266,195,599	667,157	-	7,578,463

Figure 25 shows the replicated supply chains for commodity 9 “Electronics, Electrical and Precision Equipments” as an instance of the output of Supplier Selection Model. However, it should be noted that the commodity flows presented in this figure are aggregated to the zone level.

Figure 25 Simulated supply chains for “Electronics, Electrical and Precision Equipments”.

To compare the replicated supply chains with FAF zone level supply chains at a more detailed scale, simulated commodity flows and FAF commodity flows that start or end in Illinois State are depicted for both models in Figure 26.



(a)

(b)

Figure 26 Simulated and FAF flows for "Electronic, Electrical, and Precision Equipments". (a) Simulated flows. (b) FAF flows.

As the figure shows, the simulated supply chains are different from the FAF commodity flow patterns. Although the commodity flow ranges are similar, the distribution of flows within this range is completely different. While there are more small size (KTON) flows presented in FAF data, the simulated commodity flows in this study are more evenly distributed by size (KTON).

This difference can be due to the exclusion of small production and consumption values from this study. These small production and consumption values generate small size commodity flows. However, since they are excluded from this study, small size commodity flows are underestimated. Also, this study uses a behavioral optimization supply chain model to simulate commodity flows while the FAF model employs a data mining-based flow matrix construction technique (Oak Ridge National Laboratory, 2011b) to estimate commodity flows. The different approaches used for estimation of commodity flows in the two models can be another source of dissimilarities between FAF and simulated commodity flows in this study. The same analysis was performed for other 13 types of commodities which resulted in similar outcome and conclusion.

7.4. Mode Choice and Shipment Size

The copula-based joint mode choice and shipment size model is used to simulate mode of transportation and shipment size of the simulated commodity flows between firms. As it was discussed before, 4 modes of transportation including truck, rail, air and courier and 5 discrete shipment size classes including (0-200 lbs), (201-1,000 lbs), (1001-4,000 lbs), (4,000-30,000 lbs) and (more than 30,000 lbs) are considered in the model and simulated in this study.

The 2007 FAF data (FHWA, 2007) is used to compare the simulated mode choices in this study with the reported mode choices in FAF and validate the models at aggregate level. The total amounts of simulated shipments transported by 4 considered modes are presented in Table 22 and compared with reported values in FAF data. It should be noted that the 2007 Commodity Flow Survey (CFS) (U.S. Census Bureau, 2014) is used to calibrate the developed mode choice model. The table shows relative percentages of total weight (kilo tons) and value (million dollars) of transported goods from this model simulation and compares them with the respective values from 2007 CFS and FAF data.

Table 22 Modal Splits Based on Weight of Transported Commodities.

Mode	Proposed Model	CFS 2007	FAF
Truck	82.9%	82.22%	82.58%
Rail	11.4%	17.43%	13.03%
Air	2.0%	0.03%	0.08%
Courier	3.7%	0.32%	4.32%

7.5. Shipping Chain Configuration

Since required information regarding the geographical location and characteristics of intermediate handling facilities are not available, this part of the framework is not operational yet. The choice of shipping chain including the number of stops and stop types are not simulated for the commodity flows in this study. However, the developed Shipping Chain Configuration model, explained in Chapter 6, is incorporated into the operational framework's code and once the required network information for running the model and simulating the shipping chain choice are available, the model can be used to simulate shipping chain configuration for individual shipments.

7.6. Network Analysis

After that the mode choice and shipment size of Commodity flows are simulated, the network related attributes of these shipments have to be simulated to prepare the input data for network assignment. The network analysis model requires several input information including the OD trip tables and the detailed network for all considered modes. This part of the framework is not fully operational yet. Only the simulated truck flows are converted to truck trips and assigned to the network.

The conversion process used in FAF model to transform commodity volumes to truck payload (Oak Ridge National Laboratory, 2011a) is used in this study to generate truck OD trip table. In this procedure, first the commodity volumes (tonnage) are allocated between 5 primary truck configuration. Table 23 introduces the primary truck classes and their descriptions and Table 24 presents their shares of allocated commodity volumes based on the traveled distance. These table are directly used from FAF network analysis model.

Table 23 Truck Configurations

Truck Class	Definition
SU	Single Unit Trucks
TT	Truck plus Trailer Combinations
CS	Tractor plus Semitrailer Combinations
DBL	Tractor plus Double Trailer Combinations
TPT	Tractor plus Triple Trailer Combinations

Table 24 Trucks' Commodity Volume Allocation Factors By Traveled Distance

Distance Range (miles)	SU	TT	CS	DBL	TPT
0-50	0.793201	0.070139	0.130465	0.006179	1.67E-05
51-100	0.577445	0.058172	0.344653	0.019608	0
101-200	0.313468	0.045762	0.565269	0.074434	0.000452
201-500	0.142467	0.027288	0.751628	0.075218	0.002031
501-10000	0.06466	0.0149	0.879727	0.034143	0.004225

Once the annual commodity flow are allocated between different classes of truck, the truck equivalency factors are used to convert tonnage to truck trips for all truck classes. Table 24 presents the truck equivalency factors used in this study to convert commodity flows to vehicle trips. These factors are calculated based on the type of transported commodity. The provided factors in FAF model are not used directly in this study, but they are used to obtain the average values in Table 25.

