A Preliminary Roadmap for the American Freight Data Program (DRAFT)

November 26, 2003
(Modified December 2004 for TRB)

Prepared by
Frank Southworth
Senior R&D Staff
Oak Ridge National Laboratory
P.O. Box 2008
Oak Ridge, TN 37831

Prepared for
Bureau of Transportation Statistics
United States Department of Transportation
Washington D.C. 20590

Contents

* Not included in this version of the report

Listing of Terms and Acronyms Used in This Document

Term	Definition
AAMVA	American Association of Motor Vehicle Administrators
AASHTO	American Association of State Highway and Transportation Officials
AFDP	American Freight Data Program
ACE	Automated Commercial Environment
BTS	Bureau of Transportation Statistics
CBP	Customs and Border Protection
Census	Census Bureau (DOC)
CFS	Commodity Flow Survey
CVISN	Commercial Vehicle Information Systems and Networks
DHS	United States Department of Homeland Security
DOC	United States Department of Commerce
DOD	United States Department of Defense
DOE	United States Department of Energy
DOT	United States Department of Transportation
EDI	Electronic Data Interchange
ERS	Economic Research Service (USDA)
FAF	Freight Analysis Framework
FHWA	Federal Highway Administration
FMCSA	Federal Motor Carrier Safety Administration
FAA	Federal Aviation Administration
FRA	Federal Railroad Administration
GPS	Global Positioning System
IRP	International Registration Plan
ITDS	International Trade Data System
ITS	Intelligent Transportation System
Link-OD model	A type of computer model used to combine survey – based origin-to- destination (O-D) data with link count data
MARAD	Maritime Administration
NHTSA	National Highway Traffic Safety Administration
O-D	origin-to-destination
OMB	Office of Management and Budget
ORNL	Oak Ridge National Laboratory
OSC	Operation Safe Commerce
OST	Office of The Secretary of Transportation

PIERS	Port Import Export Reporting Service
RIC	remote intelligent communication
RFID	Radio frequency identification device
SAL	Smart active label
STB	Surface Transportation Board
TRB	Transportation Research Board
USACE	US Army Corps of Engineers
USDA	United States Department of Agriculture
US ITC	United States International Trade Commission
VIUS	Vehicle Inventory and Use Survey
WIM	Weigh-in-Motion

A Preliminary Roadmap for the American Freight Data Program

1. Purpose

The purpose of this white paper is to describe a roadmap, or prototype action plan, for the American Freight Data Program (AFDP). The AFDP is a major freight data collection and analysis initiative by the Bureau of Transportation Statistics (BTS) within the United States Department of Transportation (DOT). The program covers the collection and dissemination of freight movement data for each of the major modes of transportation: truck, rail, water, air, pipeline and intermodal combinations. Data collection activities within the program target truck and intermodal transport, as the areas currently most in need of improvement. The proposed program supports the DOT's desire to develop an improved freight database, as part of its delivery of a comprehensive Freight Action Plan that improves the transportation system's capacity to accommodate significant future growth in freight operations.

Information has been pulled together from a number of sources and past efforts within and outside the BTS. This roadmap represents a first attempt to bring all of these components together in one place. The main purpose of the paper is to identify the important freight movement data needs of decision-makers, to link these to the technical issues to be addressed in developing the AFDP, to describe the most promising options available for providing these data, and to propose a set of actions to be undertaken. Scheduling and funding issues are also addressed at a fairly high level of program activity. The paper emphasizes the first six to ten years of AFDP activities: while recognizing the need for a sustained, 20 to 30 year effort to develop the sort of comprehensive and cost efficient freight movement data program that will satisfy the needs of the vast majority of data users and the decision-makers they support. The paper describes an evolutionary approach to freight movement data collection, transitioning from current data collection methods towards new approaches. Such an approach is seen as minimizing the risk of losing, or interrupting, the current time-line of existing data products, while allowing the development of new ways to collect data as well as creating new forms of data product.

2. Background

2.1 Freight Movements and the US Economy

The efficient movement of freight is key to the nation's economic vitality. More than 6 million business establishments depend on safe and efficient freight

transportation as part of their operations. By the end of the last century the US transportation system was moving more than 15 billion tons of freight valued at more than $9 trillion dollars. The "freight bill" for moving these commodities exceeded $560 billion, or 6 percent of gross domestic product [1]. By 2020 the volume of this freight is estimated to grow an additional 70 percent, building on three decades of similarly rapid growth [2].

In responding to this growing demand for freight services, while also helping to fuel it, freight operations have made use of cost-saving advances in real time and increasingly global information systems for scheduling goods movements. The emergence of just-in-time delivery services has been one result, either reducing or pushing inventory-holding costs closer to the supplier in return for more frequent (and necessarily more reliable) transportation services. This has led, in turn, to significant increases in truck and airfreight traffic and to the introduction of new and innovative inter-modal transportation options. As a result, we have a rapidly expanding, highly dynamic and multimodal freight transportation industry, and one that continues to adopt new technologies, find new markets, and evolve new business arrangements as each year passes. We also have an increasingly congested transportation infrastructure that is struggling to keep pace with the demands being placed on it. New infrastructure is costly. Governments need accurate and timely data on both current and anticipated freight movement patterns and volumes in order to invest wisely in new system capacity. Current data sources are often inadequate to this task.

2.2 Calls For Better Freight Data

A good deal of freight movement data are being collected around the United States. These data are being used at many different levels of spatial, sectoral and temporal aggregation, in both operational and planning contexts, and in support of both public and private sector policy development and decision-making. However, a sizeable literature now exists pointing out the many gaps in coverage and the significant challenges involved in piecing these different datasets together [1,3-7]. In response to these growing concerns over freight data availability and quality, a number of agencies, including the BTS, sponsored a November 2001 conference hosted by the New York Department of Transportation titled "Data Needs in the Changing World of Logistics and Freight Transportation" [3]. A major conclusion of this conference was that there exists a considerable and unmet need for freight movement data that is accurate, timely, sufficiently comprehensive and detailed in its scope. In particular, rapid improvement in geographic information technology has increased the demand for spatially detailed freight movement data within both the private and public sectors. It was concluded that effective freight planning at all levels of

government (metropolitan, state, regional and national) needs additional support from federal data collection efforts.

2.3 Outline of This Document

The rest of this document is organized as follows:

Section 3: Needs Assessment Approach and Findings
Section 4: Proposed Program Activities
Section 5: Example Activity Timeline
Section 6: Bibliography

3. Freight Data Needs Assessment

3.1 Needs Assessment Approach

Between 1999 and 2003 BTS was involved in a number of activities that shed light on who the DOT's major freight data customers are and what sort of freight data they are looking for. This included the following efforts:

- A review of critical freight data sources and data needs (1997-99)
- November 2000 CFS Freight Data Round Table in Washington DC
- November 2001 Saratoga Springs, New York Conference on Freight Data Needs
- TRB Expert Freight Data Panel Meetings and TRB Special Report 276 (2002-2003)
- TRB Expert Panel Review of the US Commodity Flow Survey Program (2002)
- A User-Solicited Transportation Data Needs/Data Gaps Study (2002)
- An In-House Assessment of Program Purpose, Scope and Major Data Products (2002-2003)
- Discussions with major freight data users within and outside the US Department of Transportation

The next two sections summarize the major findings from these review and assessment activities.

3.2 Existing Federal Freight Movement Data Programs

The federal government collects a broad range of ground-based freight movement data, including the following datasets [see 3,5-7]:

- Commodity Flow Survey (BTS/DOT)

- Transborder Surface Freight Data (BTS/DOT)
- Highway Performance Monitoring System (FHWA/DOT)
- Rail Carload Waybill Sample (STB/DOT)
- International Registration Plan (FMCSA/DOT)
- Commercial Vehicle Information Systems and Networks (FMCSA/DOT)
- Maritime Statistics (MARAD/DOT)
- Foreign Waterborne Transportation Statistics (USACE-MARAD/DOD-DOT)
- Waterborne Commerce of the United States (USACE/DOD)
- Vehicle Inventory and Use Survey (Census/DOC)
- Automated Commercial Environment (CBP/DHC)
- International Trade Data System (multiple federal agencies)
- Motor Carrier Financial and Operating Statistics (BTS/DOT)
- Office of Airline Information Air Freight Statistics (BTS/DOT)
- Office of Pipeline Safety 's Crude and Petroleum Products Pipeline Database (OPS/DOT)
- Natural Gas Pipelines Database (EIA/DOE)
- Detailed Hazardous Materials Incident Reports (FMCSA/DOT)

Between these various databases it is possible to create a multimodal picture of goods moving through the major freight activity gateways and corridors within the United States, and to map the major trade routes and markets linking the US to the rest of the world. However, considerable effort is required to do so. While the existence of a large number of freight data sources appeared to indicate a wealth of information, analysts using these data often run into significant problems: in particular, incomplete coverage of freight movements, variations in reporting, lack of geographic detail, differences in collection methods that complicate comparison and interpretation, lack of statistical information on data reliability, and difficulties in combining databases collected for transportation purposes with datasets collected for other reasons (e.g. for economic forecasting) [4].

In considering the strengths and weaknesses of these federal freight movement databases, and how they might be improved through the AFDP program, three points stand out:

- Major improvements to domestic freight movement data will require the elimination of a number of well known data gaps, in part through the collection of new data, and in part through the fusion, or synthesis, of two or more existing data sets.
- Major improvements in international freight movement data will, in contrast, come from greater access to, and use of, detailed microdata

records within current and rapidly evolving international trade and traffic databases.

- Complete coverage of freight movements within, across, into and out of the United States will require the ability to merge these domestic and international sources of freight movement data. This is likely to require a gradual evolution in data collection and synthesis, requiring a number of years to achieve: and will depend heavily on both the cooperation of the private sector and the ability to implement non-intrusive data collection and reporting methods.

Domestic Freight Movement Data Gaps. Of the three primary ground transportation modes covered by the AFDP, detailed and nationwide samples exist for both rail and waterborne station-to-station movements and their routings over the nation's railways and waterways. Both are based on carrier surveys. The US Army Corps of Engineers provides a detailed, and roughly 100% sample coverage of waterborne freight shipments by detailed commodity and vessel types. The Surface Transportation Board's Railcar Waybill Sample similarly provides, for most purposes, statistically robust estimates of the nation's major rail freight O-D and railroad specific movements. In contrast, data on truck freight movements is much more limited, and developing O-D matrices or assigning traffic to specific highway routes requires piecing together data from a number of different sources, notably:

- The US Commodity Flow Survey (1993, 1997, 2002). This is a mail-out/mail back shipper establishment survey that reports the weight, value, and true origin and destination of freight shipments by any combination or sequence of truck, rail and water (and air) modes, down to the zip code level. The survey is part of the Economic Census and covers mainly the mining, manufacturing and wholesale sectors of the economy. A BTS commissioned study using 1997 CFS and other data sources estimated these sectors to cover no more than three quarters of all the tons moved annually within the US (see Table 1). The CFS does not measure the volume of empty or partially empty truck movements associated with backhauling and repositioning. To reduce respondent burden the survey also does not collect data on truck routes, travel distances, times or costs, using routing algorithms to infer the routes taken [8].

