

NCFRP 20: Developing Sub-National Commodity Flow Data Subtask Report

Subtask Report - Review of Sub-National Commodity Flow Data Development Efforts and National Freight-Related Data Sets

literature review

prepared for

National Cooperative Freight Research Program

prepared by

Cambridge Systematics, Inc.

with

Ken Casavant, Pullman, WA
Anne Goodchild, Seattle, WA
Eric Jessup, Freiburg, Germany
Catherine T. Lawson, Albany, NY

literature scan

NCFRP 20: Developing Sub-National Commodity Flow Data

Task 1 and 2 Report - Review of Sub-National Commodity Flow Data Development Efforts and National Freight-Related Data Sets

prepared for

National Cooperative Freight Research Program

prepared by

Cambridge Systematics, Inc.
730 Peachtree Street, NE, Suite 1050
Atlanta, GA 30308

with

Ken Casavant, Pullman, WA
Anne Goodchild, Seattle, WA
Eric Jessup, Freiburg, Germany
Catherine T. Lawson, Albany, NY

date

April 2010

Table of Contents

1.0	Introduction	1-1
2.0	Methodology.....	2-1
3.0	Subnational Commodity Flow Disaggregation Techniques and Databases.....	3-1
3.1	Description of Research-Oriented Disaggregation Techniques	3-5
3.1.1	Development of a Computerized Method to Subdivide the FAF2 Regional Commodity O-D Data to County-Level O-D Data	3-5
3.1.2	Securing Truck Travel Data in Texas.....	3-11
3.1.3	Development of Nation-Wide Freight Analysis Zones.....	3-13
3.1.4	Development of County-to-County Freight O-D Matrix from 2002 FAF2 Freight Data for Subsequent Network Assignment.....	3-15
3.1.5	Commodity-Specific Disaggregation of the 2002 FAF2 Data to the County Level for New Jersey	3-17
3.1.6	University of Alabama Huntsville Commodity Flow Disaggregation Method	Error! Bookmark not defined.
	Eastern Washington Intermodal Transportation Study	3-19
3.1.7	Use of Truck GPS Data For Freight Forecasting.....	3-21
3.1.8	University of Alabama Huntsville Commodity Flow Disaggregation Method	3-22
3.1.9	University of Alabama Huntsville Commodity Flow Disaggregation Method	Error! Bookmark not defined.
3.1.10	University of Alabama Huntsville Commodity Flow Disaggregation Method	Error! Bookmark not defined.
3.1.11	University of Alabama Huntsville Commodity Flow Disaggregation Method	Error! Bookmark not defined.
3.1.12	University of Alabama Huntsville Commodity Flow Disaggregation Method	Error! Bookmark not defined.
3.2	Description of Planning-Oriented Disaggregation Techniques	3-26
3.2.1	Texas NAFTA Study Update – Final Report	3-26
3.2.2	LACMTA Cube Cargo Model.....	3-29
3.2.3	Virginia Statewide Multimodal Freight Study, Phase I	3-31
3.2.4	Central Coast California Commercial Flows Study	3-33

3.2.5	Southern California Association of Governments Global Positioning Survey Commodity Flow Data.....	3-37
3.2.6	California Commodity O-D Database Disaggregation	3-39
3.2.7	Oregon DOT Commodity Flow Database Development.....	3-39
3.2.8	Strategic Freight Transportation Analysis	3-40
3.3	Commodity-Specific Disaggregation Techniques.....	3-44
3.3.1	Potato Flows in Washington State.....	3-44
3.3.2	Diesel Distribution in Washington State	3-47
3.3.3	Hazardous Materials Commodity Flow Data and Analysis, Ongoing, Texas A&M University	3-50
3.4	International Disaggregation Techniques and Freight Survey Efforts	3-50
3.4.1	Sweden Commodity Flow Survey.....	3-50
3.4.2	International Conference on Transport Survey Methods (Part 1)	3-55
3.4.3	International Conference on Transport Survey Methods (Part 2)	3-58
3.4.4	2000 Calgary Commodity Flow Survey Report.....	3-59
3.4.5	Edmonton Region Truck/Commodity Survey	3-60
3.4.6	A Shipper-Based Survey of Goods and Service Movements in the Greater Gold Horseshoe (GGH) Report I: Survey Design and Implementation.....	3-62
3.4.7	Edmonton Commodity Flow Study.....	3-63
3.4.8	Canadian Trucking Commodity Survey	3-64
3.5	Lessons Learned from Disaggregation Techniques.....	3-65
4.0	Review of Freight-Related Data Sources.....	4-1
4.1	Description of Core Freight-Related Data Categories	Error! Bookmark not defined.
4.2	Multimodal Freight Flow Databases.....	4-6
4.2.1.	TRANSEARCH	4-6
4.2.2	Commodity Flow Survey (CFS).....	4-6
4.2.3	Freight Analysis Framework (FAF)	4-14
4.3	Mode-Specific Freight Flow Databases	4-27
4.3.1	Vehicle Inventory and Use Survey (VIUS).....	4-27
4.3.2	Carload Rail Waybill Sample	4-30
4.3.3	U.S. Waterborne Commerce.....	4-32
4.3.4	U.S. Ports and Waterway Facility Database.....	4-34

4.4	Multimodal International Gateway Freight Flow Databases	4-37
4.4.1	BTS Transborder Surface Data	4-37
4.4.2	PIERS	4-39
4.5	Economic and Establishment Data	4-42
4.5.1	U.S. Census Bureau	4-42
4.5.2	County Business Patterns	4-45
4.5.3	Woods & Poole County Economic and Demographic Projections	4-47
4.5.4	InfoUSA Establishment Data	4-48
4.5.5	Dun & Bradstreet (D&B) Establishment Data	4-51
4.6	Summary of Core Data Sets Strengths and Weaknesses	5-1
5.0	Implications for Guidebook Development	5-1
A.	Expanded List of Commodity Flow-Related Databases	A-1
A.1	Oil Pipeline Data	A-1
A.2	Air Traffic Statistics	A-4
A.3	Massachusetts Institute for Social and Economic Research (MISER)	A-5
A.4	Global Insight World Trade Service (WTS)	A-6
A.5	IMPLAN Data Files	A-7
A.6	U.S. MARAD Maritime Statistics	A-9
A.7	Foreign Trade U.S. Census Bureau	A-9
A.8	Caltrans Weigh-in Motion Data	A-10
A.9	U.S. Department of Agriculture (USDA)	A-11
A.10	U.S. Energy Information Administration (EIA)	A-12
A.11	Quarterly Census of Employment and Wages	A-12
A.12	FleetSeek National Motor Carrier Directory	A-14
A.13	FleetSeek Private Fleet Directory	A-16
A.14	FleetSeek Owner-Operator Database	A-18
A.15	ATA North American Truck Fleet Directory	A-20
A.16	Vehicle Travel Information System (VTRIS)	A-21
A.17	Regional Economic Accounts	A-23
A.18	Bureau of Labor Statistics: Productivity Statistics	A-26
A.19	TRB E-Circular E-C080, Freight Data for State Transportation Agencies – A Peer exchange, July 11, 2005	A-28

List of Tables

Table 3.1	Summary Matrix Tables of Subnational Commodity Flow Databases.....	3-3
Table 3.2	Original FAF ² Flow from Jacksonville to Atlanta.....	3-10
Table 3.3	Disaggregation Method Descriptions	3-17
Table 3.4	Examples of Applications That the Database Can Be Used For.....	3-43
Table 4.1	Freight-Related Data Sources	4-3
Table 4.2	Freight-Related Data Sources	4-5

List of Figures

Figure 2.1	Supermarket Supply Chain Example.....	2-3
Figure 2.2	Integrated PetroleumSupply Chain Example	2-3
Figure 3.1	Steps to Creating the Final Commodity Flow Database for the Study	3-36

1.0 Introduction

Information about commodity flows is a critical data resources for freight planning. Commodity flow data describe the quantity (usually measured in dollar value or tons) of commodities (products) that are shipped between origins and destinations and typically provide information on what transportation modes are used. In other words, commodity flow data describe what moves, where it moves from and to, how much of it moves, and by what modes it moves. This information is critical to freight planning for the following reasons:

- Commodity flow data provide a direct link between the inputs and outputs of an economy and the freight flows that the economy gives rise to. Commodity flows measure what is produced and consumed in an economy and can be used to generate estimates of demand for freight transportation. Commodity flow information also can be used to understand which industries generate demand for freight, and therefore benefit from freight investments. This is a critical input to decisions that prioritize freight investments.
- Commodity flow information – the types of commodities, the amount being shipped, and the distances being shipped – is a critical determinant of modal choice and commodity flow data are critical in modal diversion studies. In order to optimize multimodal transportation systems, state departments of transportation (DOTs), metropolitan planning organizations (MPOs), and Federal transportation agencies are often interested in examining the costs and benefits of competing modal investments. Commodity flow information is critical to determining how much freight is potentially divertible from one mode to another.
- Commodity flow data are often used as inputs to state and multi-state models that forecast freight transportation demand. Commodity flow data are critical in multi-state corridor studies because they provide the best representation of freight transportation demand at this level of geography. Commodity flow data also are increasingly being used to estimate air quality impacts of the transportation inputs to various industrial activities.

There are a number of public commodity flow data resources that are collected consistently at the national level, but no comparable data sources are collected at the sub-national level. In addition to these data sets, there are a number of other specialized data bases that provide partial commodity flow information. There also are private, proprietary commodity flow databases that utilize a combination of the publicly available databases, proprietary establishment data, and data provided by the private sector.

While these databases provide a wealth of useful information for national freight transportation planning, their application to sub-national planning often falls short of the needs of many state DOTs and MPOs. There are a number of well-documented problems facing state DOTs and MPOs in their use of national commodity flow data sets and the objective of this research is to address as many of these problems as possible. Some of these problems are as follows:

- **Lack of understanding of existing commodity flow databases** – Often state and local planners are not familiar with the existing databases and they may misinterpret the data or utilize data from different data sets in ways that lead to data errors.
- **Lack of publicly-available, fully-disaggregated commodity flow data for use in state and local studies** – State DOTs and MPOs often want to use commodity flow data to analyze commodity movements in specific corridors or within an MPO region. Getting accurate commodity flow data at the county or TAZ level using existing publicly available data sources is very difficult.
- **Inaccurate data at the local level when national data sets are applied** – Since national data sets are often developed from surveys and sampling procedures that are not sufficient for capturing local patterns, the results obtained from disaggregating these databases are often at odds with these local patterns.
- **Inability to provide explicit linkages of intermodal or multimodal cargo moves all the way from point of production to point of consumption** – Many of the commodity flow data sources and survey techniques do not track cargo moves in a set of linked modal transfers from point of production of a product to point of consumption. This does not directly align with the desire of many state DOTs and MPOs who want to use commodity flow data as the primary freight demand input to travel demand models and to study their impact on specific highways or rail lines. Commodity flow data often do not explicitly account for intermediate handling of cargo in specifying origins and destinations. Additionally, they also may not include all local distribution activity within an urban area and they do not account for movements of empty vehicles or service trucking.
- **Lack of route-specific commodity flow data** – The current set of techniques used to route commodity flow data to mode-specific infrastructure is often complex and resource-intensive. For example, travel demand models are commonly used to estimate truck volumes on the highways, but these models typically lose most, if not all, of the commodity detail that is included in commodity flow databases. This problem is even more acute in urban areas, where routing patterns are difficult to estimate, and commodity flow data are typically less accurate.

As an important step to addressing some of these issues, this report is a literature review that describes existing examples of efforts to construct sub-national commodity flow databases. This literature also describes some of the major

national commodity flow data that is used for freight planning activities. This literature review is part of a larger study to assist freight planners in developing sub-national commodity flow data for freight planning purposes. The overall study consists of the following eight tasks:

- **Task 1** – Describe and review examples of recent or current efforts at the sub-national level to compile and use commodity flow information for transportation planning and analysis.
- **Task 2** – Describe national data sets and their use and limitations for application at subnational levels. Evaluate disaggregation techniques that have been developed.
- **Task 3** – Develop procedures and methodologies for conducting commodity flow surveys at subnational levels including development of a sample survey instrument(s).
- **Task 4** – Demonstrate the application of the procedures and methodologies, including the survey instrument developed in Task 3 by applying it in a test case to address such issues as modal diversion, air quality, and/or public freight investment prioritization.
- **Task 5** – Prepare an interim report providing the results of Tasks 1 through 4. In addition, the interim report should include a list of 40 prospective invitees for the Task 7 workshop and a detailed draft outline of the guidebook.
- **Task 6** – Prepare a draft guidebook that will illustrate how the data should be collected and used in models for decision-making as well as provide guidance for compiling commodity flow data sets appropriate for subnational analysis.
- **Task 7** – Present the draft guidebook at a one-day TRB freight data workshop in Irvine, CA, in the fall of 2010.
- **Task 8** – Based on project panel review of the Task 7 revised guidebook, prepare a final version of the guidebook. In addition, prepare a final report for the overall project with an executive summary that documents the entire research effort.

This report serves as the deliverable for Tasks 1 and 2. There are three primary goals for this report:

1. Assemble background information that will be used to identify data gaps in the development of subnational commodity flow databases;
2. Develop background information to identify case studies for Task 3 of this study; and
3. Provide background information that will be used in the guidebook to be developed in Task 6 through Task 8 of this study.

This report is structured into the following five sections:

- Section 1.0: Introduction – Description of the literature review process.
- Section 2.0: Methodology – Description of the process used to identify, describe, and summarize the documents included in the literature review.
- Section 3.0: Subnational Commodity Flow Disaggregation Techniques and Databases – Description of literature and processes that describe the process of disaggregating national commodity flow data to local commodity flow data. This section also describes some local data collection efforts that have the potential to assist in subnational commodity flow database development.
- Section 4.0: Review of Freight-Related Data Sources – Description of the three freight flow databases most commonly used for freight analysis (the Commodity Flow Survey, Freight Analysis Framework, and Transearch) along with their development methodologies, weaknesses and limitations. Additionally, eight other supporting freight data sets are described in terms of their contents and development methodologies.
- Appendix A: Expanded List of Commodity Flow-Related Databases.

2.0 Overview and Methodology

2.1 WHAT IS A FREIGHT FLOW?

For purposes of this report, there are several definitions that have been assumed and are carried through the review of databases and literature as follows:

- Freight flows, in the simplest sense, are the movement of goods from one location to another
- Commodity flows are similar to freight flows, except that one or more commodities are specified in the movement of goods
- Commodity flow databases, therefore, describe the movement of one or more commodities between a set of origins and destinations
- Mode-specific commodity flow databases are commodity flow databases that also describe the mode used for moving goods. These modes can include truck, rail, truck-rail intermodal, air, and water. However, these modes can also have submodes such as truckload, less than truckload, and private trucking for the trucking industry or intermodal and carload for rail. Modal descriptions can also include multimodal categories such as truck-rail intermodal, water-rail movements, or truck-air movements.

It is also important to note the information that is not included in commodity flow databases. Commodity flow databases typically do not contain any information about vehicle routing. Vehicle routing tends to occur as a post-processing activity for commodity flow databases. Commodity flow databases also typically do not include information on the number of vehicles that are utilized to move the goods. This is true even for mode-specific commodity flow databases. When needed, this information is also typically estimated using some sort of post-processor. Finally, these databases tend to measure flows in fixed time periods, typically annually. Detailed day-of-week and month-of-year data are needed to make these annual estimates relevant for shorter time periods.

2.2 COMPARING FREIGHT FLOWS TO SUPPLY CHAINS

Commodity flows are different from supply chains. Supply chains generally follow a specific good from an initial origin through to final consumption. Along this supply chain path, the good may be transformed from a raw good to a finished good, thereby altering its commodity classification. The good may also pass through several freight facilities from its initial origin and destination points, change ownership multiple times, and be carried across multiple modes.

Figure 2.1 shows an example of a supply chain for supermarkets. Figure 2.2 shows an example of petroleum refining and distribution. It is important to note how each of these supply chains differ from commodity flow databases. In the supermarket supply chain example, a commodity flow database will capture some of the supply chain movements but not all of them. Additionally, the method used for each of the captured movements can differ greatly. For example, in the supermarket example in Figure 2.1, a commodity flow database may cover finished goods from food processor plants and food processor distribution centers to regional distribution centers. But the same commodity flow database may not track goods from a supermarket's regional distribution center to local grocery stores. This same commodity flow database may track goods from food processor plants using a detailed survey instrument that covered a high percentage of the facility operators. But then, the movement of goods from local vendors to grocery stores might be estimated based on national or regional economic or population data. This will have tremendous implications for the accuracy of the data, and the potential sources of error as these data are used for various freight planning purposes.

The integrated petroleum refining and distribution example (Figure 2.2) highlights potential modal and geographic complexities of commodity flow databases. Crude oil can come from multiple geographic sources in the United States, Middle East, Africa, or Central and South America. Depending on the location of crude oil, it will be shipped using various modes to get to the refinery location, such as pipeline and oil tanker ships. Commodity flow databases that are stronger in particular modes are likely to miss some of the shipments of this good. Similarly, commodity flow databases that are stronger on domestic moves rather than international flows will also only tell part of the story. Therefore, it is critical to understand the methodologies used for each of the major databases to understand which components of supply chains are captured, how well it is captured, what the implications for using the database are for freight planners, and what types of local data collection would be most useful to provide more suitable commodity flow databases.

Figure 2.1 – Supermarket Supply Chain Example

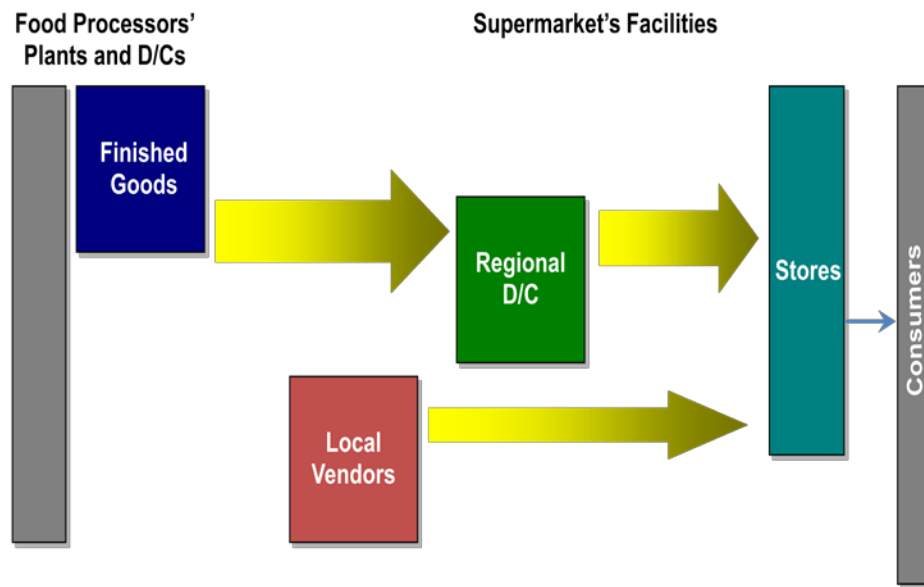
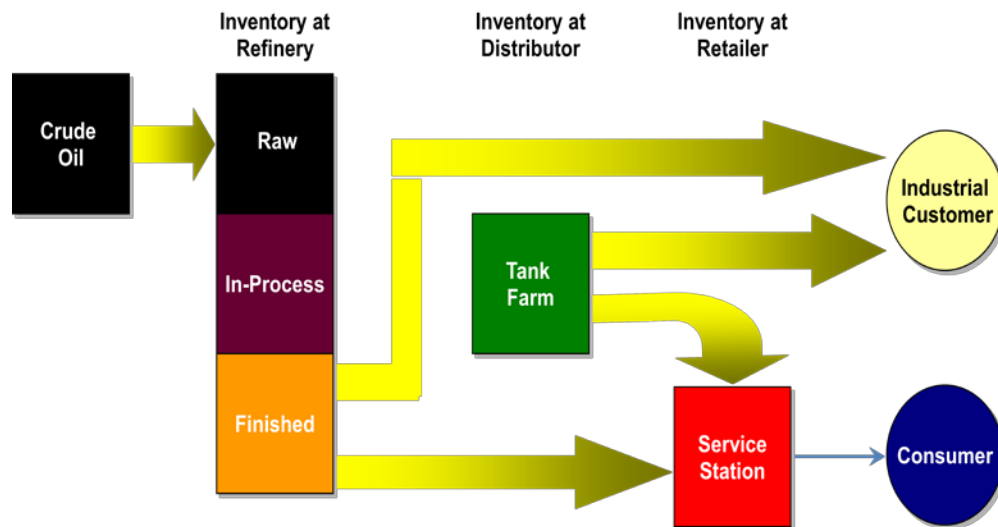


Figure 2.2 – Integrated Petroleum Supply Chain Example



2.3 METHODOLOGY FOR LITERATURE REVIEW

Review of Subnational Commodity Flow Database Techniques

For the review of subnational commodity flow database techniques, projects and documents were selected based on a combination of the research team experience with the subject matter and specific projects identified by the NCFRP Project Panel. This generated a list of several dozen projects that had developed subnational commodity flow databases. Many of these projects were filtered out of the review to avoid repetition of similar techniques. The remaining projects were selected to identify a broad range of techniques that utilized a broad set of data.

These remaining projects were categorized based on whether or not the technique was developed primarily for research purposes or if they were developed for specific freight planning studies sponsored by state and/or local transportation agencies. A separate set of techniques were identified for processes that were developed outside of the U.S. Each technique was described in terms of several factors including:

- The types of national data that were used;
- The types of local freight data that were used;
- The types of local socioeconomic data were used;
- The amount of geographic detail in the final subnational database;
- The approximate commodity detail in the final subnational database;
- The mode-specificity in the final subnational database; and
- Specific applications for which the database was applied

This consistent type of description across techniques allowed for the identification of gaps in existing procedures along with gaps in final databases.

Review of National Commodity Flow Data

This literature review begins with a description of the three primary commodity flow databases used for freight planning purposes:

- The Bureau of Transportation Statistics Commodity Flow Survey
- The Federal Highway Administration Freight Analysis Framework2 database
- IHS/Global Insight Transearch Database

The methodologies used to develop these three databases are defined along with a description of the final databases. Additionally, the limitations and weaknesses of each of these databases are described. Ultimately the goal of describing these databases is to determine how freight planning activities that

utilized one or more of these databases might be impacted by specific features of each database. In later tasks in this study, the consultant team will develop surveys that could potentially enhance the usefulness of applying these data to sub-national freight planning activities.

This literature review also highlighted a small set of other national commodity flow data that are commonly used to support freight planning purposes at the state and local level that are not multi-modal commodity-specific freight flow databases such as the three described above. A broader set of less commonly used freight-related databases are provided in Appendix A. For the core freight-related databases, the following information was captured:

- Description of how shipments are defined;
- Commodities and flows covered in the databases;
- Method of data collection for major data elements in the databases;
- Estimation and modeling used to develop the database; and
- Issues and limitations.

The broader set of freight databases were described in general terms with a focus on the following items:

- Type of data (commodity flow, count data, economic output data, employment, etc.);
- Modal coverage;
- Level of geographic detail;
- Accessibility (public or proprietary);
- Qualitative assessment of accuracy; and
- Miscellaneous key features of the database.

This description is used to provide documentation of how each database is constructed, what types of commodities and movements are included and excluded, and a description of the types of errors that can be introduced during disaggregation using specific databases.

3.0 Subnational Commodity Flow Disaggregation Techniques and Databases

This section describes subnational commodity flow disaggregation techniques and commodity flow databases that have been created with these techniques. Table 3.1 shows a matrix that describes the characteristics of several of these efforts. This table can be used by State DOTs and MPOs to understand which techniques to apply based on the type of freight planning effort they are pursuing, the type of data that they have available, and the level of resources they have available to devote to the study.

Each of the studies in Table 3.1 incorporates both a disaggregation technique and a resulting commodity flow database. However, some studies emphasized the disaggregation technique, while other studies were focused on developing a commodity flow database that was useful for a specific freight planning effort. Each of the studies has been characterized accordingly in Table 3.1. In the following sections, we discuss the disaggregation techniques that have been utilized in various efforts. Then, we describe a sample of commodity flow databases that were selected to enable a description of the range of planning efforts that can be supported with these databases, and to provide breadth in terms of applications of the various disaggregation techniques. An exhaustive list of commodity flow databases is provided in Appendix A.

In Table 3.1 and the remainder of this section, these disaggregation techniques are divided into three U.S.-based methods and one international method as follows:

1. **Research-Oriented Disaggregation Techniques** – These are multi-commodity techniques that were developed without specific applications required.
2. **Planning-Oriented Disaggregation Techniques** – These are multi-commodity techniques that were developed for specific applications.
3. **Commodity-Specific Disaggregation Techniques** – These are techniques that were designed to estimate a commodity flow database for a single commodity or a very narrow group of commodities.
4. **International Disaggregation Techniques** – These are techniques developed for regions outside of the U.S.

Table 3.1 Summary Matrix Tables of Subnational Commodity Flow Databases – National, Local and Economic Data Used

Subnational Commodity Flow Database or Disaggregation Technique	Types of National Data Used							Types of Local Freight Data Used				Types of Local Economic Data Used					
	FAF	FAF2	CFS	VUS	BTS Transborder	TRANSEARCH	Other	Roadside O-D surveys	Establishment Surveys	Local Truck Counts or VMT	Other	TAZ Employment	Census	Supply Chain Data	Population	IMPLAN	Other
Research-Oriented Disaggregation Techniques																	
FHWA Method to Subdivide FAF2 Regional Commodity to County Level		●											●				
Securing Truck Data For Texas			●									●				●	
Development of Nationwide Freight Analysis Zones			●														
Development of County-to-County Freight O-D Matrix from 2002 FAF2 Freight Data for Subsequent Network Assignment		●			●					●			●		●		Income
UAH Commodity Flow Disaggregation Method																	
Planning-Oriented Disaggregation Techniques																	
Texas DOT NAFTA Highways	●				●	●		●									
LAMTA Cube Cargo and Dynasim Models						●					ITMS	●			●		
Virginia DOT Statewide Multimodal Freight Study						●		●									
Central Coast AMBAG Commodity Flow Database	●							●								●	
SCAG GPS-Equipped Vehicle Commodity Flow Data											Land-Use and GPS Data						
California O-D Database Disaggregation		●					Nat'l Ag Stats, NETL CPPDB						●				
Oregon DOT Commodity Flow Database Development		●					Waterborne Commerce				OR Energy Report					●	
SFTA								●	●								
Commodity-Specific Disaggregation Techniques																	
Supply Chain Examples – Potato Manufacturing							U.S.D.A.				WA Potato Commission						AC Nielsen
Supply Chain Examples – Distribution of Diesel											WADEP, Petrocard						WA Dept. of Ecology
Hazardous Materials Commodity Flow Data and Analysis																	
International Disaggregation Techniques and Freight Survey Efforts																	
Swedish Commodity Flow Survey									●								
2000 Calgary Commodity Flow Survey Report									●								
International Conference on Transport Survey Methods																	

Table 3.2 Summary Matrix Tables of Subnational Commodity Flow Databases – Geographic, Commodity and Mode Detail

Subnational Commodity Flow Database or Disaggregation Technique (continued)	Geographic Detail in Final Database					Approx. Commodity Detail in Final Database			Mode Detail in Final Database					Other Items			
	TAZ	County	Region	Zip Code	Other	2-Digit	1-digit	None	Truck	Rail	Air	Water	Pipeline	Cost	Complexity	Level of Accuracy	Applications of Methodology
Research-Oriented Disaggregation Techniques																	
FHWA Method to Subdivide FAF2 Regional Commodity to County Level		•				•			•					8	9	5	None
Securing Truck Data For Texas	•						•		•					8	9	5	None
Development of Nationwide Freight Analysis Zones					•				•	•							
Development of County-to-County Freight O-D Matrix from 2002 FAF2 Freight Data for Subsequent Network Assignment									•	•							
UAH Commodity Flow Disaggregation Method																	
Planning-Oriented Disaggregation Techniques																	
Texas DOT NAFTA Highways									•								Corridor planning for cross-border highways
LAMTA Cube Cargo and Dynasim Models	•			•					•								Used to estimate corridor-level and regional truck flows in Southern California
Virginia DOT Statewide Multimodal Freight Study		•						•	•	•	•	•					Truck-rail diversion analysis, corridor planning, economic impact analysis
Central Coast AMBAG Commodity Flow Database		•				•			•	•	•	•	•				Intent to use for updating California Goods Movement Action Plan and to inform the California Marine and Intermodal Transportation Council
SCAG GPS-Equipped Vehicle Commodity Flow Data					Facility-level	•			•								
California O-D Database Disaggregation									•	•	•	•					
Oregon DOT Commodity Flow Database Development									•	•	•	•	•				
WSDOT Strategic Freight Transportation Analysis					City, State	•			•								Corridor studies, POE analysis, commodity profiles
Commodity-Specific Disaggregation Techniques																	
Supply Chain Examples – Potato Manufacturing					Corridor-level			Potatoes only	•								
Supply Chain Examples – Distribution of Diesel					Corridor-level			Diesel only									
Hazardous Materials Commodity Flow Data and Analysis (TBD)																	
International Disaggregation Techniques and Freight Survey Efforts																	
Swedish Commodity Flow Survey					3,100 local units		•		•	•	•	•					Populate Swedish national model for goods estimation
2000 Calgary Commodity Flow Survey Report																	
International Conference on Transport Survey Methods																	

3.1 DESCRIPTION OF RESEARCH-ORIENTED DISAGGREGATION TECHNIQUES

3.1.1 Development of a Computerized Method to Subdivide the FAF2 Regional Commodity O-D Data to County-Level O-D Data

General Information

This study was developed for FHWA. A draft report was produced in January 2009. The purpose of this study was to develop a method to divide FAF2 regional commodity O-D data for all modes and all commodities into county-level O-D data by commodity.

Commodity Flow Data Inputs

The primary data source used for the task, in addition the FAF2 database, is the County Business Patterns (CBP) published by the U.S. Census Bureau. In addition, data from the 2002 Agriculture Census conducted by the U.S. Department of Agriculture and U.S. Department of Energy's 2007 National Energy Technology Laboratory (NETL) Coal Plant database are used to develop the regression equations. Data from the U.S. Corps of Engineers is used to determine total tonnage by port and the U.S. DOT's Bureau of Transportation Statistics (BTS) border crossing data is used to determine rail and truck flows at the U.S. borders with Canada and Mexico. Data from the U.S. Department of Commerce STAT-USA's USA TRADE ONLINE is used to determine air cargo import and export.

Methodology

The methodology utilizes the relationship between employment by industry and the commodities which those industries produce and consume. For certain commodities, farm acreage, livestock and electricity generation information is used instead of or with NAICS employment data. While FAF2 data is available only at a regional level, employment by industry is more readily available at smaller levels of geography. The U.S. Census Bureau provides CBP employment by county by NAICS industry. Commercial and state data sources provide employment by NAICS or comparable industry classifications at smaller levels of geography. This employment data is aggregated to develop mathematical relationships between the FAF2 commodity shipments to and from a FAF2 region and the employment by industries in that FAF2 region. The availability of employment data by industry is used with these equations to estimate the expected production and attraction of freight tonnage in a FAF2 region and the units of smaller geography in that FAF2 region. The shares of the smaller units of geography tonnage to the regional tonnage is then used to disaggregate the freight flows from FAF2 regions to the smaller units of geography within those FAF2 regions.

The steps to disaggregate the data from FAF² regions to counties is summarized below:

- Step 1: The first step is to determine employment at the three-digit NAICS county level.
- Step 2: The next step is to build a bridge between the commodities in the FAF² reported using the two-digit SCTG and three-digit NAICS and other factors using regression methods. These regressions guide the development of factors for each commodity for the disaggregation of freight flow productions and attractions.
- Step 3: The third step involves determining the share of the originating and terminating tonnage by industry for each of the counties within a specific FAF² zone and is applied to the reported FAF² regional tonnages, to obtain the disaggregated FAF² O-D database for domestic freight shipment. Trip distribution is based on the share of productions and attractions within each zone that is allocated to each county.
- Step 4: Subsequent to step 3, the modal share by county for trips to and from each Port of Entry is determined. These shares are allocated to the county(ies) where the POEs are located. In the case of all modes, the allocation is aggregated together before proceeding to the next step.
- Step 5: The domestic origins and destinations for these exports and imports respectively are disaggregated in a manner similar to the flows (Step 3) that take place entirely within the United States.
- Step 6: The disaggregated import and export data (from POE origins and destinations) is then merged with the domestic flow data and provides the final FAF² disaggregated dataset.

Final Regression Analysis

The regression analysis conducted in Step 2 revealed that certain industries were more easily estimated using employment and other factors than others. Table 3.2 shows the R-squared for the tonnage production estimates. Table 3.3 shows the R-squared for the tonnage attraction estimates. These tables are sorted from lowest to highest R-square values to demonstrate the range of fit for the dependent variables based on the explanatory variables.

Table 3.3 Tonnage Production Regression Results

SCTG	Dependent Variable	Explanatory Variable(s)	R2
20-23	Various*	Chemical mfg	11%
10-15	Various**	Mining (except oil and gas)	13%
9	Tobacco Products	Beverage and Tobacco Product mfg	15%
1	Live Animals/Fish	Support Activities for Agriculture and Forestry	17%
16	Crude Petroleum	Oil and Gas Extraction	21%
38	Precision Instruments	Miscellaneous mfg	34%
24	Plastics/Rubber	Plastics and Rubber Products mfg	43%
2	Cereal grains	Food mfg, farm acres	48%
8	Alcoholic Beverages	Beverage and Tobacco Product mfg	50%
39	Furniture	Furniture and Related Product mfg	56%
4	Animal Feed	Support Activities for Agriculture and Forestry	60%
	Nonmetallic Mineral		
31	Products	Nonmetallic Mineral Product mfg	61%
6	Milled grain products	Food mfg	62%
	Other Coal &		
19	Petroleum Products	Oil and Gas Extraction, Petroleum & Coal Products mfg	62%
34	Machinery	Fabricated Metal Product mfg, Machinery mfg	63%
	Misc. Manufactured		
40	Products	Miscellaneous mfg	64%
	Other Agriculture		
3	Products	Food mfg, farm acres	65%
33	Articles of Base Metals	Fabricated Metal Product mfg	65%
		Forestry and Logging, Support activities for Ag and Forestry,	
25	Logs	Wood Product mfg	70%
		Machinery mfg, Computer and Electronic Product mfg,	
35	Electronic & Electrical	Electrical Equip and Appliance and Component mfg	70%
27	Newsprint/paper	Forestry and Logging, Printing and Related Activities	73%
30	Textiles/Leather	Textile Mills, Textile Product Mills	73%
36,37	Various	Transportation Equipment Manufacturing	74%
7	Other Foodstuff	Food mfg, chemical mfg	75%
26	Wood Products	Wood Product mfg	75%
32	Base Metals	Primary Metal mfg, Machinery mfg	75%
18	Fuel Oils	Petroleum & Coal Products mfg	77%
28	Paper Articles	Paper mfg, Printing and Related Activities	81%
17	Gasoline	Petroleum & Coal Products mfg	83%
29	Printed Products	Paper mfg, Printing and Related Activities	85%
5	Meat/seafood	Food mfg	86%
41	Waste and Scrap	NAICS 115, 221, 321-327, 331-339	86%
43	Mixed Freight	NAICS 321-327, 481, 483-488, 492-493	86%

* SCTG 20-23 is Basic Chemicals, Pharmaceutical Products, Fertilizers, and Chemical Products and Preparations n.e.c.

** SCTG 10-15 is Monumental or Building Stone, Natural Sands, Gravel and Crushed Stone, and Non-metallic Minerals n.e.c.

Table 3.4 Tonnage Attraction Regression Results

SCTG	Dependent Variable	Explanatory Variable(s)	R2
22	Fertilizers	Farm acres, 2000 population	19%
21	Pharmaceuticals	2000 Population	21%
9	Tobacco Products	Beverage and tobacco prod mfg, 2000 population	22%
10	Building Stone	Construction	25%
10	Building Stone	Total employment	29%
20	Basic Chemicals	Chemical mfg	29%
14	Metallic Ores	Primary Metal Mfg	31%
37	Transportation equipment	Total employment	41%
2	Cereal grains	Food mfg, Farm acres	54%
38	Precision Instruments	Total employment	57%
11	Natural Sands	Total employment	59%
13	Nonmetallic Minerals	Nondurable goods	60%
11	Natural Sands	Construction	62%
34	Machinery	Total mfg	64%
4	Animal Feed	Farm acres, livestock sold, 2000 population	65%
12	Gravel	Total employment	66%
19	Other Coal & Petroleum Products	Petroleum & Coal Prod Mfg, 2000 population	66%
3	Other ag. Products	Food mfg, Farm acres	67%
12	Gravel	Construction	67%
1	Live animals/fish	Food mfg, Farm acres, 2000 population	69%
33	Articles of Base Metals	Construction, fabricated metal prod mfg	69%
33	Articles of Base Metals	Fabricated Metal Prod mfg, wholesale	69%
16	Crude Petroleum	Oil and Gas Extraction, Petroleum & Coal Prod Mfg	72%
31	Nonmetallic Mineral Products	Nondurable goods, 2000 population	72%
24	Plastics/Rubber	Plastics and Rubber Products mfg, 2000 population	73%
34	Machinery	Construction, Machinery mfg	75%
31	Nonmetallic Mineral Products	Construction, nondurable	76%
39	Furniture	2000 Population	78%
24	Plastics/Rubber	Construction, Chemical mfg, Plastics and Rubber Prod Mfg	79%
18	Fuel Oils	Petroleum & Coal Prod Mfg, Total employment	80%

		Wood Prod mfg, Paper mfg, Printing and related support activities	
27	Newsprint/paper		80%
36	Motorized vehicles	Transportation equip mfg, Total employment	80%
6	Milled grain products	Food mfg, food services and drinking places	81%
	Misc. Manufactured		
40	Products	2000 Population	82%
35	Electronic & Electrical	Total employment	83%
26	Wood Products	Wood Prod mfg, Furniture and Related Prod mfg	84%
41	Waste and Scrap	2000 Population	84%
17	Gasoline	Petroleum & Coal Prod Mfg, Transportation	85%
23	Chemical Products	Chemical mfg, 2000 population	85%
26	Wood Products	Construction, Wood Prod mfg, Furniture and Related Prod mfg	86%
30	Textiles/Leather	Textile Mills, Textile Product Mills, Population 2000	86%
32	Base Metals	Primary Metal Mfg, Fabricated Metal Prod Mfg, Machinery mfg	86%
		Mining (except coal & gas), Petroleum and Coal Prod mfg,	
15	Coal	Electricity Generation KWH	87%
5	Meat/seafood	Food mfg, food services and drinking places	89%
7	Other Foodstuff	Food mfg, food services and drinking places	89%
8	Alcoholic Beverages	Food mfg, accommodation and food services	89%
43	Mixed Freight	2000 Population, Total employment	89%
		Printing and Related Support Activities, Publishing Industries	
29	Printed Products	(except Internet), Oil Services (except Public Administration)	92%
25	Logs	Forestry and logging, Paper mfg, Farm acres	93%
		Food mfg, Paper mfg, Printing and Related Support Activities,	
28	Paper Articles	Population 2000	93%

Final Commodity Flow Database

An example of the disaggregation results is shown in Table 3.4. It shows the tonnage and value of flows by truck for motorized vehicles (SCTG 36) and transportation equipment (SCTG 37) originating in FL Jacksonville (FAF ID 19) and destined to GA Atlanta (FAF ID 24).