Table 25 Truck Equivalency Factors By Commodity Type

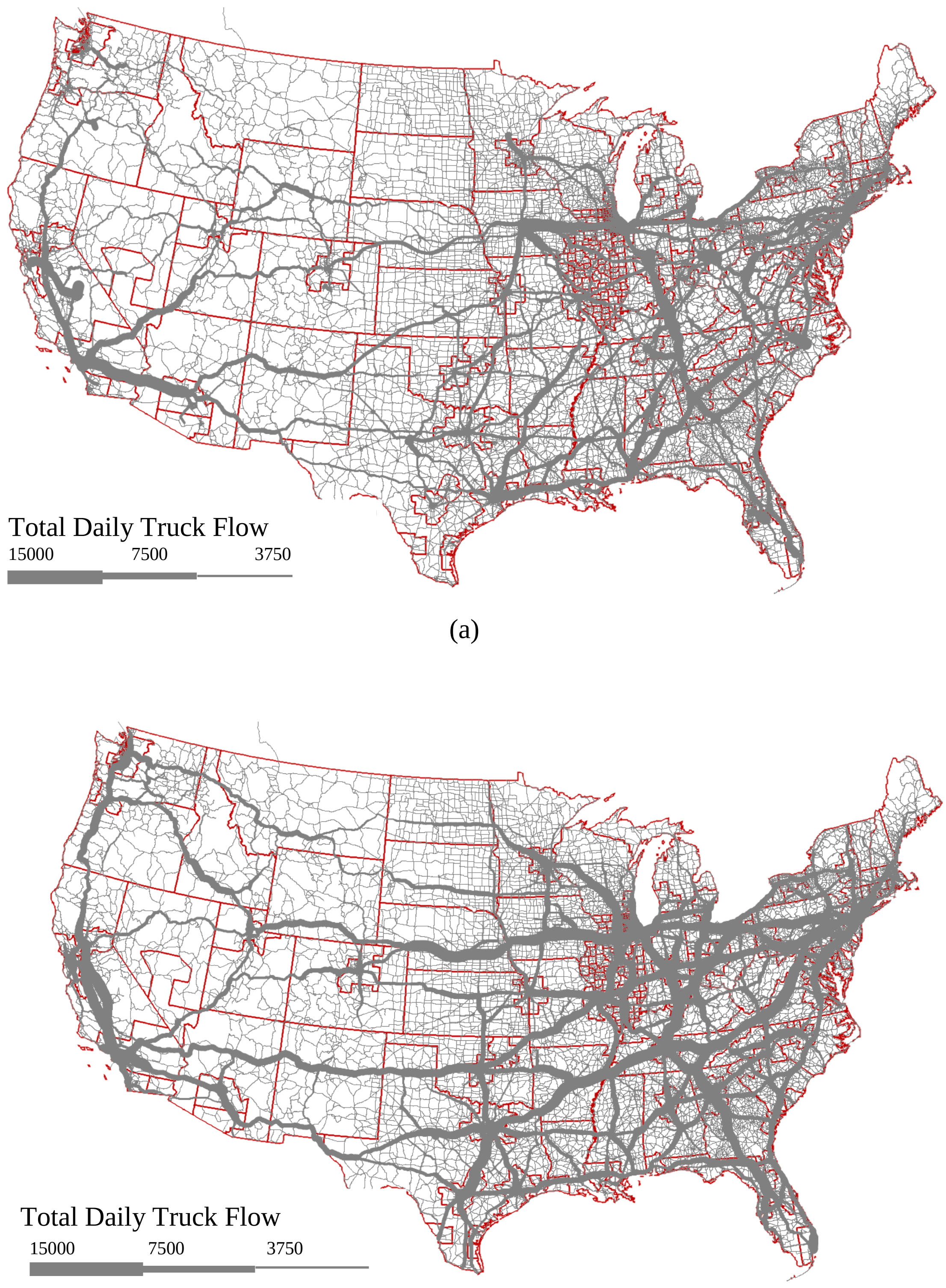
Commodity Class	SU	TT	CS	DBL	TPT
1	0.051516	0.109901	0.026273	0.024751	0
2	0.022989	0.017601	0.019358	0.008256	0.000454
3	0.02965	0.024416	0.01456	0.010713	0
4	0.046947	0.045674	0.022239	0.009492	0
5	0.036827	0.055336	0.019506	0.011691	0
6	0.045381	0.077916	0.012759	0.00153	0
7	0.016211	0.016083	0.022827	0.007747	0.002423
8	0.057142	0.087194	0.028627	0.04243	0.000649
9	0.075919	0.109759	0.00938	0.006536	0
10	0.026951	0.051787	0.022033	0.010582	0.001103
11	0.071067	0.047747	0.012723	0.003848	0
12	0.072267	0.062287	0.027553	0.013847	0
13	0.063028	0.048309	0.0117	0.015326	0
14	0.029667	0.029307	0.020111	0.00217	0

Finally, the percentage of empty trucks of each truck class is estimated using the empty truck factors from FAF model. Table 26 presents the calculated factors that are used in the network assignment model.

Table 26 Empty Truck Factors

	SU	TT	CS	DBL	TPT
Empty Truck factor	0.84	0.7	1.54	0.824444	0.053333

To convert each record of simulated commodity flows to truck trips, first the total tonnage of commodity flow is allocated between truck classes using Table 24. Next, based on the type of commodity of the simulated flow the allocated volumes are converted to the truck trips using Table 25. Finally, the estimated truck trips are increased by the percentages presented in Table 26 to account for empty vehicles in the traffic network. The results of assigning commodity flows to the traffic network using TransCAD software are depicted in Figure 27 and compared to the network assignment results of FAF model. As it can be seen there is a significant difference between this model's results and FAF data. This difference is due to several issues. First, as it was mentioned before, the small production and consumption values are excluded in the supplier selection model to reduce the problem size. These small production and consumption values generate small commodity flows. However, since they are excluded from this study, small size commodity flows are underestimated and are not simulated in the model. Also, this model only simulates truck flows and unlike FAF model do not cover other combination modes of truck shipments such as intermodal shipments, truck-air, and truck-rail. However, this significant difference needs further investigation.



(a)

(b)

Figure 27 Network assignment results for simulated and FAF commodity flows.
(a) Simulated commodity flow assignment. (b) FAF flow assignment.

8. CONCLUSION

8.1. Summary

The remarkable increase in freight movements and their significant impacts on transportation system, regional wellbeing, and economic growth provide sufficient motivation to develop reliable analysis tools to estimate commodity flows between zones and forecast the future demand and trends of goods movements among regions. While the need to develop freight demand model to better facilitate infrastructure planning and policy development has been clearly recognized for some time, the current state of knowledge and understanding regarding the movement of freight and behaviors of freight actors (suppliers, receivers, shippers and carriers) lags behind those of passenger travel by a considerable margin.

It can be concluded from this discussion in Chapters 2 and 3 that only a handful of agent-based freight transportation models have been proposed and developed in the past studies, and as a result the role of key decision making agents in the freight system have been ignored in current freight models. It also emphasizes on the need to develop agent-based models that incorporate supply chain relationships and logistics components in their framework to better capture the decision-making procedures. Agent-based microsimulation logistics models can more precisely capture the complex interactions among various agents and markets in the freight system by simulating the behavior of these decision making agents. In addition, they can be integrated with micro-level models for passenger transportation to better estimate and forecast traffic volumes on networks. They would also provide a platform to test how various policy measures and major infrastructure investments can alleviate the negative impacts of freight transportation and how implementing efficient policies would make the freight transportation system more sustainable.

This study outlines a behavioral freight transportation modeling framework that will address some of critical technical and conceptual hurdles that have challenged past efforts by applying agent-based framework in which firm-level decision making processes, including supply chain formation and selecting logistics choices, are simulated. The study demonstrates the use of disaggregate, behavioral-based modeling approaches for forecasting freight movements at disaggregate firm-level and evaluating freight policy impacts at the national/regional scale.