- The Vehicle Inventory and Use Survey (VIUS) collects data on both personal and commercial use trucks registered (or licensed) in the United States as of July 1 of the survey year. Also conducted by Census every 5 years, VIUS collects data on highway vehicle characteristics and use based on a stratified sample of vehicle owners, with the larger commercial carriers receiving a longer questionnaire than smaller operators. VIUS

provides a basis for estimating total nationwide truck-miles and ton-miles of travel by short, medium and long distance intervals, but lacks any details on the geography of such moves, and contains very limited commodity detail. [1] The owners of some 136,000 trucks were sampled in the 2002 VIUS, with approximately 360 vehicles sampled per week, spread out evenly across each week in the year. From this data state-specific, Washington DC, and US national statistics are generated.

Table 1. US Annual Freight Activity Estimates (for 1997)*

A)Total Shipments:		
Value (Billions of $)	Weight (Millions of Tons)	Ton-Miles (Billions)
$8,567	14,800	3,851
B) CFS In-Scope Shipments:		
Value (Billions of $)	Weight (Millions of Tons)	Ton-Miles (Billions)
$ 6,945	11,090	2,837
(81.1%)	(74.9%)	(73.7%)
C) CFS Out-of-Scope Shipments:		
Value (Billions of $)	Weight (Millions of Tons)	Ton-Miles (Billions)
$1,622	3,710	1,014
(18.9%)	(25.1%)	(26.3%)

* Source: "Freight USA" (reference [11]). Note numbers do not include truck freight moves in the retail and service, household, utility, construction and forestry sectors, and the domestic delivery of import shipments from a US port of entry to an in-scope CFS establishment.

A further drawback of the current CFS and VIUS is their 5-year attachment to the Economic Census: the source of their mandatory collection authority. Given the speed with which the trucking industry is not only growing but changing, more regular (more "continuous") types of activity survey have been suggested. Depending upon pending legislation, this may need to occur in the context of either mandatory or voluntary reporting requirements.

Other sources of truck activity data offer additional truck movement information: notably FHWA's Highway Performance Monitoring System (HPMS) of State-supplied AADT truck counts, and the within and between State truck operating mileage data collected by the AAMVA's International

[1] The VIUS excludes vehicles owned by Federal, state, or local governments; ambulances; buses; motor homes; farm tractors; unpowered trailer units; and trucks reported to have been sold, junked, or wrecked prior to July 1 of the year preceding the survey. http://www.census.gov/econ/www/viusmain.html

Registration Plan (IRP)[2]. Neither can be used to build statistically robust origin-destination movement patterns for anything below the State-to-State level of spatial resolution, and for limited truck size classes. BTS also collects Financial and Operating Statistics (F&OS) data from the larger[3] for-hire trucking companies, from which annual ton-mileage estimates have also been estimated. Again, these data does not supply any data on the geography of these movements, and carriers often fail to report useable ton-mileage data. Therefore, while over 70% of all tons reported by the 1997 CFS moved all or in part by truck (and over 73% by dollar value) there is currently no national dataset from which these truck movements can be generated, or from which national statistics can be drawn that show the working relationship between truck activity (e.g. vehicle miles and ton-miles driven, by vehicle type) and the volumes and spatial patterns of the different commodities being traded. This makes it difficult to understand and quantify, among other things, the relationship between truck use, future highway infrastructure needs, highway safety, and the growth in economic activity.

Given the within-mode, terminal-to-terminal nature of the above water and rail movement datasets, there is also not currently a national dataset covering the important operation of trucks in their role as the local access and/or egress mode in intermodal (truck-rail, truck-air and truck-water) transportation. This also means that current railcar waybill and waterborne vessel movement reporting systems also lack true shipment origin and destination information. Based on the above, a pressing area for AFDP attention is the development of better truck freight movement data, leading to more geographic detail on origin-to-destination vehicular traffic flows involving both truck-only and truck-intermodal movements.

International Freight Data Needs. In 2001 the U.S. transportation system carried merchandise exports worth an estimated $731 billion and merchandise imports valued at $1.1 trillion (in current dollars). Transporting this merchandise required a significant amount of equipment: using over 936,000 aircraft, 215,000 maritime vessels, and 19 million vessel, truck, and rail container entries into the United States.[4] With such a large amount of transportation equipment involved in U.S. international trade there is also the possible threat of using freight vessels

[2] http://www.aamva.org/irp/

[3] The program collects balance sheet and income statement data along with information on tonnage, mileage, employees, transportation equipment, and other related items from motor carriers with annual revenues of $3 million dollars or more (http://www.bts.gov/mcs/desc.html)

[4] http: //www.bts.gov/publications/us_international_trade_and_freight_transportation_trends /2003/html/introduction.html#1

and vehicles for terrorist activity. To understand how this freight affects the economy and security of the nation, government needs to know where it moves from and to as well as how (which modes, routes) it gets there.

A number of federal agencies are involved in the collection, processing and dissemination of international trade and transportation data. U.S. merchandise trade statistics are processed and released by the U.S. Census Bureau Foreign Trade Division. Census international merchandise trade data are captured from administrative documents required by the Departments of Commerce and Treasury. The U.S. Customs Service collects these documents at the port of entry or exit unless the information is filed electronically using the Automated Broker Interface (ABI) on imports or the Automated Export Reporting Program (AERP) on exports. Census also releases overall trade and transportation statistics that include data elements on the value, commodity, weight, country of origin and destination, U.S. port, and so forth. Many agencies obtain special extractions and tabulations from the Census Bureau, and then perform additional quality assurance reviews and analyses to meet the needs of their own customers. These include: data on North American land trade (released to BTS and disseminated as the Transborder Surface Freight Data); data on U.S. international maritime trade (released to the Maritime Administration and the Army Corps of Engineers and disseminated in multiple formats); and data on U.S. transportation related goods and overall trade data (released to BEA/DOC and disseminated in multiple formats, including balance of payments information). A popular private sector product based on Customs data is PIERS (Port Import Export Reporting Service)[5]. Several challenges arise when using these multiple data sources, including variations in accuracy, reliability, time series, and data field definitions.

U.S. merchandise trade statistics have a number of problems. Specifically, gaps in data arise from the following:

1. The absence of actual shipping weight measurements for exports is a critical problem with existing U.S. international trade data. Weight data are currently only collected for imports due to the specifications in the reporting requirements of the U.S. Census Bureau and U.S. Customs Service.
2. Data on the domestic origin and destination of international trade is not always reported correctly. This is often due to problems with data definitions and with the collection procedures used by the Census Bureau and the U.S. Customs Service. For transportation purposes the foreign trade origin of movement series (OM) is used to identify the place where

[5] http://www.piers.com/default2.asp

exports begin their journey. Data in the OM series is supposed to represent the physical transportation origin of export shipments. A major limitation of these data are the result of reporting errors made by intermediaries, such as freight forwarders or logistics providers who complete the documentation for a particular shipment. These intermediaries sometimes list their headquarters location as the point of origin or specify the location of the port of exit, which is not always where the goods began their journey. This mainly has an effect on the state distribution of non-manufactured exports, where intermediaries are more common. The impact is greatest on the allocation of exports of farm products, minerals, and other bulk commodities.

3. The lack of intermodal data. In many cases there is a change in mode of transportation with internationally traded goods from origin to final destination. However for merchandise trade statistics, the export mode of transportation is defined as the mode used when the U.S. international border is crossed. On the import side, the mode of transportation is defined as the last mode used when the freight was transported to the U.S. port of clearance or entry. Because of reporting requirements, merchandise trade statistics do not distinguish goods moved by intermodal combination.
4. The dubious accuracy of port statistics and the inability to identify the actual port of physical infrastructure. In many cases, the reported port is the port of duty filing and not the physical port of entry or exit. Electronic filing has increased the amount of administrative port filings thereby reducing the ability to accurately ascertain where goods are physically entering and exiting the U.S.

Since 1993 the BTS has contracted with the Bureau of the Census to provide surface transportation data for U.S. import and export trades with Canada and Mexico. This dataset is referred to as the Transborder Surface Freight Data set. Approximately 98% of imports and 80% of exports are collected, measured by value of goods shipped. Shipments are also reported by weight for imports (but not for exports) from Canada and Mexico. These data are supplemented by Customs data on the number of trucks and trains entering (but not leaving) the US through these landed border crossing ports (monthly and annually since 1994)[6]. BTS also publishes an annual update on the relationship between the nation's multimodal freight transportation system and the international trade it facilitates, drawing data from sources at the Bureau of Economic Analysis in the U.S. Department of Commerce, the Bureau of Customs and Border Protection in the Department of Homeland Security, and the U.S. International Trade Commission [9].

[6] http://www.bts.gov/itt/cross/

International freight data collection in the United States is currently undergoing a major change and improvement. In the near future international trade data will flow to the BTS and US DOT, and some 104 other federal agencies (at last count) via the International Trade Data System (ITDS)[7]. ITDS is a federal government information technology initiative (Initiative IT06) to coordinate, standardize, and ultimately simplify our Federal border clearance and other international trade and transportation processes. It will enhance and replace the Automated Commercial System (ACS) currently used by the Bureau of Customs and Border Protection within the Department of Homeland Security. Traders will submit standard electronic data for imports or exports only once: with the ITDS system serving as a "single window" system through which trade transactions data can flow between private traders and over 100 federal agencies involved in international trade.

Not only is the ITDS data reporting and recording system a much needed time saver for private sector traders, it is a timely response to the federal government's need for much greater visibility of incoming foreign cargos in this time of heightened concern over terrorist actions. To this end ITDS is being set up to provide each federal government agency with interests in trade transactions only the information that is relevant to its mission. Each agency submits a list of data elements it deems key to its operations. The US DOT is developing five different portals into the ITDS: one each for MARAD, FMCSA, NHTSA, FAA, and a BTS supported and maintained portal that will supply freight movement data to other modal administrations within the US DOT (FHWA, FRA).

Linking Data on Domestic and International Freight Movements

A comprehensive freight movement database will need to combine domestic and international datasets. Perhaps the ultimate expression of this idea is the creation of a universal electronic manifest (UEM). Such a manifest would provide essential information for shippers and carriers to manage inventories and logistics, and for meeting the documentation needs of domestic and international trade, hazardous cargo movement, and the growing security needs of domestic transportation. It would modernize existing paper-based waybills and allow cargo tracking across all modes.[8]

---

[7] http://www.itds.treas.gov/sitemap.html

[8] The Surface Transportation Board has regulations (49 CFR 1035) that specify certain requirements for bills of lading on rail and water shipments, and the Federal Motor Carrier Safety Administration has requirements for bills of lading on truck shipments (49 CFR Part 373). These include the weight of the shipment, but may provide no information on the type of commodity or on the true origin or destination points, or on the routing, or on the mode or modes of

Given the institutional as well as logistical requirements for such a system, this must be seen in the context of the current AFDP action plan as a longer term goal. Much of what follows discusses ways to create a federal freight movement database that substitutes for such a UEM in a cost-effective manner that is sensitive to private sector concerns over data reporting and use.

3.3 TRB Special Report 276

One outcome of the 2001 Saratoga Springs conference referred to above was the formation of an expert panel, selected and organized by the Transportation Research Board with funding from the BTS, and charged with developing "A Concept for A National Freight Data Program". This Concept was published by TRB in mid-August of 2003 as Special Report 276 [4]. The panel identified the need for a sustained, multi-year effort in freight data collection at the national level, recommending a central role for the BTS as a data collector/coordinator of program activities. Their proposed data collection program would draw together participants and data sources from a wide variety of federal, state and metropolitan planning agencies, the private sector and academia, and tap into new and potentially less expensive sources of freight data, including the use of informatics-based data collection devices using EDI[9], ITS[10] and other automated data collection methods.