Table 3.5 Original FAF² Flow from Jacksonville to Atlanta

Origin		FAF2 O Zone	Destination	FAF2 D Zone	SCTG 36 and 37 (Tons)		SCTG 36 and 37 (Dollars)	
FL Jacksonville		19	GA Atlanta	24	17,580		\$1,379,110,000	

OFIPS	OCounty	OZone	OZoneName	DFIPS	DCounty	DZone	DZonename	SCTG 36 and 37 (Tons)	SCTG 36 and 37 (Dollars)
12019	Clay	19	FL Jacksonville	13255	Spalding	24	GA Atlanta	1.40	\$106,877.51
12019	Clay	19	FL Jacksonville	13285	Troup	24	GA Atlanta	1.48	\$114,574.26
12019	Clay	19	FL Jacksonville	13293	Upson	24	GA Atlanta	0.08	\$5,985.75
12019	Clay	19	FL Jacksonville	13297	Walton	24	GA Atlanta	0.25	\$19,666.61
12031	Duval	19	FL Jacksonville	13013	Barrow	24	GA Atlanta	33.22	\$2,722,649.47
12031	Duval	19	FL Jacksonville	13015	Bartow	24	GA Atlanta	496.25	\$37,995,287.06
12031	Duval	19	FL Jacksonville	13035	Butts	24	GA Atlanta	42.25	\$3,314,302.28
12031	Duval	19	FL Jacksonville	13045	Carroll	24	GA Atlanta	146.40	\$11,718,748.41
12031	Duval	19	FL Jacksonville	13057	Cherokee	24	GA Atlanta	111.73	\$9,233,317.60
12031	Duval	19	FL Jacksonville	13063	Clayton	24	GA Atlanta	431.59	\$34,090,551.63
12109	St. Johns	19	FL Jacksonville	13013	Barrow	24	GA Atlanta	5.33	\$436,803.96
12109	St. Johns	19	FL Jacksonville	13015	Bartow	24	GA Atlanta	79.62	\$6,095,713.66
12109	St. Johns	19	FL Jacksonville	13035	Butts	24	GA Atlanta	6.77	\$531,724.83
12109	St. Johns	19	FL Jacksonville	13293	Upson	24	GA Atlanta	3.39	\$265,883.14
12109	St. Johns	19	FL Jacksonville	13297	Walton	24	GA Atlanta	11.14	\$873,578.03
Total								17,583.51	\$1,379,109,986.03

Applications of Methodology

This methodology has not yet been applied to any specific freight planning efforts.

Applicability to Other States/Regions

This methodology seems to be durable, but requires a lot of resources to repeat.

Issues and Lessons Learned

This methodology showed that local economic data can be used to estimate many of the commodities in the FAF² database. Many of the variables have R-square values well above 50 percent indicating that these local economic data points are appropriate explanatory variables for production and attraction of freight. However, because so few of the R-square values were above 90 percent, it is clear that there are other factors that also impact the amount of tonnage that are produced and attracted to a given county. Also, there are a few commodities

that have very poor statistical fits such as chemicals and many of the mined commodities.

It is interesting to note that several of the commodities correlated well with the county-level population in 2000. This demonstrates the importance of final consumption in impacting freight demand. A good portion of freight flows to destinations that are close to where people live.

Additionally, it should be noted that some of the FAF2 commodities that were regressed in this methodology were not products of the shipper survey conducted for the CFS, but were themselves modeled based on local economic factors and truck fleet database characteristics. Therefore, high degrees of fit between the explanatory and dependent variable do not necessarily indicate a match with actual onroad vehicle activity. The methodologies used for the CFS and FAF2 are described in more detail Section 4.1.

Additionally, this methodology does not take into account travel impedances when conducting trip distribution of trips that have been disaggregated to the county-level. While this is consistent with current travel demand modeling techniques on the passenger side, it may not correlate to actual truck trips which may be more likely to travel to distances that are closer rather than further, even if both trip ends are within the same FAF2 zone.

3.1.2 Securing Truck Travel Data in Texas

General Information

In Technical Report 0-4713-01, entitled *Development of Sources and Methods for Securing Truck Travel Data in Texas*, a multinomial logit (MNL) approach was proposed to estimate county-level truck travel data from the publicly available 1997 Commodity Flow Survey (CFS) and IMPLAN data. In 2004, TxDOT awarded the CTR at The University of Texas at Austin with an implementation project to explain how to use the calibrated MNL models to generate county-level truck flows for Texas and to present a detailed explanation of the required steps to calibrate the MNL models in the future. Specifically, the research team was tasked with a) embedding the MNL models into Microsoft Excel, b) populating a truck travel database with the newly released 2002 Commodity Flow Survey (CFS) data for Texas, c) developing step-by-step instruction materials on how to use the MNL models embedded in Excel, and d) piloting the instruction materials by providing training on their use to district and Transportation Planning and Programming (TPP) personnel. The instruction materials were piloted during a one-day workshop on August 25, 2006, at TxDOT's Riverside facility. This study documents the information presented to the workshop participants. For detailed step-by-step instructions on how to apply and calibrate the MNL models developed, the reader is referred to TxDOT product 5-4713-P3, entitled *Manual for the Computation of Disaggregate County-Level Truck Flows and Explanation of Model Calibration*.

Commodity Flow Data Inputs

The data inputs used for this analysis were 1997 Commodity Flow Survey data and IMPLAN economic output and input-output data.

Methodology

This study utilized a multinomial (MNL) logit model, to estimate commodity-specific freight flows at the county level. Logit models are a class of econometric models that are used to evaluate the relationship between a set of independent variables and a binary dependent variable. The predicted dependent variable is expressed as a ratio of the probability of an outcome divided by the probability of that outcome not occurring. Multinomial logit models are an expansion of the basic, binary Logit models. They are used to evaluate the relationship between a set of independent variables and a set of dependent variables that represent mutually exclusive, discrete alternatives. These alternatives can be alternate transportation routes, consumer choices, or voting patterns to name just a few.

The movement of commodity flows between zones was modeled as production flows from an origin zone (e.g., state, county) and attraction flows to a destination zone (e.g., state, county). Specifically, the annual truck flows (tonnage) was estimated from each production state to the 50 attraction states for each commodity. The production flow distribution of commodities was modeled as a function of the generalized cost of transportation and the relative attraction level of the destination zones. Similarly, the attraction flow distribution of commodities was modeled as a function of the generalized cost of transportation and the relative production level of the origin zones. Owing to a lack of generalized cost data, centroidal distances between zones were employed as the impedance measure affecting freight flow distribution.

The CTR research team calibrated two MNL models for nine of the 11 commodity groups included in the truck component of the Texas DOT statewide travel demand model. The calibrated MNL models were subsequently used to estimate county-level truck flows for Texas. The research team compared the CTR model estimates with the TRANSEARCH data for Texas. The paired sample *t* test was used to determine whether there was any statistically significant difference between the model estimates and the Transearch data used in the SAM. A confidence interval of 95 percent was specified. For six of the nine commodity groups, the mean of the model estimates for Texas county-to-state truck flows was statistically similar to the mean of the TRANSEARCH data for these flows. In the case of state-to-Texas county truck flows, the model estimates were statistically similar to the TRANSEARCH data for four of the nine major commodity groups, and for Texas county-to-county truck flows, the model estimates were statistically similar to the TRANSEARCH data for five of the nine commodity groups. It is thus believed that the MNL approach proposed by the research team provides a cost-effective alternative for obtaining Texas county-level truck flows for at least some of the commodity groups in the short term.

It should be noted that the correlation between the MNL process and the Transearch data could also be the result of the fact that both databases utilize CFS as a core dataset in its compilation.

Applications of Methodology

This methodology has not yet been applied to any specific freight planning efforts. However, it was submitted to the Texas DOT as the recommended approach for developing county-level flows for the Texas DOT Statewide Analysis Model.

Applicability to Other States/Regions

This methodology seems to be durable, and does not require a lot of resources to repeat.

Issues and Lessons Learned

While the MNL model has sufficient strengths in terms of ease of model calibration, predictive power of the model, and cost-effectiveness among others, it does have its flaws. It does not report interstate truck flow data when the number of truck flows when one unit of measure is not reported to the CFS and due to disclosure issues it does not have information about specific companies or employees. It also does not take into account contribution of other truck trip generation factors such as

Additionally, as noted in the research paper, centroidal distance is a crude proxy for the impedance measure for interstate truck flows. A number of other factors can impact truck flow distribution between states including the economy, land use patterns, supply and logistics strategies, trade agreements, and modal options.

3.1.3 Development of Nation-Wide Freight Analysis Zones

General Information

To overcome the geography limitations of the current CFS data provisions, this study suggests a finer, more disaggregate, nationwide zonal system for commodity data. Using zip code level business pattern data, this study assumes that the number of business establishments weighted by employment size classes is a proxy of freight-related economic activities. For simplicity, the 48 continental states and the District of Columbia are considered. The specific objectives in this project are as follows:

- Review the studies on TAZ development and adapt for the freight context;
- Create alternative FAZ systems;

- Compare the FAZ systems; and
- Evaluate the minimum number of zones needed for reasonable coverage.

Methodology

Chin and Hwang¹ developed freight demand models for 27 industry sectors covered by the 2002 CFS. The equations developed by them could be used to disaggregate CFS data to smaller geographies such as county or zip codes. Shin and Aultman-Hall² propose using five-digit Zip Code Business Pattern (ZBP) from the U.S. Census Bureau to develop Freight Analysis Zones (FAZs) to make the case for the CFS to be released with higher spatial resolution without loss of shipper confidentiality. The number of business establishments by zip codes was weighted by employee size and utilized as proxy of freight activity.

The primary issues related to the CFS zone design are problematic centroid locations for each zone and inappropriate aggregation of economies leading to higher needs to suppress data, and lower ability to be used for regional freight modeling. Freight analysis zones were developed based on three factors: 1) metropolitan areas, economic activity, and the highway network. The results were evaluated using three performance criteria – minimum distance, loss of information, and correlation between economic activities and zone aggregation.

According to the researchers of the study, the results indicate that freight analysis zone systems performed well, especially, the metropolitan area-based FAZ. Overall, this paper demonstrates the possibility of a more disaggregate CFS zone system that could be used for public data provision without violating confidentiality.

Applications of Methodology

This methodology has not yet been applied to any specific freight planning efforts.

Applicability to Other States/Regions

This methodology seems to be durable, but requires a lot of resources to repeat.

¹ Chin, S., and H. Hwang. *National Freight Demand Modeling – Bridging the Gap between Freight Flow Statistics and U.S. Economic Patterns*. Presented at 86th Annual Meeting of the Transportation Research Board, Washington, D.C., 2007.

² Shin, H., and L. Aultman-Hall. *Development of Nationwide Freight Analysis Zones*. Presented at 86th Annual Meeting of the Transportation Research Board, Washington, D.C., 2007.

3.1.4 Development of County-to-County Freight O-D Matrix from 2002 FAF2 Freight Data for Subsequent Network Assignment

General Information

Battelle Institute, Development of County-to-County Freight O-D Matrix from 2002 FAF2 Freight Data for Subsequent Network Assignment, August 17, 2006.

Commodity Flow Data Inputs

The commodity flow data inputs for this methodology are: VIUS data for payload factors, total truck VMT in the county, and total truck VMT activity within each FAF2 zone.

Methodology

The Battelle Institute³ has developed a methodology that allocates FAF² data to smaller geographies. In this method, disaggregated production and attraction freight trip in tonnage for each county is developed using the FAF² commodity tons with 114 FAF² freight O-D zones. The O-D matrix disaggregation is based on all commodities combined. The FAF² disaggregation model is applied to both equivalent truck trip and freight tonnage. The truck trip conversion factor was developed using the Vehicle Inventory and Use Survey (VIUS) 2002 database. The tonnage in the FAF2 database was first disaggregated into cells based on truck type and distance travelled for goods as shown in the table below.

Table 3.6 Tonnage Allocation Distribution Matrix by Distance

Distance (miles)	Single-Unit	CS	Tractor-Trailer	Double	Triple
<= 200	0.1	0.75	0.02	0.12	0.01
> 200	0	0.88	0	0.099	0.021

The next step in the process was to develop freight activity centers to which the FAF2 data could be allocated. Each freight activity center was represented by a network node for trip generation and attraction purposes. This was done by using information about high-volume freight activity sites within U.S. counties, aggregating spatially separated activity sites into a set of centers, adding special freight generators within each center, and finally assigning each activity to the nearest node on the FAF2 network. These techniques were primarily required in

³ Battelle Institute, *Development of County-to-County Freight O-D Matrix from 2002 FAF2 Freight Data for Subsequent Network Assignment*, August 17, 2006.

major urban areas, where county-level detail was not sufficient. For most locations freight activity centers matched with county boundaries.

The next step of this method calculates freight trucks/ton generated by all commodities combined in county $c(s)$ is a function of freight trucks/tons generated by all commodities zone (s) , total truck VMT within the County $C(s)$ and the total truck VMT activity within zone (s) and is given in equation 1.

$$(T)_{c(s)} = (T)_{(s)} \times (E_{c(s)} / E_{(s)}) \quad (1)$$

Where:

$(T)_{c(s)}$ = freight trucks/tons generated by all commodities combined in county $c(s)$

$(T)_{(s)}$ = freight trucks/tons generated by all commodities combined in FAF² zone (s)

$(E)_{c(s)}$ = Total truck VMT within county $c(s)$

$(E)_{(s)}$ = Total truck VMT within zone (s)

The Battelle method uses a TransCAD matrix disaggregation procedure. A lookup factor table containing the share of freight received by each county using the truck VMT distribution is developed. The truck VMT table for each FAF² zone and county were developed using the highway link-specific average annual daily truck travel (AADTT) flow data.

The Truck VMT approach allocates freight truck tons based on the share of total truck VMT. However, since truck VMT includes other uses (local delivery trucks, service trucks, utility trucks, construction trucks, etc.) and is by definition greater than freight truck VMT, for any zone and overall, the allocation of total truck VMT will rarely match the expected freight truck allocation. Furthermore, the truck VMT in a zone will include truck travel to, from, and within the zone, and while this may be appropriate for allocating freight truck tons, it also includes truck VMT which is merely passing through the zone and this truck VMT is totally unrelated to the allocation of freight to or from that zone. The VMT approach does not necessarily say how much freight is allocated to zones and inclusion of through trucks will lead to an over estimation of freight flow to and from a zone. The method also does not provide any disaggregation by commodity, which while not an issue for total congested assignment, may be an issue for other analysis (e.g., benefits to a particular industry/commodity).

Applications of Methodology

This methodology has been used to allocate FAF² flows to the national network to provide rough estimates of the amount of trucks (and tonnages) for specific roadways.

Applicability to Other States/Regions

This methodology seems to be durable, but requires a lot of resources to repeat. In particular, developing freight zones and acquiring zone-level establishment

data is most difficult in the dense urban locations that are already most problematic for truck travel demand models.

3.1.5 Commodity-Specific Disaggregation of the 2002 FAF2 Data to the County Level for New Jersey

General Information

This paper presents methods for disaggregating the FAF² data to the county level by developing different disaggregation factors for different commodity types. These new methods are also compared to the other disaggregation methods that were previously presented. The objective is to enable state and local governmental agencies to utilize FAF² commodity O-D data for a quick desktop analysis and to devise further strategies in collecting and acquiring local commodity data. The focus area of this study is the State of New Jersey.

Commodity Flow Data Inputs

The data inputs considered for this research included: employment, population, truck vehicle miles traveled, and the number of trucks entering and exiting a county.

Methodology

The study developed and applied different methods to disaggregate FAF² commodity data down to the New Jersey county level. The results of the disaggregation were then compared to Global Insight's TRANSEARCH Database and other disaggregation methods previously developed and presented as part of this study. In total, there were 12 disaggregation methods that were analyzed as shown below in Table 3.7. As an example, the disaggregation method labeled "D1" disaggregated FAF² data based on employment for domestic origins and population for domestic destinations. The ability of these regression variables to match the Transearch data was compared for each of the 13 disaggregation methods.

Table 3.7 Disaggregation Method Descriptions

Disaggregation Method	Domestic Origin (Production) Factors	Domestic Destination (Attraction) Factors
D1	Total Employment	Population
D2	Total Employment	Population of 25- to 54-year-olds
D3	Total Employment	Income adjusted population of 25- to 54-year-olds
D4	Truck Vehicle Miles Traveled (TVMT).	Truck Vehicle Miles Traveled (TVMT)
D5	Employment and TVMT combined	Income adjusted population of 25- to 54-year-olds and TVMT
D6	Number of trucks exiting county	Number of trucks entering county

Disaggregation Method	Domestic Origin (Production) Factors	Domestic Destination (Attraction) Factors
D7	TVMT of trucks entering county	TVMT of trucks entering county
D8	NAICS 6-digit commodity-specific industry employment	NAICS 6-digit commodity-specific industry employment
D9	NAICS 6-digit commodity-specific industry employment	Total Employment
D10	NAICS 6-digit commodity-specific industry employment	Population
D11	NAICS 3-digit commodity-specific industry employment	NAICS 3-digit commodity-specific industry employment
D12	NAICS 3-digit commodity-specific industry employment	Total Employment
D13	NAICS 3-digit commodity-specific industry employment	Population

Applications of Methodology

This methodology has not yet been applied to any specific freight planning efforts.

Applicability to Other States/Regions

The findings of this study indicate that no one disaggregation method produces the best results for trip productions and attractions for all commodities. Disaggregating each commodity using commodity-specific industry employment data yielded the best results in matching the TRANSEARCH database for flow origins. However, simple noncommodity-specific factors, such as truck vehicle miles traveled, total employment, or adjusted population data generally yielded better results in disaggregating flow attractions.

It should also be noted that the analysis conducted in this study was done for New Jersey. Relationships established in this study may not be the same for other states.

Issues and Limitations

This study compares methodologies to estimate Transearch data from FAF2 data. While this is a good measure to try to recreate the Transearch database, the usefulness of this analysis is somewhat limited because both FAF2 data and Transearch rely on CFS data for some of their flows. Therefore, correlations between these methods and the Transearch database may not reflect upon accuracy in predicting actual freight flows.

Also, the explanatory variables may change when allocating freight flows from the county-level to TAZ-level. Therefore, caution must be taken in applying these results to other types of analyses.

3.1.6 Eastern Washington Intermodal Transportation Study

General Information

Title: Eastern Washington Intermodal Transportation Study (EWITS). Various Reports from 1994 to 1999, especially in EWITS Reports Numbers 3, 4, 8, 9, and 21.

Source: <http://ewits.wsu.edu/>.

Date: 1993-1999.

Contact: The Principal Investigator, Project Director, and common author for the overall project was Dr. Ken Casavant, Professor of Economics, Washington State University. E-mail address is casavantk@wsu.edu

Agency: EWITS was a six-year study funded jointly by the Federal government and the Washington State Department of Transportation as a part of the Intermodal Surface Transportation Efficiency Act of 1991, under a contract with the Washington State Department of Transportation. The following were the key goals and objectives for the Eastern Washington Intermodal Transportation Study:

- Facilitate existing regional and statewide transportation planning efforts;
- Forecast future freight and passenger transportation service needs for eastern Washington;
- Identify gaps in eastern Washington's current transportation infrastructure; and
- Pinpoint transportation system improvement options critical to economic competitiveness and mobility within eastern Washington.

Commodity Flow Data Inputs

The dominant research effort was a statewide freight truck origin and destination study in April of 1993, originally proposed by the EWITS research team for eastern Washington but broadened to the statewide model with additional funding from the Washington State Department of Transportation.

The Washington study was the first in the United States to collect statewide freight truck origin and destination data through direct personal interviews of truck drivers. Questionnaire design was an important element of the successful methodology for roadside truck driver interviews. The Washington statewide truck driver survey collected information on time-of-day movements, vehicle configuration, trucking company location, origin and intended destination and

land use at that destination/origin cargo/product type, vehicle and cargo weight, use of intermodal facilities and the specific road and route traveled.

Approximately one-half of the questions (for example number of axles, trailer style, time of day, hazardous material placard) were filled out by direct observation of trained interview personnel. Questions to be asked directly to the truck driver focused on cargo, weight, use of intermodal facilities and route of travel. To the extent possible, check boxes were utilized to enable rapid completion of each interview by minimizing necessary writing. A map of major Washington State highways was attached to each questionnaire. Utilizing this map, the interviewers are able to quickly highlight Washington highways utilized by drivers traveling between their stated origin and destination.

Methodology

The statewide study involved over 300 persons conducting personal interviews at 28 separate locations, determined by the traffic flow in the area. A total of 28,000 truck drivers, in four distinct periods throughout the year, were interviewed, providing Washington with an extensive database on statewide freight and goods movements. The interviews, which were conducted on major highways and ports of entry for up to 24 hours, lasted up to three minutes for each one, with an acceptance rate by the truckers of over 95 percent.

The detail of shipments was very specific, based on using the maps in interviewing the trucker. Specific routes and roads, including those internal to Seattle and port areas were identified. Each load was characterized by commodity being moved, the size of shipment, the empty weight of the vehicle, the time of year, the time of day, the origin, the destination, whether hazardous material or not, etc. Since the interviews were conducted at four times during the year, seasonality could be evaluated, by commodity/weight/origin/destination, etc.

The numbers of interviews taken at each location were statistically determined for the necessary minimum but, depending on traffic flow, often more than the minimum was taken. The level of detail on routing allowed numerous subsequent studies on, e.g., corridor flows, city bypasses, port of entry density, regional traffic composition and weights and commodity profile, to be undertaken and produced.

The overall EWITS project also did surveys of the transportation needs and characteristics of various commodities in the state of Washington. Mail surveys were conducted of the major shippers of industries moving fruit, grains, vegetables hay and forest products. These surveys had response rates of 60 percent to 90 percent of the tonnage moved in the industries; the high response rate was due to the Principal Investigator's familiarity to the respondents as well as the respondents' confidence in the University. These surveys dealt with transportation needs as articulated by the shipper, storage capacity, volume moved, timing of movement during the year, road and route

used, mode of shipment, rates or tariffs for shipments, configuration of shipments, etc. Reports were written and are available on the EWITS website above.

Applications of Methodology

These data were used for 10 years, until they were refreshed under the new Strategic Freight Transportation Analysis (SFTA) in 2004. The survey techniques and the resultant databases are directly applicable to development of subnational commodity flow data. They served as input to corridor planning studies, such as U.S. 395 by Spokane and I 5 through and around Seattle, for trade flow studies, to analyze the impact of NAFTA on roadway utilization and investment needs, to specify traffic movement when analyzing the impact of proposed Columbia-Snake River dam draw downs on energy, pollutants, and shipper costs in the region. EWITS origin and destination flows, and data of this source and structure, can and have served as the basic framework for the state rail plan and general freight transportation plan.

Applicability to Other States/Regions

The earlier discussion indicated the types of applications that this type of data can be used for. Since this original origin and destination study several other state or partial state studies have been produced, notably the SFTA origin and destination study that refreshed the data from the original EWITS survey. Conducting this study required about \$300,000 in funding, with \$200,000 being for the conducting of the field survey itself, and the other funds being for the subsequent analysis and numerous studies that drew from the data over the next five years.

Prior to undertaking this methodology of data development it is necessary to have a good understanding of the traffic flow in the region/state used to determine appropriate sites, an understanding of the probable use of the data, a well trained interview staff, good collaboration with the State Patrol, and the solid backing of the State Department of Transportation or sponsoring agency.

3.1.7 Use of Truck GPS Data For Freight Forecasting

General Information

October 2009, Bassok, McCormack and Outwater

Puget Sound Regional Council, Washington State DOT, and the University of Washington

This document describes a method for using truck GPS data to forecast freight flows. It describes a specific activity in the Puget Sound, and briefly discusses some other uses of GPS data.

Commodity Flow Data Inputs

The GPS dataset being used contains no commodity data. However, once origins and destinations have been identified, it may be possible to deduce the commodity from the origins and destinations visited. For example, land use data at trip ends can be used to deduce information about commodities carried and industries served. This is not covered in the reviewed paper, however.

Methodology

The work described determines truck origins and destinations at the TAZ level. First, the region served is partitioned into TAZs, and each truck stop associated with one of these zones. A dwell-time threshold is used to differentiate intentional stops (origins and destinations), from unintentional stops (due to congestion or traffic management device).

Once the TAZ origins and destinations are identified, truck trip generation and consumption rates were estimated for each TAZ. Grocery store locations were identified and therefore truck trip generation rates could be estimated for this facility type. Similar analysis can be performed for any employment sector. One challenge is understanding what proportion of all trucks are represented by the trucks with GPS units.

Applications of Methodology

Truck trip generation rates for grocery stores. Other potential uses suggested include average trip and tour lengths (disaggregated by employment sectors, land use types, and times of day), corridor and inter-zonal travel speeds, and route choice.

Applicability to Other States/Regions

The use of GPS data for freight planning requires significant data management and analysis skills. Data volumes are very large. Analysis required the use of Geographic Information Systems, and therefore analysis with spatial analysis skills. Complementary GIS data, including land-use and the transportation network in the area of analysis, are required.

The collection of GPS data from approximately 2,500 trucks traveling in the Puget Sound region, and the analysis of this data, cost approximately \$200,000, of course, there are many other outcomes from this project.

3.1.8 Issues and Approaches to the Utilization of Highly Aggregated Databases in Freight Planning and Modeling

General Information

This paper is one of the four original pilot studies into disaggregation of the Freight Analysis Framework Version 2 program funded by the Federal

Highway Administration. The research team at the University of Alabama in Huntsville has experienced many of the issues and problems associated with the use of such a highly aggregated database. The university team developed approaches, processes and methodologies to utilize available data, and develop additional data sources to provide the freight information needed to model transportation systems at the local Metropolitan Planning Organization (MPO) level and at the statewide level. This paper describes the issues encountered throughout this process and the approaches that have proved beneficial and have improved the modeling performed at both levels.

Commodity Flow Data Inputs and Methodology

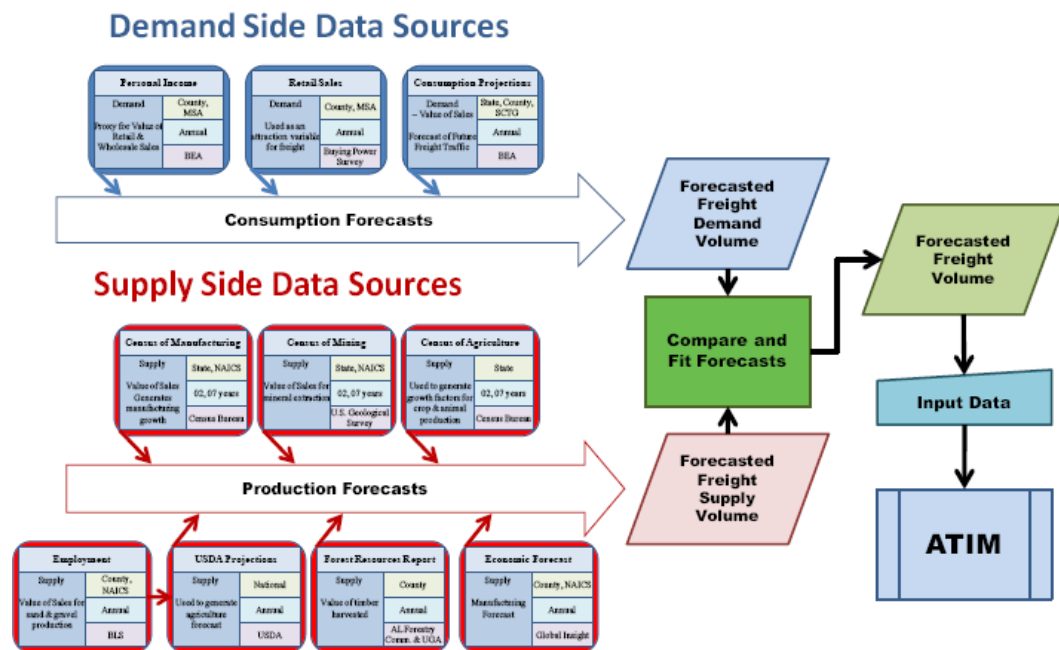
The base year for the sub-state economic database is 2002, the year corresponding to the FAF2 O-D matrices. The year 2002 is also when the US Census Bureau surveyed industries for its series of state economic censuses including the *Census of Manufacturing*, the *Census of Agriculture*, and the *Census of Mining* (Figure 2). The base year will change after the 2007 O-D matrices are released.

Until now, the disaggregation of freight from national levels for use in local areas has been primarily based on the relative employment in the local area to the total employment in the zone. The research team cited that this disaggregation technique has limitations in that productivity improvements which allow manufacturers to produce more products that require more freight shipments using fewer employees are often undetected. The research team developed a new methodology using a national freight origin/destination database and various socio-economic factors to perform disaggregation to the local level. The factors considered in this research include Value of Shipments (or Sales), Personal Income, Population, and Employment.

Thus, the Alabama sub-state economic database includes the value of sales from goods-producing industries. Using Value of Sales instead of employment factors allows for consideration of future productivity improvements and should provide a better forecast of future freight traffic.

Personal income was chosen to proxy the value of retail sales to households and wholesale sales to businesses in a sub-state region (Figure 3.2). The growth of personal income is highly correlated with the growth of household consumption expenditures and consequently should give a more accurate forecast of freight traffic than either population or employment growth.

Figure 3.2 Demand and Supply Data Sources for Freight Flow Estimates



The value of sales for manufacturing are published in the *Census of Manufacturing* for each state, metropolitan area, and county that contains manufacturing enterprises. If there are only a few manufacturers or one or two dominant firms, the value of sales data will be suppressed to protect the privacy of the firms. In Alabama the value of sales data was suppressed in 19 of 67 counties – nearly all of them sparsely populated rural counties with a single dominant company. A Value of Sales estimate must be prepared in these cases. The Census Bureau provides a range of employment for the plant(s) in these counties. Multiplying the mid-point in the employment range by the average value of sales per employee for the industry within the state provides a reasonable proxy for the actual Value of Sales in these counties. The Value of Sales in each county, including estimates that had to be made, can then be summed and compared to the actual total value of sales for the state. If the published total is larger or smaller than the total containing the estimates, the estimates, can be increased or reduced until they equal the published state total.

The *Census of Agriculture* (Figure 2) provides detailed Value of Sales data for each type of crop or animal sold from a particular county. The US Geological Survey (USGS) periodically publishes a state geological survey which includes the Value of Sales for the mineral industry. 2003 is the most recent USGS survey for Alabama. Production and sales data are provided by geological area rather than by county in this publication so it must be supplemented by information from the *Census of Mining* (Figure 2) to allocate the value of mineral extraction to each county in the state.

Sand and gravel operations are located in almost every county in Alabama and can be found using *County Business Patterns* (Figure 2) where the publication lists total employment by county. Allocating sand and gravel sales by employment give an estimate of the contribution of the sand and gravel industry to total sales in each county.

The quantity of logs harvested in each county is released in an annual report from the Alabama Forestry Commission. The data are provided by type of log and by volume in board feet. The value of these logs was determined by translating board feet into tons and using 2002 pricing data for the South published by the Daniel B. Warnell School of Forestry Resources, University of Georgia. Personal income by county is released annually by the Bureau of Economic Analysis (BEA), US Department of Commerce and is a part of the Regional Economic Accounts database.

Applications of Methodology

The methodology was used in the development of the truck component of the Mobile MPO travel demand model. The methodology identified nine possible freight movements in the area. The freight movements identified were:

1. External-External Trips (through trips)
2. Port to the US (non Alabama)
3. Alabama to the US (non Mobile)
4. Port to Alabama (non Mobile)
5. Mobile, AL and the rest of Alabama (non Birmingham, Alabama area)
6. Mobile, AL to Birmingham, Alabama area
7. Port to Mobile, AL
8. Mobile, AL to the US (non Alabama)
9. Internal to Mobile

The disaggregation of the FAF2 data to the counties is performed using county level population, personal income and value of shipment. The truck trips between the Port of Mobile and locations in Alabama, non Mobile, were determined using a statewide level network and disaggregation of the two FAF2 zones for Alabama to the county level and the port data from FAF2. These truck trips were determined using a disaggregation of the FAF2 data for zones 1 and 2 using the county population, personal income and value of shipment allocation factors. The disaggregation was performed in Microsoft Excel and only the trips to and from the Port of Mobile was distributed to the 67 counties in Alabama. It is important to note that trips were distributed to

Mobile County, these trips will be used as the basis for the Port to Mobile trip purpose.

Applicability to Other States/Regions

The application of this methodology to the Mobile MPO indicates that the methodology is portable to other locations. However, there are considerable resources needed to manipulate the databases to prepare them for incorporation into the truck model.

Issues and Limitations

This research did not seek to identify the set of economic variables that best predicts present freight movements into and out of these FAZs and consequently increase the accuracy of predicting future freight movements. This was left as a future research activity.

This disaggregation technique starts with FAF2 data. Therefore, the limitations of the FAF2 database will remain with the disaggregated database as well. These issues are discussed in greater detail in

3.2 DESCRIPTION OF PLANNING-ORIENTED DISAGGREGATION TECHNIQUES

3.2.1 Texas NAFTA Study Update – Final Report

General Information

Texas NAFTA Study Update – Final Report, February 2007, prepared for Texas Department of Transportation by Cambridge Systematics, Inc., with Alliance Transportation Group; Delcan, Inc.; C&M; LopezGarcia Group; Robert Harrison; and Bomba Associates

ftp://ftp.dot.state.tx.us/pub/txdot-info/library/reports/gov/tpp/nafta_study.pdf

Report is an update of the *Effect of the North American Free Trade Agreement on the Texas Highway System*, December 1998, prepared for Texas Department of Transportation by Louis Berger and Associates, Dye Management Group, and Street Smarts

Overview

- Objective – Estimate NAFTA freight flows on the Texas highway system.
- Methodology – Statewide Analysis Model (SAM) assignment of TRANSEARCH 2003 and 2030 domestic and international truck flows to the current highway network.
- Note – “NAFTA” connotes trade between the U.S. and Mexico that utilizes the Texas highway or rail system for at least a portion of its movement.

Commodity Flow Data Inputs

TRANSEARCH domestic flows, O&D in the U.S. and utilizing Texas highways at least partially:

- Intrastate flows between Texas counties, provided at the county level.
- Interstate flows between Texas and other U.S. states, provided at the county level for Texas and the state level for other states.
- Pass-through flows between U.S. states passing through Texas.

TRANSEARCH international flows, O&D in Mexico:

- Flows between Texas and Mexico, provided at the county level for Texas and the state level for Mexico.
- Pass-through flows between U.S. and Mexican states passing through Texas.
- Data reconciliation was necessary to obtain an acceptable base year (2003) highway assignment.
- TRANSEARCH data was compared/reconciled with the Bureau of Transportation Statistics’ Border Crossing and Transborder data, SAM truck trip tables (1998), and FHWA’s Freight Analysis Framework (FAF).
- Global Insight corrected the variation between TRANSEARCH data and aforementioned data sets.

TRANSEARCH data was summarized as commodity flows according to SAM commodity classes: Agriculture, Building Materials, Chemicals/Petroleum, Food, Hazardous, Machinery, Miscellaneous Mixed, Raw Material, Secondary, Textiles, and Wood

Roadside truck surveys were conducted at 12 border locations to calibrate/validate the assignment of flows: Origin and destination, Cargo type, Intermediate transfer points, Truck physical attributes.

Final Commodity Flow Database

TRANSEARCH database was organized and analyzed using Freight Tools_{CS}, which was used to process and describe the commodity flows

Commodity flow data were loaded into SAM and the Texas-North American Freight Flow Model (Tx-NAFF), which is used with SAM to extend analysis capabilities into Mexico; both models were updated to the 2003 base year

Validation process included a number of data sources selected to confirm the reasonableness of the flows and led to adjustments to the trip table and routing; subsequent model runs were required for a satisfactory assignment on the statewide network

Resulting assignment showed:

- Order of magnitude flows on major NAFTA corridors;
- Flows on rural corridors, especially north-south, may be underestimated; and
- Correctly estimated flows on most primary NAFTA corridors, although some manual assignment changes were made.