The proposed agent-based modeling approach is unique in the focus on multiple aspects of individual firm behavior which leads to disaggregate commodity movements and ultimately to freight vehicle flows. The approach is analogous to the activity-based modeling approach to evaluating passenger travel demand, in that the freight goods and vehicle movements are modeled as derived demand arising from the needs and behaviors of individual firms. The proposed technology will be a state-of-the-art modeling system implementing advanced behavioral-based freight modeling concepts for looking at freight demand at the disaggregate firm level.

The cutting-edge supply chain and logistics choice models including a behavioral supplier selection model, a joint model of mode choice and shipment size and a decision tree clustering model for intermediate handling facility usage, are developed and incorporated in the proposed framework to enhance the precision of the model in forecasting individual shipments and their attributes. In addition, disaggregate three-dimensional crosswalks for identifying industry use-and-make shares of different commodity classes are developed to improve the accuracy of deriving industry-to-industry freight flows. Finally, the simulated truck freight flows are assigned over traffic network.

8.2. Contributions

This study puts forward the development of a reliable microsimulation freight transportation modeling framework that posses some unique characteristics that distinguish the proposed model from other freight modeling efforts. The model is based on sound behavioral theories to represent firm-level decision making in relation to commodity flows and shipment logistics. The key research issues that have been addressed in this study include:

- *The freight data scarcity for development of a reliable analysis tool*. The development of the framework and its model components is greatly based on the publicly available freight data sources to alleviate data collection cost as much as possible. However, for the development of the disaggregate and behavioral logistics models, an online establishment freight survey was conducted. The survey provided valuable information on the logistics decision making process at firm level and individual shipment information that can be used in further studies.

- *The lack of micro-level freight analysis tools.* The framework is designed to synthesize firms as the decision making agents in the freight transportation system. It micro-simulates these agents' behavior in the economic activities and logistics decision making process within the supply chains, with the goal of providing more accurate and more detailed long-range forecasts of freight traffic

- *The lack of detailed commodity-industry crosswalks.* Current commodity-crosswalk methods tend to be highly aggregated and do a poor job of explaining the production and consumption patterns in terms of observable characteristics of firms. In addition, the relationship between the economic linkages among various industry sectors and flow of freight is not well understood. The study develops detailed and comprehensive commodity-industry crosswalks linking produced and consumed commodities by industry class. The crosswalks determine type and amount of commodities that are made and used by each industry.

- *Highly aggregated commodity input-output tables.* Current practice in freight forecasting uses aggregate commodity input-output tables derived from national sources to replicate commodity flows. Using the developed commodity-industry crosswalks developed in this study, the framework generates commodity input-output tables at firm level.

- *The lack of information about behavioral supplier evaluation and selection process.* The study develops a behavioral optimization supplier evaluation and selection model which forms supply chains between supplier and receiver firms based on their characteristics and transportation costs. It is one of the few freight modeling frameworks that includes an explicit component for forecasting supply chain configuration at firm-level.

- *The lack of disaggregate logistics choice models.* Most of the freight transportation models lack the explicit treatment of logistics choices. This framework is designed to incorporates advanced disaggregate logistics choice models for several logistics decisions including mode choice, shipment size, use of intermediate handling facilities and vehicle choice.

- *Little consideration for simultaneous logistics decision-making process.* A major drawback of current freight demand models is their step-wise structure in which logistics choices are modeled separately in sequential order. However, as it is discussed in the freight literature, the logistics choices, particularly mode and shipment size choice, are highly inter-connected and should be modeled jointly. This study proposes an advanced joint modeling framework for mode and shipment size choice. The developed model utilizes a copula-based modeling approach which has been never used in freight transportation modeling efforts to simulate mode choice and shipment size simultaneously.

- *Little consideration for multimodal disaggregate freight movements.* Most freight studies only focuses on road freight movements. The proposed model is designed to cover several modes of transportation including truck, rail, air, and courier.

Development of the agent-based microsimulation freight model provides a satisfactory analysis tool that can be used to:

- To better capture the complex interactions among decision makers in the freight system,
- To have a more realistic understanding of regional scale freight transportation,
- To have more reliable demand forecast and policy assessments tool,

- To better address adverse impacts of freight movements on the socio-economic systems

In addition, the approaches used to develop individual components of the proposed model, e.g. freight generation, supplier selection, and logistics choice can be adopted to improve the existing tools such as Freight Analysis Framework. It is expected that the disaggregate construct of the model make it possible to feed the output of this national model to various regional freight models.

8.3. Future Direction

The proposed agent-based freight transportation model in this study is a step toward improving the existing analysis tools for freight movements. There are some aspects of the proposed framework that require further exploration and improvement including:

- *The improvement of the Freight Generation Model.* The current model uses the business patters and employment data to estimate commodity generation rates at firm-level and allocate FAF production and consumption values between firms. The model can be improved by including land-use data and other economic factors and developing a more advanced economic model for estimation of commodity production and consumptions. Such models can capture the effects of economic changes at firm level and forecast production and consumption changes in response to new policies.

- *Simulating shipping chain configuration.* Since the required information on the geographical location and type of intermediate handling facilities cannot be easily obtained, the shipping chain configuration was excluded from the simulation. Once that the data become available this logistics choice can be simulated for shipments.

- *Considering role of other freight agents.* Current model considers producer and receiver firms as the only decision makers in the freight system. However, other freight agents such as carriers and third party logistics (3PL) play an influential role in making logistics decisions that have to be considered in the model to more precisely simulates logistics choices of shipments.

- *Expanding the application of current model for forecasting future freight flows.* The current model simulates the disaggregate freight flows for the base year of 2007. Applying the model to forecast future freight flows depends on the availability of input data such as business patterns. Therefore, to use the model for forecasting purpose the

- *The development of micro-level network analysis tool for freight movements.* There is still the need to develop detailed networks for different modes of transportation and simulates disaggregate commodity flows in the networks which will provide transportation performance measures that can be used for multi-modal traffic analysis purposes.

- *Integrating the current domestic microsimulation model with international freight transportation model.* The proposed framework only focuses on the domestic part of commodity flows. However, it could be improved to incorporate effects of global changes on commodity flows inside the U.S.

- *Incorporating other modes of transportation in the mode choice model.* The current mode choice model includes truck, rail, air and courier. Other modes of transportation such as water, intermodal, etc. can be included in the framework to improve the model power.

CITED LITERATURE

Abdelwahab, W. M., and M. A. Sargious. Modeling the demand for freight transport: new approach. *Journal of Transport Economics and Policy* Vol. 26, No. 1, pp. 49-72, 1992.