Concluding that an attempt to provide all data to all users of freight data is beyond the scope of any federal data collection program, Finding 4 in TRB Special Report 276 states that "The following data would capture important characteristics of freight movements and meet the major data needs of a wide variety of data users in the public and private sectors:

- origin and destination
- commodity characteristics, weight, and value
- modes of shipment
- routing and time of day, and
- vehicle/vessel type and configuration"

---

transportation employed. The carrier with whom the shipper contracts may be shown on the bill of lading, but that carrier may subcontract all or part of the carriage to other carriers, and those other carriers would not be shown on the bill of lading.

[9] Electronic Data Interchange (EDI) is the electronic exchange of business information between organizations in a structured format.

[10] Intelligent Transportation Systems

Data on intermediate stops and cargo transfers en route (notably for intermodal shipments) is also mentioned as being important, while a commodity characteristic worthy of special attention is the hazardous materials code.

These findings are consistent with other BTS-collected evidence on this topic, and support the notion that AFDP data products should be targeted primarily at transportation planners, analysts and policy makers in federal, state and metropolitan planning agencies, and at the private sector consultants and academics who often process these data for them. The above data should also be useful to private sector businesses and trade associations who track and analyze trends in freight data markets, movements, and their associated costs.
Other recommendations in Special Report 276 found valuable in AFPD development include:

1. the need to focus on the most significant data gaps, especially in the early years of the program – and notably on motor carrier (i.e. trucking) data, and data on intermodal shipments, including the domestic and international movement of containers;
2. the need for a careful transition from current BTS-based data collection efforts, notably the five-year repetition of the US Commodity Flow Survey, to more continuous survey data collection efforts, including the development and testing of different pilot survey instruments;
3. the need to develop an action plan for transitioning gradually to a data program that makes more use of passive data collection methods (i.e. EDI, ITS, informatics);
4. the need to address the thorny issue of data confidentiality in both public and private sector freight databases;
5. the need to assess the role of private sector data providers, notably those who provide a "value added" product by building freight movement data matrices around one or more publicly available datasets (and notably around the US Commodity Flow Survey)[11];
6. the use of specialized statistical and mathematical programming techniques to combine data from two or more of these data sources (termed data synthesis in the rest of this paper); and
7. the provision of training and education in the use of the program's freight data products, and where possible the development of a set of standard survey methods for use by states, MPOs and other organizations considering freight data collection: as a way to help local planners and analysts combine federal databases such as the US Commodity Flow Survey with their own regional data collection efforts.

[11] This topic is developed a little further in Section 4.3 under the topic of Data Synthesis

The panel recommended separate freight survey instruments targeted at freight carriers, shippers, receivers and distributors, as well as at traffic informatics sources. The idea here is to ensure that data are collected from each of the major organizational types involved in freight supply chains, allowing in the process for access to data from currently under-represented freight movement sectors such as retailers, terminal operators, freight distributors and warehousers, and major importers and exporters.

4. Principal Program Activities (Proposed)

4.1 Overview

The AFDP consists of the following five program elements:

- Data Collection
- Data Synthesis
- Data Quality
- Data Protection
- Data Dissemination and Outreach

Major activities under each heading are described below, emphasizing the first six years of the Program.

4.2 Data Collection Activities

The largest expenditure of funds under the AFDP will be in data collection. The first six years of the program will focus on improving the existing US Commodity Flow Survey, researching new approaches to collecting this shipper data, and developing a supporting domestic motor vehicle (truck) activity survey. The use of non-intrusive, informatics based forms of vehicle activity data collection will also be researched. In parallel to these efforts, data on imports, exports, and foreign in-transit freight shipments will be obtained from the Bureau of Customs and Border Protection. Figure 1 shows the major data collection activities proposed. [12]

[12] The dashed lines indicate the potential for collecting some data on import and export shipments as part of the proposed domestic surveys. For example, the current CFS is the only national survey that collects data on export shipment weights as well as their value. Similarly, a survey of intermodal facilities (such as selected seaports) will yield data on both domestic and foreign goods movements. These sort of overlaps should be seen as an opportunity for some cross-checking of different data collection methods.

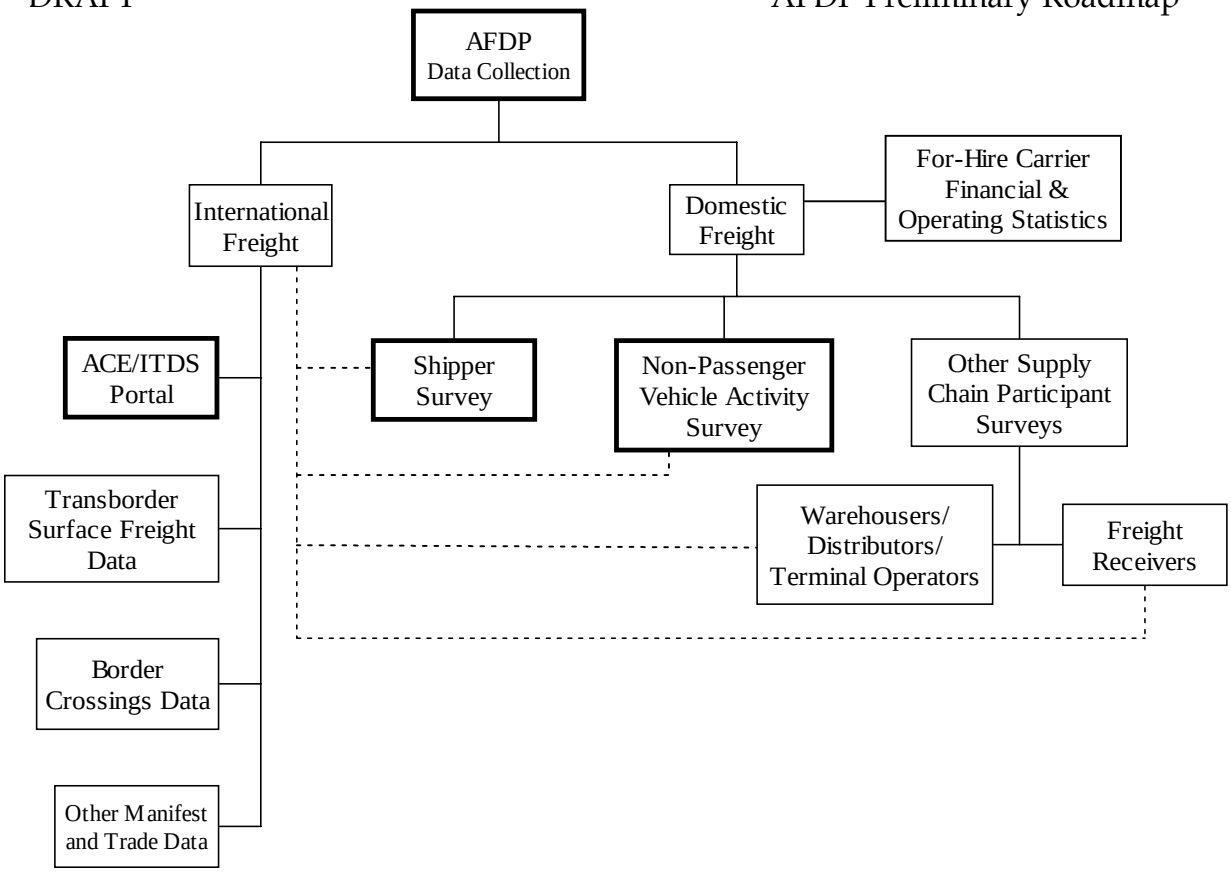


Figure 1. Proposed Principal Data Collection Activities Under the AFDP

The following sections describe the principal options available for, and obstacles to, achieving each of these goals, the technical issues BTS will face in charting the best route forward, and the approach proposed under the AFDP. In each instance nomination of the preferred survey approach is based on:

1. the ability to obtain accurate and robust responses to specific data elements
2. the required sample size
3. the expected sampling response rate
4. the expected survey costs

The movement of freight from source to final consumer can vary from a simple, one stage transfer from supplier to customer (as in the transport of coal from mine to electric utility), to a complex logistics operation involving two or more end-on modal transfers, using a combination of private and for-hire carriage, with possible storage en route. The freight may also "change hands" more than once en route, not just physically but also in terms of ownership. Key players in both single and intermodal shipments not covered by shipper or carrier surveys are freight receivers, wholesale distributors and warehousers, as well as port

(seaport, inland port, airport) and intermodal terminal (truck-rail, truck-water, rail-water) operators. See Figure 2. Also playing an increasingly important role in freight movements are freight intermediaries such as freight forwarders, freight brokers and third party logistics (3PL) operators who may not be involved in any physical movement of goods, but who play a major role in arranging carrier-shipper contracts and shipment schedules.

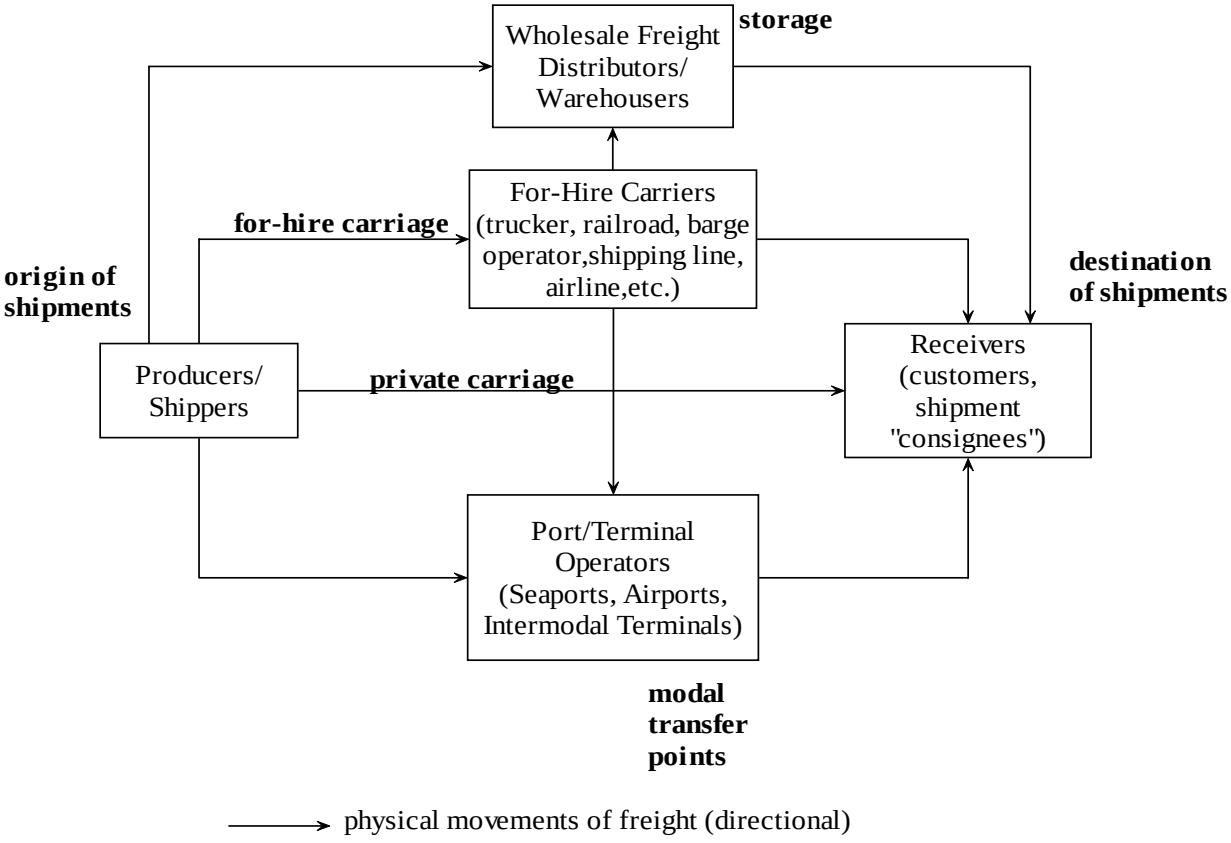


Figure 2. Different Agents Involved in the Physical Movement of Freight

4.2.1 Domestic Freight Movement Surveys

Proposed domestic freight activity surveys under the AFDP, as described below, fall into three categories:

- A regularly scheduled multimodal shipper survey
- A regularly scheduled motor carrier (truck) activity survey (funding permitting)
- Periodic surveys of other participants in the freight supply chain (funding permitting)

The key to accurate and statistically robust data collection begins with the sample frame and the nature and size of the sample taken. Two issues that need to be resolved before data collection can begin are a) whether a survey will be mandatory or voluntary in nature, and b) what the most effective method of sampling will be, given the answer to a). These issues are interrelated since different types of survey lead to different sample response rates.