To estimate the effect of future trade volumes on the Texas highway system (i.e., corridors), three scenarios were developed using SAM and 2030 TRANSEARCH commodity flow data as the basis for truck travel demand (TRANSEARCH forecasted data are based on proprietary forecast assumptions):

- No-Build – no capacity expansion to 2030, i.e., SAM base (2003) network;
- Highway Investment – how truck flows would react to new capacity, i.e., SAM base (2003) network with new projects; and
- Global Trade and Growth – Highway Investment network plus higher than forecast global trade (same SAM base network as the Highway Investment scenario).

Level of accuracy is as previously described for 2003, i.e., order of magnitude flows.

Types of Applications that Have Been Implemented

Study's focus was corridor planning informed by NAFTA highway demand:

- Freight flows depict the effects of NAFTA trade growth on the Texas highway network given several possible outcomes by 2030; and
- The range of flows on the highway system include 2003 base year, 2030 with no expansion, 2030 with new highway construction, and 2030 with accelerated trade growth with Mexico.

Applicability to Other States/Regions

The methodology appears to be directly transferrable to other regions. There are similar data sources available to implement similar steps to this study. However, the cost to assemble and analyze the data would require significant resources.

3.2.2 LACMTA Cube Cargo Model

General Information

Initial Development of an Innovative Commodity Flow and Truck Model for the Greater Los Angeles Basin, Michael Clarke, Citilabs, and Luke Cheng, LACMTA, presented at the International Chinese Transportation Professionals Association Annual Conference, January 2005, Washington, D.C.

http://www.nacota.org/htdocs/Annual_Meetings/2005/Presentations/1.1.4-cargotechnicalpaper.pdf

Initial Steps to the Building of an Innovative Commodity Flow/Truck Forecasting Model for the Greater Los Angeles Region, Michael Clarke, Citilabs, and Luke Cheng, LACMTA, presented at the 2005 Meeting of the Transportation Research Board, January 2005, Washington, D.C.

http://www.nacota.org/htdocs/Annual_Meetings/2005/Presentations/1-1-4Presentation%20on%20Cube%20Cargo.pdf

Overview

- Objective – Develop and implement an innovative freight forecasting model to estimate commodity and truck flows for the greater Los Angeles Basin.
- Methodology – Use of the Cube Cargo freight forecasting software/model.
- Client – Los Angeles County Metropolitan Transportation Authority (LACMTA).
- Note – Phase 1 described here developed updated Cube Cargo software using available data; Phase 2 involved model enhancement, calibration, and validation, for which specific reports (URLs) were not located to date.

Commodity Flow Data Inputs

Inputs to Cube Cargo were:

- Zone-level socioeconomic data such as population and households and employment by type;
- Zone-to-zone level-of-service information such as door-to-door travel times and costs for each mode (road, rail, air); and
- Matrices of existing commodity flows to use as a base for projection.

Note that “zones” were zip codes within the Southern California Association of Governments (SCAG) five-county region (now six counties), counties in the rest of California, states in the rest of the U.S., and four external zones (two each for Canada and Mexico)

Intermodal Transportation Management System (ITMS) commodity flow data developed by Caltrans (1996) were used (“matrices of existing commodity

flows”); originally coded at a county zonal level, the data were allocated to zip codes. ITMS data is a close derivative of Transearch data developed for the entire state of California.

The data were reviewed, analyzed, and aggregated to 16 commodity classes.

Final Commodity Flow Database

Cube Cargo was implemented for the SCAG five-county region:

1. Large-area networks were created for road and rail, a zone system was created, socioeconomic data were summarized, and base year commodity flow matrices were developed;
2. These data were input to Cube Cargo, which was applied to estimate annual commodity flow matrices by mode and commodity class, and annual and daily truck matrices by large and small truck class; and
3. Truck matrices were assigned to the network.

Authors stated that “more work is required to enhance, calibrate, and validate this preliminary model before it can be used for planning purposes”. This was done as part of Phase 2. Specifically, the work in Phase 2 was to prepare the model for application by showing how you can improve a subnational commodity flow database using local data. This was done by implementing the following tasks:

- Base the number and characteristics of the commodity classes on the quality of the observed data;
- Review and improve the ITMS zip code-level data;
- Estimate generation models from new employment classifications using the improved ITMS data with final commodity classes;
- Calibrate distribution and mode choice models against ITMS revised data; and
- Integrate the assignment process with car traffic and calibrate the results against counts.

Types of Applications that Have Been Implemented

Pending improvements, Cube Cargo will be used to estimate regional commodity and truck flows.

Applicability to Other States/Regions

Cube Cargo requires the following data:

- Model estimation data – Data used to estimate/update the default parameters within Cube Cargo; model estimation was not an objective of this study, hence, no data were used in this regard.

- Model validation data – Data used as the basis for comparing the model estimations for the base year (e.g., 2001) to see how well the model can simulate what is currently observed; model validation was not an objective of this phase, hence, no data were used in this regard.
- Model application data – Inputs to the model for the base year and future year scenarios; this includes a zone system with socioeconomic data, commodity flow matrices, and logistics nodes, whereby the model routes a portion of the commodity flows through warehousing/intermodal points in the system, as well as through the ports (Long Beach, Los Angeles, and Hueneme), and LAX and Ontario airports.

The methodology appears to be directly transferrable to other regions. Most states and regions have similar data sources available to implement similar steps to this study. However, the cost to assemble and analyze the data would require significant resources.

3.2.3 Virginia Statewide Multimodal Freight Study, Phase I

General Information

Virginia Statewide Multimodal Freight Study, Phase I, February 2008 (?), prepared for Virginia Department of Transportation by Cambridge Systematics, Inc. Populated template is based on the Executive Summary and Appendix B.

Phase I focused on outreach, data collection, baseline forecasting, system inventory/analysis, and freight improvement opportunities; Phase II develops analysis tools, analyzes corridor and regional freight needs and alternatives, and evaluates infrastructure and policy alternatives.

Overview of Phase I

- Objective – Collect and analyze data, inventory conditions, and identify freight demand and critical freight needs.
- Methodology – Compilation of available information into a “geospatial” freight database, and data analysis using available methods and tools; inventory and comparison of available forecasts with development of “synthesis” baseline future forecasts.

Commodity Flow Data Inputs

TRANSEARCH data:

- 2004, with forecasts to 2035;
- By mode;

- By commodity type (e.g., secondary traffic; nonmetallic minerals; coal; clay, concrete, glass, and stone); and
- Inbound and outbound at the county level.

Roadside truck O-D surveys were conducted at two weigh station locations on Interstate 81 to gauge the magnitude of pass-through truck trips versus Virginia-oriented trips (trucks moving into, out of, and within the State); more surveys are planned on other major truck routes. These flows were compared to the Transearch database and adjustments were made to freight flows, as needed.

Final Commodity Flow Database

Data analysis was performed on the TRANSEARCH data using Freight Tools_{cs}, among other tools.

Data depict:

- Tonnage, all modes and commodities, excluding through traffic, 2004;
- Truck tons, excluding through traffic, by county, 2004;
- Commodities by tonnage by county, 2004;
- Originating counties for forecast tonnage to 2035;
- Destination counties for forecast tonnage to 2035;
- Value by mode and commodity, excluding through traffic, 2004;
- Truck value, excluding through traffic, by county, 2004;
- Commodities by value by county, 2004; and
- Flow maps, by corridor:
 - Truck tonnage, excluding through traffic, 2004;
 - Through truck tonnage, 2004;
 - Truck tonnage, excluding through traffic, 2035; and
 - Through truck tonnage, 2035.

Types of Applications that Have Been Implemented

Study's focus is multifaceted, including corridor planning, Interstate 81 truck-rail diversion analysis, economic impact analysis, and policy review.

Critical issue is how the State will deal with projected doubling of truck tonnage, which will occur alongside growing urban congestion.

Refer to imminent Phase II results for specific applications

Applicability to Other States/Regions

Requirements for developing the database:

1. Existing freight flows;
2. Data analysis (tools); and
3. Baseline future forecasts (possibly synthesized).

3.2.4 Central Coast California Commercial Flows Study

General Information

The Central Coast California Commercial Flows Study is an ongoing effort to conduct a regional commodity flow profile. This project is led by The Association of Monterey Bay Governments, (AMBAG), as well as its partner agencies the San Benito Council of Governments (SBtCOG), the San Luis Obispo Council of Governments (SLOCOG), the Santa Barbara County Association of Governments (SBCAG), the Transportation Agency for Monterey (TAMC) and Caltrans District 5. Cambridge Systematics and Economic and Planning Systems (EPS) are the consultants on the project.

The overall purpose of the study is to identify ways to improve truck and rail operations and reduce the negative impacts of these operations on the transportation system throughout the Central Coast. This project is intended to directly benefit California by detailing an important and growing aspect of the State's overall goods movement system by providing important information to help update the State's Goods Movement Action Plan (GMAP). The project represents an essential transportation planning effort which establishes a basis for mega-regional goods movement strategies and priorities, including the following:

- Improved understanding of the producer export corridor aspect of the State's overall goods movement system, complementing and expanding on the work of the GMAP and California Marine and Intermodal Transportation Council (CALMITSAC).
- Centralized, multimodal distribution centers which help to rationalize truck VMT, increased rail reliability and service levels, mitigation of diesel and particulate matter emissions, improved capacity of regional port complexes, and foundation for economic development along alternative export corridors.
- Improved multiregional, government-to-government relations through the joint efforts, participation and support by the Central Coast MPOs (AMBAG, SLOCOG, SBCAG).

- Establishment of a broad stakeholder participation platform for private sector companies responsible for commercial flows, thereby greatly increasing public participation in this area.
- Enhanced understanding of the character of present and future commercial traffic flows in and through the Central Coast.⁴

Project stakeholders also include private-sector growers, manufacturers, shippers, logistics service providers, truck operators, railroads, and impacted community representatives. Responses from interviews with these stakeholders are used to help the project team better understand the primary transportation issues that impact a wide variety of stakeholders in the goods-movement process. Additionally, these interviews help the project team better understand industry logistic patterns, which supplements commodity flow data.

Commodity Flow Data Inputs

National data – describe geographic detail, commodity detail

National data sets being used include FAF2 database, the BTS CFS, and Census Data.

Local Commodity Flow Data and Economy/Industry Data

The following key data sources will be used for the data development effort:

- Federal Highway Administration’s (FHWA’s) Freight Analysis Framework (FAF2) database (disaggregated to the county level);
- IMPLAN Group, Inc’s IMPLAN economic output data and input-output ratios;
- Local commodity production and consumption data;
- Modal data; and
- Truck O-D surveys.

A regional commodity flow profile will be created from this data. This profile will be a detailed, multimodal assessment of all freight moving into, out of, within (local traffic) and through (long-haul traffic) the region. Seasonal traffic, mostly caused by agricultural industry patterns, also will be considered specifically in this profile. The resulting commodity flow O-D data will both paint a clear picture of the type, quantity, and value of the region’s commodity movements, as well as feed easily into the commodity flow forecast to be undertaken by the Project Team. The process on how the local commodity-flow database was constructed is shown in Figure 3.2.

⁴ Introductory description taken from a San Luis Obispo staff report describing project background.

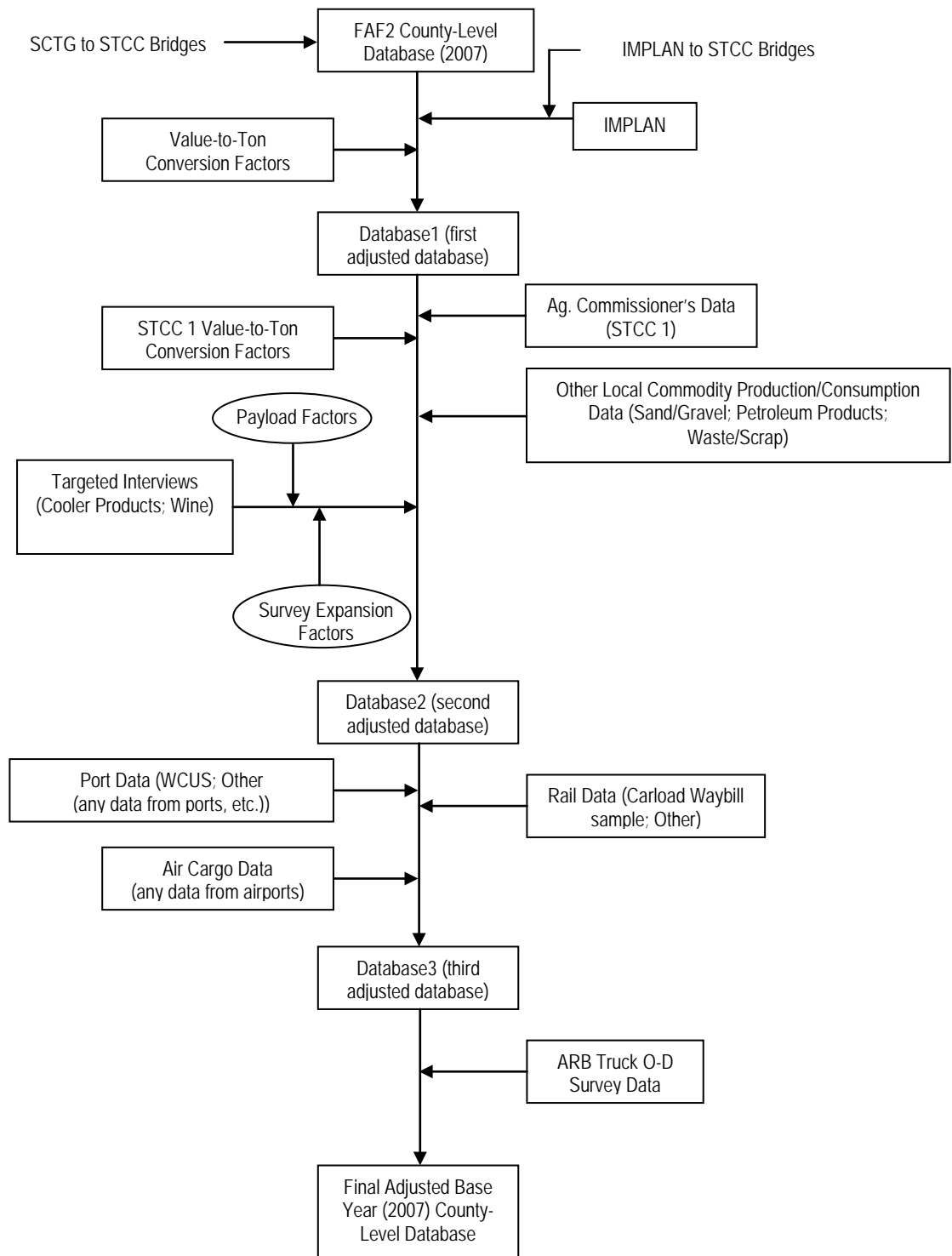
Further data was used for this project, but was used to understand economic/industry trends for the region. The above datasets include those being used for the CF piece of the study.

A commodity flow analysis will be completed for the five-county Central Coast region, including the counties of Santa Cruz, San Benito, Monterey, San Luis Obispo, and Santa Barbara. The STCC classification will be used to present the final analysis data.

Further specifics regarding the final commodity flow output:

- **Commodity classification:** The database will be based on a standard commodity classification scheme, allowing for comparisons of the database with other standard data sources as part of data validation and improvement efforts. It is proposed that for the current effort, the Standard Transportation Commodity Code (STCC) classification system be used, at the two-digit level of detail.
- **Modes:** The database will include all the major modes used for goods movement in the region, including truck, rail carload, rail intermodal, sea, air, and pipeline, as applicable
- **Shipment volume:** The database will provide commodity flow data by weight (tonnage) and value (dollars) of shipments.
- **Market areas:** As indicated above, the database will include trade flow data for both domestic as well as international trade flows. Domestic trade flows will include intra-county and inter-county flows within the study area, as well as flows between counties and external regions (outside the study area) within the U.S. International trade flows will include NAFTA trade flows, as well as overseas trade by air and sea, as applicable. The data development effort will also account for through freight flows (as applicable) in the region.
- **Geographic detail:** The database will provide commodity flow data at the county-level of detail. The structure of the database, and the commodity classification used will be such as to support any future efforts undertaken to disaggregate the data to higher levels of geography, such as zip codes or traffic analysis zones (TAZs).

Figure 3.2 Steps to Creating the Final Commodity Flow Database for the Study



- **Time periods:** The initial focus of the data development effort will be to generate an accurate base year database for the region, which can then be used to develop forecasts, and for developing truck trip tables. The base year considered for this effort is 2007.

Types of Applications that Have Been Implemented

The intent is to use this data to update the California's Goods Movement Action Plan (GMAP). Results are also intended to inform the California Marine and Intermodal Transportation Council (CALMITSAC).

Applicability to Other States/Regions

The data methods and practices in the project demonstrate a medium level of complexity, which is based on this project synthesizing a number of available public databases to create a freight profile for a 5-county region in California. The data creation and analysis portion of this study is estimated to cost approximately \$50,000.

FAF2 data must be disaggregated to the county level in order to present a commodity flow profile at the county level. Other options include purchasing Global Insight data commodity flow data, which is available at the county level.

3.2.5 Southern California Association of Governments Global Positioning Survey

General Information

SCAG Regional Goods Movement Plan, forthcoming in fall of 2011), Southern California Association of Governments/Cambridge Systematics, <http://www.scag.ca.gov/goodsmove/>. Includes updates to the region's heavy duty truck (HDT) model, using various data sources including TRANSEARCH, port gate surveys, and GPS units deployed on regional truck fleets.

Southern California Association of Governments (SCAG)

Length: n/a

Commodity Flow Data Inputs

Deployment of Global Positioning Systems (GPS) units on truck fleets operating in the SCAG region to help estimate truck trip generation and distribution.

Final Commodity Flow Database

Data collected via deployment of GPS units on truck fleets will be used to develop trip generation rates for freight by land use. The GPS units provide latitude/longitude data for each stop made by a truck. The latitude/longitude at each trip end will be matched with an establishment/land use database to get land uses at trip ends. The resulting data can then be expanded using estimates

of the truck population to develop trip generation rates per employee for the region or subregions.

GPS data will also be used to develop a land use to land use interchange matrix and trip lengths for estimating new gravity model parameters. The interchange matrix will describe the percentage of trips produced by each land use that will move to each other land use. New model parameters are primarily trip length distribution and average trip length.

Geographic Detail

GPS latitude/longitude data is accurate down to 0.111 meter. The SCAG region encompasses Los Angeles, Orange, Ventura, Riverside, San Bernardino, and Imperial Counties.

Commodity Detail

N/A – Fleets are categorized by industry rather than commodity. Two-digit Standard Industrial Classification (SIC) level of detail.

Level of Accuracy

The new data is expected to be statistically valid, meaning data from the sample can be used to infer characteristics of the universe (i.e., all trucks operating in the region) with a reasonable degree of certainty. The conclusions thus drawn will therefore be statistically defensible. As envisioned in the Model Development Plan/Data Collection Plan, the new trip rates by land use and truck type will have a 90 percent chance of being within 15 percent of the actual value.

Types of Applications that Have Been Implemented

None as of yet, but the model will achieve the improvements mentioned above. In addition, the data will also support enhancement of the mode in the following four areas: 1) truck activity at manufacturing facilities, 2) truck activity at warehouse and distribution centers, 3) truck activity regarding local pickup and delivery/urban goods movement, and 4) service truck activity.

The model will also be used to support development of the next SCAG Regional Transportation Plan (RTP), as required under SAFETEA-LU and California law.

Applicability to Other States/Regions

This effort is part of a very sophisticated and costly update to the SCAG HDT model. It is estimated that the cost of GPS data alone will be about \$180,000; total data collection costs are estimated at nearly \$500,000.

This project assumes the existence of a fully developed and calibrated HDT model for the region under analysis.

3.2.6 California Commodity O-D Database Disaggregation

General Information

Agencies involved in the research or study – U.S. Federal Highway Administration, California Department of Transportation, and California Air Resources Board.

Commodity Flow Data Inputs

National data – FHWA's FAF2 National Data

Local data:

- Census population data at the county and zip code area level for California.
- Census County Business Patterns (CBP) and Zip Business Patterns (ZBP) data.
- Farm acreage, and livestock information from the National Agricultural Statistics Service (NASS) database.
- Electricity generation data from the National Energy Technology Laboratory (NETL) Coal Power Plant Database (CPPDB).

Final Commodity Flow Database

What analysis was done: Disaggregate National FAF2 data down to county level and air basin level using regression equations tied to employment by industry, population, power plant generation, and farm acreage and livestock data.

Geographic Detail: Data broken down to counties and air basins in California.

Commodity: STCG2 commodities from FAF2.

Types of Applications that Have Been Implemented

None yet, in the process of being completed.

Applicability to Other States/Regions

It would take approximately 2 months of work from CS, Contract value for this case was \$77,000.

To develop this dataset for other states we would only need the local datasets listed above.

3.2.7 Oregon DOT Commodity Flow Database Development

General Information

In order to transform the coarse FAF2 zone flows into smaller geographic regions within Oregon, the data had to be disaggregated. The final database desired was

a county-level freight flow database to include all freight modes and 2-digit commodity code classification for the entire state.

Methodology

Since the FAF2 dataset contains the whole United States, flows with at least one trip ends within Oregon were disaggregated from FAF2 zones to Oregon counties. In the case of truck flows, this was done based on county employment and IMPLAN inter-industry coefficients. For rail flows, the FAF2 flows were compared to the Surface Transportation Board's Rail Carload Waybill data set which contains county-level detail of origin and destinations. The overall numbers were found to be comparable, so the Waybill data for 2002 was used as the base, and the FAF2 growth rates were applied to obtain the forecast. The other modes relied on local data to allocate FAF2 flows to specific Oregon facilities (rail stations, airports, marine ports, or pipeline terminals), including U.S. Corps of Engineers Waterborne Commerce data and the Oregon Energy Report. Zones outside of Oregon were aggregated from FAF2 zones to "Other Domestic" and "Other International" categories. Special consideration was made for air mail and fish commodities. The Oregon CFF methods and data are described further in Appendices A through D, with Appendix E providing a summary table of the Oregon CFF by commodity, mode, and direction.

Final Commodity Flow Database

What analysis was done: Disaggregate National FAF2 data down to county level and air basin level using regression equations tied to employment by industry, population, power plant generation, and farm acreage and livestock data.

Geographic Detail: Data broken down to counties and air basins in California.

Commodity: STCG2 commodities from FAF2.

Types of Applications that Have Been Implemented

This work was done to support the development of a freight plan for the Oregon DOT. It should be noted that the routed flows developed from this analysis were not completely reconciled with Oregon DOT expectations.

Applicability to Other States/Regions

This work was done through a consultant contract over two months at an estimated cost between \$50,000 to \$100,000. To develop this dataset for other states, only the local datasets listed above would be needed.

3.2.8 Strategic Freight Transportation Analysis

General Information

Title: Strategic Freight Transportation Analysis (SFTA).

Source: <http://www.sfta.wsu.edu/>.

Date: 2003-2008.

Contact: The Co-Principal Investigators for the overall project were Dr. Ken Casavant and Dr. Eric Jessup, Professor and Associate Professor of Economics, respectively, Washington State University. E-mail addresses are casavantk@wsu.edu and eric_Jessup@wsu.edu.

Agency: The two Co-Principal Investigators, with Washington State University receiving the grant from the Washington State Department of Transportation, served as project leaders on different studies within SFTA and were the common authors on most of the reports. The SFTA research and implementation project, with its collaborative partnerships and integrated dynamic freight data warehouse, was designed to provide economic information for strategic infrastructure investment choices, including transportation support for economic development, responding to NAFTA impacts, and other emergent issues. Specific goals were to contribute to:

- Improving knowledge about freight corridors;
- Assessing the operations of roadways, rail systems, ports (marine and inland) and identifying freight chokepoints;
- Conduct case studies of public/private transportation costs; and
- Evaluate the opportunity for public/private partnerships.

Description of the Data

SFTA was a broad research effort, with over 25 research reports and numerous working papers. But, the dominant research effort was a statewide freight truck origin and destination study, with results published in SFTA Research Report Number 10 on October 2004, and in SFTA Research Report Number 20 in October 2006, further analyzed and contrasted to the earlier EWITS survey.

This second Washington study collected statewide freight truck origin and destination data through direct personal interviews of truck drivers.

The Washington statewide truck driver survey collected information on date, time-of-day movements, vehicle configuration and trailer style, trucking company name and location, origin and intended destination and land use at that destination/origin cargo/product type, vehicle and cargo weight, use of intermodal facilities and the specific road and route traveled.

Methodology

The statewide study involved over 100 persons conducting personal interviews at 28 separate locations, determined by the traffic flow in the area. A total of 28,000 truck drivers, in four distinct periods throughout the year, were interviewed, providing Washington with an extensive database on statewide freight and goods movements. The methodology was the same with the

interviews being conducted on major highways and ports of entry for up to 24 hours, lasting up to three minutes for each one, yielding an acceptance rate by the truckers of over 90 percent.

The overall SFTA project also did surveys, some new and some updated from the EWITS study, of the transportation needs and characteristics of various commodities in the state of Washington. Mail surveys were conducted of the major shippers of industries moving grains, hay and livestock products. A detailed mail survey was done of the warehousing industry in the state. These surveys had response rates of 60 percent to 90 percent of the tonnage moved in the industries. These surveys dealt with transportation needs as articulated by the shipper, storage capacity, volume moved, timing of movement during the year, road and route used, mode of shipment, rates or tariffs for shipments, configuration of shipments, etc. Reports were written and are available on the SFTA website above.

Relevance for Commodity Flow Database Development

These data serve very well for sub national commodity flow database development. The detail of shipments was very specific, as a result of the maps used in interviewing the trucker. Specific routes and roads, including those internal to Seattle and port areas were identified. Each load was characterized by commodity being moved, the size of shipment, the empty weight of the vehicle, the time of year, the time of day, the origin, the destination, whether hazardous material or not, etc. Since the interviews were conducted at four times during the year, seasonality could be and was evaluated, by commodity/weight/origin/destination, etc.

The numbers of interviews taken at each location were statistically determined for the necessary minimum but, depending on traffic flow, often more than the minimum was taken. The level of detail on routing allowed numerous subsequent studies on, e.g., corridor flows, city bypasses, port of entry density, regional traffic composition and weights and commodity profile, to be undertaken and produced.

General Types of Applications That the Database Can Be Used For

These data were used for many studies since that time and these can serve as applications for the future. Selected examples include the following:

Table 3.8 Examples of Applications That the Database Can Be Used For

Agency	Data Requested	Purpose
Freight Strategy and Planning: WSDOT	Truck frequencies and volumes on selected corridors.	Supplement reports, presentations and freight policy plans.
	Specific truck frequencies near Toppenish, WA.	Address private citizen concerns regarding freight traffic and safety issues.
Planning and Data: WSDOT	All data collected from the 1994 EWITS study and 2002 SFTA.	WSDOT Planning and Data office is developing a planning data depository where multiple types of data would be stored.
City of Reardan	Frequency of freight traffic moving through Reardan, WA.	The data was utilized to compare and contrast changes in freight flows, by vehicle type and commodity, between 1994 and 2002 for the town of Reardan, WA.
East Region Project Definition Engineer: WSDOT	All Origin and Destination data from 1994 and 2002.	The data was being utilized to help develop and design a process for sighting future weigh-stations and weigh-in-motion locations based upon freight vehicle frequencies, truck type, commodity, etc.
Puget Sound Regional Council	All Origin and Destination data from 1994 and 2002.	The data were utilized to validate truck travel demand models used by the MPO's.
Spokane Consulting Firm	Truck frequencies from the O-D data for all freight traffic starting, ending, or passing through Lind, WA.	The consultants were hired by the City of Lind, WA to profile changing freight truck travel through Lind, WA.
U.S. DOT	Freight movement for selected cities, including Seattle-Portland, Canada-Portland, and Canada-Seattle.	The information aided "The West Coast Corridor System" Phase I report, funded as part of the Borders and Corridors budget of U.S. DOT. The analysis and data was used to confirm and in some cases to establish levels of freight activity within the corridor system.
Benton-Franklin Council of Governments	Information related to freight movements in the Tri-Cities Region.	Data utilized to aid RPO's and MPO's planning efforts.
Washington Wheat Commission	Database of Eastern Washington Grain Elevator Survey	Data was utilized to evaluate industry changes and shifts over the last 10 years.
Transportation Ministry, Seoul, Korea	EWITS and SFTA questionnaire design and survey implementation.	Information was utilized to help in the development of freight collection approaches for Seoul, Korea.
ODOT	Methodology and questionnaire types	Information is utilized to develop an urban and metropolitan freight data collection technique.
Cambridge Systematics	Identification of data needs for transportation modeling.	Development of survey design and sample frame for Port of Portland.

The Principal Investigators for this study have been approached over five times for details of how to conduct the study and the detail produced by the survey itself. Similar to the EWITS survey, about \$300,000 in funding was required, with

\$200,000 being for the conducting of the field survey itself, and the other funds being for the subsequent analysis and numerous studies that drew from the data over the next five years.

Prior to undertaking this methodology of data development it is necessary to have a good understanding of the traffic flow in the region/state used to determine appropriate sites, an understanding of the probable use of the data, a well trained interview staff, good collaboration with the State Patrol, and the solid backing of the State Department of Transportation or sponsoring agency. Many of the individuals serving as road interviewers and field coordinators in the EWITS study were also available for the SFTA survey.

3.3 COMMODITY-SPECIFIC DISAGGREGATION TECHNIQUES

3.3.1 Potato Flows in Washington State

General Information

July 2009

Andreoli, Goodchild, Jessup

University of Washington, Washington State University, Washington State DOT

Commodity Flow Data Inputs

This project independently generated commodity flow data for Washington State, beginning with agricultural data from the U.S. Department of Agriculture, plus industry data obtained from the Washington State Potato Commission (for example product loss and conversion rates and mode split, locations of potato processing facilities) and, market data from AC Nielsen (such as potato consumption rates and market prices).

Final Commodity Flow Database

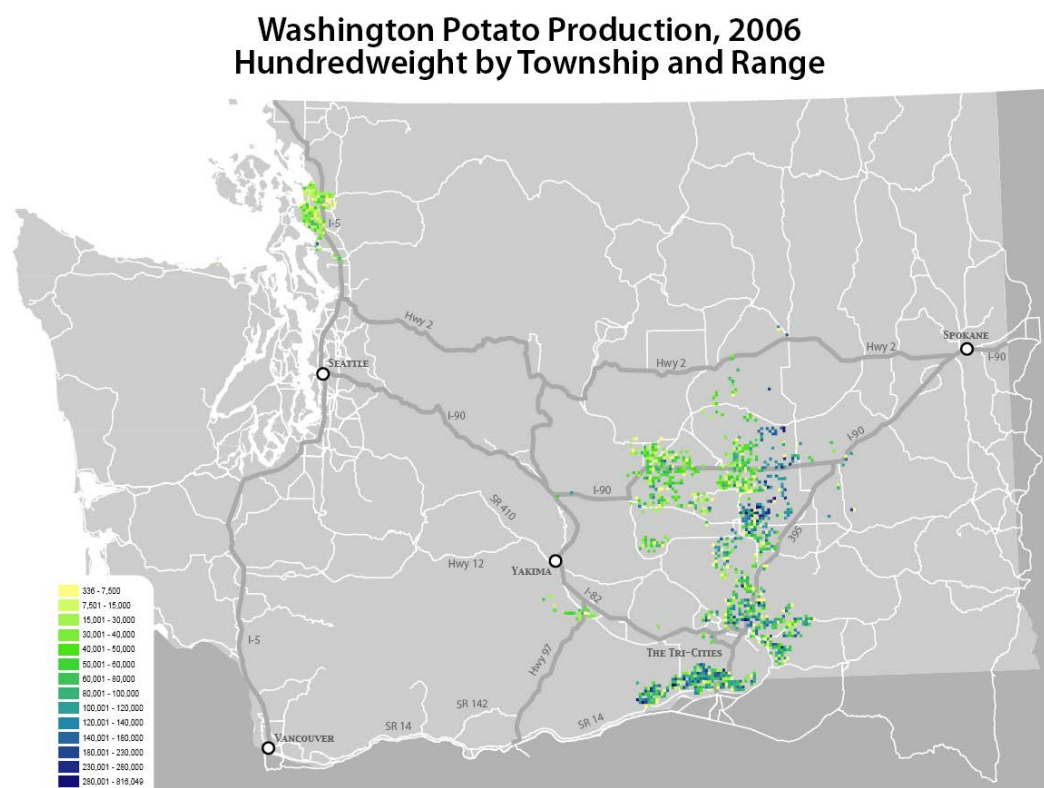
The resultant commodity flow database includes the number of daily truck movements (and potato volumes) of potato products (fresh, processed, dehydrated, and chips) on all roads in Washington State. This includes movements from potato fields to potato processors, and then on to markets around the state, or locations of export from the state.

Methodology

The methodology used was to estimate potato distribution across four categories in Washington: fresh, frozen, dehydrated, and potato chips. Additionally, the use of intermediate facilities was to be estimated. Potato production data from the USDA and the Washington State Potato Commission was used to identify

potato production in 2006 by township by hundredweight as shown in Figure 3.3. This figure shows that there were two primary potato farming locations in the state – the Lower Basin and the Skagit Valley. Conversations with the Washington State Potato Commission indicated that 94 percent of the potatoes that were grown were ultimately sold into the market. Interviews of potato production locations were used to estimate the range of potato movements from their farm location to processing plants.

Figure 3.3 Potato Production in Washington



Source: “Development and Analysis of a GIS-Based Statewide Freight Data Flow Network”, WSDOT, 2009.

For both production regions, the percentages of potatoes sold fresh or processed were estimated on the basis of a year-end comparison from the Washington State Potato Commission. The Commission was also used to identify the location of the 16 potato processing facilities in the state, and the number of trucks of processed potatoes relative to the ratio of fresh potatoes at each facility. An establishment survey conducted by the Commission was used to determine the distribution of potatoes by destination region. Interviews of potato processors and the Commission were used to complete the origin-destination patterns, develop the mode share for these trips, and routes used for truck trips. These were flows were mapped as shown below in Figures 3.4 and 3.5. Ultimately, this

data was used to determine the number of truck trips per day for specific origin-destination combinations and routes as shown in Table 3.8 below.

Table 3.9 Daily Truck Trips on Cross-Cascade Routes for Potato Products

Route and O-D Combination	Truck Trips By Potato Type				
	Total	Fresh	Frozen	Dehydrated	Chips
Hwy 2 East	4.3				
Skagit to Moses Lake		0.13			
Skagit to Spokane		1.40			
Skagit to Warden		0.13			
Skagit to other U.S. states (excluding OR and CA)		2.62			
Hwy 2 West	1.8				
Upper Basin to Stanwood		0.40	1.32	0.06	0.06
I-90 East	1.2				
Skagit to Kennewick		0.50			
Skagit to Yakima		0.37			
Skagit to Grandview		0.37			
I-90 West	32.4				
Upper Basin to Seattle		0.56	1.83	0.08	0.09
Upper Basin to Tacoma		0.47	1.52	0.07	0.07
Upper Basin to Auburn		0.56	1.83	0.08	0.09
Upper Basin to Port of Seattle		3.70	9.04	0.52	0.59
Lower Basin to Seattle		1.28	4.15	0.17	0.20
Lower Basin to Stanwood		0.92	2.99	0.12	0.15
Lower Basin to Port of Seattle		0.35	0.84	0.05	0.05
410 West	10.6				
Lower Basin to Tacoma		1.06	3.46	0.14	0.17
Lower Basin to Auburn		1.28	4.15	0.17	0.20

Source: "Development and Analysis of a GIS-Based Statewide Freight Data Flow Network", WSDOT, 2009.

Types of Applications That Have Been Implemented

The data is applied to a GIS network of state highway routes so that truck traffic on specific routes can be estimated (assuming vehicles use the shortest travel time route). Disruption scenarios can be evaluated by changing the transportation network, and evaluating new routing. Increased travel distances are estimated, and resulting costs (this requires additional industry economic data). When completed for other commodities, infrastructure use patterns can be compared, and the relative impact of infrastructure changes evaluated.

Applicability to Other States/Regions

Data is required regarding fixed infrastructure, such as transportation guideways and terminals used. Also data on movements, for example product origins, destinations, and volumes. Commodity data is required to understand any changes to the product that occur along the supply chain.

This method requires significant time in collecting, analyzing, and manipulating data from multiple data sources. While the data exists, it is held in many different locations, and is often not intended for this purpose (therefore not made readily available to the analyst). It is difficult to evaluate the accuracy of the final flows, but this is true of all commodity flow estimates.

This method is relevant for any other region, and for any commodity for which the relevant data can be obtained.

3.3.2 Diesel Distribution in Washington State

General Information

July 2009

Andreoli, Goodchild, Jessup, Rose

University of Washington, Washington State University, Washington State DOT

Commodity Flow Data Inputs

This project identifies diesel distribution routes used in the truck transportation of diesel (between terminal racks and cardlock facilities) in Washington State. Unlike the potato case study, specific numbers of trucks or volumes of diesel are not generated. However, the number of cardlock-terminal pairs that use a specific road or route in Washington is estimated. Data is obtained from a variety of public sources, primarily the Washington Department of Ecology which regulates terminal racks, the Washington Department of Environmental Protection, which regulates all diesel storage facilities, and Petrocard, which operates the cardlock facilities.

Final Commodity Flow Database

The resultant database includes the number of O-D pairs (terminal rack – cardlock facility) on all roads in Washington State.