Abdelwahab, W.M. Elasticities of mode choice probabilities and market elasticities of demand: evidence from a simultaneous mode choice/shipment-size freight transport model, *Transportation Research part E*, pp. 257-266, 1998.

Akarte, M.M., Surendra, N.V., Ravi, B., and N. Rangaraj, Web based casting supplier evaluation using analytical hierarchy process. *Journal of the Operational Research Society*, Vol. 52 (5), pp. 511–522, 2001.

Arunotayanun, K., and J. W. Polak, Taste heterogeneity in freight shippers' mode choice behaviour. Presented at the 86th Annual Meeting of the Transportation Research Board, Washington DC, 2007.

Baumol, W. J., and H. D. Vinod. An inventory theoretic model of freight transport demand. *Management Science*, Vol. 16, No. 7, pp. 413-421, 1970.

Ben-Akiva, M., and. G. De Jong. The Aggregate-Disaggregate-Aggregate (ADA) Freight Model System, In: M.E. Ben-Akiva, H. Meersman and E. van de Voorde eds. *Recent Developments in Transport Modelling: Lessons for The Freight Secto'*, Emerald press, 2008, pp. 117-134, 2008.

Ben-Akiva, M., and S. R. Lerman. *Discrete Choice Analysis: Theory and Application to Travel Demand*. The MIT Press, 1985.

Bhat, C.R., and N. Eluru. A copula-based approach to accommodate residential self-selection effects in travel behavior modeling. *Transportation Research Part B*, Vol. 43, No. 7, pp. 749-765, 2009.

Bhutta, K.S. and F. Huq, Supplier selection problem: A comparison of the total cost of ownership and analytic hierarchy process approaches, In: *Supply Chain Management*, Vol. 7 (3), pp. 126–135, 2002.

Biggs, D.B., de Ville, B., and E. Suen. A Method of Choosing Multi-way Partitions For Classification and Decision Trees. Journal of Applied Statistics, 8(1), pp. 49–62, 1991.

Boerkamps, J.H.K., van Binsbergen, A.J., and P.H.L. Bovy. Modeling Behavioral Aspects of Urban Freight Movement in Supply Chains, In: *Transportation Research Record*, 1725, pp. 17-25, 2000.

Bovenkerk, M. SMILE+, The New and Improved Dutch National Freight Model System. In: European Transport Conference, Strasbourg, 3-5 October, 2005.

Brög, W. and A. Meybur. Influence Of Survey Methods On The Results Of Representative Travel Surveys, In: Transportation Research Part A, 17 (2), pp. 149-156, 1983.

Bureau of Transportation Statistics (BTS): North American Transborder Freight Data, 2009. Available at: http://www.bts.gov/programs/international/transborder/, accessed 2014.

Bureau of Transportation Statistics, 2007 Commodity Flow Survey: Standard Classification of Transported Goods (SCTG), United States Department of Transportation, 2007.

Bureau of Economic Analysis, Industry Economic Accounts Data, 2007 Make/Use Tables, U.S. Department of Commerce, 2013. http://www.bea.gov/industry/index.htm#benchmark_io Accessed August 2013

Cambridge Systematics, Inc. Final Report: Validation and Sensitivity Considerations for Statewide Models. American Association of State Highway and Transportation Officials (AASHTO) Standing Committee on Planning, 2010.

Catalani, M., A model of shipper behaviour choice in a domestic freight transport system. Presented at the 9th World Conference on Transport Research, Seoul, North Korea, 2001.

Chai, J., Liu J. N.K. and E. W.T. Ngai, Application of decision-making techniques in supplier selection: A systematic review of literature, *Expert Systems with Applications,* Vol. 40, pp. 3872–3885, 2013.

Chan, F.T.S. Interactive selection model for supplier selection process: An analytical hierarchy process approach. *International Journal Production Research,* Vol. 41 (15), pp.3549–3579, 2003.

Chanas S., Kuchta D., Fuzzy goal programming – One notion, many meanings. *Control and Cybernetics, Vol* 31 (4), 871– 890, 2002.

Chiang, Y.S., Robert, P., and M. Ben-Akiva, A short-run freight demand model: joint choice of mode and shipment size. Center for Transportation Studies, MIT, Cambridge, Mass. 1980.

Chou, S., and Y. Chang, A decision support system for supplier selection based on a strategy-aligned fuzzy SMART approach. *Expert Systems with Applications*, Vol. 34(4), pp. 2241–2253, 2008.

Chow, J.Y.J, C.H. Yang, and A.C. Regan. State-of-the Art of Freight Forecast Modeling: Lessons Learned and The Road Ahead, In: *Transportation*, 37(6), pp.1011–1030, 2010.

Combes, F., An empirical evaluation of the EOQ model of choice of shipment size in freight transport. Presented at the 91st Annual Meeting of the Transportation Research Board, Washington, D.C., 2012.

Couper, M. P., M. W. Traugott, and M. J. Lamias, Web Survey Design and Administration, The Public Opinion Quarterly. Vol. 65, No. 2, pp. 230-253. 2001.

Cunningham, W. H. J., Freight modal choice and competition in transportation: A critique and categorization of analysis techniques, *Transportation Journal*, Vol. 21, No. 4, pp. 66-75, 1982.

Dahel, N.E. Vendor selection and order quantity allocation in volume discount environments, In: *Supply Chain Management*, Vol. 8, No. 4, pp. 335–342, 2003.

De Boer, L., Labro, E. and P. Morlacchi, A review of methods supporting supplier selection, *European Journal of Purchasing & Supply Management*, Vol. 7, pp. 75–89, 2001.

De Jong, G.C., A model of mode and shipment size choice on the Swedish commodity flow survey. Presented at the Universities Transport Study Group annual conference, Harrogate, U.K., 2007.

De Jong, G., and M. Ben-Akiva. A Microsimulation Model of Shipment Size and Transport Chain Choice, In: *Transportation Research Part B,* 41, pp. 950–965, 2007.

De Jong, G., H.F. Gunn, and W. Walker. National and international freight transport models: an overview and ideas for further developments, In: *Transport Reviews,* 24(1), pp. 103-124, 2004.

De Jong, G.C., and D. Johnson, Discrete mode and discrete or continuous shipment size choice in freight transport in Sweden, Presented at European Transport Conference, Netherlands, 2009.

Donnelly, R. A Hybrid Microsimulation Model of Freight Flows. Proceedings of the 4th International Conference on City Logistics, July 11-13, 2007. Taniguchi, E. and R.G. Thompson (Ed.). Crete, Greece, Institute for City Logistics, pp: 235-246, 2007.

Dulmin, R. and M. Valeria, Supplier selection using a multi-criteria decision aid method. In: *Journal of Purchasing & Supply Management,* 9, pp. 177–187, 2003.