Another important issue is the frequency with which data are to be collected. How often we need to update certain data elements depends upon how quickly things are changing in the real world. An effective solution to this problem is to move gradually to a system of "continuous" surveys that collect data from sub-samples of respondents on a regular (weekly, monthly, quarterly, annual) basis: allowing at least some statistics to be kept up to date on a frequent basis. Doing so in a cost-effective manner that does not put ongoing statistical programs in jeopardy will require a significant research component to the AFDP. As described below, this involves the design and field-testing of a number of pilot surveys.

4.2.1.1 U.S. Commodity Flow Survey (Expanded)

Purpose: A survey of shipping establishments should be a central component of an AFDP. Currently, the principal freight data collection and analysis program within the BTS is the US Commodity Flow Survey (CFS). The CFS is designed to provide data on the annual tons, ton-miles and dollar value of goods and materials moved within the US, by mode of transportation and commodity class. Emphasis is placed on determining the true origin and final destination of shipments. Modes covered are air, highway, rail, water (inland, intra-coastal and deep sea), petroleum pipeline and intermodal combinations of these. Data on the hazardous nature of the goods shipped as well as shipment size, by weight and value, is also collected. User discussions of CFS data limitations focus on a) its scope, b) the level of detail reported, notably its geographic detail, and c) its timeliness. [13]

Areas for Improvement. The CFS was conducted for BTS in 1993, 1997 and 2002 by the Census Bureau, under Title 13 reporting as a part of the Economic Census. A recent (2002-3) review of the CFS program by a TRB-appointed panel found the CFS to be in widespread use and to play a unique role in providing US commodity flow data [10]. However, it also noted a number of user identified data gaps in the survey and made a series of recommendations for addressing

[13] An additional disaggregation of value would be a measure of seasonal variability in the movement of selected commodities, notably agricultural products.

them. Consistent with these recommendations, the following issues should be addressed by BTS in seeking to improve the value of the CFS to its users:

1. An assessment of the size of the CFS sample
2. An assessment of new ways to survey respondents that are less costly and also reduce respondent burden
3. An assessment of the scope of the CFS and the potential for expansion to cover other industrial sectors
4. An assessment of the costs, benefits, and other factors associated with reducing the time interval between successive surveys
5. An assessment of alternative ways to access and make use of CFS micro-data records
6. An assessment of the potential for new statistical tables and other output supported by the CFS
7. An assessment of supporting methods for collecting data on shipment costs
8. Creation of a forum for CFS data users to provide comments on CFS use and to express their data needs.

The issues of sample size, survey scope and timing are closely linked. The size of the CFS sample has been successively reduced from 200,000 establishments in 1993, to 100,000 in 1997, and 50,000 in 2002. For the 2002 CFS each establishment was assigned a 1-week reporting period during each quarter, for a total of 4 weeks of shipment sampling for the calendar year. Response rates for each CFS have been around 70% (by establishment). All three surveys used a mailout/mailback survey form. Issues to be investigated as part of the sample size debate include reasons for non-response, respondent burden, and method of data collection. Web-based reporting is one potential cost-saving option for further investigation. In particular, the use of a web-based tool that automates the selection of sampled (shipment) records may be one way to improve data quality, as well as reduce respondent burden.

The scope of the CFS is also of considerable interest to current and potential users. The present survey covers mainly the mining, manufacturing and wholesale industries. Today this appears to represent less than three quarters of all goods moved within the country [11]. The CFS does not currently cover establishments involved in farming, forestry, fishing, construction, crude oil production, households, governments and foreign establishments. It also does not cover the rapidly growing and increasingly transport intensive service sector, and misses most retail establishments[14] [8]. The CFS also does not cover the first

[14] In scope retail establishments are limited to the following types: Automotive Parts and Accessories Stores, Tire Dealers, Floor Covering Stores, Other Building Material Dealers, Nursery

stage delivery of imported goods (since it does not survey non-US establishments). To complete the freight picture this means either combining CFS data with data from other sources, or expanding the survey to capture some or all of these missing elements. Both options should be considered under the AFDP.

The 5-year time interval between successive shipper surveys also requires re-evaluation. Could a more frequent survey provide better tracking of changes in the types and spatial patterns of goods shipped, and modes used? Also, could a more continuously applied survey instrument help to spread the substantial costs of CFS data collection ($15 million in 1993, $19 million in 1997, $13 million in 2002) more evenly across a multi-year planning horizon?[15]

A common request by CFS users, especially state and metropolitan planners, is for increased geographic and commodity specific detail. Spatial detail is currently limited to State-to State and top 50 Metropolitan Area-to-Area movements, with confidentiality and statistical robustness of the estimates limiting the release of O-D data to total volume, or at best 2-digit commodity classes. The only cost-effective way to address this problem in most instances is to combine the CFS data with other data sources. In particular, the desire for detailed O-D data suggests the use of county-based freight traffic generation an attraction data to disaggregate commodity and mode based CFS flows, using the latter as State-to-State control totals. Ways in which CFS data can play a more effective role in supporting these synthetic, multi-source data products need to be explored. For example, more detailed CFS trip length tabulations by 3-or 4-digit commodity class, but with limited geographic specificity, would be useful for modeling commodity flows when combined with county-specific economic or freight activity totals.

Similarly, greater commodity specific detail in value/weight ($/ton) statistics would also prove useful. This sort of synthetic commodity flow modeling is likely to increase in the future, as States and metropolitan regions develop their own freight analysis models.[16] Two recent examples of this sort of county-based flow modeling already exist on a national level, using a combination of public

---

and Garden Centers, Office Supplies and Stationery Stores, and Electronic Shopping and Mail-Order Houses.

[15] It is not clear that these costs capture the full costs of computing ton-miles: as simulated by Oak Ridge National Laboratory for the 1993 and 1997 Surveys, and by ORNL/BTS staff for the 2002 CFS.

[16] State-level studies using CFS data and other sources to produce inter-county commodity flow include studies for Indiana [14], Ohio [15] and Wisconsin [16]. A number of recent State and sub-State regional studies, including studies in Florida and Virginia, have also used the Reebie Asssociates' TRANSEARCH database in this context, which also incorporates CFS data.

and private data sources to generate commodity trade volumes [12] and both commodity and equivalent vehicle flows [13]. Ways in which the CFS can improve its support for such data products deserve further study. This includes an examination of alternative ways to gain access to both the latest and historical CFS data, for users both within and outside federal government. [17]

To reduce respondent burden the CFS also does not collect data on routes, travel distances, times or costs. Instead, it relies on least impedance routing algorithms to infer the routes taken [8]. These impedances are modeled as a weighted combination of distance, time and shipment costs. There is currently no single, nationally sampled data source from which to generate representative truck transportation costs that can be broken down on the basis of origin-destination channel and commodity class (at least, not in the public domain). There are, however, a large number of data sources on vehicle capital, operation and maintenance costs, including many that are readily accessible via the World Wide Web. These sort of cost data are often used in freight flow models to carry out project-specific economic impact analyses. They are also used in flow modeling to influence the estimation of both base year and forecast year commodity flows (see section 4.3 below). A useful AFDP product would be one that either provides directly, or provides a method for computing, representative estimates of such O-D /commodity/vehicle equipment based movement costs, along with their statistical properties.

Data Collection Methods: Two options for future shipper surveys recommend themselves. Both retain the form of self-administered questionnaires as the preferred approach to surveying freight shipping establishments [17]. One option is to continue the current approach of collecting a CFS every five years, with the obvious benefit of its being tied to the mandatory reporting and concurrent data collection activities associated with the Economic Census. The other is to develop a CFS based on a continuously applied annual sample of shippers. The value of considering such an approach lies in its potential for addressing a number of the above-discussed issues, i.e.

- the potential for more timely updates to data products
- the ability to target a smaller sample size each year, allowing for a more streamlined annual budget allocation
- the ability to create an ongoing time series of continuously updated freight movement trends and statistics

[17] For example, the Surface Transportation Board on occasion grant access to the micro data underlying the Railcar Waybill Sample for within DOT studies, subject to strict disclosure rules on its subsequent use.

- the ability to revisit past data in light of subsequent findings to produce more accurate, revised estimates;
- the potential for economies of scope, using relevant, supplementary surveys at relatively marginal cost: including the funding of such supplements by other federal agencies or branches of DOT when interest is peaking on a particular topic (such as on the movement of freight within a particular industrial sector or geographic corridor)
- the potential for developing robust time series data for more aggregate (geographic, sectoral, modal) movement statistics, while gradually building up a more up-to-date representation of the data for use at more disaggregate, notably O-D specific, levels.
- the potential for introducing more timely improvements in data collection and database management

The increased flexibility offered by a continuous sampling approach, it must be noted, leads to more complex data management issues, and requires greater technical sophistication in developing the data expansion process. Technical issues BTS will need to resolve include the design of a sampling frame that covers multiple years of data in order to capture the complete set of US freight movements.[18] An expanded scope for the CFS will both benefit from and complicate this task. One benefit of using a larger sample for the 2007 CFS (see below) is that it can provide a solid base from which to kick off such a continuous sampling procedure.

Proposed Activities. Given the time it takes to establish a new survey instrument, CFS activities under the AFDP over the next six years are usefully separated into two supporting lines of effort:

- Preparations for and implementation of the 2007 CFS,
- Pilot surveys and other research and development activities in support of a continuous and expanded, post- 2007 shipper survey instrument.

A challenge for program development will be to create a funding profile that can handle both of these activities. Section 5 below (cf Figures 5 and 6) contains an example tasking profile for both CFS 2007 and a post-2007 continuous survey replacement.

[18] A review of recent continuous survey instruments both within the US and abroad, and for passenger as well as freight movement surveys will shed further light on this. Battellino and Peachman, 2001, for example, discuss a 3-year continuous sampling frame involving the successive reduction in relative standard errors on trip estimates at 4 different levels of geography [18].

2007 CFS Activities: These activities center on a CFS whose scope (i.e. economic sectors covered by the survey) remains the same as in 2002, with the exception of any non-government hazardous materials shippers who are currently out of scope:

- sample size and sample design issues
- an improved understanding of non-response: causes, cures and potential for biased results
- the possible use of new, including mixed data collection methods (electronic filing of questionnaires, mail-based filing, etc.)
- a possible expansion of the sample (or an "over-sampling") to capture a higher percentage of hazardous materials shipments, as well as the addition of questions about the vehicles used and routes taken in transporting these materials.