Methodology

Information describing the diesel distribution network within Washington state is not readily available at a single source, but rather is held in a multitude of locations by a number of agencies. The researchers made several inquiries at multiple agencies as follows:

- Washington Transportation Plan Update Freight Movement (2008c) – provided general information on the locations of refineries, pipelines, and terminals, and outlines the distribution of petroleum products on an aggregated level within Washington state and to other locations.
- State of Washington Energy Facility Site Evaluation Council report (2007) – verifies much of the WSDOT information and describes the pipeline and barge supply chains in greater detail
- The Washington State Freight and Goods Transportation System (FGTS) 2007 Update – provided an understanding of the roadway network within the state as it applies to the movement of freight
- The Washington State Department of Agriculture (DOA) – regulates the accuracy of the quantity and quality of fuel delivered at gas stations in the state. The locations that it monitors are based on a list of businesses that have fuel meters registered with the Washington State Department of Licensing (DOL). The data provided by this source was locations by geography not fuel type.
- Washington State Department of Ecology – regulates active underground storage tanks in the state. This department provided a databases of all active USTs by tank volume, fuel type, geographic coordinates, and physical addresses for over 10,000 USTs, of which over 2,000 were classified as holding fuel.
- The US EPA - regulates above-ground storage tanks (ASTs). The EPA provided a database containing all current year ASTs with tank volumes, fuel types, and physical addresses for the 67 identified tanks containing diesel.
- The Washington State Department of Revenue – assesses and collects fuel taxes at terminal locations throughout the state. The tax is applied per gallon. This department provided information at the 27 diesel fuel terminal locations, including the five refineries.

The next step was to match UST and AST locations to the diesel fuel network. ECY was used to obtain information on oil product carrying vehicles. Outbound data included only exports, and inbound data were not origin-specific. The U.S.

Army Corps of Engineers Waterborne Commerce Statistics Center contains aggregated petroleum and crude waterborne imported and exported data for Washington state, but they are not disaggregated for diesel product, and are only origin-destination specific by state or country. The Washington State Energy Profile was obtained from the U.S. Energy Information Administration website. Conversations with EIA staff led to estimates of the total production capacity at each Washington state refinery, and estimates of the percent of diesel production for total Washington state refined production. Information from the Federal Energy Regulatory Commission was used to estimate pipeline flows.

The Washington Oil Marketers Association provided information on relationships between marketers and major oil companies and verified the general diesel supply chain network. It also provided names and contacts for diesel distribution marketers. These marketers were contacted to fill in missing pieces of the diesel distribution network not previously uncovered by the research.

The Strategic Freight Transportation Analysis (SFTA) at Washington State University conducted a roadside origin-destination survey of freight trucks in 2003 and 2004. Collected data includes UN placard numbers which categorizes whether or not a truck is carrying diesel products. The collected also include origin-destination pairs at the city level, and payload weight.

A comprehensive database on roadway flows for diesel movements was not obtained, but many pieces were acquired. The resulting stitching together of data, information received from interviews, and implied from other sources resulted in the following supply chain map of diesel flows with modal and geographic distributions as shown in Tables 3.9 and 3.10.

Figure 3.4.- Map of Diesel Flows – page 110.

Table 3.9 – Refinery Capacity (barrels/calendar day) – page 111

Table 3.10 – Transportation Mode Split from Washington State Refineries – page 112

Types of Applications That Have Been Implemented

When combined with a GIS network of state highway routes, O-D pairs using each route can be estimated (assuming vehicles use the shortest travel time route). This provides one indication of the relative importance of these infrastructure facilities to the diesel distribution system. Disruption scenarios can be evaluated by changing the transportation network, or by removing one of the source nodes (terminal racks) and evaluating the new routing pattern. Increased travel distances are estimated per O-D pair, and the new pattern of infrastructure usage. When completed for other commodities, infrastructure use patterns can be compared, and the relative impact of infrastructure changes evaluated.

Applicability to Other States/Regions

Unlike the Potato industry, the diesel distribution industry did not have a well organized group (such as the Potato Commission), which aggregates data. Individual diesel marketers were unwilling to share their flow data. Data was therefore only obtained on fixed infrastructure, not flow.

Significant effort was required to identify the data sources, but once the sources were identified, creating GIS layers and completing the routing analysis can be expediently performed. This method is efficient and recommended for other regions. Data on fixed infrastructure is typically easy to obtain, and even in the absence of commodity flow data, some measure of infrastructure importance can be calculated. However, true commodity flows were not identified.

3.3.3 Hazardous Materials Commodity Flow Data and Analysis, Ongoing, Texas A&M University

The objective of this project is to produce an updated, user-friendly guidebook for conducting hazardous materials commodity flow surveys to support local risk assessment, emergency response preparedness, and resource allocation and to support analyses across jurisdictional boundaries. This guidebook will be aimed at transportation planning and operations staff at the local and regional levels, as well as local and regional personnel involved in hazardous materials training, and emergency response. It is expected that all modes of transportation, all classes and divisions of hazardous materials, and the effects of seasonality on hazardous materials movements will be included.

At this time, the draft final report is under review and revision.

The project is being funded through the Hazardous Materials Cooperative Research Program (HMCPR) as HM-01 for \$300,000.

3.4 INTERNATIONAL DISAGGREGATION TECHNIQUES AND FREIGHT SURVEY EFFORTS

3.4.1 Sweden Commodity Flow Survey

General Information

Commodity Flow Survey 2001 Method Report, 2003:2004, Statistics Sweden, http://www.sika-institute.se/Doclib/Import/104/sr_2003_4e.pdf

In 2000, Statistics Sweden, (SCB), was commissioned to carry out a commodity flow survey for 2001 (VFU01) by the Swedish Institute for Transport and Communications Analysis (SIKA), the responsible agency for official statistics concerning transportation and communications in Sweden. The study also involved the Swedish National Rail Administration, the Swedish Civil Aviation Administration, the Swedish Maritime Administration, Vinnova, and the

National Road Administration. SCB had previously produced two small commodity flow surveys in 1996 (VFU 1996) and 1998 (VFU 1998). The explicit purpose of the 2001 CFS was to improve knowledge of transport of goods by the business sector, and to obtain and maintain an overall view of the need to move goods within and outside Sweden. Specifically, the CFS was to describe the locations between which different shipments of commodities are moved, using actual data.

Methodology

The survey was a combination of a sample survey and a register-based survey. The sample survey included local units involved in mining, manufacturing and wholesale sectors. This was supplemented with register data for forest and logging products, sugar beet cultivation and dairy farming. The survey population consisted of outgoing and incoming shipments at particular local units within companies. A local unit is a single physical location where business transactions occur.

The sample was selected using a stratified three-stage sampling-design. From a sample of 12,419 local units, commodity flows were to be estimated for the universe of 38,000 local units. The first stage was the selection of the local units within the companies. The second sampling stage was the selection of different time periods for reporting weeks (e.g., quarterly reporting for the largest firms). The final sampling stage was to select individual shipments for the selected local units for specific time periods. The reported data on individual shipments includes: value and weight; commodity type; modes of transport; cargo type; origin; destination; and geographic point of departure/entry from/to Sweden for foreign shipments (SIKA 2005). The local units were stratified by main type of commodity produced, geographic location and size (number of employees). To cover the desired local units, a new sampling frame was constructed every quarter comprised of approximately 3,100 local units. The reporting weeks were spread over the quarters. The 2001 Commodity Flow Report (Swedish and English) is available on the Internet http://www.sika-institute.se/Doclib/Import/101/sm_0710105.pdf

Final Commodity Flow Database

According to Werke,⁵ these data were intended to serve the needs of the SAMGODS models, a modeling platform for long-term Goods Movement at the national level. A review of the 2001 deployment revealed a 78 percent response

⁵ Werke, L., (2007). *Swedish Commodity Flow Surveys Evaluated – Statistics Sweden's Experience and Survey Adjustments since 2001*. A paper presented at the Third International Conference on Establishment Surveys (ICES -III) in Montreal, Quebec, Canada on June 18-21, 2007. www.amstat.org/meetings/ices/2007/presentations/.../LarsWerke.ppt.

rate, with 11 percent reporting electronically. Commissioners reported being satisfied with the quality and the delivery of the surveying effort. However, there were reports of respondent burden and some concern over missed commodities (e.g., sand and gravel).

The next CFS deployment occurred over a two-year period (2004/2005). Based on the findings from the 2001 CFS, the following recommendations were made: to cover more sectors; to develop a logistics model for SAMGODS; to collect additional variables; to increase the use of electronic reporting; reducing respondent burden; and improving the overall design of the survey. There were also concerns regarding controls, prompting the development of simple control tools for the deployment effort.

From a sample of 11,992 local units, commodity flows were estimated for the universe of 26,000 local units (SIKA 2006). The response rate for the 2004/2005 CFS declined to 74 percent, while electronic reporting increased to 18 percent. However, there were reports of respondent burden, with some problems with particular sectors (Werke 2007). The *2004/2005 Commodity Flow Survey Report (Swedish and English)* is available on the Internet

http://www.scb.se/statistik/TK/TK1007/ss_2006_12_rev.pdf

Commodity Flow Data Inputs

The data elements for the national surveying effort included:

- Quantitative shipment variables
 - Invoice value (excluding value-added tax and freight costs)
 - Weight (excluding packaging)
- Qualitative shipment variables
 - Address for shipment which has been received/sent from/to another address than the address of the local unit
 - Commodity code
 - Cargo type
 - Dangerous goods (cross-marked)
 - All transport modes in Sweden
 - All transport modes outside Sweden
 - The recipient's postcode (for delivery in Sweden)
 - The final destination in Sweden (border crossing point for export/import)
 - Final destination abroad
 - Recipient country
 - Place sent from

- Country sent from
- Access to and use of private siding for rail transport
- Access to and use of wharf for sea transport

Final Commodity Flow Database

The Commodity Flow Survey for 2001 produced data on movement of goods within Sweden, and with Sweden and foreign recipients/consignors. The final CFS database was used to produce the tables in the CFS Report with a 95 percent confidence interval. Commodity groups are listed in NST/R-terms (page 25) and Swedish Standard Industrial Classification – SN192 (page 26) (see the 2001 *Commodity Flow Report (Swedish and English)*)

http://www.sika-institute.se/Doclib/Import/101/sm_0710105.pdf.

The Commodity Flow Survey for 2004/2005 produced data for movement of goods within Sweden, and with Sweden and foreign consignors/consignees. The final CFS database was used to produce the tables in the CFS Report with a 95 percent confidence interval. Commodity groups are listed in NST/R-terms (page 26) and Swedish Standard Industrial Classification – SN12002 (page 27) (see the 2004/2005 *Commodity Flow Survey Report (Swedish and English)*)

http://www.scb.se/statistik/TK/TK1007/ss_2006_12_rev.pdf.

Types of Applications That Have Been Implemented

One of the major objectives of the CFS 2001, 2004/2005 deployments was to provide sufficient base-year matrices for the Swedish National Freight Model (SAMPLAN) (see *SAMPLAN The Swedish National Freight Model*)

http://www.sika-institute.se/Doclib/Import/106/sa_2004_1.pdf).

Hill et al. (no date) used the CFS data to produce improved visualizations (see http://www.citilabs.com/usergroup/files/futurapresentations/SIKA_Swedish_National_Freight_Model_part1.pdf).

Hill, P., Johansson, M. and T. Vorraa. The Swedish National Freight model. SIKA. (no date).

http://www.citilabs.com/usergroup/files/futurapresentations/SIKA_Swedish_National_Freight_Model_part1.pdf

The Swedish CFS was used in the POLCORRIDOR Logchain study, a project aimed at developing a new rail freight supply corridor that would link the Nordic countries with South-Eastern Europe. Other data sources for this effort included: Nordic databases and statistics derived from Norwegian and Swedish international trade statistics; official transport statistics on different transport modes, port statistics, private statistics from rail operations, harbors and ferry companies.

Fearnley, N., Grue, B., and J. Ludvigsen. 2003. Demand Assessment and Market Analysis for a new Trans-European Rail Transport Corridor. A paper presented at the European Transport Conference in Strasbourg, France, on October 8-10, 2003.

The authors proposed the use of the Swedish CFS to develop a logistics model for Norway and Sweden. They describe an “aggregate – disaggregate-aggregate” approach. They suggest the Swedish CFS be a starting point for the disaggregate approach, however, it was based on a sample of suppliers, and their approach focuses on the behavior of the receivers. They propose using the CFS as a “sample of supplier/receiver pairs.” Since Norway has no equivalent type dataset, they conclude that they need to develop a synthetic or hybrid approach for their modeling efforts.

De Jong, G., Florian, M., Grenland, S. and A. van de Voort. 2005. Specification of a Logistics Model for Norway and Sweden. A paper presented at the European Transport Conference in Strasbourg, France, on October 3-5, 2005.

This paper describes the continuing work on the logistics model for Norway and Sweden.

De Jong, G., Ben-Akiva, M., and J. Baak. 2007. A Micro-Model for Logistics Decisions in Norway and Sweden Calibrated to Aggregate Data. A paper presented at the European Transport Conference at the Leeuwenhorst Conference Centre in the Netherlands on October 17-19, 2007.

De Jong, G. and M. Ben-Akiva. (2007). A micro-simulation model of shipment size and transport chain choice. *Transportation Research Part B*, 41:950-965.

The Oresund region has been impacted on the opening of the Oresund Bridge connecting Denmark and Sweden. In order to model the freight flows between the two countries, the authors reviewed the Swedish CFS and other data sources (e.g., the FEMEX/COMVIC RP dataset) and concluded that they would need to apply an aggregated or weighted estimation based on original base-line O-D matrices, in conjunction with level of service matrices from assignment models.

Holmblad, R., and C. Hansen. 2008. Freight choice model for mode and crossing: A forecast model for the Oresund region. A paper presented at the European Transport Conference at the Leeuwenhorst Conference Centre in the Netherlands on October 6-8, 2008.

The authors made extensive use of the Swedish CFS dataset (referencing the models explaining the choice of mode chain and discrete shipment size for the Swedish CFS for 2004/2005 developed by E. Windisch in 2009⁶). They develop three models: one for mode choice and two models that simultaneously

⁶ Windisch, E., (2009). *A Disaggregate Freight Transport Model of Transport Chain and Shipment Size Choice on the Swedish Commodity Flow Survey 2004/2005*, M.Sc. Thesis, Delft University of Technology.

explaining mode and shipment size, using either discrete or a continuous variable for shipment size.

De Jong, G. and D. Johnson. 2009. Discrete Mode and Discrete or Continuous Shipment Size Choice in Freight Transport in Sweden. A paper presented at the European Transport Conference at the Leeuwenhorst Conference Centre in the Netherlands on October 5-7, 2009.

Applicability to Other States/Regions

Considerable research was conducted prior to the CFS 2001 deployment. The methodology appears to be directly transferrable to the U.S. Similar data sources are available to implement the steps described in this study. However, the cost to assemble and analyze the data would require significant resources.

Issues/Limitations

Overall, this methodology appears to be similar to the BTS Commodity Flow Survey. However, it would be interesting to conduct an analysis to determine how well economic data matched with the freight flow data derived from the expanded survey. This would be particularly interesting for sectors that were considered out-of-scope for the BTS CFS. This could serve to further validate the process used by FAF2 for out-of-scope industries. There do not appear to be the same concerns about confidentiality with the Swedish data, so it is possible that there are more cells that were populated.

3.4.2 International Conference on Transport Survey Methods (Part 1)

General Information

Patier, D. and J-L. Routhier. How to Improve the Capture of Urban Goods Movement Data? In *Transport Survey Methods: Keeping Up with a Changing World*, edited by P. Bonnel, M. Lee-Gosselin, J. Zmud, J. Madre, 2009. Published by Emerald, United Kingdom. Based on the proceedings of the International Steering Committee for Travel Survey Conferences (ISCTSC) 8th Annual International Conference on Transport Survey Methods in Annecy, France, May 25-31, 2008.

The authors review activities pertaining to urban goods movement data. With respect to commodity flow surveys, the Best Urban Freight Solutions (BESTUFS II) activities are reviewed. A study, *Urban freight data collection harmonization and modeling*, was conducted by the Laboratoire d'Economie des Transports in Lyon. The research team surveyed 70 European experts and received completed questionnaires from 35 respondents, in 11 European countries to gather information on urban data collection activities.

The authors also review the Canadian business surveys in Edmonton and Calgary. They found that 3,000 business establishments in the Calgary Region in 2000 and 4,500 in the Edmonton Region in 2001-2002 were surveyed.

Commodity Flow Data Inputs

Urban freight data in the country of the expert respondent:

- Name of the data-collection survey;
- Observation unit (vehicle, commodity, establishment, depot, delivery/pick-up, etc.);
- Name of organization conducting data collection effort;
- Reason for data collection;
- Use of the data for modeling;
- Frequency of data collection;
- Last recorded data collection effort;
- Type of data collected;
- Method of data collection;
- Sample size;
- Units of measurement;
- Geographical area where data was collected; and
- Determination of difficulty of extracting urban data from dataset.

Inventory of urban goods data collection in 11 European countries, indicates that commodity flow (O-D) surveys have been conducted at the national level in Belgium, Sweden, and Switzerland. The purpose of these efforts was to capture exchanges between regional areas.

The BESTUFS series can be accessed at <http://www.bestufs.net/>

Debauche, W. and D. Decock. Report on urban freight data collection in Belgium. Belgian Road Research Centre, August 2006. BESTUFS WP 3.1 Series.

http://www.bestufs.net/download/BESTUFS_II/key_issuesII/BESTUFS_II_results_datacollection/BESTUFS_II_data_collection_report_BE.pdf

Abel, H. and W. Dietrich. Report on urban freight data collection in Switzerland. August 2006. BESTUFS WP 3.1 Series.

http://www.bestufs.net/download/BESTUFS_II/key_issuesII/BESTUFS_II_results_datacollection/BESTUFS_II_data_collection_report_CH.pdf

Browne, M., and J. Allen. (2006). Report on urban freight data collection in Sweden. August 2006. BESTUFS WP 3.1 Series.

http://www.bestufs.net/download/BESTUFS_II/key_issuesII/BESTUFS_II_results_datacollection/BESTUFS_II_data_collection_report_SE.pdf

Regional data collection in Canada (Edmonton and Calgary) (as presented at the Conference)

Data elements included a full range of commodities transported (goods and services) and a description of the vehicle movements during an assigned day of travel.

The objective of the surveys was to investigate the generation of goods movements, distribution patterns, influence of employment size and type of commodities shipped, including the pattern of delivery for a given commodity flow, the use of depots, all the trips needed to accommodate a particular commodity flow and the types of vehicles used for the movement of the commodity.

The surveying effort used a sampling frame supplied by Alberta Treasury, including business name, address, and industry classification. The data required considerable “cleaning.” Recruitment began with a contact by mail, a follow-up phone call to determine number of employees and the nature of the business activities, and preliminary acceptance of participation. Finally, a formal letter was sent, designating a contact person to receive the appropriate forms for participation.

The actual collection of the data occurred at three possible levels: a person at the firm was made responsible for survey deployment, determined by a phone call, and later, visited for a review of the completed forms; a first visit is made to begin the surveying effort, with a phone call made after the survey day and a visit to review completed forms; and finally, a survey expert remained on site for the entire surveying effort.

The weighting strategy used an individual scale factor based on employment size, standard industrial category and geographical location. The final step was to match expansion results loaded on a node-and-link modeling network with actual observed vehicle classification counts.

Types of Applications that Have Been Implemented

According to the authors of the Canadian review, the data is being used to develop a novel form of tour-based micro-simulations of commercial vehicle movements.

Applicability to Other States/Regions

In Calgary, the total cost was CAD \$600,000 and \$800,000 for the Edmonton portion of the study.

3.4.3 International Conference on Transport Survey Methods (Part 2)

General Information

Zly, K., Kalinowska, D., Madre, J-L., and B. Leore. Vehicle-Based Surveys: Synthesis of a Workshop. In *Transport Survey Methods: Keeping Up with a Changing World*, edited by P. Bonnel, M. Lee-Gosselin, J. Zmud, J. Madre, 2009. Published by Emerald, United Kingdom. Based on the proceedings of the International Steering Committee for Travel Survey Conferences (ISCTSC) 8th Annual International Conference on Transport Survey Methods in Annecy, France, May 25-31, 2008.

Roadside interviews (RSI) provide a surveying methodology that can be conducted in conjunction with traffic counts (automated). The methodology requires a sample of vehicles to be pulled off the desired route, in a safe and convenient location. This is sometimes accomplished with the assistance of traffic police. After stopping, drivers are asked for their participation in a short interview, often a short survey.

Commodity Flow Data Inputs

Data elements collected include:

- Vehicle occupancy;
- Type of commodity carried for goods movement;
- Trip purpose;
- Origin and destination (O-D) of trip;
- Departure and arrival times;
- Route; and
- Driver characteristics.

Expansion factors are used to scale information to universe of trucks.

The authors discussed issues facing the use of roadside surveys due to safety concerns, making only a few locations adequate for a stopping/turn-off activity. This location may not be optimal for data collection and could add bias to the data. Drivers may feel pressured to remain on time for their delivery schedule and rush their participation. Also identified was the high cost of conducting roadside surveys and potentially low productivity of trying to operate at a temporary site.

One variation of the traditional on-site survey, is to allow the driver to take the survey instrument and complete the form at a later time. While making it possible to distribute more surveys, it is possible that the response rate will fall due to leaving the responsibility for completion to the driver.

3.4.4 2000 Calgary Commodity Flow Survey Report

General Information

2000 Commodity Flow Survey Report. International Results Group. 2001.

http://www.calgary.ca/DocGallery/BU/trans_planning/forecasting/commodity_flow_survey_report.pdf

City of Calgary Planning Policy/Data Management and Forecasting

The 2000 Commodity Flow Survey was conducted for the Planning Policy/Data Management and Forecasting at the City of Calgary and Alberta Infrastructure to support their 2001 model update project. In addition, the City of Calgary wanted to include the flow of goods and services in update of the Calgary Regional Travel Model (CRTM), in 2001. The survey effort required contacting business establishments and shipping depots to collection information on the type, amount, value and method of shipping goods. The survey was an establishment-based, shipper/receiver on-site and mailback survey. Data was collected on all outbound shipments, and inbound and outbound for depots, for a 24-hour period. For each shipment, data was collected on the major commodity, weight, value, method of transportation, origin, and destination.

An anticipated use of the commodity flow data was to be able to prioritize projects that are critical to business operations, based on “commodity mobility.” Having a database function should reduce the need for special studies to determine the right projects to accommodate commodity flow.

The study area included the City of Calgary and the surrounding region. The survey deployment began August 2000 and was completed in January 2001. For the City of Calgary, from the original 49,354 companies in the sample database, 3,107 responded to the survey and provided usable data. For the Region, from the original 11,102 companies in the sample database, 304 provided usable data.

The survey methodology included an introductory letter, with recruitment conducted using a telephone call. Recruiters made scheduled several meetings with participants. Companies were expected to control the distribution of the survey materials internally, with help available from the consulting team, if desired.

Commodity Flow Data Inputs

Regional Data Collection

Individual shipment data included: type of goods shipped, direction of shipment (in or outbound), address of destination, quantity of goods being shipped, unit of quantity, value of goods, and routing.

Additional data collection efforts included an external truck study and commodity flow studies in near-by regions.

Types of Applications that Have Been Implemented

The authors conducted a dual region study, including Edmonton and Calgary regions. In their comparison of commodity types carried on commercial vehicles in the two regions, they found Edmonton had a greater proportion of trips hauling fabricated metal, manufactured goods. Stefan, K., Brownlee, A., McMillan, J., and J. Hunt. (no date). *The Nature of Urban Commercial Movements in Alberta*.

Hunt J. and J. Stefan. 2003. Calgary Commercial Movement Model. A paper presented at the 17th Annual EMME/2 Users Conference, Calgary, Alberta, Canada, October 2003.

Hunt, J., Stehan, K., Brownlee, A., McMillan, J., Farhan, A., Tsang, K., Atkins, D., and M. Ishani. 2004. A Commercial Movement Modeling Strategy for Alberta's Major Cities. A paper presented at the Goods Movement – New Approaches Session of the 2004 Annual Conference of the Transportation Association of Canada, Quebec City, Quebec. <http://www.tac-atc.ca/english/resourcecentre/readingroom/conference/conf2004/docs/s15/hunt.pdf>

Stefan, K. and J. Hunt. (no date). Review of Urban Commodity Movement Demand Modeling Approaches. http://www.calgary.ca/DocGallery/BU/trans_planning/forecasting/review_of_urban_commodity_movement.pdf

3.4.5 Edmonton Region Truck/Commodity Survey

General Information

<http://www.edmonton.ca/transportation/RoadsTraffic/EdmontonRegionExternalTruckCommoditySurvey1.pdf>

Alberta Transportation and the City of Edmonton conducted the External Truck/Commodity Flow Survey, with support from Economic Development Edmonton, Alberta Motor Transport Association, and Alberta Capital Region Alliance.

Alberta Transportation and the City of Edmonton combined their efforts in the fall of 2001 to conduct an external truck/commodity survey in order to better understand the movement of commercial goods by truck into and out of the Edmonton Region. The purpose of the study was to assess transportation needs for both the City of Edmonton and Alberta Transportation.

Roadside interviews were conducted from September through November 2001, at 24 sites on provincial highways between 8:00 AM and 4:00 PM. Over 6,500 truckers were interviewed.

The database was assembled electronically then scaled up to represent daily trips. The information was also used to complement the expansion of the Edmonton Region Personal Travel Model and to plan for future transportation

needs. The study area includes the Edmonton Census Metropolitan Area and portions of the County of Lamont.

Analysis included the percentage of trucks by number of commodities, trucks by commodity type (including inbound and outbound), commodity origins and destinations.

Commodity Flow Data Inputs

Regional data – Edmonton Census Metropolitan Area and County of Lamont.

The survey collected the following:

- Commodities carried;
- Origins and destinations;
- Type of vehicles;
- Stops being made in the Region;
- The time of day of travel;
- Highways used to transport commodities; and
- Ownership of the vehicle.

Additional data collection efforts included commodity flow surveys and external truck studies in near-by regions.

Applications of Methodology

The authors conducted a dual region study, including Edmonton and Calgary regions. In their comparison of commodity types carried on commercial vehicles in the two regions, they found Edmonton had a greater proportion of trips hauling fabricated metal, manufactured goods. Stefan, K, Brownlee, A., McMillan, J., and J. Hunt. (no date). *The Nature of Urban Commercial Movements in Alberta*.

Applicability to Other States/Regions

The authors conducted a dual region study, including Edmonton and Calgary regions. In their comparison of commodity types carried on commercial vehicles in the two regions, they found Edmonton had a greater proportion of trips hauling fabricated metal, manufactured goods. Stefan, K, Brownlee, A., McMillan, J., and J. Hunt. (no date). *The Nature of Urban Commercial Movements in Alberta*.

3.4.6 A Shipper-Based Survey of Goods and Service Movements in the Greater Gold Horseshoe (GGH) Report I: Survey Design and Implementation

General Information

A Shipper-Based Survey of Goods and Service Movements in the Greater Gold Horseshoe (GGH) Report I: Survey Design and Implementation (2007). Roorda et al.

Partially funded by the Ministry of Transportation, Transportation Planning Section.

The study area, the Region of Peel, is largely suburban, located just west of the City of Toronto.

The Region of Peel Commercial Vehicle Survey was designed to address several specific data gaps and included commodity information (e.g., commodity type, weight, value, origin and destination of goods shipments).

The survey was conducted from October 2006 to May 2007. The mail-out mail-back component of the Region of Peel Commercial Vehicle Survey had an overall response rate of 25.3 percent (597 responses), exceeding the target of 400 surveys. These surveys were sufficiently representative of industry types and company sizes, with an intentional oversampling of manufacturing industries, and a somewhat heavier representation of mid-sized firms. The targeted population was all business establishments located within the Region of Peel.

The sample frame available from Info Canada contained 47,000 companies in the Peel Region, up to six industry classifications and the number of employees for each company. The company's industry is classified according to Standard Industry Classification (SIC) code. The industry classifications are converted into North American Industry Classification System (NAICS) and aggregated to the industry groupings.

Commodity Flow Data Inputs

Regional data collection for commodities.

Commodity type, weight and value.

In addition to the commodity data, the survey effort included a driver's survey and a GPS component (see McCabe et al., Comparing GPS and non-GPS Survey Methods for Collecting Urban Goods and Service Movements. A paper presented at the 8th International Conference on Survey Methods in Transport in Annecy, France on May 25-31, 2008).

Applications of Methodology

Use of 2006 Region of Peel Commercial Travel Survey to support the development of a logistics model.

Roora, M., Cavalcante, R., McCabe, S. and H. Kwan. (2009). A conceptual framework for agent-based modeling of logistics services. *Transportation Research Part E*, 46 (2010): 18-31.

3.4.7 Edmonton Commodity Flow Study

General Information

http://www.edmonton.ca/transportation/RoadsTraffic/CommodityFlowSurveyProjectReportNov_2003.pdf

2002 Edmonton Region Commodity Flow Study was a joint project between the City of Edmonton and Alberta Transportation, with additional support from Economic Development Edmonton (EDE), Alberta Motor Transport Association (AMTA), and Alberta Capital Region Alliance (ACRA).

The surveying effort was undertaken, beginning in October 2001, to determine the characteristics of goods and service movements. The information was to be used to assist in the assessment of regional transportation needs, particularly for trucks, and for the development of short- and long-term transportation plans. Specifically, the CFS was to determine the quantity, O-D, and types of vehicles used to move goods in the Edmonton Region. The sample frame contained 13,792 eligible establishments, with 4,324 agreeing to participate in the surveying effort. The study area was the Edmonton Census Metropolitan Area (CMA). The sample of goods and services for the one-day survey was expanded to represent the universe of goods and service activities.

The data collected included goods or service type, quantity shipped and the units used to measure the quantity and the estimated value of the shipment.

Analysis included daily vehicle trip origins by goods category,

Commodity Flow Data Inputs

Regional data – For analysis and reporting, the City of Edmonton was grouped into 14 geographic areas. The surrounding Region was grouped into eight areas.

Commodity groups were classified into 18 categories (see page 10, Table 6 – Goods shipments).

Additional data collection efforts included commodity flow surveys and external truck studies in near-by regions.

Applications of Methodology

The authors conducted a dual region study, including Edmonton and Calgary regions. In their comparison of commodity types carried on commercial vehicles in the two regions, they found Edmonton had a greater proportion of trips hauling fabricated metal, manufactured goods. Stefan, K., Brownlee, A., McMillan, J., and J. Hunt. (no date). *The Nature of Urban Commercial Movements in Alberta*.

Kruger, D., Errol, T., A. Clavelle. (2007). Results of the TAC Project on the Framework for the Collection of High-Quality Data on Urban Goods Movement (Phase 1). A resource paper prepared for the 2007 Annual Conference of the Transportation Association of Canada in Saskatoon, Saskatchewan.

3.4.8 Canadian Trucking Commodity Survey

General Information

Transborder Trucking from a Canadian Perspective: Results from the Trucking Commodity Origin and Destination (TCOD) Survey. (no date). E. Hamilton of Statistics Canada. (Source: Rob Tardiff).

Agencies involved in the research or study

Statistics Canada

The Trucking Commodity Origin and Destination (TCOD) uses a sampling frame based on Statistic's Canada's Business Register of about 2 million businesses. 67,000 are classified as NAICS 484 – in 2005 – truck transportation industry participants. In this group are all Canadian firms where trucking is the main activity (does not include private trucking activities). The frame consists of 4,000 for-hire carriers with revenue of \$1 million or more. The sampling strategy selected a 50 percent sample of eligible firms, and selected a sample of shipments from in-sample carriers (13 percent of all shipments in 2005 sample.)

The collection effort had three processes: transcription of shipping documents from 80 percent of estimated shipments; telephone collection of all shipments of respondents with a small number of “typical” shipments (14 percent of estimated shipments); electronic data collection of all shipments from a few large respondents (6 percent of estimated shipments).

Commodity used Standard Classification of Transported Goods (SCTG) at three-digit level. Origin and destination included city/place names, and imputed postal and zip codes, with distances derived using PC Miler using codes for ends. The data contains weight, derived ton-kilometers and revenue. Regions covered included all of Canada (Atlantic, Quebec, Ontario and West), United States (Northeast, Midwest, South and West) and Mexico.

Study revealed that 16 percent of freight moving in 2005 was carried across the Canadian-U.S. border. Most of the activity revolves around Ontario and Quebec in Canada, and Michigan, Ohio, Pennsylvania and New York in the U.S. The most transported commodity was logs, wood, paper products, base metal products, motor vehicles and parts.

Commodity Flow Data Inputs

National data – describe geographic detail, commodity detail

Commodity used Standard Classification of Transported Goods (SCTG) at three-digit level. Origin and destination included city/place names, and imputed

postal and zip codes, with distances derived using PC Miler using codes for ends. Data collected includes: weight, derived ton-kilometers and revenue.

Regions covered included all of Canada (Atlantic, Quebec, Ontario and West), United States (Northeast, Midwest, South and West) and Mexico.

Applications of Methodology

This methodology has not yet been applied. However, there are plans to apply this data in the future which the research team has started to make inquiries about. We will update this section as information becomes available throughout the study.

3.5 LESSONS LEARNED FROM DISAGGREGATION TECHNIQUES

The literature review identified several efforts to develop sub-national commodity flow data. These efforts were divided into four categories, and the lessons learned from each category are described below.

3.5.1 Lessons Learned from Research-Oriented Disaggregation Techniques

There have been several attempts to disaggregate FAF2 data to the county level. Most of these Some of the methods relied on multinomial logit functions. Other methods tested different local data ranging from employment, number of establishments by industry sector, sector-specific truck VMT, and population. These methods confirmed that there are a wealth of local socioeconomic data for consideration in doing this type of disaggregation. All of the methodologies were successful at developing local commodity flow databases. However, the accuracy of these databases were never checked using real-world truck activity data or freight flow databases. One potentially beneficial activity to consider in future analyses is collecting small sets of establishment or household surveys to verify the validity of these methodologies. Additionally, the small surveys could be used to test the validity of different local economic factors that are appropriate for different industry sectors or commodities.

Another problem with disaggregating the FAF2 database is that it replicates the errors that potentially exist in the original database. In particular, the CFS out-of-scope sectors are estimated based on a combination of truck activity data and local socioeconomic data. These methods were never validated in the FAF2 development. Therefore, using these same techniques to disaggregate the FAF2 data has the potential of propagating errors through a local freight flow database.

One of the more potentially useful analyses was the comparison of the FAF2 data at the regional level to regional economic data. This was helpful because it starts

to highlight particular sectors which appear to be well-modeled using socioeconomic data, and sectors that are not well-modeled using these factors. This provides an initial cut at the industries to target in terms of conducting small sample surveys, and which industries will require different types of processes for estimation at the local level.

3.5.2 Lessons Learned from Planning-Oriented Disaggregation Techniques

The planning applications showed that for some applications, the utilization of Transearch data by itself is indeed sufficient. For the Virginia DOT Freight Plan, a trade flow analysis was conducted to identify major trading partners for counties and regions within Virginia. The Transearch data was adequate for this type of analysis. However, it is interesting to note that roadside truck origin-destination surveys were also collected for this freight plan. This data was used to verify the truck commodity movements on the highway infrastructure, as a means of confirming the Transearch data, and as a means of obtaining information on truck movements directly from the field rather than a quantitative post-processor applied to Transearch data. The roadside truck O-D survey data were particularly useful in conducting truck-rail diversion analyses for specific highway corridors in the state.

One of the lessons learned is that it is important for planning agencies to be very specific about what they want to do with freight flow data in order to assess which data source or sources are best and which analytical methods would be best to apply to the data.

There have been some efforts to utilize some combination of BTS CFS data, FAF2 data, and Transearch data as the starting point in the development truck travel demand models for estimating truck movements. It should be noted that similar processes can be utilized to understand the flows for specific commodities or commodity groups as needed. The model accuracy will of course be dependent on the accuracy of the underlying commodity flow database and the process used to transform tons to trucks.

The most significant issues with utilizing the main freight flow databases (CFS, FAF2, and Transearch) were the problems estimating trip chains or tours, urban trips, and short truck trips. These are based on the methodologies used to describe these databases, and is explained in more detail in Chapter 4.

There have been some initial discussions about the potential use of GPS data for developing local commodity flow databases. GPS data have already been used in several studies to provide a wealth of data on truck activity. However, there are some issues with using GPS data to develop commodity flow data. Obtain a stratified random sample of GPS-equipped vehicles across industry or commodity types has been a challenge at both the national and local levels. Additionally, GPS data do not track information about loading and unloading. This must be inferred from stop times and land uses at stops. This carries over

into the lack of any commodity data being tracked by the GPS equipment. Again, this information would need to be estimated based on understanding (or surveying) land uses at these stops.

3.5.3 Lessons Learned from Commodity Specific Disaggregation Techniques

There have been several efforts to develop commodity specific flow data using local economic data. In Oregon, this was used to refine the Transearch database that was purchased for the state. In Washington, these databases were developed to identify the number of trucks and truck routes used for select industries in the state. The primary lessons learned from these analyses is that the amount of information at the local level varies across industries and commodities. However, there are ways that commodity flow data provided by CFS, FAF2, and Transearch can be validated, enhanced, and disaggregated using these local data sources.

The Washington examples provide the best examples of the disparity of data available. For the potato sub-industry, the Washington Potato Commission was able to provide detailed information on where the crops were produced, what percent were processed or sold raw, where processing occurred, what modes were used, and where end customers were likely to be located. For diesel commodities, disparate pieces of data were spread across several government agencies, several layers of the private sector supply chain, and specific individuals. Each potential data source had its own potential for error or bias, and many potential sources of data were unwilling to provide information. After an exhaustive review of all of these sources, only a rough picture of data could be provided. The lesson learned is that there is no standard method for developing sub-national local data from existing sources. It will vary by state, sub-state region, and industry. In future tasks, the research team will develop a list of likely sources of additional information for industries and commodities that are not well-modeled in the three major commodity flow databases, but it is important to note that additional data should be sought depending on the amount of detail desired.