Falagario, M., Sciancalepore, F., Costantino, N. and R. Pietroforte, Using a DEAcross efficiency approach in public procurement tenders. *European Journal of Operational Research*, Vol. 218(2), pp. 523–529, 2012.

Federal Highway Administration, Freight Analysis Framework (FAF), Office of Freight Management and Operations, U.S. Department of Transportation, 2007. http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/ Accessed 2013

Federal Highway Administration, Freight Facts and Figures 2011. Office of Freight Management and Operations. U.S. Department of Transportation, 2011.

Federal Highway Administration: Quick Response Freight Manual I and II. Office of Freight Management and Operations. U.S. Department of Transportation, FHWA-HOP-08-010, 2007.

Friesz, T.L., Holguı´n-Veras, J.: Dynamic game-theoretic models of urban freight: formulation and solution approach. In: Methods and Models in Transport and Telecommunications, Advances in Spatial Science. Springer, Berlin 2005.

Golob, T., Regan, A. Trucking Industry Adoption of Information Technology: A Multivariate Discrete Choice Model. *Transportation Research Part C: Emerging Technologies*, Volume 10 Issue 3. pp. 205-228, 2002.

Gray, R., Behavioural approaches to freight transport modal choice, *Transport reviews*, Vol. 2, Issue 2, pp. 161-184, 1982.

Haleh, H. and A.Hamidi, A fuzzy MCDM model for allocating orders to suppliers in a supply chain under uncertainty over a multi-period time horizon. In: *Expert Systems with Applications*, Vol. 38 (8), pp. 9076–9083, 2011.

Hall, R. W. Dependence between shipment size and mode in freight transportation. *Transportation Science*, Vol. 19, No. 4, pp. 436-444, 1985.

Hannan E. L., Linear Programming with Multiple Fuzzy Goals, *Fuzzy Sets and Systems* 6, 235-248, 1981.

Hensher, D., Figliozzi, M.A.: Behavioural insights into the modeling of freight transportation and distribution systems. Transp. Res. B 41(9), 921–923, 2007.

Hesse, M., Rodrigue, J.P.: The transport geography of logistics and freight distribution. J. Transp. Geogr. 12(3), 171–184, 2004.

Ho, W., Xu, X. and P. K. Dey, Multi-criteria decision making approaches for supplier evaluation and selection: A literature review, *European Journal of Operational Research*, Vol. 202, pp. 16–24, 2010.

Holguin-Veras, J. Revealed preference analysis of commercial vehicle choice process. *Journal of Transportation Engineering*, Vol. 128, No. 4, pp. 336-346, 2002.

Horowitz, A. and D., Farmer. A Critical Review of Statewide Travel Forecasting Practice. In: *Transportation Research Record: Journal of Transportation Research Board*, 1685, pp. 13–20, Washington, D.C., 1999.

Horowitz, A. White Paper: Statewide Travel Demand Forecasting. In: Meeting Federal Surface Transportation Requirements in Statewide and Metropolitan Transportation Planning: A Conference, American Association of State Highway and Transportation Officials (AASHTO), 2008.

Huang, S.H. and H. Keska, Comprehensive and configurable metrics for supplier selection. *International Journal of Production Economics*, Vol. 105 (2), pp. 510–523, 2007.

Hunt, J. D., K. J. Stefan, and A. T. Brownlee. Establishment-Based Survey of Urban Commercial Vehicle Movements in Alberta, Canada: Survey Design, Implementation, and Results. In: *Transportation Research Record*, 1957, pp. 75-83, 2006.

Hunt, J.D., R. Donnelly, J.E. Abraham, C. Batten, J. Freedman, J. Hicks, P.J. Costinett, and W.J. Upton. Design of a Statewide Land Use Transport Interaction Model for Oregon. Proceedings of World Conference on Transportation Research, Seoul, South Korea, 2001.

Ishani, Meheboob. Edmonton Region External Truck/Commodity Survey, Report and Results. Report for Alberta Transportation and the City of Edmonton. July 2003.

ISL and Banister Research & Consulting, Inc. 2005 Household Travel Survey: Summary Report on Weekday Travel by Residents of the Edmonton Region. Report to Alberta Infrastructure and Transportation. July 2006.

Jiang, F., Johnson P., and C. Calzada, Freight demand characteristics and mode choice: An analysis of the results of modeling with disaggregate revealed preference data. *Journal of Transportation and Statistics*, Vol. 2, No. 2, pp. 149-158, 1999.

Kass, G.V. An Exploratory Technique For Investigating Large Quantities of Categorical Data. Journal of Applied Statistics, 29, 119–127, 1980.

Langley, C. J. The inclusion of transportation costs in inventory models: some considerations. *Journal of Business Logistics*, Vol. 2, No. 1, pp. 106-125, 1980.

Lawson, C. and Strathman, J. Survey Methods for Assessing Freight Industry Opinions –Final Report. SPR 328 to Oregon DOT and FHWA. 2002.

Levary, R. R. Using the analytic hierarchy process to rank foreign suppliers based on supply risks. *Computers and Industrial Engineering*, Vol. 55(2), pp.535–542, 2008.

Lewis, K. and D. P. Widup, Deregulation and rail-truck competition: evidence from a translog transport demand for assembled automobiles. *Journal of Transport Economics and Policy*, Vol. 16, No. 2, pp. 139-149, 1982.

Liedtke, G., and H. Schepperle. Segmentation of the transportation market with regard to activity-based freight transport modelling. Int. J. Logist. Res. Appl. 7(3), pp. 199–218, 2004.

Liedtke, G. Principles of micro-behavior commodity transport modeling. In: *Transportation Research Part E: Logistics and Transportation Review*, 45(5), pp. 795–809, 2009.

Liu, F., Ding, F.Y. and V. Lall, Using data envelopment analysis to compare suppliers for supplier selection and performance improvement, *Supply Chain Management*, Vol. 5 (3), pp.143–150, 2000.

Lloret-Batlle R. and F. Combes, Estimation of Inventory-Theoretic Model of Mode Choice in Freight Transport. In: Proceedings of the 92nd Transportation Research Board Annual Meeting, 2013.

Lung Ng, W. An efficient and simple model for multiple criteria supplier selection problem. *European Journal of Operational Research*, 186, pp. 1059–1067, 2008.

Mafakheri, F., Breton, M., and A. Ghoniem, Supplier selection-order allocation: A two-stage multiple criteria dynamic programming approach. *International Journal of Production Economics*, Vol. 132(1), pp. 52–57, 2011.