Early questions to be resolved include: 1) the determination of how large a sample is needed to ensure robust O-D flow estimates at both the State-to-State and largest 50 Metropolitan Area-to-Area levels, [19] and 2) how these estimates and other published table entries were affected by the drop in the 2002 sample size? It is suggested that BTS research the use of electronic filing and web-based questionnaires for those establishments that wish to use them. Specifically, can these methods of data collection help BTS to cut survey costs, improve response rates, and reduce respondent burden?

While the CFS provides the O-D geography involved in specific truck and truck-intermodal movements of hazardous materials (hazmat), it currently allows only summary tables to be created for these sampled shipments. In these times of heightened concerns over the security of the transportation system a more in-depth look at these shipments seems warranted. Over-sampling of these shipments ought to be possible based on the identification of major hazmat shipping establishments. Given the restrictions placed on the transport of such materials, these shippers are also likely to be more informed about the routes, vehicles and containers used to move these materials. BTS should research the cost and effort involved in expanding the 2007 CFS sample to include additional hazmat shippers, and to add additional questions to the survey forms sent to these shippers. These data could then be compared and possibly combined with other hazmat truck traffic and O-D data sources, notably the FMCSA's Detailed Hazardous Materials Incident Reports, [20] as well as data on hazardous waste transportation collected through the DOE's National Transportation Program.[21]

[19] Prior to 2002, the CFS sampling frame was based on National Transportation Analysis Regions.
[20] http://www.fmcsa.dot.gov/safetyprogs/hm.htm
[21] http://www.ntp.doe.gov/

Post-2007 CFS: It is also suggested that research be carried out to determine if the current approach to data collection should be replaced with an annual and continuously collected CFS, and expanded in scope to capture freight shipments by establishments involved in farming, forestry, fishing, construction, retail, government, services and the utility industry.

In expanding the types of establishment covered by the CFS, the BTS should enter into discussions with the US Department of Agriculture (USDA), which collects sample data on farm-based transportation of goods (estimated crudely in [11] to be about 7% of tons moved and 2.3% of the freight moved within the US by dollar value), and with the US Department of Commerce (DOC) for advice on sampling the government, services, construction and possibly also the retail freight movement sectors, as defined under NAICS[22]. A first step is to prioritize these currently missing shipments by economic sector. Annual survey activity of this type would begin in calendar year 2010. A significant R&D effort is needed to develop both the sample frame and best method for collecting data from such a survey. This will require some time to develop. The size of the annual sample for such a survey remains to be determined. Funding permitting, it might be on the order of 20,000 to 25,000 establishments per year: with possible over-sampling of different types of shipments in different years.

4.2.1.2 Non-Passenger Motor Vehicle Activity Survey (New)

Purpose: As a partner with the above-described shipper survey, the main role of a motor vehicle activity survey is to gain a much better and quantifiable understanding of the relationship between truck operating characteristics, truck movement patterns, and the commodity flows they support. The survey instrument should provide detailed information on the sort of highway routes taken, the geography and timing of pick-up and drop activities, and the relationship of ton-miles to partial, fully loaded and empty vehicle miles. National and regional statistics that such a survey should be able generate, and significantly improve upon, include:

- total truck miles (including empty and partially empty miles)
- highway impact (gross vehicle weight x distance transported)
- transport output (cargo weight x distance transported)
- truck productivity (truck miles per ton-mile)
- truck trip circuity factors (by commodity class and vehicle type)
- truck emissions and fuel consumption (by commodity class and vehicle type)

[22] North American Industrial Classification Series.

- (inputs to) the per-mile and per ton-mile costs of freight movements (by commodity class and vehicle type)

The size of such a motor carrier vehicle activity survey remains to be determined. The discussion below assumes nationwide coverage of the commercial truck fleet as defined, for example, by the VIUS. Targeted annual surveys focusing on specific economic sectors, freight corridors, or vehicle types offer less costly options that should also be explored.

Data Collection Methods: The principal options for sampling truck activity patterns have been identified as:

1. Vehicle-Based Sampling
2. Vehicle/Driver Intercept Surveys
3. Vehicle Tracking Surveys

Each of these approaches has its pros and cons, as summarized in the Table shown Appendix 1. (The table also provides details from recent examples of each approach).

Vehicle based sampling, using a trip diary format to collect a sample of a vehicle's pickup, drop and repositioning activity over the course of a sampled week has a number of things to recommend it. Among these are the ability to draw a representative sample frame based on vehicle registration data, and the opportunity to get data on not only the vehicle but also its cargo and driver characteristics. The burden of filling out trip diary information can, however, present significant response problems for voluntary trip diary surveys. Since the mid 1980's mail-in/mail-back survey instruments of this type, as used by state, regional and metropolitan planning agencies, have obtained response rates in the 8% to 26% range: while telephone based, or combined telephone and mail based surveys have recorded improved but still limited response rates, in the 24% to 43% range [7,19]. The voluntary Canadian Vehicle Survey of 2000 and 2001 [20] got a similar response rate (measured in truck-days) on its medium and heavy truck trip logs (weekly dairies). In contrast, the mandatory Continuous Survey of Road Goods Transport, carried out annually by the United Kingdom Department of Transportation since 1994, achieves a response rate of around 95% (87% useable). This survey collects data on some 18,000- 20,000 vehicles per year in the form of a mail-out/mail-back weekly truck travel diary [21].

In terms of both survey design (notably the length) and the sampling approach taken, much depends on whether BTS obtains mandatory authority for the survey under ending DOT reauthorization. Of the 92 million or so vehicles classified as "trucks" in the United States, some 70% of these are primarily

personal use vehicles, notably light duty pickups, minivans and sports utility vehicles. This leaves a primarily non-passenger truck universe of around 27.6 million vehicles engaged in commercial activities: of which a little over 8 million are trucks with more than two axles (including some 2.1 million combination trucks and tractor-trailers). Sampling to cover the same 92 million non-passenger truck universe, at a rate similar to that of the 5-year VIUS, would seem to require an annual sample size of around 27,000 vehicles, if sampled over a 5 year period: or over 45,000 vehicles per year over a 3 year period. [23] A noticeably smaller sample would presumably be able to cover the 27.6 million principally commercial use vehicles: but separating out these latter trucks from the non-commercial vehicles may be difficult, i.e. which light duty trucks are actually in commercial use? Without mandatory authority to back up such a vehicle activity survey, it also seems unlikely that the response rate achieved would warrant the costs associated with a lengthy, trip diary based survey instrument[24]

An option that may be worth exploring here is a tie-in between an AFDP-supported continuous truck-dairy based survey and the VIUS, making use of the latter's sample frame. (For the 2002 VIUS the Census Bureau purchased a sample frame generated by R.L. Polk & Co). Vehicle activity diaries might be collected on an annual, rotational basis, possibly with different types of vehicles selected each year for diary completion. The characteristics of the vehicles reported in these travel diaries might then be tied to operating characteristics in the larger, 5-year based vehicle sample.

A vehicle/driver intercept survey is another data collection option. One benefit of this type of survey is a typically high response rate. A good example of this type of survey is the Canadian National Roadside Survey [22] carried out in 1991, 1995 and 1999. This is a voluntary driver survey, with a 7 day continuous truck count for a representative week, based on random intercept sampling at 238 sites on the Canadian road network. The 1999 survey achieved an 88% response rate from drivers. Benefits of this approach include the link between vehicle (and driver) characteristics, the commodities being moved, and the routes being driven: although commodity details may be less easily obtained and/or less accurately reported than from shipper surveys: while drivers may not know the true origin or destination of the freight. There is also the cost of traffic delays during driver interviews to consider (and safety issues for interviewers).

Such intercept surveys pose a number of design questions, including the selection of routes to sample from, and whether to use changeable (e.g. annual,

23 The UK vehicle sampling rate is much higher:, but with no Commodity Flow Survey to support it.

24 VIUS costs for sample frame development, data collection and report generation are on the order of $8 million dollars.

bi-annual) route sampling. Sampling on a national basis is problematic for a nation as large, and with a highway network as dense, as that in the United States. Besides the sheer size of the exercise needed, randomly selected interview locations may create safety and traffic delay costs: while data collected at weight stations or other off-road locations may introduce non-randomness and bias into the process. Establishing a mechanism for data collection, and necessarily one involving the cooperation of state and local traffic agencies, suggests that this approach is better suited to the sampling of specific high volume truck corridors within selected states/regions, as an adjunct to other regionally sampled data collection methods.

Vehicle tracking surveys represent a third data collection option. They make use of the latest in telecommunications and electronic tracking and reporting technology, notably the Global Positioning System (GPS) of satellites. [25] Implementation costs alone suggest the use of such a survey instrument on a voluntary basis in the near term, with participation by a voluntary sample of motor carriers. An example of such a survey was carried out in California in 1997-98 [23]. [26] Data elements collected per truck included:

- Truck routes and stops en route
- Timing of starts and stops
- Travel times and speeds
- Truck idling and soak times[27]

Benefits of this sort of data collection include mileage data that can be disaggregated by highway functional class, time of day, type of business use and by urban versus rural location; data that are useful for fuel consumption and air quality studies; and data that are easily tied to GIS mapping software.

The table in Appendix 1 provides additional details and lists additional pros and cons for each of the three sampling approaches.

Proposed Activities. The relative costs, respondent burden, and institutional barriers to be overcome in carrying out either a large vehicle intercept or vehicle tracking survey suggests an initial focus of activity on a vehicle license-based sample, if nationwide coverage is required. However, this pre-supposes mandatory authority to collect such trip diaries. A continuous, annually

[25] http://www.navcen.uscg.gov/

[26] Data based on a sample of some 140 trucks used by the 24 trucking firms that agreed to take part (of the 236 firms initially contacted). Data was collected from continuous vehicle tracking for 4 months, then for a further 8 months during 1997-98.

[27] " soaks" refer to periods when the vehicle's engine in turned off between operations.

sampled vehicle ownership survey seems most preferable, along the lines of the United Kingdom survey described above, and as a partner to an eventually continuous shipper survey. The sample frame for the proposed survey would be vehicle license based. Sample size should be determined on the basis of obtaining robust estimates for the above-listed motor freight activity measures. As with the proposed shipper survey, over-sampling of different vehicle types (e.g. double trailer trucks, hazmat shipments) might be employed in different years to address specific policy relevant issues. Figure 3 shows the sort of timeline required to develop a new survey instrument of this type. A four year development and first application schedule is shown. Subsequent years would continue the annual data collection exercise while also seeking to improve upon it.

Given the time, effort and level of funding required to support this new survey activity, a number of options should be researched, including:

- an alternating, bi-annual shipper/motor vehicle survey program (starting in 2008 with the first motor carrier activity survey)

- development of a functional linkage to the 5-year VIUS, with the continuous vehicle activity survey focusing on sample designs that emphasize the vehicle routing, backhauling and repositioning, and truck operating speed aspects of the data (possibly by sampling vehicles using specific high volume highway corridors or high volume freight gateways on a rotational basis)

Failure to obtain mandatory reporting authority for a vehicle activity survey suggests further exploration of the options for estimating the trucking measures of performance listed above. Options include:

- the use of a much simpler, and less burdensome, survey questionnaire , possibly targeted at specific parts of the motor carrier industry in a given year

- the use of data from either, or both, vehicle intercept and vehicle tracking surveys linked to geographically or sectorally specific project assessments, (and whether or not supported initially with AFDP funds).