3.5.4 Lessons Learned from International Disaggregation Techniques

It is interesting to note that international freight planners experience many of the same issues as those in the U.S. The lack of freight data at the local level was cited by many researchers. Additionally, the lack of standardized data collection and modeling techniques created a disparate set of processes that are used in different locations. Many of the countries have shipper-based surveys that are similar to the BTS CFS. Others have developed hybrid processes of shipper-base and carrier-based surveys which may be of interest for application in the U.S.

4.0 Review of Freight-Related Data Sources

Our review of freight-related data in Chapter 3 indicates that there are 11 core data sources that are critical for developing subnational commodity flow databases. These 11 data sources were frequently utilized in the disaggregation processes and databases described in Chapter 3. These 11 core databases can be grouped into five data categories as shown in Table 4.2.

There are also dozens of secondary freight-related sources that can be used to provide supplemental information to the subnational commodity flow database development process depending on the techniques being used and the planning activity being supported.

Table 4.1 shows a full list of the freight-related data sources that were reviewed for this study along with categorical descriptions of each data base. The remainder of this section focuses on describing the core freight-related databases. The secondary data sources are described in more detail in Appendix A.

Table 4.1 Freight-Related Data Sources

Data Source	Type of Data			Count	Economic Output	Employment	Modes Covered			Inland Water	Air Cargo	Pipeline	No modal detail	Geographic Detail Included					Accessibility		Other	
	Commodity Flow	Vehicle Flow	Fleet Data				Truck	Rail	International Marine					National	International	State	County	Other	Public	Proprietary	Frequency of Data Release	Accuracy
Core Freight-Related Databases																						
FAF	●						●	●	●	●	●			●		●		89 FAF Regions	●			
CFS	●						●	●	●	●	●			●		●		Regions	●			
TRANSEARCH	●						●	●	●	●	●			●		●	●			●	Annual	
U.S. Waterborne Commerce - Domestic	●									●				●		●		Port zones	●			
U.S. Waterborne Commerce - Foreign	●								●					●		●		Port zones	●			
PIERS	●								●									Port		●		
Carload Rail Waybill Sample	●							●						●		●	●			●		
BTS Transborder Surface Data	●						●	●						●	●	●				●		
Establishment Data - InfoUSA					●	●							●	●	●	●	●			●		
Establishment Data - Dun & Bradstreet					●	●							●	●	●	●	●		●			
Census County Business Patterns					●	●							●	●	●	●	●		●			
Woods & Poole Economic Forecast					●	●								●	●	●	●					
U.S. Economic Census					●	●							●	●	●	●	●		●			
Secondary Freight Databases																						
VIUS	●		●				●							●		●			●			
Oil Pipeline																			●			
Air Traffic Statistics	●										●			●	●				●			
U.S. Ports and Waterway Facility Database	●								●					●	●			Port	●			
VTRIS-W				●			●							●	●	●	●	Corridor	●			
Regional Economic Accounts					●	●	●	●	●	●	●								●			
BLS Productivity Data					●	●							●						●			
Freight Facts and Figures	●		●				●	●	●	●	●			●	●	●			●			
North American Transportation Statistics	●	●	●				●	●	●	●	●			●	●	●			●			
Quarterly Census of Employment and Wages						●								●		●	●					
Fleet Owner’s Magazine FleetSeek Database National Motor Carrier Directory			●				●									●				●		
Fleet Owner’s Magazine FleetSeek Database Private Fleet Directory			●				●									●				●		
Fleet Owner’s Magazine FleetSeek Database Owner-Operator Database			●				●									●				●		

Data Source	Type of Data				Economic Output	Employment	Modes Covered							Geographic Detail Included					Accessibility		Other	
	Commodity Flow	Vehicle Flow	Fleet Data				Truck	Rail	International Marine	Inland Water	Air Cargo	Pipeline	No modal detail	National	International	State	County	Other	Public	Proprietary	Frequency of Data Release	
ATA Fleet Directory			●				●							●						●		
Massachusetts Institute for Social and Economic Research (MISER) U.S. Exports by State of Origin of Movement (State Exports), Foreign Trade Database					●	●							●						●			
MARAD U.S. Imports and Exports Transshipped Via Canadian Ports (Annual Report)	●								●					●	●				●			
MISER - now WISERTrade U.S. Exports by State of Origin of Movement (State Exports), Foreign Trade Database	●				●				●					●	●	●			●			
DRI/McGraw-Hill World Sea Trade Service	●								●					●	●					●		
IMPLAN Data Files					●	●								●	●	●	●			●		
U.S. Department of Agriculture Fresh Fruit and Vegetable Shipments by Commodities, States, and Months					●									●	●	●	●		●			
U.S. Department of Agriculture Grain Transportation Report					●									●	●	●	●		●			
Caltrans Weigh-in-Motion Data				●			●									●			●			
The Colography Group, Inc. U.S. Air Freight Origin Traffic Statistics	●										●			●						●		
U.S. DOR MARAD Maritime Statistics	●								●										●			
U.S. Foreign Trade Division U.S. International Trade Statistics	●				●				●				●	●	●				●			
U.S. Energy Information Administration State Energy Profiles	●				●	●						●		●				Region	●			
U.S. Department of Energy Quarterly Coal Report, 2008	●				●	●							●	●				Region	●			
California EDD Population 1990-2008						●							●				●		●			

Table 4.2 Freight-Related Data Sources

Category	Database
Multimodal Freight Flow Database	CFS
	FAF
	TRANSEARCH
Mode-Specific Freight Flow Database	U.S. Waterborne Commerce – Domestic
	U.S. Waterborne Commerce – Foreign
	PIERS
O-D-Specific Freight Flow Database	Carload Rail Waybill Sample
Freight Establishment Data	BTS Transborder Surface Data
Economic Data	Establishment Data – InfoUSA
	Establishment Data – Dun & Bradstreet
	County Business Patterns
	U.S. Economic Census

4.1 MULTIMODAL FREIGHT FLOW DATABASES

4.1.1 Commodity Flow Survey (CFS)

Sources:

BTS website. URLs:

http://www.bts.gov/publications/commodity_flow_survey/

http://www.bts.gov/help/commodity_flow_survey.html

General Information

The Commodity Flow Survey (CFS) is a joint effort by the Bureau of Transportation Statistics (BTS) and the U.S. Census Bureau. The survey is conducted on a five-year cycle as a component of the economic census. The first CFS was conducted in 1993, followed by surveys in 1997, 2002, and 2007. The primary goal of the 2007 Commodity Flow Survey (CFS) is to estimate shipping volumes (value, tons, and ton-miles) by commodity and mode of transportation at varying levels of geographic detail (i.e., national, state, select metropolitan statistical areas (MSAs)). A secondary objective is to estimate the volume of shipments moving from one geographic area to another (e.g., flows of commodities between states, regions, etc.) by mode and commodity.

The CFS data are used by policy-makers and transportation planners in various Federal, state, and local agencies for assessing the demand for transportation facilities and services, energy use, and safety risk and environmental concerns. Additionally, business owners, private researchers, and analysts use the CFS data for analyzing trends in the movement of goods, mapping spatial patterns of commodity and vehicle flows, forecasting demands for the movement of goods, and determining needs for associated infrastructure and equipment.

Sectors Included

The CFS is a shipper-based survey, and captures data on shipments originating from select types of business establishments located in the 50 states and the District of Columbia. The 2007 survey sampled over 100,000 establishments with paid employees that were located in the United States and were classified, using the 2002 North American Industry Classification System (NAICS) in the following sectors:

- Mining;
- Manufacturing;
- Wholesale trade;
- Select retail trade industries – specifically electronic shopping, mail-order houses, fuel dealers, publishers; and

- Auxiliary establishments (i.e., warehouses and managing offices) of in-scope multi-establishment companies. An advance survey of approximately 40,000 auxiliary establishments was conducted to identify auxiliary establishments with shipping activity.

The CFS does not include establishments classified in the following sectors:

- Farms;
- Forestry;
- Fishing;
- Utilities;
- Construction;
- Government-owned entities (except government-owned liquor stores);
- Transportation;
- Most retail and services industries; and
- Foreign-based business importing to the United States. However, in theory, domestic portions of imported shipments can be captured in the CFS once arriving at a U.S.-based establishment (assuming it is an eligible shipping establishment included in the CFS).

Sampling Frame

The sample was selected using a stratified three-stage design in which the first-stage sampling units were establishments; the second-stage sampling units were groups of four 1-week periods within the survey year, and the third-stage sampling units were shipments.

The first-stage sampling frame was developed from a subset of establishment records extracted from the Census Bureau's Business Register. The resulting frame comprised approximately 754,000 establishments as shown in Table 4.3.

Table 4.3 CFS Number of Establishments and Industry Strata

Trade Area	Number of Establishments	Industry Strata
Mining	6,789	3 (four-digit NAICS)
Manufacturing	327,826	21 (three-digit NAICS)
Wholesale	356,477	18 (four-digit NAICS)
Retail	25,190	2 (NAICS 4541 and 45431)
Services	22,539	1 (NAICS 5111 and 51223 combined)
Auxiliaries	14,878	3 (combinations of NAICS 4931 and

		551114)
Total	753,699	48

The frame for the second stage consisted of 52 weeks in 2007. Each establishment was systematically assigned to report for four reporting weeks – one in each quarter of the reference year. Each of the four weeks were evenly spaced such that there were 13 weeks between each reported period.

The frame for the third stage was the shipment selection. For each of the four reporting weeks, the respondent was requested to count or estimate the total number of shipments. If an establishment made more than 40 shipments during that week, the respondent was asked to provide information only about a systematic sample selected by the shipper. If an establishment made 40 or fewer shipments, the respondent was asked to provide information about all of the shipments.

Description of Data Collected

Information collected for each outbound shipment includes the following:

- Shipment ID number;
- Shipment date (month, day);
- Shipment value;
- Shipment weight in pounds;
- Commodity code from Standard Classification of Transported Goods (SCTG) List;
- Commodity description;
- United Nations or North American number for Hazmat material shipments;
- U.S. destination (City, State, Zip Code) – or gateway for export shipment;
- Modes of transport;
- An indication of whether the shipment was an export
- City and country of destination for exports; and
- Export mode.

For a shipment that included more than one commodity, the respondent was instructed to report the commodity that made up the greatest percentage of the shipment's weight.

The full questionnaire is shown in Appendix A.

CFS Definition of a Shipment

By CFS definition, a shipment is a single movement of goods, commodities, or products from an establishment to a single customer or to another establishment owned or operated by the same company as the originating establishment (e.g., a warehouse, distribution center, or retail or wholesale outlet). Full or partial truckloads are counted as a single shipment only if all commodities on the truck are destined for the same location. If a truck makes multiple deliveries on a route, the goods delivered at each stop are counted as one shipment. Interoffice memos, payroll checks, or business correspondence are not considered shipments. Shipments such as refuse, scrap paper, waste, or recyclable materials are not considered shipments unless the establishment is in the business of selling or providing these materials.

Geographic Coverage

Geographic strata were defined by a combination of the 50 states, the District of Columbia, and 65 metropolitan areas (MAs) based on their population and importance as transportation gateways. All other MAs were collapsed with the non-MAs within the state into Rest of State (ROS) strata. When an MA crossed state boundaries, size of each part of the MA was considered relative to the MAs total measure of size when determining whether or not to create strata in each state in which the MA was defined. Six MAs had strata in two or more states.

Relevance for Commodity Flow Database Development

The CFS can be used to develop state-level commodity flow databases. However, because the sub-state regions used in the CFS are a select group of metropolitan regions, it does not always match up well with the boundaries of MPOs or other sub-state regions that are conducting freight planning activities. Therefore, there is a significant amount of work that is required to develop a sub-state commodity flow database from the CFS.

General Types of Applications That the Database Can Be Used For

Analysts and researchers in both the public and private sectors use data from the CFS for a variety of purposes, including:

- Analyzing trends in goods movement over time;
- Conducting national, regional and sectoral economic analysis;
- Developing models and analytical tools for policy analyses, management and investment decisions;
- Forecasting future demand for goods movement and associated infrastructure and equipment needs;
- Establishing benchmarks for estimating national accounts; and
- Analyzing and mapping spatial patterns of commodity and vehicle flows.

CFS data are also used as the basis for the Federal Highway Administration's Freight Analysis Framework, a model that displays by mode the movement of goods over the national transportation network. In addition, the CFS Hazardous Materials report is the sole source of hazardous materials flow data available for the highway mode.

Issues Related to Definition of a Shipment

As described earlier, if a truck makes multiple deliveries on a route, the goods delivered at each stop are counted as one shipment. However, respondents with multi-stop tours are requested to report shipments at constant intervals depending on the number of shipments at the respondent's facility. This portion of the questionnaire is shown in Table 4.4. Additionally, respondents only record the city, state, and destination of the shipment as shown in Figure 4.1, so there is no note on whether shipments are part of multi-stop tours. While this does not impact the accuracy of the freight flows, it will impact the ability to convert the CFS database into truck trips.

The complexity of this data collection technique also calls into question whether or not facilities that make multi-stop tours are filling in this information correctly. It is unclear if facilities have the ability to sort shipments that are part of multi-stop tours into individual shipments in the manner requested by the survey form.

Table 4.4 process for Recording Shipments in CFS

Number of Outbound Shipments Reported	Report every...	Mark (X) One
1-40	Report every outbound shipment	
41-80	Report every 2nd outbound shipment	
81-100	Report every 3rd outbound shipment	
101-200	Report every 5th outbound shipment	
201-400	Report every 10th outbound shipment	
401-800	Report every 20th outbound shipment	
801-1600	Report every 40th outbound shipment	
1601-3200	Report every 80th outbound shipment	
3201-6400	Report every 160th outbound shipment	
6401-12800	Report every 320th outbound shipment	
More than 12800	Call Census	

Figure 4.1 Sample CFS Form for Capturing Shipment Information

U.S. Destination or U.S. Exit Port (Complete for all shipments.)										Mode(s) of transport to U.S. destination. <i>Enter all that apply in order used. Use codes at bottom.</i>		* Intermodal shipment? (Y/N) Export? (Y/N)		Foreign Destination (for export shipments only) Note: In column (I) enter the U.S. port, airport, or border crossing of exit.		Export mode Line No.	
(I)			(J)	(K)	(L)	(M)		(N)	(O)								
City	State	ZIP Code				City	Country										
Los Angeles	CA	90040	2, 4	Y	Y	Beijing	China	6	0								
Newark	NJ	07105	4	N	N				00								
									1								
									2								
									3								
									4								
									5								

Issues Related to Industry/Commodity Coverage

As noted during the development of the 2002 FAF database, the CFS has the following gaps relative to specific industries that are included in the survey:

- **Farm Based:** CFS omits shipments of farm commodities from the farm to the first point of sale, e.g. a grain elevator or a stockyard.
- **Fisheries:** CFS omits shipments of fish and seafood from the boat at the dock to the processor or from the fish farm to the processor.
- **Crude Petroleum:** Crude petroleum shipments are completely outside the scope of the 2002 CFS.
- **Natural Gas:** Natural gas shipments are completely outside the scope of the 2002 CFS.
- **Municipal Solid Waste (MSW):** MSW shipments are completely outside the scope of the 2002 CFS.
- **Logging:** CFS omits shipments of logs from the point of harvest to the initial point of processing.
- **Construction:** CFS does not cover shipments originating from the construction sector. The construction sector includes construction companies or establishments engaged in construction of residential and

nonresidential buildings; utility systems; highway, street and bridge construction; and specialty trade contractors.

- **Services:** The survey does not cover shipments originating from establishments involved in service industries. The missing services industries are finance and insurance; real estate, rental and leasing; professional, scientific and technical services; administrative and support, and waste management and remediation services; education services; health care and social assistance; arts, entertainment and recreation; accommodation and food services; other services (e.g., repair and maintenance, personal and laundry, religious, etc); and public administration. Also, the CFS does not include management of companies and enterprise services with the exception of corporate, subsidiary and regional managing offices.
- **Publishing:** The CFS data gap on the publishing industry is primarily due to the adoption of the North American Industry Classification System (NAICS) in the 2002 CFS for selection of business establishments. In the 1997 and 1993 CFS businesses were selected based on their descriptions in the Standard Industry Classification (SIC).
- **Retail:** CFS does not cover shipments originating from retail trade stores, including motor vehicle and parts dealers, furniture and home furnishings stores, electronics and appliance stores, building material and garden equipment and supplies dealers, food and beverage stores, health and personal care stores, gasoline stations, clothing and clothing accessories stores, sporting goods, book and music stores, general merchandise stores, florists, used merchandise, manufactured home dealers, etc.
- **Household and Business Moves:** CFS does not capture freight movements by carriers that transport household and business furniture, equipment, etc.

Additionally, petroleum products are technically within the scope of the CFS. However, previous research suggested that the 2002 CFS, like earlier editions, undercounted petroleum products.

In some cases, one or more shipments in a commodity's supply chain were absent from the CFS survey. In other cases, whole categories of shipments were omitted from the survey, such as the movement of retail commodities from the point of final purchase to the home, business, etc. In yet other cases, there was evidence that the 2002 CFS undercounted some commodities and types of shipments – based on significant differences with other reliable data sources.

Earlier research suggested that previous CFS surveys undercounted total U.S. freight by a significant amount. A study by ORNL completed in 2000 estimated that the 1997 CFS captured only 75 percent of total U.S. freight shipments measured in tons, 74 percent when measured in ton-miles, and 81 percent when measured in value. Research comparing the 2002 CFS to production totals

indicate that overall, the database has captured only 54 percent when measured in tons, 67 percent in ton-miles, and 63 percent in value.

Issues Related to Timing of Data Collection

Collecting data during four weeks in the year from each of the respondents goes a long way towards removing the seasonal bias that occurs in several surveys that are taken over short periods of time. It would be interesting to conduct seasonal analyses of CFS survey data to determine the extent of seasonal variations by origin-destination combinations, mode, and industry/commodity.

However, conducting the CFS survey once every five years along with the release of the data three years after the data collection does create a significant lag impacting the usefulness of the data. This impact will be especially pronounced with the most recent 2007 CFS data due to the severe economic recession that occurred since the survey data were collected. This recession has changed the structure of many sectors, and impacted shipper-receiver relationships such that caution must be used when applying CFS to account for current freight flows patterns.

Issues Related to Missing Data

Due to the multi-dimensional nature of the final CFS database and the relatively limited number of samples, there will be several origin-destination-commodity-mode combinations that will have zero values after the CFS survey data is expanded. This makes it difficult to estimate combinations that have small amounts of shipments. Instead these flows will be misallocated to other combinations that have a large amount of flows.

This zero cell issue will be compounded by the need to suppress data for combinations that are dominated by one or two shippers due to concerns about sharing proprietary information. This will cause an even greater number of cells to not show values that can be reported in the CFS.

Issues with Coverage of International Flows

There are several well documented issues with the CFS treatments for international flows. These can be summarized as follows

- **Imports:** Imports are completely outside the scope of the 2002 CFS. However, once import commodities enter the United States and change ownership, further shipments of those "imports" are captured within the CFS.
- **Exports:** The 2002 CFS collected data from U.S. business establishments located in the United States; thus the survey included exports from the United States by all freight modes. However, analysis of the 1993 and 1997 CFS export data suggests that the CFS underestimated U.S. export shipments.

- **In-transits:** The CFS does not include shipments of commodities that originate outside of the United States, enter the United States by whatever mode, and then are shipped to some other country. Such shipments are called In-transits.

4.1.2 Freight Analysis Framework (FAF)

Sources:

FHWA Website:

http://ops.fhwa.dot.gov/freight/freight_analysis/faf/faf2faf3thoughtsnov8.htm

The FHWA has published forecasts for FAF2 that extend to 2035, the complete database and documentation are available on-line for download at:

http://ops.fhwa.dot.gov/freight/freight_analysis/faf/index.htm

General Information

The FHWA Freight Analysis Framework (FAF) includes estimates of the weight and value of commodity movements by origin, destination, commodity, and mode for the most recent Economic Census year, 30-year forecasts, and a network database in which tons are converted to truck payloads and assigned to specific routes on the highway network. FAF maps and tables are featured in publications such as the annual Freight Facts and Figures, as well as in congressional testimony and in policy studies such as the biennial Status of the Nation's Highways, Bridges, and Transit: Conditions and Performance.

Description of Data

FAF2 covers all flows to, from, and within the U.S. for 114 CFS regions (major metropolitan areas and balances of states), 17 additional metropolitan areas that serve as major international gateways, and 7 international regions. The data includes:

- Shipment Origin;
- Shipment Destination;
- Shipment Weight in Tons;
- Shipment Value;
- Shipment Mode (truck, rail, water, air, pipeline, intermodal, others);
- Commodity classification (Two-digit Standard Classification of Transported Goods (SCTG) scheme); and
- Port of Entry information for International flows.

FAF2 Methodology

The FAF2 2002 base year database is built entirely from public data sources. The starting point for the 2002 FAF is CFS Table 17 (Shipment Characteristics by Destination State, Two-Digit Commodity and Mode of Transportation for State of Origin: 2002). This table was then disaggregated from the state level to the FAF regional level by dividing shipments equally across all FAF regions that comprise each state.

At this level of detail, the majority of cells are empty or suppressed, primarily due to disclosure rules and suppression due to a lack of statistical significance. Additionally, as mentioned in the CFS section, earlier research suggested that previous CFS surveys undercounted total U.S. freight by a significant amount. A study by ORNL completed in 2000 estimated that the 1997 CFS captured only 75 percent of total U.S. freight shipments measured in tons, 74 percent when measured in ton-miles, and 81 percent when measured in value. The 2002 CFS is estimated to have captured only 54 percent when measured in tons, 67 percent in ton-miles, and 63 percent in value.

For those cells within Table 17 that were marked as zeros, the CFS data were compared to Rail Waybill, Waterborne Commerce, and air freight data to verify that neither of those datasets contradicts the zero value. In cases of contradiction, i.e. where observations are found for cells previously marked as zero, the restriction on that cell to be zero was lifted.

Log-linear modeling was then used to estimate the value of cells that have no observed value in CFS. This procedure estimates the most likely value of those missing cells, based upon statistical relationships extracted from cells with known values. For example, although CFS info is not available for fertilizer shipments from Iowa to Memphis, CFS information is available on the total fertilizer shipments from Iowa to all other FAF regions and for all commodities shipments from Iowa to Memphis. By examining the statistical relationships at higher orders of aggregation, a maximum likelihood value can be estimated for each missing cell. Log-linear modeling was used to estimate these statistical relationships among FAF regions, modes, and commodities at 2, 3, and 4 dimensional levels.

Iterative proportional fitting (IPF) was then used to ensure that the matrix developed in Step 6 is consistent with totals at higher levels of aggregation. The IPF procedure produces new estimates for each cell in the table at the 2, 3, and 4 dimensional levels such that they are in agreement with the marginal constraints, and is done so in an iterative fashion. In a two dimensional case, the elements of each row of the table are prorated so that their totals equal the corresponding marginal; then the elements of each column are prorated so their totals equal their corresponding marginal. After this initial step, the estimates in the table no longer add across the rows to agree with the first marginal. The steps are repeated iteratively until the procedure converges to the unique solution that sums to the marginals while preserving the initial relationships between the

variables in the table. The product of this step is a complete Table 17 (four dimensional) in which the initial values from Table 17 are maintained for those known cells at the 1, 2, 3, and 4 dimensional levels, including the true-zero values from Step 4.

The next step is to add in the categories of out-of-scope CFS shipments. FAF out-of-scope flows were estimated initially at the national level based on economic production values. For example, to fill in the gap for out-of-scope farm shipments, 2002 Census of Agriculture and the 2004 Agricultural Statistics were used to estimate values. These values were converted to tons based on crop-specific conversion rates. These values were converted to average distances using VIUS average truck trip length. A summary of the steps utilized for each out-of-scope CFS shipment type is shown below in Table 3.5.

These national totals were subsequently disaggregated to the FAF regional level and integrated into 2002 FAF2 Commodity Origin-Destination Database. The national estimates are first allocated to the state level by using state-level VIUS freight truck mile data at the 2-digit SCTG level. The estimate for each state is then allocated to the counties within the state based on an appropriate measure of the economic activity that generates the freight activity. The county weight was typically based on the 2002 County Business Patterns. The total amount of truck activity added in through this out-of-scope estimation process was significant as shown in Table 4.5 below. The flows estimated from the CFS out-of-scope industries represented a large percentage of the under-representation in the CFS database.

Table 4.5 Variables for State and County Allocations for CFS Out-of-Scope Shipments

Out-of-Scope CFS Business Sector	National to State Allocations	Commodity Used for State Allocation	State to County Allocation (Origin)	State to County Allocation (Destination)
Construction	VIUS Industry Sector	All	CBP Sector Employment	CBP Sector Employment
Services	VIUS Industry Sector	All	CBP Sector Employment	CBP Sector Employment
Retail	VIUS Industry Sector	All	CBP Sector Employment	Population
Farm-Based	VIUS Commodity	Animals	Value in Farm Sales (USDA)	CBP Animal Slaughtering and Processing Employment
Farm-Based	VIUS Commodity	Cereal	Value in Farm Sales (USDA)	CBP Grain and Oil Seed Milling Employment
Farm-Based	VIUS Commodity	Other Agriculture	Value in Farm Sales (USDA)	CBP Food Mfg Employment
Logging	VIUS Commodity	Logs and Other Wood	Round Wood Production (NFS)	CBP Wood Products Employment
Printing	CBP Industry Employment	Printed materials	CBP Industry Employment	Population
Fisheries	CBP Industry Employment	Live Fish	CBP Industry Employment	CBP Seafood Products Employment

Table 4.6 Estimated Size of Out-of-Scope CFS Shipments

Out-of-Scope CFS Sector	Value (\$ millions)	Tons (thousands)	Ton-miles (millions)	Average shipping distance (miles)
Farm-Based	200,646	1,051,285	40,222	38.26
Fisheries	3,181	4,714	259	54.94
Logging	7,871	350,191	16,271	46.46
Construction	924,974	591,449	62,003	104.83
Services	284,601	277,413	30,500	109.94
Publishing	98,657	32,330	13,945	431.33
Retail	1,408,236	1,050,277	94,411	89.89
Household and Business Moves	12,739	21,204	5,563	262.36
Total	2,940,905	3,378,863	263,174	142

Source: "Out-of-Scope" reports developed for FAF by Macrosys and ORNL.

Relevance for Commodity Flow Database Development

Similar to the CFS, FAF can be used for state-level commodity flow databases. However, FAF is more useful for developing sub-state commodity flow databases than CFS, because it provides 114 sub-national regions that fully cover the U.S. This is more geographic detail than a state-level commodity flow database, but still significantly less than the over 3000 counties which are included in the Transearch database product. Additionally, FAF has taken many steps fill in the gaps of the CFS survey process. Therefore, it should be considered a more complete database than CFS. Many recent efforts have been conducted to subdivide the FAF data to county-level. These are discussed in detail in Chapter 4.

General Types of Applications That the Database Can Be Used For

FAF1 was initially intended to be a policy analysis tool. FAF2 has been used to estimate O-D patterns of flows through portions of the highway network, particularly for analyses of network disruptions. FAF2 data have been used in policy analyses such as the Highway Economic Requirements System and freight bottleneck studies, providing base conditions and trend forecasts without policy interventions.

FAF2 has also been used as the starting point for many disaggregation techniques as discussed in Chapter 4. Often the goal is to develop a county-level freight flow database. This is either transformed to a county-level truck trip table for statewide travel demand models or to a Traffic analysis zone (TAZ) level truck trip table for metropolitan level travel demand models.

FAF2 data can also be an input to cost allocation and vehicle size and weight analyses. Special tabulations of FAF2 are under development for California as a key input to air quality and greenhouse gas emission studies.

FHWA will soon be developing the third generation of FAF, FAF3. FAF3 will include several enhancements including outreach to power and novice users to improve the quality and ease of use of FAF products.

Issues Related to Geographic Coverage

FAF2 covers all flows to, from, and within the US for 114 CFS regions (major metropolitan areas and balances of states), 17 additional metropolitan areas that serve as major international gateways, and 7 international regions. FAF2 does not include through shipments (a.k.a. in-transits) from foreign origin to foreign destination via the US, and the origin-destination database does not provide flows among individual counties. FHWA is initiating work with Canada and Mexico to estimate in-transit commodity flows, and increases in the number of CFS regions in 2007 thanks to doubling of the CFS sample may eliminate the need for identifying additional international gateways. However, several state DOTs and MPOs request county-to-county flows, FAF2 estimates are limited to flows among multi-county regions.

FAF2 disaggregates region-to-region flows to the county level as an intermediate step in creating the network database, but the temporary file is not published because flows among individual county pairs have significant errors that are believed to offset one another once loaded onto the published network database. Increased statistical reliability from the greatly expanded CFS sample in 2007 reopens the question of whether a standard disaggregation method for creating county-to-county flows with a prescribed set of locally collected supplemental data should be developed as an extension of FAF3. It should also be noted that only truck trips longer than 50 miles are routed onto the FAF network, so there will be significant under-counting of trucks on the FAF network, particularly in urban settings.

Issues Related to Modal Coverage

FAF2 was expanded to include all modes (truck, rail, water, air, and pipeline), as well as two categories of intermodal (truck-rail and other). Tonnage by truck in the origin-destination database includes long-distance and local, but only tonnage moving at least 50 miles is converted into freight-hauling trucks in the network database.

"Intermodal" in the FAF is based on CFS definitions, which include shipments by postal and courier services and any shipment using more than one mode. This categorization of "intermodal" is much broader trailer-on-flatcar or containerized service, sometimes leading to confusion. Can better modal definitions be developed within the confines of the 2007 CFS?

Issues Related to Commodity Coverage

FAF2 is based on the Standard Classification of Transported Goods (SCTG) at the 2-digit level. The SCTG is used in the US and Canada, and is based on the Harmonized System that underlies trade statistics throughout the world. SCTG has some comparability issues with commodity classifications based on industry of origin, such as the Standard Transportation Commodity Classification (STCC) system used primarily by railroads; however, SCTG provides a critical link between transportation and trade data, and has more useful distinctions of commodities carried by all modes (compared to the STCC's emphasis on bulk goods that move by rail). The 2-digit level of the SCTG does not break out ethanol and other commodities that have grown in importance since mid 1990's, and it does not provide a direct way to classify flows between hazardous and non-hazardous cargo. Many FAF customers would prize an origin-destination matrix for hazardous cargo, especially by hazard class.

Issues Related to Network Coverage

In the FAF network database, commodities are routed over the entire National Highway System, the entire National Network designated for conventional combination trucks, and additional highway mileage connecting freight activity centers. The network database does not identify freight-hauling truck moves between places less than 50 miles apart, nor does it relate commodity flows among FAF regions to individual rail lines, waterways, or pipelines. Early plans for FAF2 to create probability-based assignments of traffic within FAF regions and between adjacent regions to handle local traffic were not realized. FAF2 used the FAF1 strategy of disaggregating flows to counties and selected sub-county generators such as major ports and assigning the flows to individual routes. FAF assignments were matched to truck volume estimates for individual highway segments from the Highway Performance Monitoring System (HPMS), and revealed several quality problems with HPMS data.

Issues Related to How Often Database is Updated

Because FHWA data is developed from CFS data, it suffers from the same problems as CFS in terms of the lag between data collection and data availability. This will be particularly acute for the FAF3 database that will be developed from the 2007 CFS data. The severe economic recession that occurred after the CFS data were collected will have a significant impact on the ability of the data to be used to estimate freight flows in today's world.

FHWA recently initiated a program to provide provisional annual estimates of the FAF origin-destination database. This is developed by adjusting the base year estimates of weight and value by origin, destination, and mode with economic and modal growth factors. Estimates by commodity type are not made. Each year's estimates supersede rather than supplement the previous year's estimates.

Issues Related to How a Shipment is Defined

FAF2 defined shipments in a similar fashion as CFS. Therefore, the problems with handling multi-stop tours will be the same in both databases. While these problems will not impact the accuracy of the freight flow databases, they will impact the truck trips that are derived from the FAF2 database. This will be particularly relevant for less than truckload (LTL) trucking firms that aggregate smaller shipments regionally and deliver these shipments to multiple destinations. It is likely that LTL firms have deliveries and pickups across multiple FAF regions, such that understanding the sequencing and arrangement of these shipments will significantly impact truck routing.

Issues Related to How Missing Data are Estimated

FAF2 developed a very rigorous technique to estimate the zero and suppressed cells in the 2002 CFS database. For the zero cells that are the result of CFS survey coverage, the use of log-linear methods seems appropriate to fill in these gaps.

However, for zero cells that are the result of data suppressed for proprietary concerns, this methodology could be problematic. This is because the reason that the data are suppressed is because the freight flows are uniquely dominated by a small set of shippers. Therefore, estimating these values by using distributions of other origin-destination-commodity combinations would not necessarily lead to a great fit. This could lead to potentially significant deviations of the FAF2 estimates from actual freight flows.

The process used for FAF2 to estimate industries that were considered out-of-scope CFS industries was to develop production statistics by industry and to disaggregate that national data to states using VIUS data, then allocate it to county using county-level economic activity data. While this is a logical estimation process, it does rely on correlations between these existing data sources and truck activity that have not yet been verified using actual field data. Therefore, there is the potential for errors to occur based on a lack of these relationships in actual truck activity. For example, if State A had more truck VMT per shipments of a particular product than State B, then this process would over-allocate this product to State A. Similarly, if County A had a lower cost per pound process of cultivating potatoes than County B, then this methodology would under-allocate potato shipments to County A. This top-down allocation methodology is also vulnerable to data weaknesses and errors that may be present in VIUS and county-level economic databases. These will be discussed later in this chapter. FAF did mention the need to verify the methodology

developed for out-of-scope shipments using small sample surveys for specific industries at the local level.

4.1.3 TRANSEARCH

General Information

TRANSEARCH is a privately-maintained nationwide database of freight traffic flows between U.S. county or zip code markets, with an overlay of flow across infrastructure. This database combines primary shipment data obtained from 22 of the nation's largest freight carriers with information from public sources (Federal, state, provincial agencies, trade and industry groups, and a sample of motor carriers) and is accompanied with 30-year forecasts consistent with Global Insight's macro forecasts. TRANSEARCH is compiled and produced annually.

Information on the database has been largely extracted based on the Global Insight Development of the Transearch Database Report prepared in 2008. This documentation describes different steps in the database development process at different levels of detail. However, the information is still useful for initiating a discussion of using Transearch as a sub-national commodity flow database for freight planning purposes.

Description of Data

The database includes information describing commodities (by Standard Transportation Commodity Classification (STCC) code), tonnage, origin and destination markets, and mode of transport. Forecasts of commodity flows for up to 25 years also are available.

The following variables are included in the TRANSEARCH database:

- Outbound, inbound, intra, and through shipments by geography
 - Geographies include 172 Business Economic Areas (BEAs)
 - Over 3,000 counties
- Routed volumes along individual trade lanes or corridors
- Tonnage, value, and units of shipments
- Seven major transportation modes, including truck, rail, waterborne, and air
- Detail for over 340 commodities
- Canada and Mexico cross-border flows

Methodology

Freight Production Estimates. The Transearch U.S. commodity flow database begins by establishing state-level production volumes by industry or commodity based on the Census Bureau's Annual Survey of Manufactures and the Census of

Manufactures. This information is updated to the base year using industrial production indices, and supplemented by trade association and industry reports. This data source reports production in dollars, which are then converted to tons using commodity value/weight relationships maintained by Global Insight. Once the production values are established, tonnages moving by rail, water, air and pipeline are netted from the totals using mode-specific data from other sources. The remaining freight volumes are allocated to the truck mode.

Shipments are localized to the level of counties using street-address employment and activity information, population data, industry reports, and proprietary traffic information from freight carriers. Relationships between industries are determined with input-output data. The chief sources of production and shipment estimates along with the modes they influence are shown in Table 4.1.

Table 4.7 Data Elements Used in Developing Local Production/Consumption Factors

Database	Modal Flows Estimated with Database
US Dept. of Commerce Census/Survey of Manufactures	Truck, water, air
GII Industrial Production Indices	Truck, water, air
Trade Association Production & Shipment Reports	Truck, water, air
US Geological Survey Mineral Industry Reports	Truck, water
GII/InfoUSA Street-Address Industrial Employment & Activity	Truck
County Population Data	Truck
Inter-Industry patterns (Input-Output Table)	Truck, Air
Motor Carrier Industry Financial & Operating Statistics	Truck
Railroad Industry Proprietary Traffic Factors	Truck
Private Port Directories	Water

Specifically, rail tonnages are estimated using the Surface Transportation Board annual Railroad Waybill Sample. US/Canada flows are largely included in this sample, while the US/Mexico flows are derived from the BTS border crossing statistics and from routings suggested in the Waybill, interpreted with a rail network routing model.

For waterborne traffic, Transearch utilizes various components of the US Army Corps of Engineers (ACE) Waterborne Commerce Data Series. This source provides state-to-state annual volumes of broad commodity groupings which are disaggregated to county level using originating and terminating volumes by port and more specific commodity type, which are also provided by ACE.

Air cargo data are constructed using BTS Airport Activity Statistics. BTS enplanement data report the total tonnage originating at each airport, while a separate data series, BTSS T-100, reports cover airport-to-airport flow volumes. The data are then translated to county based on airport location information. Adjustments are made to account for international traffic, and the use of hubs, such that the final Transearch database represents traffic from source airport market to consuming market with hub facilities not depicted. The BTS Commodity Flow Survey is used to provide a broad set of commodity types for air cargo, which is refined based on production and consumption patterns at the origins and destinations respectively.