McCabe, S. The Region of Peel Commercial Travel Survey: Mailout–Mailback Survey Design and Preliminary Analysis. Masters of Applied Science Thesis, Department of Civil Engineering, University of Toronto, Canada, 2007.

McFadden, D., and C. Winston. Joint estimation of discrete and continuous choices in freight transportation. Presented at Meeting of the Econometric Society, Washington D.C., 1981.

McFadden, D., Winston, C., and A. Boersch-Supan. Joint estimation of freight transportation decisions under nonrandom sampling. In Daughety, A. F. (Ed.) *Analytical Studies in Transport Economics*, Cambridge University Press, Cambridge, UK, pp. 138-157, 1985.

Nam, K., A study on the estimation and aggregation of disaggregate models of mode choice for freight transport. *Transportation Research Part E: Logistics and Transportation Review*. Vol. 33, No. 3, pp. 223-231, 1997.

Narasimhan R., Goal programming in a fuzzy environment. *Decision Science*11, 325–336, 1980.

Narasimhan, R., Talluri, S. and S.K. Mahapatra, Multiproduct, multicriteria model for supplier selection with product life-cycle considerations. *Decision Sciences*, Vol. 37, No. 4, pp. 577–603, 2006.

National Cooperative Freight Research Program Report 8: Freight-Demand Modeling to Support Public-Sector Decision Making. Transportation Research Board of The National Academies. Washington D.C. 2010.

National Cooperative Highway Research Program Report 388: A Guidebook for Forecasting Freight Transportation Demand. Transportation Research Board, National Research Council. Washington D.C. 1997.

National Cooperative Highway Research Program Synthesis 358: Statewide Travel Forecasting Models, A Synthesis of Highway Practice. Transportation Research Board of The National Academies. Washington D.C. 2006.

National Cooperative Highway Research Program Report 606: Forecasting Statewide Freight Toolkits. Transportation Research Board of The National Academies. Washington D.C. 2008.

Nelsen, R.B. *An Introduction to Copulas.* Springer Science and Business Media, Inc., 2006.

Norojono, O., and W. Young, A stated preference freight mode choice model. *Transportation Planning and Technology*, Vol. 26, No. 2, pp. 195–212, 2003.

Oak Ridge National Lab, FAF3 FREIGHT TRAFFIC ANALYSIS, Submitted by Battelle, 2011a.

Oak Ridge National Lab, The Freight Analysis Framework Version 3 (FAF3): A Description of the FAF3 Regional Database And How It Is Constructed, U.S. Department of Transportation, 2011b.

Oum, T.H. A cross sectional study of freight transport demand and rail-truck competition in Canada. *Bell Journal of Economics*, Vol.10, pp. 463-482, 1979.

Patterson, Z., Ewing, G. O., and M. Haider, How different is carrier choice for third-party logistics companies? mixed-logit approach. Presented at the 87th Annual Meeting of the Transportation Research Board, Washington DC, 2008.

Patterson, Z., Ewing, G. O., and M. Haider, Shipper preferences suggest strong mistrust of rail: results from stated preference carrier choice survey for quebec city–windsor corridor in Canada. *Transportation Research Record*, Transportation Research Board of the National Academies, Washington, D.C., pp. 67-74, 2007.

Portoghese, A., Spissu, E., Bhat, C., Eluru, N., and I. Meloni, A Copula-Based Joint Model of Commute Mode Choice and Number of Non-Work Stops during the Commute. *International Journal of Transport Economics*, Vol. 38, Issue 3, pp. 337-362, 2011.

Pourabdollahi, Z., Mohammadian, A. and Kawamura, K. 2012. A Behavioral Freight Transportation Modeling System: An Operational and Proposed Framework. In: Proceedings of the 14th Annual International Conference on Electronic Commerce, 196-203, Singapore, 07 - 08 August, 2012.

Pourabdollahi, Z., Javanmardi, M., Karimi, B., Mohammadian, A. and K. Kawamura, Mode and Shipment Size Choice Models in the FAME Simulation Framework. In: The 92nd Annual Meeting of the Transportation Research Board, Washington, D.C., 2013.

Quesada-Molina, J.J., Rodriguez-Lallena, J.A., and M. Ubeda-Flores, What are copulas? *Monograf´ıas del Semin. Matem. Garc´ıa de Galdeano*. Vol. 27, pp. 499–506, 2003.

RAMON Classification Detail List, Statistical Classification of Products by Activity in the European Economic Community, 2008 version, http://ec.europa.eu/eurostat/ramon/index.cfm?TargetUrl=DSP_PUB_WELC, Accessed 2013

Rich, J., Holmblad, P.M., and C.O. Hansen, A weighted logit freight mode-choice model, *Transportation Research Part E,* Vol. 45, pp. 1006–1019, 2009.

Richardson, A., Ampt, E. and A. Meyburg. Survey Methods for Transport Planning, 1st Edition, Eucalyptus Press, 1995.

Roberts, P. O., Presented at the 3rd World Conference on Transport Research, Rotterdam, Netherlands, 1977.

Roorda, M., McCabe, S., and Kwan, H. A Shipper-Based Survey of Goods and Service Movements in the Greater Golden Horseshoe (GGH). Report to the Ministry of Transportation of Ontario, Region of Peel. 2007.

Roorda, M.J., Cavalcante, R., McCabe, S., and H. Kwan. A Conceptual Framework for agent-based modeling of logistics services. In: *Transportation Research Part E: Logistics and Transportation Review* 46 (1), pp.18-31, 2011.

Saen, R.F., Suppliers selection in the presence of both cardinal and ordinal data. *European Journal of Operational Research*, Vol. 183 (2), pp. 741–747, 2007.

Saen, R. F., Developing a new data envelopment analysis methodology for supplier selection in the presence of both undesirable outputs and imprecise data. *International Journal of Advanced Manufacturing Technology*, Vol. 51(9–12), pp. 1243–1250, 2010.

Samimi, A., Mohammadian, A. and K. Kawamura. A Behavioral Freight Movement Microsimulation Model: Method and Data, In: *Journal of Transportation Letters: The International Journal of Transportation Research* 2 (1), pp. 53-62, 2010.

Samimi, A., Mohammadian, A., K. Kawamura. An Online Shipment Survey in US: Lessons Learnt and a Non-response Bias Analysis. In: Proceedings of the 89th Annual Meeting of the Transportation Research Board, Washington, D.C. January 2010.

Samimi, A., Pourabdollahi, Z., Mohammadian, A., and K. Kawamura. An activity-based freight mode choice microsimulation model. In: Proceedings of the 91st Annual Meeting of the Transportation Research Board, Washington, D.C., 2012.

SAS Institute Inc., SAS software, Version 9.2 of the SAS System for Windows. Copyright © 2002-2008 by SAS Institute Inc., Cary, NC, USA.