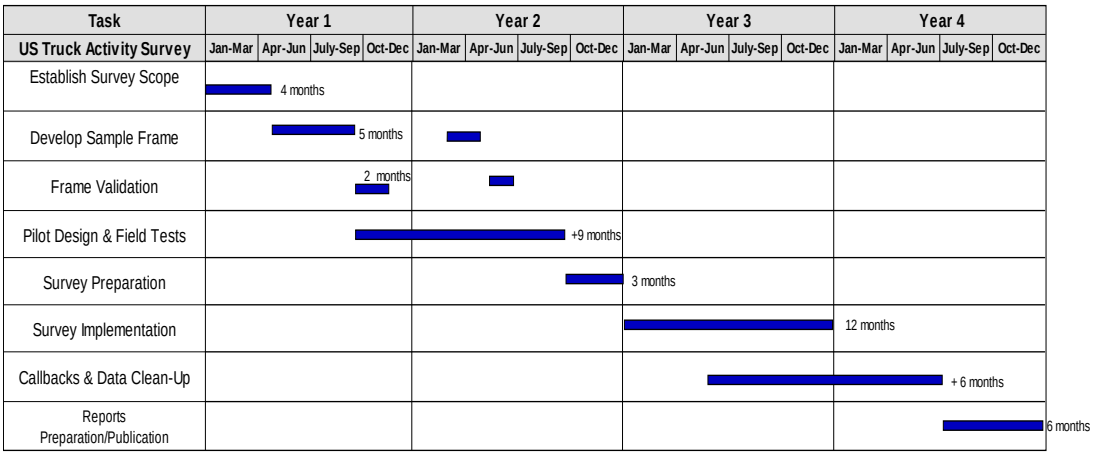


Figure 3. Example survey development timeline: motor carrier activity survey.

4.2.1.3 Surveys of Other Supply Chain Participants

Purpose. While AFDP data collection efforts are focused largely on shippers and motor carriers as the best combined sources of data on the physical characteristics of the majority of freight movements, additional freight agents can also provide useful data on freight throughputs, delivery times and costs (cf Figure 1). In addition to freight receivers, freight warehousers, freight distributors and terminal operators can all play an important role in shaping the spatial patterns of goods movement.

Data Collection Methods. Surveying any of these agents presents a challenge. While a survey of freight receivers can reveal a good deal of information about the freight service being offered, there are often many receivers for the products of a single shipper, while a single carrier often serves many receivers. Identifying and sampling them is therefore an added challenge. Many receivers also may not be aware of the means (origins, routes, equipment) used to deliver the goods. Obtaining operational data on freight terminals is also a challenge, with frequent turnover in the size, operation and even location of terminals on even an annual basis. The Census Bureau collects nationally representative statistics on warehousing and for-hire transportation industries, but expansion of this sample frame to capture operational details from specific facilities would represent a new survey activity. Obtaining terminal volume (freight throughput) and capacity (maximum throughput) is similarly problematic given the range of functions and variety of terminal types used. Extensive data on US seaport and airport activity is available from the US Army Corps of Engineers and the Office of Airline Information within the BTS, respectively. BTS has also recently enhanced and expanded the truck-rail and truck-water intermodal terminals database associated with the 1997 and 2002 Commodity Flow Surveys.

An alternative approach that may also be worth considering is the use of "supply chain surveys", involving a mixed sample drawn from each of the above types of establishments, i.e. from shippers, carriers, distributors, terminal operators, receivers and also freight forwarders. More than one US experiment in the tracking of complete O-to-D supply chains is currently underway, making use of recent developments in electronic reporting. For example, various branches of the DOT (FAA, OST, FHWA) along with the State of Illinois, the American Transportation Research Institute and participating manufacturing, trucking, and airline companies are testing the use of an Internet-based electronic supply chain manifest (ESCM) for transferring information between each of the players involved in source-to-destination freight moves through Chicago's O'Hare and New York's JFK Airports.[28]

An experiment with this type of data collection was also recently tried in Europe, based on more traditional data collection methods involving both face to face and telephone interviews with the different supply chain participants [24]. This sort of survey offers a greater understanding of the full logistics costs involved in moving freight from source to destination. However, it also requires a potentially complex, multi-stage sample design. Both a benefit (and an added complication) of such a survey is the collection of data on who is responsible for moving the freight at each stage, and what institutional arrangements (e.g. between carrier and shipper, shipper and broker, shipper and receiver) have been made to facilitate this. Recent European (French and Dutch) experience with this sort of survey identifies a number of problems with achieving good response rates, and with higher costs per successful response than with traditional shipper surveys [24]. In particular, the probability of getting a complete description of complex, multi-actor supply chains proved to be low. How such an approach might stand up under different sampling designs and under mandatory reporting requirements is currently unclear.[29]

Proposed Activities. With the above points in mind, it is suggested that initial AFDP efforts follow two lines of development:

- Continue existing BTS efforts to develop a nationwide intermodal terminals database[30] focused on the identification of such terminals' ownership, geographic location, operating functions, types of freight handled (by commodity, modes, carriers, origins and destinations), freight

[28] http://ops.fhwa.dot.gov/freight/pp/Air%20cargo.pdf

[29] Even if the average response rate is 75% at each stage in the supply chain (a little higher than the rate experienced by the CFS) a three-interview chain would yield only a 42% successful completion rate.

[30] http://www.bts.gov/gis/ntatlas/facilities.html

throughput capacities and reported annual and seasonal freight throughputs.
- Research methods for capturing the physical and institutional components of complete freight movement supply chains.

The terminals database should include all major terminals within US seaports, airports, and all large truck-rail intermodal terminals. Large truck-only transfer (i.e. break-bulk, freight consolidation) terminals should also be included. Determining what size of terminal to include in the database is an issue, given the number and cost of identifying and characterizing these facilities. Distinguishing separate terminals within, for example, a large seaport, will require identification of terminal operator and/or terminal owner. Co-operation with State DOT's in this effort would prove valuable, and probably essential, to success.

Methods for supply chain tracking should focus on the use of web-based informatics and EDI technologies. Attention should also be paid to the most successful means of developing public-private agency partnerships in the collection and use of such data. In terms of the broader AFDP goal of understanding and quantifying US freight movements, this activity should focus on the analysis of freight passing through large ports, intermodal terminals and warehousing/breakbulk facilities (e.g. "truck villages").

4.2.2 International Freight Movement Data

Purpose. The BTS publishes and disseminates a number of data products dealing with international trade and the freight movements this involves.[31] The focus of the AFDP in the area of international freight movements should be on getting more out of ongoing federal data collection efforts. Unlike domestic freight data sources, the sources of international freight data are closely tied to trade data, with US Customs collecting information on effectively all of the goods entering or leaving the nation and bound from or to other countries. With concerns over both tariff fees and, increasingly, national security, there exists an established tradition of collecting data from and about each of the participants in the freight movement supply chain: shipper, carrier, broker, forwarder and receiver.

The challenge for the AFDP is to access and make the most use of this bill of lading and freight manifest data, for the purposes of developing national freight flow matrices in terms of tons, dollars and vehicle/vessel units, broken down by mode, commodity, route, border port, origin and destination. There are a

[31] http://www.bts.gov/itt/

number of gaps in current international freight movement datasets that make this a difficult proposition. These include less than complete coverage of export shipment weights and vehicle counts, limited geographic detail on the true origins and destinations of merchandise trades, limited modal detail on the domestic portions of international shipments, and no consistent means of linking the domestic (i.e. within-US) and foreign portions of these shipments. These problems make it difficult to construct both aggregate, national statistics on freight movement volumes, as well as build a more detailed picture of origin-to-destination freight movements by either mode or commodity.

Data Collection Method: The AFDP will draw most, and perhaps eventually all, of its international freight data from the Bureau of the Customs' International Trade Data System (ITDS). These data will come into a BTS data portal via the Customs Service's Automated Commercial Environment (ACE)[32] software system. ACE is an ongoing, multi-million dollar upgrade in electronic trade data collection and reporting. BTS will be one of over 100 participating government agencies (or PGAs), and one of five agencies within the US DOT, drawing specific data elements from ACE. It is planned to use this data portal to collect near real time freight manifest information based on 104 different data elements covering the commodity, tonnage, value, timing, initial origin and final destination of each shipment, as well as data on the type of shipper, carrier, freight forwarder, and any stops at specific US and foreign seaports/airports, Foreign Trade Zones or border crossing stations en route.

Digital data on ground-based transborder freight movements from/to Canada and Mexico are expected to begin coming into ITDS in the Fall of 2004, with waterborne and airborne freight imports and then exports becoming available by the end of fiscal year 2007 [25]. Principal AFDP activities associated with acquisition of these data involve the initial set-up of the ACE/ITDS data portal, and subsequent, ongoing operation and maintenance activities associated with regular extractions of specific data elements using the ACE software.

Proposed Activities: Current activities involve ongoing acquisition and processing of the Transborder Surface Freight Data set and the Border Crossings data BTS receives from Customs and the Census Bureau. Once ITDS comes online the principal AFDP activity is to extract and make the most use of the data supplied through the ACE portal. This involves the following activities:

- transitioning from current databases to sources of data within the ACE/ITDS data environment
- developing methods for filling the existing gaps in current data

[32] http://www.customs.treas.gov/

- developing publicly disseminated reports based on ITDS data
- making this information readily available to the other modal agencies within DOT
- using the data to examine the supply chains associated with selected international freight movements

An example of an ongoing study of international supply chains is the Transportation Security Administration's Operation Safe Commerce (OSC) initiative. [33] This and other initiatives should be monitored, and joint activities proposed where innovative data collection and analysis opportunities exist.

4.2.3 Informatics: Use of Passive Data Collection Methods

Purpose. The growing presence of many different kinds of real time information gathering technology means that future traffic data collection is going to make greater use of non-survey based approaches. These methods will hopefully one day prove less expensive than traditional survey methods.

Data Collection Methods: Existing technologies include the following [26-29]:

- *Active Roadway Sensors*: Fiber Optic Sensors, Inductive Loop Detectors, Magnetic Sensors, Piezoelectric Sensors, Pneumatic Road Tube, Weigh-in-Motion (WIM) Sensors;
- *Passive Roadway Sensors*: Infrared Sensors, Microwave Radar, Passive Acoustic Array Sensors, Ultrasonic Sensors, Video Image Sensors
- *On-Board Sensors:* Bar-Code Scanners, Microchip-Based Smart Cards, Radio Frequency Identification Devices (RFID) and Remote Intelligent Communication (RIC), Smart Active Labels, Satellite/GPS-Based Vehicle Tracking.
- *Wide-Area Sensors*: IKONOS satellite imagery, Light Detection and Ranging (LIDAR), small plane, helicopter and uninhabited autonomous vehicles/micro aerial vehicles (UAVs/MAVs)

New technology such as Radio Frequency Identification Devices (RFIDs) promises considerable advances in vehicle, vessel, container and package tracking [30]. RFID employs radio frequency communications to exchange data between a portable memory device and a host computer. An RFID system typically consists of a tag or label containing data storage, an antenna to

[33] The initial initiative tracked a constantly monitored container shipment between Eastern Europe and New Hampshire, fitted with onboard tracking sensors and door seals. A second phase of OSC will distribute grants to the Ports of Los Angeles, Long Beach, Seattle, Tacoma and the Port Authority of New York and New Jersey to identify specific supply chains along particular trade routes and to analyze every aspect of the supply chain from packaging to delivery for vulnerabilities. http://www.tsa.gov/public/display?theme=88&content=531

communicate with the tag, and a controller to manage the communication between the antenna and the computer. An RFID tag can be embedded in a package or placed on a person. Combined with remote intelligent communication (RIC) technology, RF-based wireless reporting can be used to track the location, condition, and content of goods at every stage in a product's supply chain, and do it in near real time. This includes the emerging technology of Smart Active Labels (SALs) which use RFID tags containing an internally powered microchip linked to an antenna for wireless reception and transmission purposes [31]. A read/write mode, suitably powered RFID can be used as a dynamic, electronic cargo manifest. The potential for increased cargo security alone is going to bring this sort of "smart tag" technology into the mainstream for freight and inventory management.