The truck mode is sub-allocated to three sub-modes. For-hire and private trucking is allocated based on relative volumes reported in the CFS. The for-hire segment is split between Truckload and Less Than Truckload (LTL) components using industry data on the level of LTL shipments, and prior Transearch patterns.

Domestic Truck Trip Distributions. The Transearch database uses its proprietary Motor Carrier Data Exchange as the starting point and main driver for developing domestic truck trip distribution patterns. The Motor Carrier Data Exchange is based on primary shipment data obtained from 22 of the nation's largest freight carriers. Participating carriers are primarily large truckload and LTL operators with average lengths of haul over 500 miles. However, the sample also takes in owner-operator business, portions of private carriage and dray activity, and significant amounts of regional (under 500-mile) traffic.

Carriers that participate in the exchange program submit a summary of their annual traffic flows that include origin, destination, and volume. Commodity information is captured through SIC codes, carrier commodity codes, or equipment types. Traffic is reported at the 3-digit or 5-digit zip code level. The total number of shipments included in that database is roughly 70 million. The sampling rate for the exchange is reported to be 7 percent overall, 3 percent under 500 miles, and 1 percent under 100 miles. The data are not collected by stratified random sample, but they are obtained for a broad cross section of the trucking market, including diverse industrial and geographic segments.

Sample rates are then calculated by dividing the amount of traffic reported by the Data Exchange carriers by the amount of relevant truck traffic determined in the earlier modal processing. These sample rates are then used to determine the degree to which flow pattern development will rely on either the carrier-reported patterns or historical patterns, including those from the CFS. For longer-haul shipments and commodities that are moved in dry-van trailers, the Data

Exchange is more heavily relied upon. For shorter haul truck volumes, the Transearch database relies more heavily on CFS. In addition, certain kinds of non-manufactured goods transported by truck begin with this source. For example, CFS is used as the starting point for ores and non-metallic minerals.

Specialized Truck Flow Development. Secondary truck traffic is also incorporated into the truck component of the database. For warehouse and distribution center truck traffic, volumes are estimated based on three sources:

1. Transearch commodity shipments inbound to markets combined with input-output tables;
2. Locations of warehouse facilities compiled from street-address establishment data and information provided by the Public Warehouse Association; and
3. Portions of the Motor Carrier Exchange program.

Agriculture-related truck shipments are estimated based on data from the US Department of Agriculture (USDA). Fresh produce is modeled using production data and distribution patterns historically gathered by the USDA. Other types of agriculture are estimated using county-level production data by crop, product or livestock also obtained from USDA data. County consumption volumes are based on establishment level factors for relevant facilities (e.g. grain elevators). Distribution is estimated with a two-step national model that employs a gravity algorithm in its first stage, followed by iterative proportional fitting.

Coal volumes are estimated in Transearch using information available from the Department of Energy. Coverage of forests, fisheries, and haulage of waste and scrap is not covered in the standard Transearch data set, but are available on a custom basis.

For purposes of vehicle routing, Transearch also estimates empty trucks. This is done by first observing imbalances of inbound and outbound loads, by category of trailer on a nationwide basis. Then, empty vehicle mileage factors from VIUS are extracted, and checked against industry factors and market conditions.

Domestic Flows of International Movements. Export volumes are developed using the export component of the Census Bureau's Annual Survey of Manufactures and the Census of Manufactures. Import volumes are drawn from U.S. Department of Commerce data, and are subsequently combined into the traffic flows at the point of importation.

Export traffic is embedded in Transearch through the use of production statistics mentioned above. For shipments moved by road, production for export remains blended with domestic production in outbound truck volume. Truck movements of import volume are handled in Transearch as outbound flows from the seaport and are based on foreign trade data and inland distribution patterns originally created for LATTS, and adjusted for present-day import volumes and contemporary economic geography.

Issues Related to Coverage

For local truck shipments, Transearch relies much more heavily on the CFS than it does the Motor Carrier Exchange due to the fact that the exchange is more accurate for long haul shipments. However, as noted earlier, the CFS methodology excludes several out-of-scope industries. Transearch has developed a methodology to add in farm-based industries and the Transearch secondary shipments are designed to cover the retail industries. However, it is not clear that there was any coverage for the other out-of-scope industries including fisheries, logging, construction, services, publishing, and household and business services. Many of the trucking moves from these industries will be local, so it is unlikely that the Motor Carrier exchange covered them in sufficient fashion.

TRANSEARCH does not report international air shipments through the regional gateways. Additionally, specific origin and destination information is not available for overseas waterborne traffic through marine ports. Overseas ports are not identified and TRANSEARCH estimates the domestic distribution of maritime imports and exports. TRANSEARCH data also does not completely report international petroleum and oil imports through marine ports. Finally, TRANSEARCH assigns commodity data only to the truck, rail, air, and water modes, though a large percentage of foreign imports (by weight) consist of oil and petroleum products – commodities that are frequently shipped via pipeline to storage and distribution points.

Issues Related to Sampling and Data Collection

The primary data collection conducted for the Transearch database is the Motor Carrier Data Exchange. The developers of Transearch acknowledge that the sample of carriers is not a stratified random sample. The Transearch data attempts to balance this by ensuring that these carriers include data across a broad cross section of the trucking market, including diverse industrial and geographic segments. Without more information on specifically which trucking companies participated in the survey, it is difficult to determine the impact of this sample on the trucking information provided in Transearch.

Issues Related to How Often Data Are Updated

The Transearch database is updated annually. However, its reliance on CFS for portions of its database development leaves it vulnerable to the same time lag issues as the CFS itself. As mentioned earlier, this is particularly relevant given the current severe economic recession that occurred after the most recent CFS data collection effort.

Issues Related to How a Shipment is Defined

Intermodal shipment definition is different in Transearch than it is in the CFS or FAF. Transearch breaks up intermodal shipments into their modal components. For example, a truck-rail intermodal shipment will be recorded three times in

Transearch – two truck dray trips and one long-haul rail intermodal trip. The truck dray components are noted as a secondary shipment in Transearch, but it is not possible to link it with a specific origin-destination movement in the rail database. Similarly, air cargo and inland waterway intermodal traffic are broken into their component modal pieces. This makes it more difficult to identify end-to-end shipments in the Transearch database.

Issues Related to How Missing Data are Estimated or Modeled

Transearch relies rather heavily on estimating freight movements based on economic data. The developers of Transearch have access to a wealth of proprietary economic data through the larger IHS/Global Insight company that go above and beyond what is available in the public sector. However, there are several freight movements that do not involve an economic transaction such as inter-company truck moves. It is often difficult to estimate these moves based on economic data alone, and there was not stated process for validating these economic relationships in the Transearch literature. Similar to FAF2, it would be ideal to verify the relationship between the economy and freight flows using small survey samples. This is particularly true for short-distance truck movements.

General Types of Applications that the Database Can Be Used For

Transearch can be used as a planning tool for understanding freight flows to support the development of strategic marketing, operating, and/or investment plans. Transearch is also often used as the starting point to developing truck trip tables for travel demand models at both the state and regional level. The TRANSEARCH database also allows users to conduct trade flow analysis to identify primary trading partners at the commodity-specific level.

4.2 MODE-SPECIFIC FREIGHT FLOW DATABASES

4.2.1 Vehicle Inventory and Use Survey (VIUS)

Source

U.S. Census Bureau website:

URL: <http://www.census.gov/svsd/www/vius/products.html>

General Information

The Vehicle Inventory and Use Survey (VIUS) provides data on the physical and operational characteristics of the nation's private and commercial truck population. Its primary goal is to produce national and state-level estimates of the total number of trucks. The first survey was conducted in 1963. It was then conducted every five years beginning in 1967 and continuing to 2002. Prior to 1997, the survey was known as the Truck Inventory and Use Survey (TIUS).

VIUS has not been included in the budget for the 2007 Economic Census. VIUS has been discontinued and the 2002 VIUS is the last survey available.

Public use microdata files are available for years 1977 and later. Publications are available for all years. Visit <http://www.census.gov/svsd/www/vius/products.html> to access these files and publications.

Description of Data

The physical and operational characteristic data provided by VIUS include, but are not limited to, the following:

- Vehicle Make, Model, Year;
- Vehicle Body Type;
- Length of Vehicle;
- Percent of Miles Driven classified by commodity, and locations;
- Kind of Business;
- Base home state;
- Type of Cab;
- Type of Fuel;
- Hazmat Information;
- Number of cylinders;
- Axle Configuration; and
- Gross Vehicle Weight.

Methodology and Limitations

The Vehicle Inventory and Use Survey (VIUS) is a probability sample of all private and commercial trucks registered (or licensed) in the United States. The sample for each year is shown in Table 4.8.

Table 4.8 VIUS Sample Sizes by Truck Model Years

Year	Sample Size
2002	136,113
1997	131,083
1992	153,914
1987	135,290
1982	120,000
1977	116,400

1972	113,800
1967	~ 120,000
1963	~ 115,000

The VIUS excludes vehicles owned by Federal, state, or local governments; ambulances; buses; motor homes; farm tractors; and nonpowered trailer units. Additionally, trucks that were included in the sample but reported to have been sold, junked, or wrecked prior to the survey year (date varies) were deemed out of scope.

The sampling frame was stratified by geography and truck characteristics. The 50 states and the District of Columbia made up the 51 geographic strata. Body type and gross vehicle weight (GVW) determined the following five truck strata:

1. Pickups;
2. Minivans, other light vans, and sport utilities;
3. Light single-unit trucks (GVW 26,000 pounds or less);
4. Heavy single-unit trucks (GVW 26,001 pounds or more); and
5. Truck-tractors.

Therefore, the sampling frame was partitioned into 255 geographic-by-truck strata. Within each stratum, a simple random sample of truck registrations was selected without replacement. Older surveys were stratified differently: for the 1963-1977 TIUS the survey was stratified by “small trucks” and “large trucks.”

Relevance for Commodity Flow Database Development

VIUS data are often used to develop payload factors for converting tons to trucks to transform commodity flow databases into vehicle flow databases. However, there is also significant activity and commodity information in this database as well which makes it useful for allocating tonnage to different vehicle types, vehicles making different trip lengths, and vehicles carrying specific commodities. Additionally, it can be used to estimate truck VMT by commodity truck type, and truck trip length bin, along with estimating empty truck percentages. The VIUS database is a critical input for both the FAF and Transearch databases. Additionally, the database is often used to develop vehicle flow databases that can then be turned into truck trip tables for importing into a truck component of a travel demand model. With the discontinuation of VIUS, other sources will be needed to perform this function.

General Types of Applications that the Database Can Be Used For

VIUS data are of considerable value to government, business, academia, and the general public. Data on the number and types of vehicles and how they are used are important in studying the future growth of transportation and are needed in

calculating fees and cost allocations among highway users. The data also are important in evaluating safety risks to highway travelers and in assessing the energy efficiency and environmental impact of the nation's truck fleet. Businesses and others make use of these data in conducting market studies and evaluating market strategies; assessing the utility and cost of certain types of equipment; calculating the longevity of products; determining fuel demands; and linking to, and better utilizing, other datasets representing limited segments of the truck population.

4.2.2 Carload Rail Waybill Sample

Source

Surface Transportation Board website:

URL: http://www.stb.dot.gov/stb/industry/econ_waybill.html

General Information

The Carload Waybill Sample (Waybill Sample) is a stratified sample of carload waybills for all U.S. rail traffic submitted by those rail carriers terminating 4,500 or more revenue carloads annually. A waybill is a document issued by a carrier giving details and instructions relating to the shipment of a consignment of goods. Typically, it will show the names of the consignor and consignee, the point of origin of the consignment, its destination, route, and method of shipment, and the amount charged for carriage.

Because the Waybill Sample contains sensitive shipping and revenue information, access to this information is restricted. Access is granted when the Waybill Sample is the only single source of the data, obtaining the data from other sources is burdensome or costly, and/or the data are relevant to issues pending before the Board. There is also a group designated as "other users." The rules for release of waybill data are codified at 49 CFR 1244.9.

Because the Master Waybill File contains sensitive shipping and revenue information, access to this information is restricted to: railroads; Federal agencies; the states; transportation practitioners, consultants, and law firms with formal proceedings before the STB or State Boards; and certain other users. Rules governing access to Waybill Data are described in 49 CFR 1244.9.

Anyone can access the non-confidential data in the Public Use File by sending a written request to: OEEAA, Surface Transportation Board, 1925 K Street, NW, Washington, D.C. 20423-0001.

Description of Data

Data fields include:

- Names of Consignor and Consignee;
- Shipment Origin;
- Shipment Destination;
- Route;
- Method of Shipment; and
- Amount charged for the Carriage.

Methodology and Limitations

Railroads may file waybill sample information by using either: 1) authenticated copies of a sample of audited revenue waybills (the manual system); or 2) a computer-generated sample containing specified information (the computerized system or MRI). The waybill submissions from these two methods are combined in a 900 byte Master Record File containing a movement-specific Confidential Waybill File and a less detailed Public Use Waybill File. The content of waybill requests are described in 49 CFR 1244.9.

The Waybill Sample is a continuous sample that is released in yearly segments. For the past several years, the sample has contained information on approximately 600,000 movements. It includes waybill information from Class I, Class II, and some of the Class III railroads. The STB requires that these railroads submit waybill samples if, in any of the three preceding years, they terminated on their lines at least 4,500 revenue carloads. The Waybill Sample currently encompasses over 99 percent of all U.S. rail traffic.

Relevance for Commodity Flow Database Development

The Waybill data provides detailed commodity flow data for the rail mode. A summarized version of the Rail Waybill data is included in FAF2. It is also included in Transearch data purchases with the permission of the Surface Transportation Board. The detail Waybill sample can provide facility-specific commodity flow information. However, reporting of these data is not allowed due to proprietary concerns.

General Types of Applications that the Database Can Be Used For

Data from the Waybill Sample are used as input to many STB projects, analyses, and studies. Federal agencies (Department of Transportation, U.S. Department of Agriculture, etc.) use the Waybill Sample as part of their information base. The Waybill Sample also is used by states as a major source of information for developing state transportation plans. In addition, nongovernmental groups seek access to waybill sample data for such uses as market surveys, development

of verified statements in STB and state formal proceedings, forecast of rail equipment requirements, economic analysis and forecasts, academic research, etc.

4.2.3 U.S. Waterborne Commerce

Sources

U.S. Army Corps of Engineers (Waterborne Commerce Statistics Center) website:

<http://www.iwr.usace.army.mil/ndc/wcsc/wcscmiss.htm>

<http://www.iwr.usace.army.mil/ndc/data/datawcus.htm>

General Information

The primary function of the Waterborne Commerce Statistics Center, under the authority of the Rivers and Harbors Act of 1922, is to collect, process, distribute, and archive vessel trip and cargo data. These statistics are used to analyze the feasibility of new projects and to set priorities for new investment, and for the operation, rehabilitation and maintenance of existing projects. The Waterborne Commerce Statistics Center's standard publications, *Waterborne Commerce of the United States*, is issued in five parts (one to cover each coast and a national summary).

The United States Army Corps of Engineers (USACE) used to publish the Waterborne Databanks and Preliminary Waterborne Cargo Summary reports that contained foreign cargo summaries, including value and weight information by type of service on U.S. waterborne imports and exports. These statistics were based on the U.S. Bureau of the Census trade data matched to the U.S. Customs vessel entrances and clearances. Starting in 2006, the Corps discontinued the generation of the monthly and annual Waterborne Databanks and Preliminary Waterborne Cargo Summary reports. Historical Waterborne Databanks will still be available for purchase from the Maritime Administration (MarAd). The Corps will continue to produce and enhance its manifest-based data products (at their Waterborne Commerce Statistics Center) which serve as the official statistics on U.S. Foreign Waterborne Transportation Statistics. These products are available on the Corps web site at: <http://www.iwr.usace.army.mil/ndc/wcsc/wcscmiss.htm>

Description of Data

Domestic and foreign vessel trips and tonnages by commodity for ports and waterways are covered in the *Waterborne Commerce of the U.S., Parts 1-5*. Foreign waterborne commerce between the U.S. and foreign countries are summarized by U.S. port, foreign port, foreign country, commodity group, and tonnage. Data summaries include: origin to destination information of foreign and domestic waterborne cargo movements by region and state, and also

waterborne tonnage for principal ports and state and territories. This database covers deep sea marine activity, not inland waterways. Internal waterway tonnage indicators are updated monthly on the NDC website.

Under Federal law, vessel operating companies must report domestic waterborne commercial movements to the Corps. The types of vessels include: dry cargo ships and tankers, barges (loaded and empty), fishing vessels, towboats (with or without barges in tow), tugboats, crew boats and supply boats to offshore locations, and newly constructed vessels from the shipyards to the point of delivery. Vessels remaining idle during the monthly reporting period are also reported.

Movement data acquired by the Center is primarily for the use of the Corps and other government agencies; however, summary statistics, which do not disclose movements of individual companies, are also released to private companies and to the general public.

Methodology and Limitations

Domestic Commerce

Contiguous and noncontiguous states and territories constitute the geographical space upon which domestic commerce may be transported. This includes Hawaii, Alaska, the 48 contiguous states, Puerto Rico and the Virgin Islands, Guam, American Samoa, Wake Island, and the U.S. Trust Territories.

The waterborne traffic movements are reported to the Corps of Engineers by all vessel operators of record on ENG Forms 3925 and 3925b (or equivalent) approved by the Office of Management and Budget under the Paperwork Reduction Act (44 U.S.C. 3510(a)). The reports are generally submitted on the basis of individual vessel movements completed. For movements with cargo, the point of loading and the point of unloading of each individual commodity must be delineated. Cargo moved for the military agencies in commercial vessels is reported as ordinary commercial cargo; military cargo moved in Department of Defense vessels is not collected.

In summarizing the domestic commerce, certain movements are excluded: cargo carried on general ferries; coal and petroleum products loaded from shore facilities directly into bunkers of vessels for fuel; and insignificant amounts of government materials (less than 100 tons) moved on government-owned equipment in support of Corps projects.

Beginning in 1996, fish are excluded from internal and intraport domestic traffic. The fish landing data in Tables 4.3 and 5.3 are furnished by the National Marine Fisheries Service.

In tables containing domestic tonnage totals for the United States, the commodity “Waterway Improvement Materials” is not included. “Waterway Improvement Materials” tonnage is included in domestic ports, waterways, and waterway systems. This is the same procedure that has been used in years prior to 1990;

therefore, the tonnages for years 1990 and later are comparable to tonnages in years prior to 1990.

Foreign Commerce

Foreign commerce is waterborne import, export, and in-transit traffic between the United States, Puerto Rico and the Virgin Islands, and any foreign country. These statistics do not include traffic between any foreign country and the United States Territories and Possessions (American Samoa, Guam, North Mariana Islands, and U.S. outlying islands).

Beginning with the calendar year 2000 publication, foreign waterborne import, export, and in-transit cargo statistics are derived primarily from data purchased from the Port Import Export Reporting Service, a division of the Journal of Commerce and supplemented by data furnished to the Corps of Engineers by the U.S. Bureau of the Census and the U.S. Customs Service. Foreign cargo is matched to the vessel moves to improve geographic specificity.

The Republic of Panama is considered a foreign country. However, individual vessel movements with origin and destination at U.S. ports traveling via the Panama Canal are considered domestic traffic. Alaskan crude oil (origin at Valdez, Alaska) shipped via the Panama pipeline (west to east) and destined for gulf and east coast ports also is considered domestic commerce.

Import and export shipments for use of the United States Armed Forces abroad are not reported to the Waterborne Commerce Statistics Center (WCSC). Beginning with calendar year 1989, shipments under the military assistance program of the Department of Defense are included in the statistics under the appropriate commodity code. In prior years, these cargoes were given as commodity code 9999.

Relevance for Commodity Flow Database Development

The U.S. Waterborne Commerce data provides detailed commodity flow data for the water mode. A summarized version of this data is included in FAF2. It is also included in Transearch data purchases. The data provide facility-specific commodity flow information.

General Types of Applications That the Database Can Be Used For

These statistics are used to analyze the feasibility of new projects and to set priorities for new investment, and for the operation, rehabilitation and maintenance of existing projects.

4.2.3 U.S. Ports and Waterway Facility Database

Source

USACE NDC website:

<http://www.iwr.usace.army.mil/ndc/ports/ports.htm>

General Information

U.S. Army Corps of Engineers' (USACE) Ports and Waterways database contains a national inventory which delineates the Nation's principal coastal, Great Lakes and inland port and waterway terminal and transfer facilities. Their primary mission is to collect, review, analyze, prepare, edit, and publish data on the physical and intermodal characteristics of these ports and waterways in the United States. The data consists of complete listings of port area's waterfront facilities, including information on including detailed information on berthing accommodations, petroleum and bulk handling terminals, grain elevators, warehouses, cranes, transit sheds, marine repair plants, fleeting areas, floating equipment, and docking and storage facilities. As the only Federal agency that has the primary mission to collect, review, analyze, prepare, and publish data on the physical and intermodal characteristics of commercial port facilities, this database the "clearing-house" for all commercial port information in the Federal government.

Collection of data is performed on a rotational basis to ensure on-site accuracy at each facility. The information is used to analyze the use and improvement of existing terminals and the planning and development of new ones. Primary users are Federal, state and municipal agencies; port and waterway development authorities.

The data are available as a series of 56 reports which include general descriptive information of the area, details of each facility, summary tables, photographs and an aerial overview identifying each facility and the commodity or service it provides. Archived records of facilities are available from the 1920s.

Description of Data

The Port Series Reports describe the physical and intermodal (infrastructure) characteristics of the coastal, Great Lakes, and inland ports of the United States. Facility data include, but are not limited to:

- Location (latitude/longitude, mile, and bank);
- Operations (name, owner, operator, purpose, handling equipment, rates, and details of open and covered storage facilities);
- Type and dimension of construction (length of berth space for vessels and/or barges, depth, apron width, deck elevation, and details of rail and highway access); and
- Utilities available (water, electricity, and fire protection).

In addition, the published reports include extensive descriptive material such as project authorization, bridge/tunnel/railroad infrastructure data, meteorological information, and anchorage descriptions. All viable commodity handling and

maritime service wharves are included, as are aerial maps and photos of individual facilities.

Methodology and Limitations

Data is collected at the pier, dock, and wharf site by professional engineers on a rotational basis to ensure on-site accuracy at each facility. The time span of on-site data source collection varies from 1 to 10 years. Other data from facility operators, port organizations, transportation companies, Waterborne Commerce Statistics Center, and other Federal, state and local agencies are collected, analyzed, and updated on an ongoing basis. Once collected, the data is analyzed and the PWS is then updated. The data is extracted for the 56 Port Series Reports published at intervals of approximately 10 years, covering approximately 200 individual port areas. Summary and historical data are frequently made available upon request. The data may be extracted from the PWS and processed through locally developed programs as input to special studies.

Relevance for Commodity Flow Database Development

The detailed facility-level information provided in this database can be used to supplement other marine sources to determine the allocation of commodities to facilities in wider commodity flow databases.

General Types of Applications That the Database Can Be Used For

The information is used to analyze the use and improvement of existing terminals and the planning and development of new ones. Primary users are Federal, state and municipal agencies; port and waterway development authorities. The Federal governments participation is to encourage commerce, facilitate competition and hence contribute to the public's well-being

Users of PWS data include, but not limited to:

- Corps of Engineers;
- OCE: Mobilization, Emergency Management, Navigation, Dredging, and Planning. WRSC: NDC, WCSC, and IWR;
- Divisions/Districts Regulatory Functions, Port Mobilization: Plans, Project Management, Reservoir Regulation, and Emergency Management;
- DoD Agencies: MTMC, DCA, Naval Control of Shipping, Naval Sealift Command, ONI, TRANSCOM, GCAC, DISA; and
- Dept of Transportation, State/Local Government, Private Industry: MARAD, BTS and USCG, Planners and regulatory functions, Pilots, Ocean Shipping Agencies, Port Authorities, and Commodity Shippers.

4.3 MULTIMODAL INTERNATIONAL GATEWAY FREIGHT FLOW DATABASES

4.3.1 BTS Transborder Surface Data

Source

BTS website:

<http://www.bts.gov/programs/international/transborder/>

General Information

The North American Transborder Freight Database, available since April 1993, contains freight flow data by commodity type and by mode of transportation (rail, truck, pipeline, air, vessel, and other) for U.S. exports to and imports from Canada and Mexico. The database includes two sets of tables; one is commodity-based while the other provides geographic detail. The purpose of the database is to provide transportation information on North American trade flows.

Description of Data

- Aggregate shipping charges on imports in U.S. dollars.
- Commodity type (two-digit commodity code: Schedule B or HTSUSA code).
- Container Code Distinguishes whether the merchandise is containerized
“1” = containerized shipment.
- Country of Origin or Destination.
- District and Port of Entry
- Domestic/Foreign Code (distinguishes whether the Code merchandise was produced in the U.S.) Disaggregated Method of Transportation (MOT) (vessel, air, mail, truck, rail, pipeline, other, unknown).
- Freight Costs.
- Shipping Weight.
- Statistical Month.
- Statistical Year.
- Direction of Trade (export or import).
- Value of Shipments.

BTS provides access to the data through an interactive searchable interface called North American Transborder Web. This allows users to create multivariable cross-tabulations on port, geography and commodity for all modes of transportation. Search results can be viewed on-line and then downloaded.

Additionally, the monthly and annual North American Transborder Freight Data can be downloaded in raw table formats. Users with a need to customize and manipulate these statistics for various purposes may choose to download these files instead of using the interactive searchable interface.

Methodology and Limitations

The North American Transborder Freight Dataset is extracted from the Census Foreign Trade Statistics Program. Import and export data are captured from administrative records required by the Departments of Commerce and Treasury. Historically, these data were obtained from import and export paper documents that the U.S. Customs Service (Customs) collected at a port of entry or exit. However, an increasing amount of import and export statistical information is now being captured electronically.

Imports

For imports from Canada and Mexico, over 96 percent of entries are collected electronically. U.S. imports of merchandise is compiled primarily from automated data submitted through the U.S. Customs' Automated Commercial System. Data are compiled also from import entry summary forms, warehouse withdrawal forms and Foreign Trade Zone documents as required by law to be filed with the U.S. Customs and Border Protection. Data on imports of electricity and natural gas from Canada are obtained from Canadian sources.

Exports

U.S exports of merchandise are compiled primarily from the Automated Export System (AES), paper Shipper's Export Declarations (SEDs), and Canadian data provided by Statistics Canada. The United States is substituting Canadian import statistics for U.S. exports to Canada in accordance with a 1987 Memorandum of Understanding signed by the Census Bureau, U.S. Customs and Border Protection, Canadian Customs, and Statistics Canada. This data exchange includes only U.S. exports destined for Canada and does not include shipments destined for third countries by routes passing through Canada.

Reliability

Import and export data are a complete enumeration of documents collected by U.S. Customs and Border Protection and are not subject to sampling errors. However, while quality assurance procedures are performed at every stage of collection, processing, and tabulation, the data are still subject to several types of non-sampling errors. The most significant of these include reporting errors, undocumented shipments, timeliness, data capture errors, transiting goods, and underestimation of low-valued transactions.

The North American Transborder Freight Dataset is the best publicly available approximation for analyzing North American Transborder transportation flows. However, as was noted in previous sections, the North American Transborder Freight Data are a subset of these statistics. Users should be aware that trade

data fields (such as value, commodity classification) are typically more rigorously reviewed than transportation data fields (i.e., mode of transportation and port of entry/exit). Users should also be aware that the use of foreign trade data to describe physical transportation flows might not be direct. For example, this dataset provides surface transportation information for individual Customs districts and ports on the northern and southern borders. However, because of filing procedures for trade documents, these ports may or may not reflect where goods physically crossed the border. This is because the filer of information may choose to file trade documents at one port while shipments actually enter or exit at another port.

Users should also note that the North American Transborder Freight Dataset represents Census' first attempt to disaggregate the various surface modes of transportation in U.S. foreign trade statistics. Since the dataset was first made available in April, it has gone through several refinements and improvements. When improbabilities and inconsistencies were found in the dataset, extensive analytical reviews were conducted, and improvements were made to the dataset based on these reviews. Therefore, the overall reliability of the dataset is generally very good. However, accuracy does vary by direction of trade and individual data field. For example, import data are generally more accurate than export data. This is primarily due to the fact that the Customs uses import documents for enforcement purposes while it performs no similar function for exports.

Relevance for Commodity Flow Database Development

BTS Transborder data is useful for estimating U.S.-Mexico and U.S.-Canada cross-border flows at the state level. It is often necessary to disaggregate this data to estimate cross-border flows at the facility-specific level.

The Transborder data can also be used to compare with the international flow data in other commodity flow databases such as FAF, CFS and Transearch.

General Types of Applications That the Database Can Be Used For

This type of information is being used to monitor freight flows and changes to these since the signing of the North American Free Trade Agreement (NAFTA) by the United States, Canada and Mexico in December 1992 and its entry into force on January 1, 1994. The database is also being used for trade corridor studies, transportation infrastructure planning, marketing and logistics plans and other purposes. It allows users to analyze movement of merchandise by all land modes, waterborne vessels, and by air carriers.

4.3.2 Port Import Export Reporting Service (PIERS)

Source

PIERS website:

<http://www.piers.com/>

General Information

PIERS was developed 30 years ago by the Journal of Commerce as the Port Import Export Reporting Service. Today, PIERS is a source of U.S. waterborne trade information. PIERS is commonly referred to for container trade volume estimates used by the maritime industry to plan capacity and set rates and by government agencies to track cargoes and traffic. PIERS trade data is also relied upon by over 6,000 private- and public-sector clients in over 40 countries as the source for actionable intelligence.

PIERS maintains a database of import and export information on the cargoes moving through ports in the U.S., Latin America, and Asia. PIERS collects data from over 42,000 bills of lading every day.

Description of Data

PIERS data supplies a detail on nearly 100 individual fields for imports and exports, including:

- Shipper;
- Consignee;
- Detailed commodity description;
- Container number;
- Twenty-Foot Equivalent Units (TEUs);
- Vessel name;
- Port origin, destination and transshipment; and
- English and local language data.

Methodology and Limitations

PIERS has access to import/export data due to the Freedom of Information Act along with U.S. Customs Regulations which authorize press organizations to copy certain shipping documents available to the public.

U.S. Import Products: Import information is obtained from vessel manifests and U.S. Customs Automated Manifest Systems (AMS) data tapes from all U.S. ports. Reporters throughout the country (including Alaska, Hawaii and Puerto Rico) collect import information from all U.S. ports. PIERS information is verified by the PIERS quality control staff, who audits the records every month against a list, supplied by U.S. Customs, of vessels arriving at U.S. ports. In the event of a discrepancy, PIERS requests the appropriate documentation from either U.S. Customs or the ship line. In addition, ship lines and importers that subscribe to

PIERS verify their own shipments in the system and notify them of any discrepancies.

U.S. Export Products: Reporters throughout the country (including Alaska, Hawaii and Puerto Rico) gather export information from bills of lading at all U.S. ports. The ship lines are obliged by law to submit documentation for every item on board the vessel. The PIERS quality control staff verifies PIERS export information against a list of vessels, supplied by U.S. Customs, exiting U.S. ports. Ship lines and exporters that subscribe to PIERS verify their own shipments and notify us of any disagreements among the shipment records.

Global Imports/Exports: PIERS obtains Brazilian, Chilean, Colombian, Ecuadorian, Peruvian and Venezuelan import and export data by forming partnerships with companies that specialize in manifest collection. They enhance this data to meet their conventions through a rigorous quality control process that includes codification and standardization.

One limitation of using PIERS data is that the data include information on the companies that are shipping the goods rather than specific origins and destinations of goods being shipped. Origins and destinations are generally derived based on company address information that is also provide in the shipping documents. This leads to two specific potential problems. The first problem is that the company shipping the good is not always the shipper or receiver of the good. In some cases the company that handles the shipping is a third-party logistics provider and none of its facilities are either an inland origin or destination. The second issue is that even if the shipping documents do properly describe the shipper or receiver of the goods, the address shown on the shipping document may be a national headquarter or a regional administrative office rather than the actual inland location of the shipment. This issue can often be corrected through discussions with port operators that are familiar with supply chain operations of specific companies, but this can be a tedious process depending on the number of shipments that need to be verified.

Relevance for Commodity Flow Database Development

PIERS data can be used to provide detailed commodity-specific and origin-destination specific data at the facility level. This makes it useful for disaggregating marine data that is provided at more aggregate levels such as in CFS or FAF.

FHWA is currently determining the effectiveness of using PIERS data to develop a component of the inbound inland flow data for the next update to the FAF database.

General Types of Applications That the Database Can Be Used For

PIERS data on U.S. waterborne trade is used by Federal, state and local government agencies – including the U.S. Departments of Agriculture, Commerce, Defense, Homeland Security, Justice and Transportation – to set

policy and conduct operations. PIERS provides cargo intelligence solutions for Government agencies to accomplish their missions of risk assessment, enforcement, policy-making, and program administration.

4.4 ECONOMIC AND ESTABLISHMENT DATA

4.4.1 U.S. Census Bureau

Source

Annual Survey of Manufactures:

<http://www.census.gov/manufacturing/asm/index.html>

County, Metro or Zip Code Business Patterns:

<http://www.census.gov/econ/cbp/index.html>

U.S. Economic Census Information:

<http://www.census.gov/econ/census07/>

Date: 2007

Contact: General Census Bureau Inquiries

Phone: 301-763-2547

E-mail: econ@census.gov

Agency: The U.S. Census Bureau collects and compiles a large amount of national data that is available to the general public. A significant amount of this data is relevant for freight and commodity flow analysis at the subnational, state, and national level. Two of these databases, the Commodity Flow Survey (CFS) and the Vehicle Inventory and Use Survey (VIUS) are explained elsewhere in this report and will not be addressed in this section. Instead three other data sources will be described and evaluated for addressing commodity flow information needs at the subnational level. These include the Annual Survey of Manufactures and the broader U.S. Economic Census information (of which ASM is a part).

Description of Data

Annual Survey of Manufactures (ASM):

This annual survey of manufactures in the U.S. provides a very broad level (national) understanding of manufacturing activity. This information includes statistics on employment, payroll, supplemental labor costs, cost of materials consumed, operating expenses, value of shipments, value added by manufacturing, detailed capital expenditure, fuels and electric energy used and inventories. There are basically three separate databases available from the

Annual Survey of Manufactures including, 1) Statistics for Industry Groups and Industries, 2) Value of Product Shipments and 3) Geographic Area Statistics which are available at the state-level.

U.S. Economic Census Information:

In addition to the data and information described above, the U.S. Economic Census is conducted every five years with the aim of providing detailed information regarding the United State's economy, available at various geographic detail (national, state, local). This information is utilized by a variety of government agencies, business and academic researchers to assess economic performance, business market analysis, evaluate industry growth and understand business competitiveness across different industries. There are a wide variety of different data series available from the economic census, including:

- Core Business Statistics Series – Primarily data organized by two- and three-digit NAICS for the U.S.
- Industry Series – Data organized by NAICS and product at the U.S. level.
- Geographic Area Series – Industry information at the U.S., state, county, place and metro level.
- Subject and Summary Series – Summarized information on product lines, concentration ratios and firm size for U.S. and state level.
- Zip Code Statistics Series – Number of establishments by sales size and industry at zip code level.
- Economy Wide Key Statistics – Four key statistics for all industries and geographic area. These statistics are 1) number of establishments, 2) value of sales, shipments, receipts, revenue or business done, 3) annual payroll and 4) number of employees.

Methodology

Annual Survey of Manufactures (ASM):

The information collected for the Annual Survey of Manufactures is the result of a survey of approximately 50,000 manufacturing establishments. The sample list is updated every five years and as of 2002 represented about 14 percent of the 346,000 active manufacturing establishments in the U.S. The survey is conducted annually, except for years ending in 2 and 7 for which the manufacturing sector information is part of the economic census. There are about 10,000 establishments that are repeatedly selected with certainty as part of the sample selection while the other 40,000 are selected with a probability proportional to their establishment size.

County, Metro or Zip Code Business Patterns:

The national Business Register is the primary source for all the county, metro and zip code business pattern data. This database contains a record of each known single-unit business establishment that has employees and is located in the U.S. or Puerto Rico. Single-unit businesses are identified using their Employer Identification Number (EIN) issued by the Internal Revenue Service (IRS) for reporting of payroll taxes. Multi-unit businesses and firms may contain multiple EINs representing all the different establishments associated with the multi-unit business. Thus there is a one-to-many relationship between each EIN and all the associated establishments for multi-unit business. The data and information for single-unit establishments originates mostly from administrative records and some census survey sources. Multi-unit business information primarily originates from the Census Bureau's Economic Census conducted every five years and the annual Company Organization Survey (COS). Any business that does not have an EIN or any employees is excluded from the county, metro and zip code business pattern data.

U.S. Economic Census Information:

The methodology for capturing and compiling economic data through the economic census has changed periodically since its beginning. Most of information is compiled from the sample survey of all employer establishments in the U.S., but information on some sectors are comprised from sample surveys to a subset (sample) of the population as is the case with the construction sector. Prior to 1967, the time interval for the economic census varied but since has been implemented every 5 years. The range of industries covered has broadened over time, especially beginning in 1992 with the inclusion of transportation, communications and utilities, financial, insurance and real estate industries. In addition, the classification system changed in 1997 from the earlier Standard Industrial Classification (SIC) to the current North American Industry Classification System (NAICS). As with all census data, the information is subject to non sampling errors and in some situations sampling errors as well.

Relevance for Commodity Flow Database Development

Employment data is the most commonly used source for disaggregating commodity flow data from larger geographies to smaller geographies. Census data in particular can be used to disaggregate data from state or regional level to county or zip code level.

Employment data and value of sales data can also be used to provide a level of validity to commodity flow data. Commodity movements should match with corresponding employment in economic sectors that match the inputs and outputs of the commodity movements.