SAS PublicationsDecision Trees: What Are They? Decision Trees for Business Intelligence and Data Mining: Using SAS Enterprise Miner. http://support.sas.com/publishing/pubcat/chaps/57587.pdf, Accessed October 2013.

Schaefer, R. and D.A. Dillman . Development of a Standard E-Mail Methodology: Results of an Experiment, *Public Opinion Quarterly*, Fall, Vol. 62, pp. 378-397. 1998.

Seydel, J., Data envelopment analysis for decision support. *Industrial Management and Data Systems*, Vol. 106 (1), pp.81–95, 2006.

Seydel, J., Supporting the paradigm shift in vendor selection: Multicriteria methods for sole sourcing, *Managerial Finance*, Vol. 31 (3), pp. 49–64, 2005.

Sklar, A. Random Variables, Joint distribution functions, and copulas. *Kybernetika*, Vol. 9, pp. 449-460, 1973.

Stefan, K. J., J. D. P. McMillan, and J. D. Hunt. An Urban Commercial Vehicle Movement Model for Calgary. In: *Transportation Research Record*, 1921, pp. 1-10, 2005.

Sturm, K., Pourabdollahi, Z.,and A. Mohammadian, Descriptive and non-response bias analyses for a nationwide freight establishment survey. In: The 92nd Annual Meeting of the Transportation Research Board, Washington, D.C., 2013.

Talluri, S. and R.C. Baker, A multi-phase mathematical programming approach for effective supply chain design. *European Journal of Operational Research*, Vol. 141 (3), pp. 544–558, 2002.

Tavasszy, L.A. Freight Modeling– An Overview of International Experiences, In: *Transportation Research Board Conference, Freight Demand Modeling: Tools for Public Sector Decision Making Summary of Conference*, 40, pp. 47-55, 2008.

Tavasszy, L.A., Smeenk,B., and C.J. Ruijgrok. A DSS for Modeling Logistics Chains in Freight Transport Policy Analysis, In: *International Transportation Operation Research*, 5 (6), pp. 447-459, 1998.

Train, E. K. *Discrete choice methods with simulation*. Cambridge University Press, 2003.

Train, K. *Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand*. The MIT Press, Cambridge, 1986.

Trivedi, P.K., and D.M. Zimmer. Copula Modeling: An Introduction for Practitioners. *Foundations and Trends in Econometrics*, Vol. 1, Issue 1, Now Publishers, 2007.

U.S. Census Bureau, Population and Housing Unit Estimates 2011, Historical data, U.S. Department of Commerce. http://www.census.gov/econ/cbp/, Accessed August 2013

U.S. Census Bureau, Historical Income Tables: Households 2011, Table H.6, U.S. Department of Commerce. http://www.census.gov/hhes/www/income/data/historical/household/index.html, Accessed August 2013

U.S. Census Bureau: Statistics of U.S. County Businesses Patterns (CBP-2007). U.S. Department of Commerce, 2009. http://www.census.gov/popest/, Accessed August 2013

U.S. Census Bureau, Commodity Flow Survey (CFS-2007). U.S. Department of Commerce, 2009. http://www.census.gov/econ/cfs/, Accessed August 2013

Wadhwa, V. and A.R. Ravindran, Vendor selection in outsourcing. *Computers and Operations Research*, Vol. 34 (12), pp.3725–3737, 2007.

Weber, C.A., Current, J.R. and W.C. Benton, Vendor selection criteria and methods, *European Journal of Operational Research*, 50, pp. 2–18, 1991.

Weber, C.A., Current J.R. and A. Desai, Non-cooperative negotiation strategies for vendor selection, *European Journal of Operational Research*, 108, pp. 208–223, 1998.

Weber, C.A., Current J.R. and A. Desai, An optimization approach to determining the number of vendors to employ, *Supply Chain Management*, Vol. 5, No. 2, pp. 90–98, 2000.

Windisch, E., De Jong, G. C., van Nes, R. and Hoogendoom, S. P. A Disaggregate Freight Transport Model of Transport Chain and Shipment Size Choice. In: Proceedings of European Transport Conference, Glasgow, Scotland, 11-13 October 2010.

Wisetjindawat, W., and K. Sano. A Behavioral Modeling in Microsimulation for Urban Freight Transportation, In: *Journal of the Eastern Asia Society for Transportation Studies*, 5, pp. 2193–2208, 2003.

Wisetjindawat, W., Sano, K., Matsumoto, S., and P.Raothanachonkun. Microsimulation Model for Modeling Freight Agents Interactions in Urban Freight Movement. In: Proceedings of 86th Annual Meeting of the Transportation Research Board, Washington DC, 21–25 January 2007.

Wu, T., and J. Blackhurst, Supplier evaluation and selection: An augmented DEA approach. *International Journal of Production Research*, Vol. 47(16), pp.4593–4608, 2009.

Yang, C. H., J. Y. J. Chow, and A. C. Regan. State-of-the Art of Freight Forecasting Modeling: Lessons Learned and the Road Ahead. In: Proceedings of the 88th Annual Meeting of the Transportation Research Board, Washington DC, 11–15 January 2009.

Yu, M., Goh, M. and H. Lin, Fuzzy multi-objective vendor selection under lean procurement. In: *European Journal of Operational Research*, Vol. 219(2), pp. 305–311, 2012.

VITA

Education

University of Illinois, Chicago (UIC)

PhD Candidate in Transportation Engineering, August 2010 – November 2014 GPA: 3.85/4.0

Thesis Proposal: An Agent-based Freight Transportation Modeling Framework

Sharif University of Technology, Tehran, Iran

Master of Science in Transportation Engineering, February 2010 GPA: 3.7/4.0

Thesis: Travel Demand Modeling under Congestion Pricing, 2010

Sharif University of Technology, Tehran, Iran

Bachelor of Science in Civil Engineering, May 2007 GPA: 3.63/4.0

Teaching Experience

University of Illinois, Chicago, IL

Civil and Materials Engineering Department

Graduate Teaching Assistant

- Highway Geometric Design class Fall Semester 2014

Research Experience

University of Illinois, Chicago, IL **08/2010-12/2014**

Civil and Materials Engineering Department

Graduate Research Assistant

- CFIRE: Behavioral microsimulation model of multimodal freight movement (2010–2014)
- Goods Movement Study in Illinois: Application to Freight Transportation and Logistics
  http://ict.illinois.edu/publications/report%20files/FHWA-ICT-13-010.pdf (2011-2013)
- GPS Based Pilot Survey of Freight Movements in the Midwest Region
  http://www.wistrans.org/cfire/documents/FR_CFIRE0413.pdf (2011-2013)
- Modeling Seniors' Activity-Travel Data
  https://www.ideals.illinois.edu/handle/2142/45748 (2011-2013)