Another, increasingly widely used tracking technology is GPS. The commercial component of the Global Positioning System (GPS) is a worldwide radio-navigation system formed by linking together 24 orbiting satellites and their network of ground stations. Vehicle or cargo tracking down to a few meters is already possible, with further spatial refinements (down to centimeters) under development.

A potentially significant addition to these different types of real-time data collection, notably for congestion analysis, is the use of micro aerial vehicles (MAVs), which seem likely to bring down the price of surveillance at major traffic intersections or along major traffic corridors. Here the potential for freight flow analysis would be in the combination of traffic count data from these aerial devices with surveyed of O-D movements: notably as joint inputs to the "link O-D models" discussed in Section 4.3 below

Proposed Activities: The potential for largely automated freight data collection seems obvious, given enough time and resources to develop these informatics technologies. Less obvious, and in need of study, is the use to which this information can and should be put by public agencies. US experiments with an electronic supply chain manifest (ESCM) discussed above suggest that information technology (IT) applications that benefit private freight movement agents will advance quickly. Similar experiments with different freight IT applications are ongoing in Europe [32]. AFDP activities should focus on identifying how these different technologies might be used to develop aggregate statistics for use by participating public agencies, notably in the estimation of hourly, daily and (through aggregation) seasonal O-D freight flow volumes. There is also considerable potential for monitoring and measuring the travel speeds and en route delays associated with location specific truck, rail and waterborne commerce movements.

In researching these technologies the AFDP should keep track of the latest development at other federal agencies: and in particular the Federal Motor Carrier Safety Administration's (FMCSA) Commercial Vehicle Information Systems and Networks (CVISN) program [34]. A goal for CVISN is the integration and automatic processing of data on the carrier, vehicle, driver and cargo associated with domestic or international commercial truck trips. It involves accessing data from a number of existing sources, including data from trucks moving between States under the International Registration Plan (IRP), and systems for reporting of oversize/overweight movements, carrier registrations, and hazardous material credentials [33]. A CVISN link to the Automated Commercial Environment (ACE) discussed in section 4.2.2 is also planned. How the AFDP can benefit from this data integration activity needs to be explored.

4.3 Data Synthesis Activities

Purpose: Given the impossibility of funding all of the surveys for which data would be useful, full advantage must be taken of methods for combining data from existing sources, to produce reliable and robust estimates of otherwise missing data elements. A dictionary definition of a synthesis is " a complex whole made up of a number of parts united".[35] State and metropolitan area transportation planners have for decades made use of so-called "synthetic" freight and commodity flow matrices, using the sort of modeling techniques discussed below. And at the national level the need to combine different data sets is even more of a requirement as soon as detailed O-D flow geography is required.

To support the AFDP's various data users, data synthesis needs to occur at more than one level of data resolution. First, there is a need to develop and maintain (update annually) an internally consistent and complete estimate of annual US domestic and international freight movements, by major commodity and mode of transportation. This involves the piecing together of a complex mosaic of databases, including the different modal datasets listed at the beginning of Section 3.2. As a minimum, this effort should be able to produce estimates of the annual, nationwide number of tons, dollars, and ton-miles moved by each of the major modes of transportation. Such an effort would build on previous BTS-funded research in this area [11].

In tackling this issue, the AFDP can also provide a service to its users by promoting, and where necessary developing, the use of accurate "cross-walks"

[34] CVISN is the commercial vehicle operations part of the DOT's National Intelligent Transportation Systems architecture. It includes standards for communications technologies such as electronic data interchange (EDI) and dedicated short range communication (DSRC).
[35] The Oxford Universal Dictionary on Historical Principles (3rd ed., 1955)

between the different commodity coding schemes shown in Table 2. This activity includes continuation of BTS's International Transportation Data Harmonization (ITDH) website, as an aid to the exchange of information between different countries.

Table 2. Principal Commodity/Industry Classification Systems:

Classification System:	Data Base (Agency):
Standard Transportation Commodity Codes (STCC)	Rail Waybills (Association of American Railroads)
Standard Classification of Transported Goods (SCTG)	1997 Commodity Flow Survey (US DOT/Bureau of Transportation Statistics)
Waterborne Commerce Statistics Center (WCSC) and Lock Performance Monitoring System (LPMS) commodity codes	Waterborne Commerce Statistics (US DOD/Army Corps of Engineers)
Harmonized Commodity and Coding System (HS) and Harmonized Tariff System Annotated (HTSA)	Foreign Trade Statistics (US ITC/ Office of Tariff Affairs and Trade Agreements)
North American Industry Classification Series (NAICS)	County Business Patterns (US DOC/Census Bureau)
Standard International Trade Classification (SITC)	United Nations

At the more disaggregate level at which specific O-D flow volumes are measured, some level of synthesis is also required within each of the major transportation modes: and most notably within truck freight movement matrices. Here synthesis activities may also take on the added job of providing estimates of missing data items, notably cell values within multi-dimensional movement tables.

Often the use of two or more data sources to create freight or commodity flow matrices involves trying to add either spatial or sectoral (industry, commodity class) detail by breaking down more aggregate matrices. A good example of where data synthesis can and should play an important role is the recent BTS Section 5115 International Trade Traffic Study [34]. This study combined data from the CFS, the Public Use Rail Carload Waybill, US Foreign Waterborne Transportation Statistics , US Imports and Exports of Merchandise, Transborder Freight Data, County Business Patterns and the Port Import and Export Reporting Service(PIERS). The data was used to develop annual truck freight movement estimates between US border crossings, US seaports and US counties, by weight and value: resulting in State specific annual highway ton-mile and value-mile estimates that it was hoped would be useful in the development of highway funding apportionment formulas.

In the future, such a study of freight flows could be enhanced in a number of ways, by making use a number of the data improvement proposals discussed above. These include:

- adding freight movements from economic sectors not covered in the CFS (cf section 4.2.1.1);
- using better motor carrier data, such as that proposed in section 4.2.1.2 above, on truck pickup and drop practices, partial and empty loads, types of vehicles used, and the use of preferred routes: to better estimate ton-miles;
- providing a more detailed set of O-D shipment data for imports and exports via the ITDS system (cf section 4.2.2);
- using the latest highway specific truck count/ITS data to verify major routes taken (cf Section 4.2.3).

All of these improvements will be needed if DOT is to establish a set of statistically accurate highway freight O-D flow estimates for the US, on a county basis for the purposes of future planning studies. These matrices could then serve as either a starting point, or point of reference, for many State and inter-metropolitan area studies.

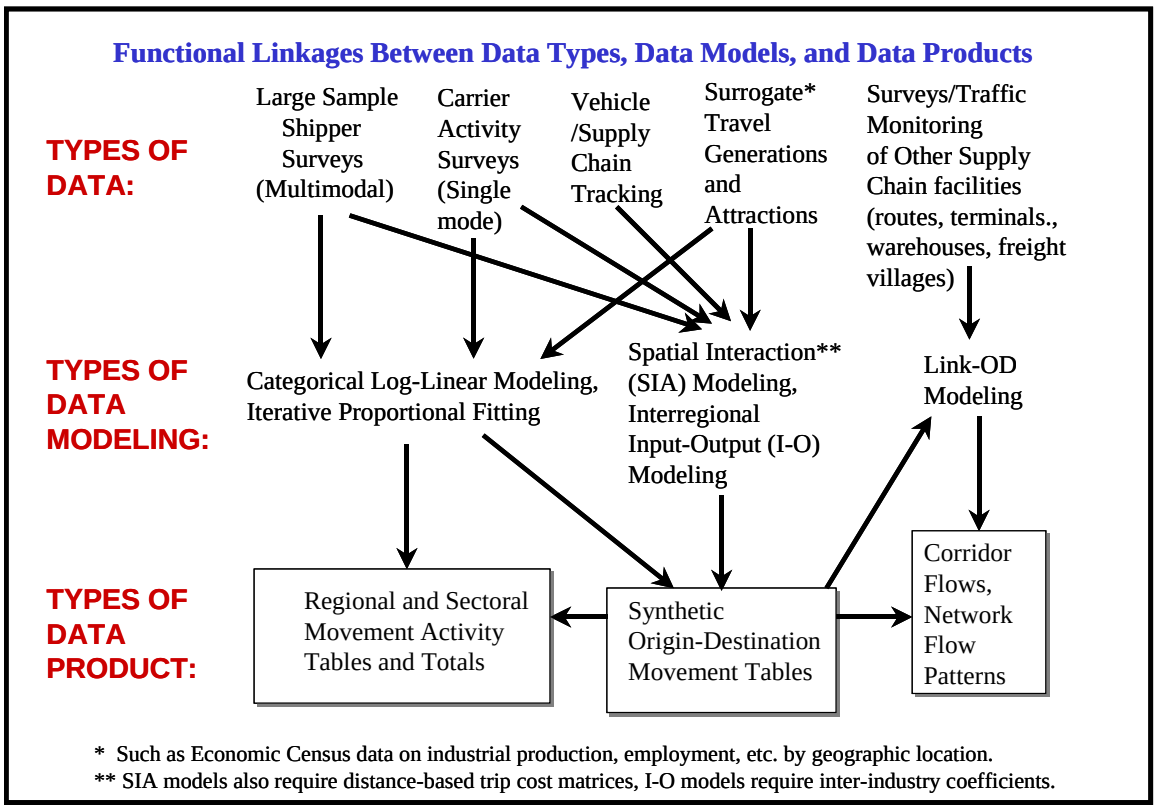
Candidate Analysis Methods. A number of methods are useful for either filling missing elements in movement tables or for allowing further disaggregation of existing movement matrices, reflecting a well-established tradition in spatial data analysis. In most cases this involves developing joint probability distributions from two or more marginal distributions, with different datasets providing the different margins. These methods include [5, Chapter 4]:

- log-linear/maximum likelihood modeling to fill gaps in categorical data tables
- generation of synthetic O-D tables using spatial interaction models
- disaggregation of goods movement tables using inter-regional input-output models, and
- enhancement of O-D tables (synthetic or otherwise) using link-OD models

A common situation facing the would-be freight movement analyst is that he/she has access to either survey data or regression-model based estimates of the volume of freight moved into or out of specific facilities, or specific traffic analysis zones. What is missing is an estimate of the flows between these locations. Survey-based nationwide O-D flow data exists only at levels of aggregation (e.g. State-to-State) and lack the necessary spatial detail for the user's purpose. Given the very large number of possible solutions to what are typically rather large O-D matrices, two techniques likely to see increasing use are interregional input-output (I-O) and link-OD modeling. The I-O approach

builds on an extensive economic and spatial analysis literature and the techniques is used today within a number of different federal agencies. The link-OD approach allows the analyst to combine flow data based on shipper or carrier establishment surveys with informatics-based link traffic count data. Surrogate travel demand data, such as population, employment and value of production data (possibly generated by an I-O model), and disaggregated by region and industrial sector, can also be added to the mix. Figure 4 shows the linkages between the different types of data collection, data modeling (i.e. data synthesis) and data products envisioned.

Finally, in developing multi-modal freight movement matrices, a subsequent conversion to vehicle or vessel equivalents is often required, for each mode in the table. This is also a task that requires further data collection and statistical analysis.