General Types of Applications That the Database Can Be Used For

These data can also be used to understand the relative sizes of business by sector and by geography. Similar to other data on economic and industry performance, these data help provide greater context and knowledge regarding the key economic drivers of a given region or subset of industries and may be most helpful when developing broad business and economic policy at the national, state and local level. These data can also be used to provide control totals by industry and geography for freight flows in and out of regions.

4.4.2 County Business Patterns

Source

U.S. Census Bureau website:

<http://www.census.gov/econ/cbp/intro.htm>

General Information

County Business Patterns is an annual series that provides subnational economic data by industry. It covers most of the country's economic activity. The series excludes data on self-employed individuals, employees of private households, railroad employees, agricultural production employees, and most government employees.

This series has been published annually since 1964 and at irregular intervals dating back to 1946. The comparability of data over time may be affected by definitional changes in establishments, activity status, and industrial classifications.

County Business Patterns is useful for studying the economic activity of small areas; analyzing economic changes over time; and as a benchmark for statistical series, surveys, and databases between economic censuses. Businesses use the data for analyzing market potential, measuring the effectiveness of sales and advertising programs, setting sales quotas, and developing budgets. Government agencies use the data for administration and planning.

Description of Data

County Business Patterns reports are available approximately 18 months after each reference year. Data are available for each state, county, metropolitan area, the District of Columbia, and the Commonwealth of Puerto Rico (including each municipio), plus a U.S. summary. Data include number of establishments by employment size class and six-digit NAICS, quarterly and annual payroll, and employment during the week of March 12. In addition, Zip Code Business Patterns data are available shortly after the release of CBP. The Zip Code Business Patterns include the number of establishments by NAICS industry.

All CBP data are released in hypertext tables (html), in comma-delimited format for spreadsheet or database use, and in the American FactFinder interactive system.

To avoid disclosure of information about individual employers, all data products reflect the infusion of noise (new in 2007) as well as the suppression of values in certain data cells.

Data Reliability

Payroll and employment data are tabulated from administrative records for single-unit firms and a combination of administrative records and survey-collected data for multi-unit firms. They are not subject to sampling error, but are subject to nonsampling errors, which can be attributed to several sources: inability to identify all cases that should be in the set; definition and classification difficulties; errors in recording or coding the data obtained; and other errors of coverage, processing, and estimation for missing or misreported data.

Methodology and Limitations

CBP data are extracted from the Business Register, the Census Bureau's file of all known single and multi-establishment companies. The Annual Company Organization Survey and quinquennial Economic Censuses provide individual establishment data for multilocation firms. Data for single-location firms are obtained from various programs conducted by the Census Bureau, such as the Economic Censuses, the Annual Survey of Manufactures, and Current Business Surveys, as well as from administrative records of the Internal Revenue Service (IRS), the Social Security Administration (SSA), and the Bureau of Labor Statistics (BLS).

The Census Bureau does not create a multi-unit company structure in the Business Register for very small employers (less than 10 employees) identified in the Economic Census. In addition, the COS is an annual mail survey that includes all multi-unit companies with 250 or more employees. Companies with less than 250 employees are only selected for the COS when administrative record sources indicate the company may be undergoing organizational change and are adding or dropping establishments. Establishments for smaller companies may be missed, as well as establishments for companies not responding to the Economic Census or the COS. The Census Bureau takes much effort to get establishment information for large companies because of their importance to the economy.

As mentioned above CBP data do not include information on the self-employed, employees of private households, railroad employees, agricultural production employees and most government employees. Therefore, freight that is moved by these sectors will be missed if this is the basis for generating potential survey respondents. That will be most important for the agricultural sector where goods movement is a large part of the business.

Relevance for Commodity Flow Database Development

County Business Patterns is actually a specific type of Census data. Therefore, it can be used similarly to other Census data for developing sub-national commodity flows. CBP employment data can be used to disaggregate commodity flow data from the state or regional level to the county level.

Employment data can also be used to provide a level of validity to commodity flow data. Commodity movements should match with corresponding employment in economic sectors that match the inputs and outputs of the commodity movements.

General Types of Applications That the Database Can Be Used For

The data are the standard reference source for small area (county) economic data. They are used to benchmark public and private sector business statistical series, surveys and databases between economic census years. Federal agencies also use the data to determine employee concentrations and trends by industry.

Private businesses use the data for market research, strategic business planning, and managing sales territories. State and local government budget, economic development, and planning offices use the data to assess business changes, develop fiscal policies, and plan future policies and programs.

4.4.3 Woods & Poole County Economic and Demographic Projections

General Information

County Economic and Demographic Projections published by Woods & Poole Economics, Inc.

Home page: www.woodsandpoole.com

Products: www.woodsandpoole.com/purchasing/cart.cgi

Samples: www.woodsandpoole.com/pdfs/SP10.pdf (“State Profile”)
www.woodsandpoole.com/pdfs/MSA10.pdf (“MSA Profile”)
www.woodsandpoole.com/pdfs/DP10.pdf (“Data Pamphlet”)

Description of Data

Type of Data: Economic and demographic projections (see details below)

Modes Covered: NAICS

Geographic Detail: States, regions, counties, and Metropolitan Statistical Areas (MSAs) in U.S.

Accessibility: Proprietary – available for purchase (see details below)

Frequency of Data Release: Annual

Data Detail: Historical data from 1970 and projections through 2040 for 900 variables for all states, regions, counties, and MSAs

Purchase Detail: Available in print or CD (spreadsheet files), variously priced according to product (*Complete Economic and Demographic Data Source*, State Profiles [includes counties and MSAs in a state], MSA Profiles, Data Pamphlets [one county or MSA], et al.)

Methodology

The Woods & Poole methodology is proprietary. However, their marketing material mentions that county population growth is a function of both projected natural increase and migration due to economic conditions. Although not an indicator of future accuracy, the average absolute percent error for Woods & Poole's 10-year total population projections has been ± 8.4 percent for counties, ± 5.7 percent for MSAs, and ± 4.2 percent for states."

Relevance for Commodity Flow Database Development

Woods & Poole is another source of employment data similar to the Census data. It is provided at the county-level. Therefore, its relevance for commodity flow database development is similar to the Census County Business Patterns employment data. It can be used to disaggregate commodity flow data from the state or regional level to the county level.

One advantage of Woods & Poole data is that forecast data are also available. This allows for disaggregation of future year commodity flows from state and regional levels to the county level.

Data Elements

Population data by age, sex, and race; employment and earnings by major industry; personal income by source; household data by size and income bracket; retail sales by kind of business

4.4.4 InfoUSA Establishment Data

Source

InfoUSA website:

<http://www.infousagov.com/index.asp>

General Information

InfoUSA is a provider of sales and marketing support for products for all types of businesses, from small Mom and Pop shops to large corporations. The company compiles databases of 14 million U.S. businesses and 200 million U.S. consumers, and 1.2 million Canadian businesses and 12 million Canadian

consumers under one roof in Omaha, Nebraska. InfoUSAGov is a division of InfoUSA is dedicated to supplying government officials with the most accurate data available for investigative purposes, research, planning, education, and recruitment efforts.

Description of Data

Data elements include the following:

- Business Name;
- Full Address;
- Telephone Number;
- Fax Number;
- Geo Code (Longitude and Latitude);
- Type of Business (Yellow Page Title/SIC Code/NAICS Code);
- Name and Title of Key Executives;
- Gender of Key Executives;
- Headquarters, Branch, and Subsidiary Information;
- Credit Rating Codes;
- Estimated Annual Sales Volume;
- Zip Code, including ZIP+4;
- Carrier Route Codes ;
- Franchise and Brand Information ;
- Professional Specialties;
- Year Established;
- Size of Yellow Page Ad;
- Internet Web Site Addresses (URL's);
- Stock Exchange and Ticker Symbol;
- News Items;
- Number of Employees; and
- Lawsuits, Judgments, Bankruptcies.

Methodology and Limitations

Data is gathered from multiple sources and the information is verified by telephone to ensure accuracy. Nearly every single business is included, no matter how small, how large, or how newly established.

InfoUSA, Inc's information is continuously updated from new directories, and is also telephone-verified. It is possible to find a small percentage of errors and "out of business" names. It is not uncommon to have 5-10 percent undeliverable names, especially in industries with high levels of turnover.

Sources:

- 5,200 Yellow Page and Business White Page Directories.
- 20 Million phone calls to verify information. Every business is called one to four times a year.
- County Courthouse and Secretary of State Data.
- Leading business magazines and newspapers.
- Annual Reports.
- 10Ks and other SEC filings.
- New business registration and incorporations.
- Postal service information including National Change of Address, ZIP+4 carrier route and Delivery Sequence Files.

Business information can be identified and categorized by:

- Location: Zip Code, Neighborhood, City, Metro Area, County, Area Code, State.
- Type of Business: Yellow Page Heading, Major Industry Group, SIC Code or Professionals (doctors, dentists, etc.).
- Business Size: Number of Employees, Sales Volume.
- Credit Rating.
- Location Type: Corporate Headquarters, Headquarters of a Subsidiary, Branch.
- Phone and Fax Numbers.
- Key Decision-makers/Executive Names.

Relevance for Commodity Flow Database Development

Establishment data is also often used for disaggregating commodity flows from national, regional or state level to smaller geographies. Establishments can be aggregated for specific economic sectors based on the geography of interest. Then, commodity flows that are directly related to these economic sectors can be grouped and then allocated to the desired geography based on their proportion of economic activity in each geographic unit.

InfoUSA data allows for this disaggregation to occur based on either the number of employ. The geographic specificity is limitless, because specific addresses are provided for each establishment.

General Types of Applications That the Database Can Be Used For

Small businesses, small-office/home-office operators, and salespeople, use InfoUSA products and services to find new customers, increase revenue, cut selling costs, and increase their profits. The large corporations use the database for direct marketing, database marketing, analyzing their customers and making their sales and marketing more productive and efficient.

InfoUSA has recently supplied databases to:

- The U.S. Bureau of the Census who selected both the business and residential databases to serve Census 2000; and
- The U.S. Department of Labor, and multiple state labor departments, who selected the business database to serve as the national Employer Database to assist workforce development and welfare-to-work programs that help the unemployed find jobs.

4.4.5 Dun & Bradstreet (D&B) Establishment Data

Source

Dun & Bradstreet Website: www.dnb.com

General Information

D&B is a source of commercial information and insight on businesses across the world. D&B's global commercial database contains more than 140 million business records.

Description of Data

D&B has up to 1,500 data elements on any given business, depending on the information available for that company. These data elements range from the obvious company name and address to not-so-obvious industry profile, NAICS codes, SIC codes, financial statements, small business indicator, commercial credit score, financial stress score, geocode, domestic and global linkage, payment experiences, and any special news and events.

Methodology and Limitations

The steps for constructing this database are as follows:

- Global Data Collection brings together data from a variety of sources worldwide.
- D&B integrate the data into the database through their patented Entity Matching, which produces a single, more accurate picture of each business.
- D&B applies the D-U-N-S® Number as a unique means of identifying and tracking a business globally through every step in the life and activity of the business.
- D&B uses Corporate Linkage to enable customers to view their total risk or opportunity across related businesses.
- Finally, their Predictive Indicators use statistical analysis to rate a business' past performance and to indicate how likely the business is to perform that same way in the future.

Both InfoUSA and D&B data are establishment information databases. They are primarily used by parties interested in developing lists of companies to include in marketing campaigns. These private companies are therefore interested in information about industry sector, size of company in terms of employees and amount of sales. They are also interested in having contact information for each company, specifically mailing addresses and phone numbers. However, there is the potential for headquarter bias in both of these datasets. Companies may only be listed in terms of where their primary operations are such as a headquarters or an administrative building. In many instances this can be different than where goods are being stored, processed, and shipped. Additionally, the number of employees and sales are tied to companies rather than establishments. This can impact the estimate of freight activity at a specific address or an entire region.

Relevance for Commodity Flow Database Development

Establishment data is also often used for disaggregating commodity flows from national, regional or state level to smaller geographies. Establishments can be aggregated for specific economic sectors based on the geography of interest. Then, commodity flows that are directly related to these economic sectors can be grouped and then allocated to the desired geography based on their proportion of economic activity in each geographic unit.

Dun&Bradstreet data allows for this disaggregation to occur based on either the number of employ. The geographic specificity is limitless, because specific addresses are provided for each establishment.

General Types of Applications That the Database Can Be Used For

Customers use D&B products to: mitigate credit and supplier risk, increase cash flow and drive increased profitability; increase revenue from new and existing customers; and convert prospects into clients faster by enabling business professionals to research companies, executives and industries. Government products and solutions offered by D&B can be viewed at: <http://www.dnbgov.com/>

It should be noted that both InfoUSA and Dun & Bradstreet collect similar types of establishment data. These databases are primarily developed to service marketers in identifying new potential customers to contact and attempt to sell their products and services. Therefore, they obtain as much information as they can from as many different sources as they can to create, update, and expand their establishment data. Their methods do not appear to be standardized or publicly available in any detail. Therefore, it is difficult to determine, if one database is better than the other for freight planning purposes.

4.5 CONCLUSIONS REGARDING NATIONAL FREIGHT DATA

This review has demonstrated that there is a wide variety of data that exist to support the development of commodity flow databases at the local level. The Transearch database is currently the only one-stop shop for these types of databases. Both FAF2 and CFS provide data at the sub-state level. These three databases are the most commonly used freight flow databases for freight planning. However, they each have their strengths and weaknesses as discussed in Section 4.1.

For the CFS has a large number of “zero cells” due to the great number of origin-destination-commodity-mode combinations that it covers, and the limited sample size collected in the database. The CFS also has a large amount of suppressed data due to the need to protect proprietary information of shippers. The CFS has several out-of-scope industries which are not captured in the database at all. The CFS also does not have a methodology for capturing information on trip chaining or multi-stop truck tours.

The FAF2 database fills in the holes in the CFS database using a combination of economic and truck activity data for out-of-scope industries and suppressed data. However, the methodologies developed to fill in these gaps have not been verified with real-world data. As discussed in Chapter 3, the methodologies of building freight flow data from local socioeconomic data and state-level truck activity data appears to be very accurate for some commodities, and not accurate for others.

Transearch utilizes the most complex methodology in its database development. Additionally, the methodology is updated annually, so describing their process

is a bit of a moving target. Generally, the database uses a combination of economic production data, input-output relationships, mode-specific databases (e.g. Waybill, Waterborne Commerce), and CFS in its development. One of the unique features of Transearch is its proprietary motor carrier data exchange which it relies on heavily for the truck component of this database. However, it utilizes CFS for more local truck shipments, and is therefore subject to the same issues as CFS for these trips in terms of out-of-scope industries and trip chaining.

Many freight planners have compared information in the CFS, FAF2, or Transearch to local economic or vehicle flow information and found differences. These differences could have been the result of weaknesses in the databases or misinterpretations of the information presented in the databases. The information in this chapter is designed to describe these databases with sufficient detail to determine why these differences may occur.

Additionally, there is a plethora of data that can be used to enhance or develop commodity flow databases at the local level. However, as demonstrated in Chapter 3, a great deal of work is often required to properly incorporate and combine these local data sets. In Chapter 5, we will discuss the major findings regarding the national freight databases and their implications for developing commodity flow databases at the local level. We will also discuss the implications of these findings for the development of the guidebook for this study.

5.0 Conclusions and Implications for Guidebook

This report has described several processes for developing sub-national commodity flow data and dozens of freight-related data sets. Ultimately, this review of databases and previous studies will be used to develop the guidebook to assist transportation agencies obtain and/or develop sub-national commodity flow data that is useful for freight planning purposes. This chapter summarizes the key findings of the scan conducted in the first two tasks. It also describes options for incorporating this information into the guidebook.

5.1 CONCLUSIONS FROM LITERATURE SCAN

5.1.1 Applications of Sub-National Commodity Flow Data

The most common application for developing sub-national commodity flow data was for creating inputs for the truck component of a travel demand process. In particular, truck trip tables for travel demand models are typically developed by starting with a statewide, regional or county-level commodity flow database, and disaggregating it to the traffic analysis zone level using employment data from various sources. These truck trip tables are often calibrated by adding local truck trips using socioeconomic data similar to the trip generation process used for automobiles.

Typically, when freight flow data is routed to freight infrastructure, the commodity information is of less importance than the overall number of vehicles that are moving. However, there has been some recent interest in understanding commodity distribution of truck traffic on specific routes.

The second most common application for developing sub-national commodity flow data was for identifying data for trade flow purposes. This includes the development of commodity-specific data for improving pre-existing commodity flow databases (e.g. the Oregon Commodity Flow Database). There have been a handful of transportation agencies that have purchased the TRANSEARCH commodity flow database, and have undertaken efforts to improve the database by adjusting it to better match with local economic or truck count data.

Another trade flow application of sub-national commodity flow data is the development of supply chain data for a specific industry (e.g. Washington potato supply chain). This application is typically focused on identifying critical infrastructure, bottlenecks, and reliability issues for a specific industry. It is also used to determine the resilience of the freight infrastructure for a particular industry. It can be used to answer “what if” questions such as:

- What if a particular roadway had a major accident and was closed for several hours in a day?
- What if a particular roadway is closed and under construction for a few months?
- What if a particular region loses rail service?

This commodity-specific data is typically developed using a bottom-up process combining local economic data with industry interviews with supply chain decision makers.

The first two applications mentioned above are typically heavily reliant on obtaining TRANSEARCH data as a starting point for commodity flow database development. Therefore, the strength of the analysis is intimately tied to the strength of the TRANSEARCH database.

It should also be mentioned that several freight planners have compared the information in the CFS, FAF2, and Transearch to local economic data and found what was perceived to be discrepancies in the databases. There is a need to reconcile these databases with local economic production databases. There is also a need to develop processes to enhance these databases to incorporate information from local economic data and vehicle activity data.

5.1.2 Recent Research in Sub-National Commodity Flow Data

With the emergence of the FHWA FAF and FAF2 databases over the last ten years, an alternative commodity flow database became available for freight planning purposes. The key missing step to transform the data into a truck trip table for travel demand models was the disaggregation of the data from the 50-state level for FAF or the 114-zone level for FAF2 to the traffic analysis zone (TAZ) level. These processes have largely relied on local employment data available at the county and zip code level from public sources, and employment data available at the TAZ level in the travel demand model. There have been a few examples of validated truck components of travel demand models that have been developed using a disaggregation process that started with FAF2 data. The model developed for the South Alabama Regional Planning Commission, the Metropolitan Planning Organization for the Mobile MPO, by the University of Alabama at Huntsville is a good example of this application of disaggregating FAF data. However, none of these models were validated at the commodity-level.

There are also examples of disaggregating FAF data for more research-oriented studies. This includes the work done by Oak Ridge National Laboratories and the New Jersey Department of Transportation. These studies successfully disaggregated FAF data, but the final disaggregated commodity flow databases have not yet been used for freight planning purposes by transportation agencies.

The following challenges have been encountered when attempting to develop and utilize disaggregated FAF data to develop local commodity flow databases:

1. Several of the methodologies mentioned thus far have required utilizing complex processes and large amounts of resources. It is unclear that these disaggregation processes are more cost-effective than purchasing proprietary commodity flow databases that are already available at the county or zip code level.
2. The CFS data that is used for much of the FAF2 database does not include sufficient detail to characterize multi-stop tours. CFS captures each stop on these multi-stop tours as individual shipments. As FAF2 is disaggregated to finer geographies, there is a greater likelihood of these multi-stop tours having stops in multiple regions. Therefore, detailed information on these tours is needed to incorporate multi-stop tour trips into disaggregated databases.
3. These disaggregation methodologies rely on local employment and land use data. However, local employment data from public sources are problematic due to the suppression of data at the geographic level desired for developing truck trip tables. Land use data are often vague and not frequently updated to reflect actual activities that are occurring at the facility level.
4. The relationship between socioeconomic data and truck trip generation and distribution has not been well established. Sources such as the FHWA Quick Response Freight Manual and the ITE Trip Generation Handbook develop estimates of truck trip generation from limited field data across a narrow set of industry and land use characteristics. It is unclear how well these relationships hold up in different regions of the country or for sub-sectors of industries.
5. There is no available process for validating sub-national commodity flow databases. There is a lack of data sources to confirm the accuracy of these databases. None of the literature identified for this review validated their processes for developing sub-national commodity flow data.

Due to many of these challenges, the FHWA has issued guidance advising transportation agencies that the FAF databases are intended to serve as policy tools, and not intended for disaggregation to the local level.

5.1.3 Locally-Collected Freight Data

Over the years, there have been several freight data collection efforts focused on acquiring information on local freight movements. These typically include:

- Roadside origin-destination surveys – These surveys are typically conducted at weigh stations and rest areas capturing trip information by interviewing truck drivers. Because of the location of these surveys, they typically capture long-haul truck traffic. A variant of these surveys are

gate surveys conducted at freight facilities which interview truck drivers as they leave or enter facilities such as ports or rail intermodal yards

- Establishment surveys – These are interviews of operators of freight facility establishments intended to collect information on truck trip characteristics of specific industries or sub-regions
- Truck-following studies – These studies are conducted by following trucks as they enter or leave a roadway or facility. Trucks are followed to either their trip end or until they leave the study area. These are typically conducted in short corridors to determine the percent of through traffic relative to local traffic

Another source of local truck trip data is global positioning data (GPS) data. The availability of this type of data has increased dramatically as GPS data has become more readily utilized in the trucking industry and third-party vendors have begun to aggregate this information across companies. The most common application of this data is to track the performance of roadways, typically by summarizing speeds on heavily traveled corridors. There has been consideration of using GPS-equipped trucks to develop origin-destination information for the truck mode as a whole. Additionally, land use information at trip ends has been considered in terms of providing insight on commodities carried by trucks.

5.2 IMPLICATIONS FOR GUIDEBOOK DEVELOPMENT

The literature review described in this report indicated that the Guidebook on Developing Sub-National Commodity Flow Data could be useful by serving the following purposes:

- Describing the methodologies utilized in the development of the most commonly used freight flow databases – the BTS Commodity Flow Survey, the FHWA FAF2 database, and Transearch, along with their limitations and weaknesses.
- Describing the freight planning applications that benefit from use of sub-national commodity flow data
- Describing the strengths and weaknesses of the current state-of-the-practice in disaggregating national and state commodity flow data for development of truck trip tables and trade flow analyses. This would include describing methods to supplement the databases to minimize their weaknesses.
- Providing summaries of the types of freight data that are available and how they can be useful for developing local commodity flow data. This would include highlighting examples of how local commodity flow data have been used for specific freight planning efforts.

- Identifying a range of existing and new methodologies for conducting commodity flow surveys at the sub-national level. This will occur in Task 3 of this study.
- Describing the application of one specific procedure to test cases to demonstrate the applicability of the methodology for developing sub-national commodity flow databases. This will occur in Task 4 of this study.

This report represents the first step in answering these questions by documenting the national commodity flow databases and techniques used to develop sub-national commodity flow data. In Task 3, we will review specific data collection processes (e.g. establishment surveys, roadside O-D surveys, GPS surveys) and discuss strengths and weaknesses of these methodologies in terms of their ability to improve upon the types of techniques that are described in this report.

This report has developed a significant amount of material that can be used as reference material for the guidebook. In particular, the literature on various disaggregation methods will serve as a reference guide for alternative approaches for conducting this process. Additionally, the long-list of freight-related databases reviewed in this report will be useful in providing a comprehensive reference guide on the available data at the local and national level to support a wide range of freight planning efforts.

A. Expanded List of Commodity Flow-Related Databases

A.1 Oil Pipeline Data

Source

U.S. Energy Information Administration

http://www.eia.doe.gov/oil_gas/petroleum/info_glance/petroleum.html

General Information

The U.S. Energy Information Administration (EIA) is the statistical and analytical agency within the U.S. Department of Energy. EIA collects, analyzes, and disseminates independent and impartial energy information to promote sound policy-making, efficient markets, and public understanding of energy and its interaction with the economy and the environment. EIA is the Nation's premier source of energy information and, by law, its data, analyses, and forecasts are independent of approval by any other officer or employee of the United States Government. EIA conducts a comprehensive data collection program that covers the full spectrum of energy sources, end uses, and energy flows; generates short- and long-term domestic and international energy projections; and performs informative energy analyses. EIA disseminates its data products, analyses, reports, and services to customers and stakeholders primarily through its website and customer contact center.

Description of Data

Reports on import,/export, and movements of oil available on the website include:

- Weekly Estimates of Imports and Exports
 - Crude oil and petroleum products by U.S. and PAD District (weekly, four-week average).
- Imports by Area of Entry
 - Crude oil and petroleum products by U.S., PAD District (monthly, annual).
- Imports by Processing Area
 - Crude oil and petroleum products by PAD District (monthly, annual).

- U.S. Imports by Country of Origin
 - Crude oil and petroleum products by country of origin (monthly, annual).
- PAD District Imports by Country of Origin
 - Crude oil and unfinished oils by country of origin (monthly, annual).
- Company-Level Imports
 - Crude oil and total petroleum imports by the top 15 countries (monthly) and a summary of crude oil imports by company from the Persian Gulf (year-to-date).
- Residual Fuel Imports by State of Entry
 - by U.S., PAD District, State (monthly, annual).
- Percentages of Total Imported Crude Oil by API Gravity
 - (monthly, annual).
- Exports
 - by U.S., PAD District (monthly, annual).
- Exports by Destination
 - by country of export destination (monthly, annual).
- U.S. Net Imports by Country
 - Crude oil and petroleum products by country and region (monthly, annual).
- Movements Between PAD Districts:
 - **by Pipeline**, Tanker, and Barge, by Pipeline, by Tanker and Barge
 - (monthly, annual).
- Net Receipts of Crude Oil and Petroleum Products by Pipeline, Tanker, and Barge
 - (monthly, annual).
- F.O.B. Costs of Imported Crude Oil:
 - by Area, by Crude Stream, by API Gravity
 - (monthly, annual).
- Landed Costs of Imported Crude Oil:
 - by Area, by Crude Stream, by API Gravity

Methodology and Limitations

EIA information products are based on surveys, models, and external information sources. In disseminating non-EIA information, EIA focuses on using the best available information and ensuring the transparency of that information. Data users are encouraged to consider the initial sources of information presented in EIA's information products and to determine the suitability of such information for their purposes.

EIA analytical reports are prepared using a variety of analytical techniques, including simple tabulations with descriptive summary statistics, multivariate statistical methods, and econometric models. Analytical techniques are reviewed for their appropriateness to the data and the analysis being conducted and are clearly identified in reports.

Appropriate quality control procedures are used in all steps of preparing information products. Documentation of EIA information products is designed to improve understanding of the information so that users may assess the suitability of the information for their needs. Information products are reviewed by technically qualified staff prior to dissemination to ensure their quality.

EIA informs users of the concepts and methodologies used in collecting and processing the data, the quality of the data disseminated, and other features that may affect the use or interpretation. By providing information on methodologies and concepts to information users, EIA enables users to make judgments and verify that the data they are using are similar in conceptual framework and definitions to the data they need to complete their work.

The information also allows users to better understand possible sources of error which might restrict their uses of the data. In the area of statistical information, objectivity requires acknowledging that errors in statistical estimates are unavoidable. These errors generally fall under the categories of sampling and nonsampling errors. Sampling errors result when estimates are based on a sample and not a complete enumeration of the population of interest. EIA provides information regarding what is known about the magnitude of these errors, such as variances or coefficients of variation to quantify the magnitude of sampling errors. Though quantifying nonsampling errors is more difficult, EIA does provide information to assist users in understanding those possible error sources. EIA shall correct errors and issue revisions of previously disseminated information, as appropriate

General Types of Applications That the Database Can Be Used For

EIA's data and analyses are widely used by Federal and state agencies, industry, media, researchers, consumers, and educators. All of EIA's products can be accessed through its Web site, <http://www.eia.doe.gov>, which averages approximately 2.5 million visits per month. EIA also distributes information and data on specific topics through 47 e-mail subscription lists and seven RSS feeds.

- The Nation's leaders rely on EIA for timely and comprehensive information to formulate energy policy and programs.
- Industry looks to EIA for official estimates on energy demand, supply, prices, markets and financial indicators.
- Media and the general public rely on EIA for the most comprehensive source of current and historical data and information on all aspects of U.S. energy.
- The international community relies on EIA's products for timely information on world energy supply and demand.
- Primary and secondary school educators and students turn to the EIA Kid's Page (<http://www.eia.doe.gov/kids/>) for a wealth of energy-related information and activities.

A.2 Air Traffic Statistics

Sources

RITA BTS website:

http://www.rita.dot.gov/publications/budget_estimates/fy2008/html/section_03/reimbursable_ats.html

http://www.bts.gov/programs/airline_information/air_carrier_traffic_statistics/

General Information

RITA/BTS collects and publishes a variety of data about the operations of foreign and U.S. domestic airlines. The primary purpose of the program is to provide Congress, DOT, and other Federal agencies with uniform and comprehensive aviation data that are accurate, timely, and relevant for use in making aviation policy decisions and administering aviation-related programs.

Description of Data

- Releases traffic data monthly covering total monthly air passenger enplanements for the industry, airline rankings by air passenger enplanements, traffic volume, revenue miles, available seat miles, load factors and trip lengths.
- Releases domestic operating profit and loss data quarterly for individual airlines and by carrier groups (e.g., majors, low-cost carriers); airline domestic unit costs; and revenue yield.
- Releases on-time performance data monthly, providing the overall on-time arrival and departure performance of airlines, airports, and specific flights.
- Provides data monthly on causes of flight delays, characterized in five categories: Air Carrier, Extreme Weather, National Aviation System (NAS), Late-arriving aircraft and Security.

Methodology and Limitations

The Air Transportation Statistics Program collects, processes, and regularly releases/disseminates airline data from four primary data collections: On-Time Flight Performance (including causes of delays), domestic and international passenger and freight traffic, passenger ticket information, and airline financial and employment information. This air passenger itinerary and fare information is disseminated on the BTS website and directly from BTS as datasets and reports. The program also provides specialized reports for DOT and other Government agencies that they require in fulfilling their legislative mandates.

BTS collects and disseminates airline financial, traffic, performance and operational data from 150 U.S. airlines. Traffic data to and from the United States are collected and disseminated from 135 foreign air carriers that operate air service to the United States. Annually, BTS collects over 8,000 reports from U.S. and foreign airlines. Besides collecting and disseminating airline data, BTS continues to enhance its airline data edit and validation procedures in order to maintain a high level of data quality for DOT decision-makers

General Types of Applications That the Database Can Be Used For

DOT program uses of BTS airline data include the Airport Improvement program, Essential Air Service, monitoring the performance of the air transportation industry, and conducting status evaluations at both the individual airline and at industry levels, as well as conducting International Negotiations of air service agreements. The use and visibility of this airline data collection, which was originally mandated to enable oversight of airline competition, has grown the past five years as airlines and their markets have increased in complexity and competitiveness.

A.3 Massachusetts Institute for Social and Economic Research (MISER)

Source

<http://www.umass.edu/miser/>

General Information

The Massachusetts Institute for Social and Economic Research (MISER) founded in 1981 by the University of Massachusetts is an interdisciplinary research institute of the College of Social and Behavioral Sciences. MISER's research involves planning, strategy, and forecasting, with a focus on social, economic, and demographic issues.

Description of Data

As the Lead Agency in the State Data Center (SDC) program, MISER is the liaison between the Commonwealth and the U.S. Bureau of the Census and houses all the latter's data on Massachusetts. This includes:

- 1980 and 1990 Decennial Censuses;
- Quintennial Economic Censuses;
- Annual Current Population Surveys;
- Public Use Microdata Samples (PUMS);
- Building Permits and Demolitions;
- County Business Patterns;
- Census of Governments; and
- TIGER/Line Files.

General Types of Applications That the Database Can Be Used For

MISER is internationally known for research on foreign trade data which has been purchased by organizations around the world.

- MISER has improved the U.S. database on both the state of origin of export data, for which it is the national release point, and the exporter data base.
- MISER has developed the only non-Federal contributions to the National Trade Data Bank (NTDB).
- MISER has developed an MIS system for foreign trade data to help potential exporters locate the best foreign markets for their products.

MISER's work is designed to formulate new public policy for use by government policy-makers, as well as to develop information systems showing the relationship between public policy and the economy of Massachusetts and New England. MISER contracts with a variety of clients to perform customized, large-scale research projects and analyses.

A.4 Global Insight World Trade Service (WTS)

Source

<http://www.ihsglobalinsight.com/ProductsServices/ProductDetail2374.htm>

General Information

Global Insight's World Trade Service (WTS) provides a comprehensive view of international trade markets and commodities in the world. Each quarter, the database delivers current and projected measurements of bilateral trade between countries by commodity, value, volume, and mode of transport. Clients also

receive the quarterly publication Trends in Global Economy and Trade, which provides a detailed view of the economic and tactical issues driving U.S. trade forecasts.

Description of Data

WTS Features:

- 16 trade concepts measured in value and volume.
- 77 commodity groups.
- 54 countries and 16 world regions.
- Nine modes of transport, including:
 - Ocean Container (TEUs, tons, numbers of 20-foot and 40-foot containers);
 - Ocean Bulk (Liquid, Dry);
 - Ocean General Cargo/NeoBulk;
 - Air Cargo; and
 - Overland/Other (e.g., Pipelines).

General Types of Applications That the Database Can Be Used For

Clients use this database to analyze, size, and compare markets; uncover risks and opportunities; and better formulate strategic plans. Custom WTS Applications include:

- Port Modernizations and Concessions;
- Port-Specific Forecasts;
- Economic Impact Analysis;
- Policy Impact Analysis; and
- European Hinterland Analysis.

A.5 IMPLAN Data Files

Source

http://IMPLAN.com/v3/index.php?option=com_kb&task=article&article=108&Itemid=16

General Information

Minnesota IMPLAN Group, Inc (MIG, Inc) are the developers of the IMPLAN economic impact modeling system. IMPLAN is used to create complete, extremely detailed Social Accounting Matrices and Multiplier Models of local economies. IMPLAN® Data Files are combined with the IMPLAN® Version 3.0

software system. This enables development of sophisticated models of local economies in order to estimate a wide range of economic impacts. IMPLAN Data Files have been published annually since 1990.

Data is released annually in November, with the most current data year being 2008. Data can be purchased at a large number of regional levels, allowing you to define and work with a specific local economy. Complete coverage of the U.S. is available, as well as statewide information, county and zip code information

Description of Data

IMPLAN® Data Files include information for a set of highly disaggregated industries, sorted generally by their four- and five-digit NAICS codes (2001 and later data; earlier data sets are by three- and four-digit SIC codes). Information provided in every data set includes regional employment, income, value-added, household and government consumption. IMPLAN data files are compiled from a wide variety of sources including the U.S. Bureau of Economic Analysis, the U.S. Bureau of Labor, and Census.

Data Features:

- 440 Industrial Sectors: typically five-digit NAICS in manufacturing, two- to four-digit for other sectors.
- Regional Data available for the U.S., States, Counties, and zip codes.
- All elements balanced to the U.S. National Income and Product Accounts (NIPA).
- Conforms to I/O accounting definitions.
- Data elements are editable/customizable.

General Types of Applications That the Database Can Be Used For

Can be used to assess the following:

- Impact of New Businesses;
- Income Generate by Tourism;
- Economic Impact of Professional Sports;
- The Importance of Agriculture;
- Cost Benefits of Resource Management;
- Economic Contribution of Higher Education; and
- Permit Applications.

A.6 U.S. MARAD Maritime Statistics

Source

<http://www.marad.dot.gov/>

General Information

Collects, maintains, and disseminates data on domestic and international transportation, vessel characteristics and itineraries, port facilities, shipbuilding and repair, ship values, financial reports and vessels' operating expenses, shipping activities, and maritime employment. MARAD also publishes reports on marine transportation and cruise passenger statistics, and conducts research on maritime issues.

Description of Data

The Maritime Administration releases a wide variety of data and statistical reports on U.S. foreign waterborne commerce, with geographic (port, state, country), vessel, and value/weight information. Marad also releases U.S. Exports and Imports Transshipped Via Canada and Mexico, which summarizes U.S. maritime transshipments through Canada and Mexico, including a series of tables that summarize this traffic by U.S. Customs District, foreign trade area, and commodity. This information can be accessed at: http://www.marad.dot.gov/library_landing_page/data_and_statistics/Data_and_Statistics.htm

General Types of Applications That the Database Can Be Used For

The Marine Transportation System data inventory contains an overview of available marine transportation-related information that is either used or produced by the Federal Government. This inventory is made available to provide analysts, managers, and decision-makers with data and information useful for statistical and performance measurement. <http://marapps.dot.gov/mts/index.jsp>

A.7 Foreign Trade U.S. Census Bureau

Source

<http://www.census.gov/foreign-trade/index.html>

General Information

Foreign Trade of the U.S. Census Bureau is the official source for U.S. export and import statistics and responsible for issuing regulations governing the reporting of all export shipments from the United States.

Description of Data

Data are compiled in terms of: commodity classification, quantities, values, shipping weights, method of transportation (air or vessel), state of (movement) origin, customs district, customs port, country of destination, and whether contents are domestic goods or re-exports.

Since January 1989, commodities have been compiled under Schedule B harmonized classifications of domestic and foreign commodity exports. These transactions are classified under approximately 8,000 different products leaving the United States. Statistics are also compiled under the Standard International Trade Classification (SITC), North American Industry Classification System (NAICS), End-Use Commodity Classification, and Advanced Technology Products. These statistics include data about all 240 U.S. trading partners, 400 U.S. ports, and 45 districts.

Data are continuously compiled and processed. Documents are collected as shipments depart, and processed on a flow basis. Reports summarize shipments made during calendar months, quarters, and years. Statistics are reported monthly approximately 40 to 45 days after the end of the calendar month and on a year-to-date basis.