Sharif University of Technology, Tehran, Iran **01/2009-01/2010**

Civil Engineering Department

Graduate Research Assistant

- Travel Demand Modeling Under Congestion Pricing (2011-2013)

Professional Affiliations

Council Awards Chair; Freight Mobility Council, Institute of Transportation Engineers **2012 – 2013**

Member; Freight Mobility Council Executive Committee, Institute of Transportation Engineers **2012**

Member; Institute of Transportation Engineers Inc., **2011 - Present**

Publications

Pourabdollahi, Z., Karimi, B., Mohammadian, A., and Kawamura, K. Shipping Chain Choices in Long Distance Supply Chains: Descriptive Analysis and a Decision Tree Model, Transportation Research Records: Journal of the Transportation Research Board, Washington D.C., No. 2410, pp. 58-67, 2014

Sturm, K., Pourabdollahi, Z., Mohammadian, A., and Kawamura, K. GPS and Driver Log-Based Survey of Grocery Trucks in Chicago, Transportation Research Records: Journal of the Transportation Research Board, Washington D.C., No. 2410, pp. 31-39, 2014.

Samimi, A., Mohammadian, A., Kawamura, K., and Pourabdollahi Z. An Activity-based Freight Mode Choice Microsimulation Model, Transportation Letters: The International Journal of Transportation Research, Maney Publishing, Vol. 6, No. 3, pp. 142-152, 2014.

Pourabdollahi, Z., Karimi, B., and Mohammadian, A. A Joint Model of Freight Mode and Shipment Size Choice, Transportation Research Records: Journal of the Transportation Research Board, Washington D.C., 2378, pp. 84-91, 2013.

Conference Presentations

Pourabdollahi, Z., Karimi, B., Mohammadian, A., and Kawamura, K. Shipping Chain Choices in Long Distance Supply Chains: Descriptive Analysis and a Decision Tree Model, In the 93rd Annual Transportation Research Board Meeting, Washington D.C. 11-16, January, 2014.

Pourabdollahi, Z., Karimi, B., Mohammadian, A., and Kawamura, K. A Multi-Criteria Supplier Evaluation Model for Supply Chains, In the 93rd Annual Transportation Research Board Meeting, Washington D.C. 11-16, January, 2014.

Sturm, K., Pourabdollahi, Z., Mohammadian, A., and Kawamura, K. GPS and Driver Log-Based Survey of Grocery Trucks in Chicago, In the 93rd Annual Transportation Research Board Meeting, Washington D.C. 11-16, January, 2014.

Karimi, B., Pourabdollahi, Z., and Mohammadian, A. Nested Logit-Based Latent Segmentation Model for Examining Rhythms of Long-Distance Trips in Illinois, In the 93rd Annual Transportation Research Board Meeting, Washington D.C. 11-16, January, 2014.

Karimi, B., Pourabdollahi, Z., and Mohammadian, A. Rhythms in Shopping Participation of Seniors Versus Younger Cohort: Latent Segmentation AFT-Based Duration Model. In the 93rd Annual Transportation Research Board Meeting, Washington D.C. 11-16, January, 2014.

Pourabdollahi, Z., Mohammadian, A., and Kawamura, K. An Optimization Model for Supplier Selection, In the INFORMS 2013 Annual Meeting, Minneapolis, MN, October 6-9, 2013.

Pourabdollahi, Z., Mohammadian, K. and Kawamura K., A Hierarchical Database Model of Shipping Chain Choice for Freight Transportation, Student Freight Symposium, The University of Memphis Intermodal Freight Transportation Institute, 4-6, February 2013.

Pourabdollahi, Z., Karimi, B., Mohammadian, K. and Kawamura K., An Optimization Model of Warehouse Location and Retailer Allocation, Student Freight Symposium, The University of Memphis Intermodal Freight Transportation Institute, 4-6, February 2013.

Pourabdollahi, Z., Karimi, B., and Mohammadian, A. A Joint Model of Freight Mode and Shipment Size Choice, In the 92th Annual Transportation Research Board Meeting, Washington D.C. 12-18, January, 2013.

Pourabdollahi, Z., Javanmardi, M., Karimi, B., Mohammadian, A., and Kawamura, K. 2013. Mode and Shipment Size Choice Models in the FAME Simulation Framework:, In the 92th Annual Transportation Research Board Meeting, Washington D.C. 12-18, January, 2013.

Sturm, K., Pourabdollahi, Z., Mohammadian, A., and Kawamura, K. Nationwide Establishment and Freight Survey: Descriptive and Nonresponse Bias Analyses. In the 92th Annual Transportation Research Board Meeting, Washington D.C. 12-18, January, 2013.

Pourabdollahi, Z., Mohammadian, A., and Kawamura, K. A Behavioral Freight Transportation Modeling System: An Operational and Proposed Framework, Proceedings of the 14th Annual International Conference on Electronic Commerce, ACM, August 2012, 196-203.

Pourabdollahi, Z., Mohammadian, A., and Kawamura, K., A Conceptual Framework for a Behavioral Freight Transportation Modeling System with Logistics Choices, 13th International Conference on Travel Behaviour Research, Toronto 15-20, July 2012.

Pourabdollahi, Z., Mohammadian, A., and Kawamura, K., A. A Behavioral Freight Transportation Modeling System, In: 4th Transportation Research Board Conference on Innovations in Travel Modeling (ITM), Tampa, Florida, 29 April-2 May, 2012.

Samimi A., Pourabdollahi, Z., Mohammadian, A., and Kawamura, K. An Activity-based Freight Mode Choice Microsimulation Model, In proceedings of: 91th Annual Transportation Research Board Meeting, Washington D.C. 22-26, January, 2012.

Pourabdollahi, Z., and Moahhamadian, A. A Conceptual Framework for a Behavioral Freight Transportation Modeling System with Logistics Choices, In: The 4th METRANS National Urban Freight Conference, Long Beach, California, 11-14 October, 2011.

Pourabdollahi, Z., and Moahhamadian, A. FAME II: A Proposed Framework for a Behavioral Freight Transportation Modeling System with Logistics Choices, In: The Mid-Continent Transportation Research Symposium, Ames, Iowa, 18-19 August 2011.

Pourabdollahi, Z., and Moahhamadian, A. An Enhanced Freight Activity Micro-simulation Estimator (FAME) Model using Freight Analysis Framework (FAF3.0) dataset, In: The Mid-Continent Transportation Research Symposium, Ames, Iowa, 18-19 August, 2011.