General Types of Applications That the Database Can Be Used For

The BEA uses the data to update U.S. balance of payments, gross domestic product, and national accounts. Other Federal agencies use them for economic, financial, and trade policy analysis (such as export promotion studies and export price indexes). Private businesses and trade associations use them for domestic and overseas market analysis, and industry-, product-, and area-based business planning. Major print and electronic news media use them for general and business news reports.

Import statistics provides a principal economic indicator as well as the most complete and only official source of monthly statistics on U.S. imports.

A.8 Caltrans Weigh-in Motion Data

Source

<http://www.dot.ca.gov/hq/traffops/trucks/datawim/>

General Information

Data WIM systems provide 24-hour traffic information at key locations on California highways.

Description of Data

The information collected includes:

- Axle weights and gross weight;
- Axle spacing;
- Vehicle Classification; and
- Speed.

General Types of Applications That the Database Can Be Used For

The information gathered is essential for the following functions:

- Pavement studies;
- Highway monitoring and capacity studies;
- Accident rate calculations; and
- Analysis of truck transport practices.

A.9 U.S. Department of Agriculture (USDA)

Fresh Fruit and Vegetable Shipment by Commodity

<http://www.usda.gov/wps/portal/usdahome>

<http://www.ams.usda.gov/AMSV1.0/getfile?dDocName=STELPRDC5067296>

The annual Fresh Fruit and Vegetable Shipments Report contains domestic rail and piggyback shipments, available truck and air shipments and exports of fresh fruits and vegetables by commodity, modes of transportation, origins and months. Available domestic boat shipments from U.S. possessions and imports from foreign countries are also shown.

Grain Transportation Report

[http://www.ams.usda.gov/AMSV1.0/ams.fetchTemplateData.do?template=TemplateA&navID=AgriculturalTransportation&leftNav=AgriculturalTransportation&page=ATGrainTransportationReport&description=Grain%20Transportation%20Report%20\(GTR\)](http://www.ams.usda.gov/AMSV1.0/ams.fetchTemplateData.do?template=TemplateA&navID=AgriculturalTransportation&leftNav=AgriculturalTransportation&page=ATGrainTransportationReport&description=Grain%20Transportation%20Report%20(GTR))

The weekly Grain Transportation Report (GTR) covers developments affecting the transport of grain, both in the domestic and international marketplace. This weekly publication reports on the latest volume and price data for barges, railroads, trucks, and ocean vessels involved in the transport of grain.

A.10 U.S. Energy Information Administration (EIA)

State Energy Profiles

<http://tonto.eia.doe.gov/state/>

State Energy Profiles is EIA's gateway to all of EIA's State energy data, plus individual State and Territory Profiles that present key facts and statistics about energy markets and industries.

Features of the State Energy Profiles:

- Energy Maps display State energy infrastructure and renewable energy potential.
- Quick Facts provide the most important State energy information.
- Written Overviews explain how fossil fuels, renewables, and alternative energy sources affect each State's energy markets.
- Data Tables provide the most current State-level statistics from EIA surveys.
- State Rankings Pages display selected data series ordered by data value.
- Associated Links take users to specific web pages on which EIA's detailed data and time series can be found.
- State Data Directory Pages provide comprehensive lists by subject of EIA's State-level information plus direct links to the information.

Quarterly Coal Report

http://www.eia.doe.gov/cneaf/coal/quarterly/qcr_sum.html

The Quarterly Coal Report (QCR) provides detailed quarterly data on U.S. coal production, distribution, exports, imports, receipts, prices, consumption, stocks, synfuel, and quality. Data on U.S. coke production, distribution, exports, imports, and consumption are also provided.

A.11 Quarterly Census of Employment and Wages

General Information

Quarterly Census of Employment and Wages (QCEW) published by the U.S. Department of Labor, Bureau of Labor Statistics (BLS)

Home page: www.bls.gov

Product overview: www.bls.gov/cew/home.htm

Database details: www.data.gov/details/318

In cooperation with State Employment Security Agencies (SESAs)

Description of Data

Type of data: Employment and wages

Modes covered: NAICS

Geographic detail: U.S., states, counties, and some core-based statistical areas (CBSAs), which include Metropolitan Statistical Areas (MSAs)

Accessibility: Public. Some data cannot be disclosed

Frequency of data release: Quarterly. Employment for the pay period including the 12th of the month and wages for a calendar quarter

Other: Data are published every quarter in the *County Employment and Wages* press release, usually six-seven months after quarter end; subset of data is published through *Create Customized Tables* system (see sample below); annual bulletin *Employment and Wages, Annual Averages* is published 10 months after year end

Methodology

Quarterly tax reports on number of people employed and wages paid submitted to SESAs by employers subject to state unemployment insurance (UI) laws and Federal agencies subject to Unemployment Compensation for Federal Employees (UCFE) program, comprising 99.7 percent of all wage and salary civilian employment

Data collection method/instruments: www.bls.gov/opub/hom/homtoc.htm

Uses

Input to BLS programs as well as other Federal and state programs, e.g., administer employment security program; measure state UI revenues; measure national, state, and local employment

Data Elements

Average annual wages, average weekly wages (by county), employment and wages (number employed, average weekly wage, rankings)

Create Customized Tables Format

Bureau of Labor Statistics Data

Quarterly Census of Employment and Wages HELP

1 Select a State or MSA or US Total
Alabama

2 Select one or more Areas
Alabama -- Statewide

3 Select one or more Industries
Total, all industries
Goods-Producing
Natural Resources and Mining
NAICS 11 Agriculture, forestry, fishing and hunting
NAICS 111 Crop production
NAICS 1111 Oilseed and grain farming
NAICS 11111 Soybean farming
NAICS 111110 Soybean farming
NAICS 11112 Oilseed, except soybean, farmi

4 Select one or more Ownerships
Total Covered
Federal Government
State Government
Local Government

5 Select one or more Establishment Sizes
All establishment sizes
Fewer than 5 employees per establishment
5 to 9 employees per establishment

6 Select one or more Data Types
All Employees
Number of Establishments
Total Wages (in thousands)
Average Weekly Wage

7 Get Data
OR for Multiple Queries
Add To Your Selection ->

Your Selection: (0 series selected) NOTE: Select a maximum of 200 series.

A.12 FleetSeek National Motor Carrier Directory

General Information

National Motor Carrier Directory maintained by Fleet Owner's FleetSeek data products group

Home page: www.FleetSeek.com

Sample profile and data elements: http://fleetowner.com/fleetseek/nmcd_new.htm

Description of Data

Type of data: For-hire trucking fleet contact information, demographics, operating characteristics

Modes covered: Truck

Geographic detail: National (i.e., fleets with headquarters in U.S.)

Accessibility: Proprietary – subscription available for purchase (see details below)

Frequency of data release: Annual (2008, 2009, etc.). Changes/additions posted on-line at least once every three days

Other: Subscription (\$495) allows Internet access for one year to 50,000 for-hire fleets (with search, select, sort, and print capabilities) and printed directory of 27,000 fleets; additional users (\$39.95 per year per user); option to download priced separately

Related databases: Canadian Motor Carrier Directory, Mexican Motor Carrier Directory, Private Fleet Directory (see separate template), and Owner-Operator Database (see separate template)

Methodology

“Accurate directory compiled by an entire team of marketing professionals whose sole responsibility is updating and verifying trucking intelligence.”

“We not only research; we also verify by seeking second and third sources to make sure what we publish is correct. Plus, we make over 100,000 telephone calls annually to corroborate and update our listings.”



Uses

Businesses (shippers, suppliers, truck manufacturers, insurance providers, etc.) can find freight, truckload, and LTL carriers; local, regional, and national carriers; obtain mailing lists; track trucking insurance

Sample Profile and Data Elements

NMCD Data Elements

- Carrier Name
 - Address
 - Principle Name
 - Title
 - Telephone
 - Toll-free, Fax, Web
 - DOT
 - Motor Carrier Number
 - Safety Rating & Date
 - Revenue
 - Carrier Type
 - Trucks, Tractors & Trailers
 - GVW Classes
 - Trailer Types
 - Commodities Transported
 - Standard Industrial Classification
 - Maintenance Indicator
 - Insurance Coverage & Policy Date
- (a separate *FleetSeek* component. Carriers having insurance info are indicated by the [shield icon](#).)

[Home](#) | [Search Database](#) | [Terms of Use](#) | [FAQ's](#) | [Contact](#) | [Admin](#) | [Logout](#)

[Return to Search Results](#)
[Financial & operating data available \(Blue Book\)](#)
[Print](#)

[Logout](#)

January 31, 2007

Last Update: 07/11/06

TTS RRI: 34838

Yellow Transportation, Inc.
 PO Box 7270
 Shawnee Mission, KS 66207-0270
E-mail:
Web: www.yellowroadway.com

Vehicles

	Owned	Leased	Total
Straight Trucks	101	0	101
Tractors	7991	0	7991
Trailers	32784	0	32784

GVW Class

	182	3,485	6	7	8
Y	Y	Y	Y	Y	Y

Trailer type

	1	2	3	4
Van	-	-	-	-

Commodity transported 1: General Freight
Commodity transported 2: -
Commodity transported 3: -
Commodity transported 4: -

Principal: James L Welch
Title: Chief Exec. Officer
Fleet Revenue: \$3,421,300,000.00
Safety Rating: SATISFACTORY 05/03/95
Std. Carrier Alpha Code: YFSY
Carrier type: FOR-HIRE
Carrier/Fleet Type: General Freight
Motor carrier number: 112713
Dot Number: 65616
Toll-free: (800) 458-3323
Telephone: (913) 344-3000
Fax: (913) 344-4707
SIC Code: 4207
SIC classification: For-hire trucking, general freight
Maintenance at Facility: Y
TL LTL designation: LTL
Year established: 1950

Insurance	Required	On File	Policy Effective Dates	Company	State
Cargo:	Yes	Yes	03/01/1996	FEDERAL INSURANCE CO.	NJ
BIPD Primary:	\$5,000,000.00	\$8,000,000.00	09/07/2006	OLD REPUBLIC INSURANCE CO.	PA
Bond:	No	No			

[Return to Search Results](#)
[Financial & operating data available \(Blue Book\)](#)
[Print](#)
[FleetSeek Profiles](#)
[Logout](#)

A.13 FleetSeek Private Fleet Directory

General Information

Private Fleet Directory maintained by Fleet Owner's FleetSeek data products group

Home page: www.FleetSeek.com

Sample profile and data elements: http://fleetowner.com/fleetseek/pfd_new.htm

Description of Data

Type of data: Private trucking fleet contact information, demographics, operating characteristics

Modes covered: Truck

Geographic detail: National (i.e., fleets with headquarters in U.S.)

Accessibility: Proprietary – subscription available for purchase (see details below)

Frequency of data release: Annual (2007-2008, 2008-2009, etc.). Changes/additions posted on-line at least once every three days

Other: Subscription (\$495) allows Internet access for one year to 35,000 fleets of 10 and more vehicles operated by companies with core businesses outside of trucking (with search, select, sort, and print capabilities) and printed directory of 27,000 fleets; additional users (\$39.95 per year per user); option to download priced separately

Related databases: Canadian Motor Carrier Directory, Mexican Motor Carrier Directory, National Motor Carrier Directory (see separate template), and Owner-Operator Database (see separate template)

Methodology

“Accurate directory compiled by an entire team of marketing professionals whose sole responsibility is updating and verifying trucking intelligence.”

“We not only research; we also verify by seeking second and third sources to make sure what we publish is correct. Plus, we make over 100,000 telephone calls annually to corroborate and update our listings.”

Uses

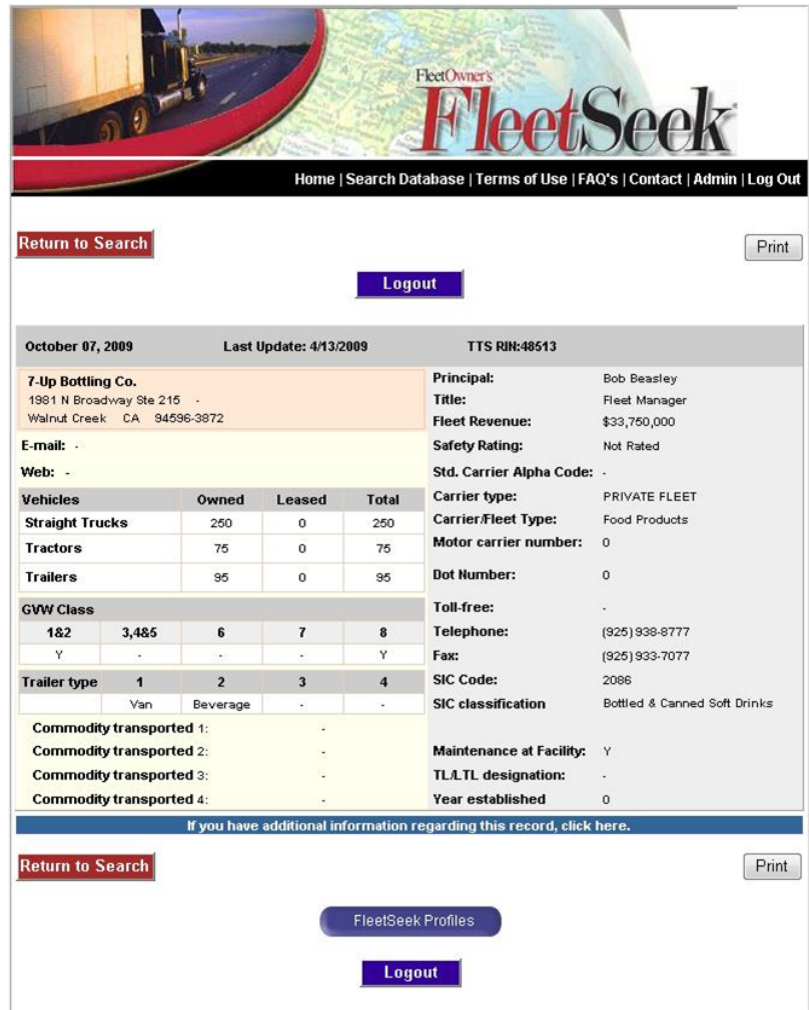
Suppliers of trucks, tires, tracking systems, tools, and other products needed by trucking operations can find private carriers/obtain contact and demographic information, obtain mailing lists

Sample Profile and Data Elements

PFD Data Elements

- Fleet Name
- Address
- Fleet Executive Name
- Title
- Telephone
- Toll-free
- Fax & Web
- Fleet Type
- Number of Trucks, Tractors & Trailers
- GVW Classes
- Trailer Types
- Industry Designation
- SIC Classification
- Maintenance Indicator
- Safety Rating & Date (if applicable)
- DOT & Motor Carrier Number (if applicable)
- Insurance Coverage & Policy Date (if applicable)

Sample FleetSeek Profile



Return to Search **Print**

Logout

October 07, 2009 Last Update: 4/13/2009 TTS RJH:48513

7-Up Bottling Co.
1981 N Broadway Ste 215
Walnut Creek CA 94596-3872

E-mail: -
Web: -

Vehicles	Owned	Leased	Total
Straight Trucks	250	0	250
Tractors	75	0	75
Trailers	95	0	95

GVW Class

182	3,485	6	7	8
Y	-	-	-	Y

Trailer type

1	2	3	4
Van	Beverage	-	-

Commodity transported 1: -
Commodity transported 2: -
Commodity transported 3: -
Commodity transported 4: -

Principal: Bob Beasley
Title: Fleet Manager
Fleet Revenue: \$33,750,000
Safety Rating: Not Rated
Std. Carrier Alpha Code: -
Carrier type: PRIVATE FLEET
Carrier/Fleet Type: Food Products
Motor carrier number: 0
Dot Number: 0
Toll-free: -
Telephone: (925) 938-8777
Fax: (925) 933-7077
SIC Code: 2086
SIC classification: Bottled & Canned Soft Drinks
Maintenance at Facility: Y
TLA.TL designation: -
Year established: 0

Return to Search **Print**

FleetSeek Profiles

Logout

A.14 FleetSeek Owner-Operator Database

General Information

Owner-Operator Database maintained by Fleet Owner's FleetSeek data products group

Home page: www.FleetSeek.com

Sample profile and data elements: http://fleetowner.com/fleetseek/oos_new.htm

Description of Data

Type of data: Owner-operator contact information, demographics, operating characteristics

Modes covered: Truck

Geographic detail: International (i.e., fleets with headquarters in U.S. and Canada)

Accessibility: Proprietary – subscription available for purchase (see details below)

Frequency of data release: Changes/additions posted on-line at least once every three days

Other: Subscription (\$995) allows Internet access for one year to 100,000 truck operators of one vehicle and no more than one trailer (with search, select, sort, print, and download capabilities); additional users (\$39.95 per year per user); regional versions (\$395)

Related databases: Canadian Motor Carrier Directory, Mexican Motor Carrier Directory, National Motor Carrier Directory (see separate template), and Private Fleet Directory (see separate template)

Methodology

“Accurate directory compiled by an entire team of marketing professionals whose sole responsibility is updating and verifying trucking intelligence.”

“We not only research; we also verify by seeking second and third sources to make sure what we publish is correct. Plus, we make over 100,000 telephone calls annually to corroborate and update our listings.”

Uses

Carriers can find drivers as replacements and new recruits; companies can market products directly to drivers

Sample Profile and Data Elements

00 Data Elements

- Carrier Name
 - Address
 - Principle Name
 - Title
 - Telephone
 - Toll-free, Fax, Web
 - DOT
 - Motor Carrier Number
 - Safety Rating & Date
 - Revenue
 - Carrier Type
 - Trucks, Tractors & Trailers
 - GVW Classes
 - Trailer Types
 - Commodities Transported
 - Standard Industrial Classification
 - Maintenance Indicator
 - Insurance Coverage & Policy Date
- (a separate *FleetSeek* component. Carriers having insurance info are indicated by the [shield icon](#).)

Return to Search Results [Print](#)

[Logout](#)

January 31, 2007 Last Update: 08/18/06 TTS RIR: 445285

2 Creeks Trucking
RR 1 Box 360
Windsor, IL 61957

E-mail:
Web:

Vehicles	Owned	Leased	Total
Straight Trucks	0	0	0
Tractors	1	0	1
Trailers	1	0	1

GVW Class

GVW Class	Owned	Leased	Total
182	3,485	6	7
			8

Trailer type

Trailer type	1	2	3	4
Reefer	-	-	-	-

Commodity transported 1: Meat
Commodity transported 2: Fresh Produce
Commodity transported 3: Grain
Commodity transported 4: Dry Bulk Commodities

Principal: Bruce Caruthers
Title: Owner Operator
Fleet Revenue: \$175,000.00
Safety Rating: NOT RATED // **Std. Carrier Alpha Code:**
Carrier type: OWNER OPERATOR
Carrier/Fleet Type: Owner Operator
Motor carrier number: 484721
Dot Number: 1225682
Toll-free: -
Telephone: (217) 644-2287
Fax: -
SIC Code: 4218
SIC classification: Owner-operator, individual
Maintenance at Facility:
TL/TL designation: TL
Year established: -

Insurance	Required	On File	Policy Effective Dates	Company	State
Cargo:	Yes	Yes	05/01/2004	LANCER INSURANCE COMPANY	NY
BIPD Primary:	\$750,000.00	\$1,000,000.00	05/01/2004	LANCER INSURANCE COMPANY	NY
Bond:	No	No			

If you have additional information regarding this record, click here.

Return to Search Results [Print](#)

[FleetSeek Profiles](#)

[Logout](#)

A.15 ATA North American Truck Fleet Directory

General Information

ATA Fleet Directory maintained by the American Trucking Associations, ATA Business Solutions

Home page: www.atabusinesssolutions.com

Product overview: www.atabusinesssolutions.com/p-60-ata-fleet-directory.aspx

Description of Data

Type of data: For-hire and private trucking fleet (over 10 trucks) contact information, demographics, operating characteristics

Modes covered: Truck

Geographic detail: National (i.e., fleets with headquarters in U.S.)

Accessibility: Proprietary – available for purchase (see details below)

Frequency of data release: Annual, with quarterly updates on CD

Other: Available in print or CD (starting at \$315), the latter with a search and reporting program for fleet searches (by state, zip code, area code, and fleet size), printing reports, creating mailing labels; free quarterly updates on CD; available in state editions and customized selections (e.g., by commodity type)

Methodology

ATA membership information and ?

Uses

Targeted at trucking industry suppliers/vendors, e.g., ATA Allied Member Directory

Marketed as the “best source of leads for your sales team in the industry.”

“In today’s competitive marketplace, you need contact names, guaranteed accurate mailing addresses, double-verified phone and fax numbers to run your business. You also need this information to remain current throughout the year.”

Data Elements

Company name, address, phone and fax numbers, commodities carried, fleet size, fleet type, trailer type, SIC codes, and more

A.16 Vehicle Travel Information System (VTRIS)

General Information

Source: <http://www.fhwa.dot.gov/ohim/ohimvtis.cfm>

Date: 2009

Contact: David Jones

Ph: 202-366-5053

E-mail: david.jones@fhwa.dot.gov

Office of Highway Policy Information

Federal Highway Administration

Agency: The VTRIS software and data reporting system was developed jointly by the Signal Corporation and the Federal Highway Administration Office of Highway Policy Information. The data that populates this information system originates from state-level Weigh-In-Motion (WIM) station data and is compiled/summarized and then distributed to each state and FHWA field office for policy planning.

Description of Data

The VTRIS data tables and summary information originates from highway-level Weigh-in-Motion and other permanent traffic recorders situated throughout the highway networks in each state. These traffic recording sites collect and compile traffic data characterizing the traffic attributes for each station located on state and Federal highways. Continuous vehicle count and classification data is available by station and highway (lane and direction), in addition to vehicle weight information for nine vehicle classifications. The VTRIS software and database was developed and is maintained by the FHWA, initially to help standardize traffic information from different states to improve pavement management programs across different pavement types and functional class of highway. VTRIS provided a standardized process for states to test for data accuracy and validation prior to being input into the national VTRIS database. An on-line query system also allows users to develop individual data queries from the VTRIS database by a variety of different data attributes including;

- Time (hourly, monthly, annually).
- State (AR, CA, GA, HI, IN, MN, MT, NC, NM, OH, PA, SD, TX, WA).
- Direction and Lane.
- Highway Functional Class.
- Station ID.
- Report Type:
 - W-2: Summary information on the number of vehicles counted and weighed;
 - W-3: Provides vehicle weight (average gross, percent loaded/empty and average weight of loaded and empty) data by FHWA vehicle class and is primarily utilized for payload distribution analysis;
 - W-4: ESAL factors by lane and direction primarily for pavement impact analysis;
 - W-5: Provides a matrix of vehicle counts that fall into the FHWA vehicle classes by 22 gross operating weight categories. This information is also primarily used for payload distribution and pavement analysis;
 - W-6: Contains the overweight vehicle report by FHWA vehicle class and axle configuration; and
 - W-7: Reports the distribution of overweight vehicles by FHWA vehicle class.

No commodity, shipment route, payload detail or O-D data and information are available from the VTRIS system. The information on vehicle classification counts and weight information is publically available at no cost for all stations from the participating states identified above.

Methodology

Each individual state Department of Transportation collects and compiles traffic data for the WIM sites within their respective state, primarily as part of their highway pavement management system. This information is then provide to FHWA and made available as part of the VTRIS on-line data query system. The information and data on traffic characteristics is generally very accurate overall, although periodic data anomalies can occur at specific stations due to collection equipment malfunction. The strength of this type of database is the ability to compare and evaluate traffic characteristics across multiple states, highway and pavement types, and by different vehicle classifications. However, no commodity, vehicle route, O-D, or shipment information is available and significantly limits the applicability of this data for subnational commodity flow analysis. This information may supplement or be utilized to validate other databases to verify freight and traffic characteristics on different corridors and highways.

Relevance for Commodity Flow Database Development

The VTRIS data is not useful for commodity flow analysis, except perhaps in validating other commodity flow information and data. This could be achieved by vehicle configuration type or station location to better understand how commodity flows match the proportion of vehicles types that are predominately utilized for given commodity shipments.

General Types of Applications that the Database Can Be Used For

This data is excellent for identifying traffic characteristics at specific locations and comparing to other locations or states. The freight-related attributes would include vehicle configurations, weight characteristics across different truck and vehicle types and the number of vehicles that fall into different weight categories across the various vehicle configurations. While no commodity information is available from this data, total freight tonnage may be obtained and can be useful for a wide variety of freight studies, especially for targeted freight corridors or specific highways.

A.17 Regional Economic Accounts

General Information

Source: Bureau of Economic Analysis:

<http://www.bea.gov/regional/about.cfm>

Date: 2008

Contact: Joel D. Platt

Associate Director for Regional Economics

Ph: 202-606-9605

E-mail: reis@bea.gov

Agency: The Bureau of Economic Analysis (BEA), part of the U.S. Department of Commerce, provides a vast amount of data and information regarding economic activity within the U.S. The Regional Economic Accounts data provides relatively detailed economic activity and growth information for state and local economies. In addition to other data attributes, gross domestic product, personal income and regional economic multipliers are available from the BEA.

Description of Data

The economic information available as part of the Regional Economic Accounts are summarized by industry level utilizing the NAICS coding system described above. Prior to 1997 the SIC coding system was utilized. There are basically three data attributes available as part of the BEA's Regional Economic Accounts. These include:

- **Gross Domestic Product (or Gross State Product)** – For each industry and state the current dollar GDP is comprised of compensation of employees, taxes on production and imports less subsidies, and gross operating surplus. This information is available for each state and metropolitan area (approximately 363 in the U.S.).
- **State Personal Income** – This is primarily an indicator of personal wealth for state residents and one means of comparing economic well-being of residents across states. This information is available on a quarterly and annual basis for each state, metropolitan area, county (3,111 counties in the U.S.), combined statistical areas (123), micropolitan statistical areas (576), metropolitan divisions (29), and BEA economic areas (179).
- **Regional Economic Multipliers** – These are economic multipliers (the percent change in output of one or more industries in a given area from a given change in output from other industries in the same area) representing the economic performance of all industries in a given geographical area. Currently these are available for a processing fee at the state and county level. These accounts-level information is utilized primarily as inputs for regional economic input-output modeling and computable general equilibrium modeling by regional economist.

Methodology

These data are compiled from a variety of different sources and government agencies (Federal, state and local), depending on the specific data component. The GDP information is primarily compiled using data from Federal, state and local agencies, BEA account information and data reported by private companies on attributes such as labor income, business taxes and capital income. A large proportion of the information used to calculate GDP originates from the

economic census that is conducted every five years and is available from 1963. GDP estimates during non-Census years are estimated using interpolation and extrapolation techniques based upon the most recent census information. For a complete description of how the gross domestic product is compiled by the BEA, see <http://www.bea.gov/regional/pdf/gsp/GDPState.pdf#page=3>. The state personal income information is also comprised from many different data source at various government bureaus and agencies. This includes a combination of administrative records, surveys and periodic censuses. Data is compiled from the Bureau of Labor Statistics, U.S. Department of Labor, Centers for Medicare and Medicaid Services, U.S. Department of Health and Human Services, Social Security Administration, Internal Revenue Service, U.S. Department of Treasury, U.S. Department of Veterans Affairs, U.S. Department of Defense, U.S. Department of Agriculture and the U.S. Department of Commerce. In aggregate, the sum of all state personal income available from the BEA differs slightly from that produced as part of the National Income and Product Accounts (NIPA) due to differences in geographic coverage and the treatment of income from U.S. citizens living and working abroad and foreign nationals living and working in the U.S. For a detail explanation of the process and methodology of compiling and estimating the state personal income by the BEA, see http://www.bea.gov/regional/pdf/spi2008/Complete_Methodology.pdf. The process for developing the regional multipliers dates back to the 1970s, utilizing a methodology developed by Daniel H. Garnick,⁷ as part of the Regional Input-Output Modeling System (RIMS or the later version RIMS II). Utilizing national BEA information on firm activity by NAICS classification, region-specific multipliers (state, county, metropolitan area, etc.) are derived using location quotients and Leontief's inversion approach. A complete description of the process for developing the regional economic multipliers may be found at <http://www.bea.gov/regional/rims/brfdesc.cfm#f1>.

Relevance for Commodity Flow Database Development

Detailed information on commodity movements is not available from these data sources, especially data attributes such as commodity type, O-D detail, route taken and any details regarding shipment attributes (payload weight, time of delivery, etc). Therefore, the relevance of these data in the development of commodity flow databases is in the broader sense related to better understanding of the underlying economic activity that may attract or generate freight activity in a region or geographic location. By obtaining a better understanding of how different businesses and industries are performing, as reflected by the GDP, personal income of area residents or the regional economic

⁷ See Daniel H. Garnick, *Differential Regional Multiplier Models*, *Journal of Regional Science* 10 (February 1970): 35-47; and Ronald L. Drake, *A Short-Cut to Estimates of Regional Input-Output Multipliers*, *International Regional Science Review* 1 (Fall 1976): 1-17.

multiplier of an industry or subset of industries, then we may be better equipped to draw inferences on regional freight flows in aggregate.

General Types of Applications that the Database Can Be Used For

These data are primarily used for understanding income, GDP and regional economic changes that occur at various geographic contexts. Similar to other data on economic and industry performance, these data help provide greater context and knowledge regarding the key economic drivers of a given region or subset of industries and may be most helpful when developing broad business and economic policy at the national, state, and local level.

A.18 Bureau of Labor Statistics: Productivity Statistics

General Information

Source: U.S. Department of Labor, Bureau of Labor Statistics

<http://www.bls.gov/bls/productivity.htm>

Date: 2009

Contact: Jay Stewart, Ph.D.

Division Chief, Bureau of Labor Statistics

Division of Productivity Research and Program Development

Ph: 202-691-7376

Agency: The Bureau of Labor Statistics (BLS), part of the U.S. Department of Labor, provides a vast amount of data and information regarding labor activity, unemployment, workplace injuries, pay/benefits and labor productivity within the U.S. This information, particularly the labor productivity data, is used for economic and policy analysis and to forecast future price, wage and productivity changes.

Description of Data

The productivity data provided by the BLS provides very detailed information on a variety of measures, including:

- **Labor Productivity:** These data include attributes such as output per hour, hourly compensation, unit labor costs, total output and hours of labor for all persons. These are available for select industries in manufacturing, mining, utilities, wholesale and retail trade and services. These measures primarily strive to relate labor production to total output for the U.S. economy and the major industry sectors.
- **Multifactor Productivity (MFP):** The MFP strives to relate total production output that results from a specified set of inputs, as a measure of productivity over time. These are provided for the U.S. business sector, the nonfarm

business sector, the manufacturing sector and 18 groups of manufacturing industries, 86 detailed manufacturing industries, railroad transportation, air transportation and utilities. These are also available quarterly or as annual time series.

- **International Comparisons:** These data allow comparison of many labor attributes for select industries across different countries. These data attributes include hourly compensation costs, productivity and unit labor costs, total labor force, employment and unemployment rates and consumer prices.
- **Productivity Research and Program Development:** This is a division of the Bureau of Labor Statistics that conducts policy analysis on labor productivity, labor economics, econometrics, industrial organization and statistics. A list of research areas and recent publications and working papers are available at <http://www.bls.gov/dpr/>.

Methodology

Similar to the Regional Economic Accounts information, the productivity data is comprised of many different data streams from a wide variety of Federal, state and local agencies. Output, as measured by GDP originates from the BEA and Department of Commerce. Labor input is primarily derived from the Bureau of Labor Statistics Current Employment Statistics (CES) program. For a detailed account and explanation of the methods used in the development of the productivity statistics, see <http://www.bls.gov/lpc/lpcmethods.pdf>

Relevance for Commodity Flow Database Development

The information provided by the BLS productivity statistics is limited as it relates to the development of commodity flow information. Similar to earlier described economic and production data at the national, state, or local level, the productivity statistics do not provide any data or information on commodity shipments, origins, destinations, mode choice, route, payload weight or vehicle configuration. Thus the application of this productivity information resides in the improved understanding of how different industries, business, sectors are performing across different geographic realms. These data are informative for understanding productivity changes over time and that may improve knowledge of general freight activity across different industry sectors.

General Types of Applications that the Database Can Be Used For

These data are primarily used for understanding productivity and economic changes that occur at various geographic contexts. Similar to other data on economic and industry performance, these data help provide greater context and knowledge regarding the key economic drivers of a given region or subset of industries and may be most helpful when developing broad business and economic policy at the national, state, and local level.

A.19 TRB E-Circular E-C080, Freight Data for State Transportation Agencies – A Peer exchange, July 11, 2005

General Information

November 2005

James Hall

<http://onlinepubs.trb.org/onlinepubs/circulars/ec080.pdf>

Transportation Research Board Statewide Transportation Data and Information Systems Committee

Document Summary

The Transportation Research Board (TRB) Committee on Statewide Data and Information Systems (ABJ20) hosted a peer exchange to discuss state transportation agency collection and use of freight data and to identify ongoing issues and trends. It was attended by state transportation agency officials and a few others from Cambridge Systematics (served as facilitators), BTS, FHWA, Universities, AASHTO and the Port of Portland.

Before to the peer exchange, state agency participants completed an extensive questionnaire on freight data practices and attended TRB's Commodity Flow Survey (CFS) Conference. The survey asked about current practices and uses of freight data. Major concerns were on data collection and access, data quality, and a lack of necessary resources and staffing. Responses also portrayed the varied and complex uses of freight data including modal program development, air quality analysis, safety investigations, hazardous materials monitoring, domestic and international trade, corridor analysis, supply chain analysis, and transportation modeling.

Other trends include the participation of local planning agencies and the acquisition of freight data at a local level. Participants identified determination of data and information needs, logistics and supply chain analysis, freight data framework models, definition of the role of state department of transportation, data management personnel, and understanding of the universe of data sources and available data as the major challenges and issues relating to freight data.

DOT representatives were asked the following questions:

1. List the types of policy questions, planning studies, project plans and designs, or other activities for which your DOT needs to analyze data on the movement of commodities, truck travel, rail freight and shipping, safety, or other aspect of freight transportation. For each type of policy question, planning study, or other issue requiring information about freight, indicate whether the questions, studies, or issues are raised by offices within your state DOT, other executive agencies, your state legislature, local

governments, or constituents. Also, indicate whether responses are needed within days, weeks, or months.

2. What kinds of freight data do you use when responding to the questions and issues in Question 1? Where do you get the data (within your DOT, from other state agencies, from local governments, from the Federal government, from consultants)?
3. Who performs the analyses of the questions and issues in Question 1? freight specialists in your Office of the Secretary or other department-wide office? planners or other generalists in your Office of the Secretary or other department-wide office? freight specialists in a modal administration? planners or other generalists in a modal administration? freight specialists or generalists from other state agencies? national consulting firms? local consulting firms? Local universities?
4. How satisfied is your DOT with data and analytical tools for answering the questions and issues in Question 1? What are the most important improvements your DOT would like to see made to freight data and analytical tools?
5. What do you do to improve the quality of freight data and analytical tools in your state?
6. Does your agency have staff expertise for answering the questions and issues in Question 1? What are the most important skills and training you would like to have for your staff?
7. How much does your state spend on freight data, analytical tools, training, and assistance in using freight data? What do you think the appropriate expenditure level should be?
8. Does your state or the local jurisdictions in your state plan to begin or expand collection of freight data? If yes, what kind of expertise are you using, and how are you funding the effort?

The report includes a section describing the activity in each of the represented states, so this summary includes a very broad set of data inputs and applications.

Commodity Flow Data Inputs

In general the document reports very little experience with commodity flow data. Traffic flow data is much more heavily used. Studies are typically regional, and use the DOTs own data for analysis on specific corridor or regional planning. States with specific experience or interest in commodity data are described below.

- **California** – State experience with commodity flows is limited, although again, significant traffic flow data is used. Use TRANSEARCH and PIERS for their Intermodal Transportation Management System for regional and statewide studies. Also use Waybill sample.

- **Colorado** – No DOT experience with commodity flow data. Some interest in CFS, FAF, and GeoFreight, but currently a lack of training.
- **Florida** – TRANSEARCH database from Reebie and Associates. Developing a freight data clearinghouse with their Seaport Office, TranStat and Office of Information Systems. The document does not provide many specifics, but they also take advantage of the CFS, FAA, Bureau of Business and Economic Research at the University of Florida.
- **Kentucky** – TRANSEARCH data used for truck O-Ds, Waybill data, and specific data collection efforts.
- **Minnesota** – Hired consultants to derive freight flow data from the TRANSEARCH database but process is not described in the document. Railroad Carload Waybill Sample is purchased on an annual basis from the STB.
- **New York** – TRANSEARCH, PIERS, Waybill Samples, NATRI, Global Insight, CFS, and the VIUS. No details are provided as to how these are manipulated or specifically used.
- **Ohio** – TRANSEARCH and waybill data. Surveys were also conducted to support statewide modeling, but no additional detail provided in the document.
- **Pennsylvania** – TRANSEARCH, waybill, surveys, sharing with other state agencies (no other specifics provided).
- **Port of Portland** – CFS, ODOT statewide commodity flow forecast (“fused data sources such as PIERS, Army Corps Waterborne Commerce, Reebie and Associates TRANSEARCH, etc.”), State of Origin International Trade data, customs district data, International Trade Administration, FAF.
- **Wisconsin** – REMI and IMPLAN models are used for economic forecasts. Global Insight provides a commodity flow database and software

Final Commodity Flow Database

Not relevant in the case of this report.

Types of Applications that Have Been Implemented

Modal program development, air quality analysis, safety investigations, hazardous materials monitoring, domestic and international trade, corridor analysis, supply chain analysis, and transportation modeling.

Applicability to Other States/Regions

The estimates of cost for this data ranges by size of the state and the level of effort. California estimates as much as \$10,000,000 historic spend on regional studies. Colorado reported a current spend of \$325,000, but \$5,000,000 on

specific freight-related studies. Most states mention the need for better employee training, and the use of consultants to perform any significant data analysis or manipulation.