Source: Adapted from [5, Figure 1]

Figure 4. Linkages Between Types of Data Collection , Data Modeling, and Data Products

In all of the above cases, and using any of the above techniques (singly or in combination) the purpose is to help federal, state and local planning agencies develop more detailed freight movement matrices based on sound statistical principles and methods by getting the most out of existing datasets.

To this end, the latest developments in these and related statistical modeling techniques should be explored, and findings made accessible to potential users of AFDP data products. This can be done through web-based directions to the latest publications, and through web-site hosting of new publications as well as prototype computer codes developed under AFDP-funded research. An increasingly important component of such a web site's content is likely to be references to private vendor products developed using AFDP supported and other (including proprietary) data sources. Popular examples of such products already exist, notably the TRANSEARCH multi-modal commodity flows data base now being used by a number of State DOTs, and developed at the county-to-county level for use within the Federal Highway Administration's Freight Analysis Framework (FAF)[36].

Proposed Activities. It is proposed that the AFDP produce the following data products based on its efforts in data synthesis:

1. An internally consistent and multimodal national picture of the annual tons, dollars, ton-miles and vehicle and vessel miles of freight moved into, out of, within and through the US, updated on an annual basis.
2. A set of annually updated county-to-county based annual truck, rail, inland and intra-coastal barge commodity flow matrices, by tonnage and dollar-value
3. The vehicle/vessel equivalences of these commodity flow matrices, including truck–rail and truck-water intermodal, and truck to airport drayage traffic volumes.
4. A set of procedures for updating these annual freight movement estimates to a set of future year forecasts.
5. Procedures to assist state, regional and metropolitan freight planners in generating disaggregate commodity and vehicle/vessel-based freight flow estimates, using generally available data sources or in combination with other, vendor-specific and/or local area data sets.
6. An interactive web-site for access to and discussions of all of the above methods and products, including the availability of private vendor products that State and regional planning agencies might use.
7. A web-site to discuss and promote both domestic and international transportation data harmonization issues

[36] http://ops.fhwa.dot.gov/freight/freight_analysis/faf/

Data products under bullets 2 , 3 and 4 will require rigorous statistical modeling and reporting of not only expected cell values but also the variances in, and standard errors associated with such O-D matrices. As data collection and reporting improves over time the national picture described in bullet 1 might also be expanded (i.e. disaggregated) to cover major regions of the country.

4.4 Data Quality

As a federal statistical agency the BTS has a requirement placed upon it to ascertain and, within its own data collection efforts, to maintain data quality. In doing so it follows both DOT and Office of Management and Budget (OMB) quality guidelines on information dissemination and good statistical practice, and periodically seeks external reviews of its data collection and dissemination programs. Of particular interest are improvements that increase the accuracy, reduce the respondent burden and also reduce the overall cost of data collection through:

- New data collection methods – research into a mixture of data collection methods, including web data entry, electronic data exchange, and scanning of firm-specific reports to reduce both the burden and the transcription errors for respondents who would prefer such reporting methods.
- Improved sample designs – including alternative within-establishment as well as across establishment sampling designs where multiple data records (e.g. multiple shipments) are concerned; and variable length reporting for shipments or establishments of different sizes, to reduce sample variances.
- Careful analysis of survey non-response – understanding why some firms fail to return surveys, and how to overcome this.
- Improved statistical estimation techniques – investigation of the effects of large firm response rates on precision and sample variance, and of the effects of imputations on published results.
- New data processing methods – including research into the use of automated and on-line data editing and quick-callback systems.

The AFDP should also be subject to carefully developed data standards for public reporting, as well as standards governing creation of the metadata files and data dictionaries associated with each of its data collection activities.

4.5 Data Protection

A problem that has long hindered public use of freight movement data is the proprietary or business sensitive nature of much private sector information. This is why access to microdata records such as those collected by the Commodity Flow Surveys is carefully restricted under Title 13 of the US Code. Specifically, AFDP disseminated data must comply with the Confidential Information Protection and Statistical Efficiency Act (CIPSEA) of 2002.[37]

The deluge of information now passed daily over the Internet raises many concerns over data security. Getting the most out of data collected, while still protecting client confidentiality, has become an important area, with some recent contributions to this literature from BTS funded research [35,36]. AFDP activities should focus on developing and using innovative statistical techniques for filling in the cells of large, multi-dimensional freight movement tables, while still complying with all restrictions on client confidentiality. A variety of techniques will need to be investigated, including aggregation, perturbation or suppression of cell values, data swapping (i.e. modifying microdata by exchanging similar respondents between data groupings), selective sampling (e.g. releasing only a scaled subset of micro data records), the use of threshold and cell concentration rules, data rounding, and the use of simulation models and other synthetic data substitution methods to mask specific cell values [37]. The goal in all cases is to protect confidential data while minimizing data loss. There are some obvious overlaps here with the data synthesis methods (notably maximum likelihood based methods) discussed in Section 4.3

Concerns over national security must also be given a high priority [38]. The AFDP should contain an ongoing task to review the potential risk associated with releasing specific freight movement data elements. Communication with the Department of Homeland Security (DHS) on this topic is advisable.

4.6 Data Dissemination and Outreach

The AFDP will produce a wide range of products, including databases, reports, and both prototype and user-tested software tools to assist analysts in manipulating freight movement data. AFDP information should be disseminated via the BTS website, through peer reviewed publications, and by attendance at conferences and workshops at which BTS products will be on display. Data should be downloadable via the BTS website and/or distributable

[37] Title V of E-Government Act of 2002, Public Law 107-347. http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=107_cong_public_laws&docid=f:publ347.107.pdf

on CD-ROM, and shipped in reply to user requests generated through visits to these various venues.

Principal data products include the results of the various AFDP surveys, notably the Commodity Flow Surveys and Motor Carrier Vehicle Activity Survey. The results of the program's various pilot tests, as well as R&D findings into the use of innovative data collection procedures should be made available to the general user: as should the results of any data synthesis and data protection procedures found to have general applicability to the freight data community. User experiences with AFDP data and data processing products should become an important part of this information gathering and dissemination process, as an aid to learning by example from prior analyses.

Outreach activities should involve periodic surveys of the current and potential users of AFDP products. This should occur through the BTS website as well as through more directed efforts, and include planners and analysts in federal, state and metropolitan government both within and outside DOT. An automated, computer-based procedure for documenting and following up on such requests and comments needs to be developed. AFDP staff should also be available to respond quickly to requests for information from US DOT management and from other decision-makers in federal government.

5. Example Program Timeline and Ball-Park Costs

The original draft of this report contained a set of estimated costs to develop the AFDP for three different timelines/costing profiles. Under review it was determined that these timelines/estimates needed further development and revision, and are not contained in this version of the report. Estimating the costs of an AFDP as described presents a number of challenges, not least of which is a limited literature on the constantly evolving costs of conducting the sort of freight establishment surveys that will consume the majority of the program's budget. Complicating such estimates, new survey methods, such as use electronic filings of questionnaires, promise savings that have not yet been quantified (or confirmed) in any consistent and statistically representative manner. In particular, the specifics of the proposed motor carrier activity survey, and the effects of expanding the scope of the Commodity Flow Survey need to be developed further (how many days reporting per vehicle diary, how many shipments sampled, etc.). In the interests of further discussion on program content, this version of the report is made available without these estimates.

6. Bibliography

1. BTS (1998) *Transportation Statistics Beyond ISTEA. Critical Gaps and Strategic Responses.* Bureau of Transportation Statistics, US Department of Transportation, Washington, DC 20590.

2. FHWA (2002) *The Freight Story: A National Perspective on Enhancing Freight Transportation. FHWA-OP-03-004*. Federal Highway Administration, US Department of Transportation, Washington, DC 20590.

3. Meyburg, A. H. and Mbwana, J.R. (eds) (2002) Conference Synthesis: Data Needs in the Changing World of Logistics and Freight Transportation. New York State Department of Transportation, Albany, NY. (http://www.dot.state.ny.us/tss/conference/synthesis/pfd)

4. TRB (2003) A Concept for a National Freight Data Program. Special Report 276. Transportation Research Board, Washington, DC 20001. (http://gulliver.trb.org/news/blurb_detail.asp?id=1730)

5. Southworth, F. (1999) The National Intermodal Transportation Data Base: Personal and Goods Movement Components. April, 1999. Report prepared by Oak Ridge National Laboratory for the Bureau of Transportation Statistics, Washington D.C. 20590.

6. Cambridge Systematics, Inc., et al (1996) Quick Response Freight Manual. Final Report. Report to the Federal Highway Administration, Washington DC 20590. Appendix K. (http://tmip.fhwa.dot.gov/clearinghouse/docs/quick/index.stm)

7. Jack Faucett Associates et al (1997) Multimodal Transportation Planning Data. NCHRP National Cooperative Highway Research Report. Project 8-32-(5). Transportation Research Board, Washington DC, 20001. March 1997.

8. Census Bureau (1999) 1997 Economic Census, Transportation. 1997 Commodity Flow Survey. Report EC97TCF-US. US Department of Commerce, Economics and Statistics Administration, Suitland, MD.

9. BTS (2003) *U.S. International Trade and Freight Transportation Trends.* BTS03-02. Bureau of Transportation Statistics, US Department of Transportation, Washington, DC 20590.

10. TRB (2003) Letter Report of March 20, 2003. Committee to Review the Bureau of Transportation Statistics' Survey Programs (Commodity Flow Survey).

(http://gulliver.trb.org/publications/reports/bts_cfs.pdf)

11. ORNL (2000) *Freight USA. Highlights from the Commodity Flow Survey and Other Sources.* Report prepared by Oak Ridge National Laboratory for the Bureau of Transportation Statistics, Washington, DC 20590. May 2000.

12. Vogt, D.P., Southworth, F., Peterson, B.E. and Rizy, C. (2002) *Estimation of Regional Economic Benefits from Improvements in Great Lakes/St. Lawrence Seaway Navigation: Phase 1 Reconnaissance Study*. Report prepared for the Ohio River Division, US Army Corps of Engineers.

13. FHWA, Freight Analysis Framework (FAF). Office of Freight Management and Operations, Federal Highway Administration. Washington, DC 20590. (http://ops.fhwa.dot.gov/freight/adfrmwrk/)

14. Black, W.R. (1997) *Transport Flows in the State of Indiana; Commodity Database Development and Traffic Assignment: Phase II.* Transportation Research Center, Indiana University, Bloomington, IN.

15. Cambridge Systematics, Inc., et al. (2002) *Freight Impacts on Ohio's Roadways, June 2002. Final Report* prepared for the Ohio Department of Transportation, Columbus, Ohio. June, 2002. (http://www.dot.state.oh.us/planning/Freight/FreightImpacts.htm)

16. Sorratini, J.A. (2000) Estimating Statewide Truck Trips Using Commodity Flows and Input-Output Coefficients. *Journal of Transportation and Statistics.3.1:* 53-68.

17. Taylor, M.A., Young, W. Wigan, M. and Ogden, K.W. (1992) Designing a large-scale travel demand survey: new challenges and new opportunities. *Transportation Research Vol 26A*: 247-261.

18. Battellino, H. and Peachman, J. (2001) The Joys and Tribulations of a Continuous Survey. *International Conference on Transport Survey Quality and Innovation, Kruger National Park, South Africa*. August, 2001.

19. Lawson, C.T. and Strathman, J.G. (2002) *Survey Methods For Assessing Freight Industry Options.* Federal Highway Administration Report FHWA-OR-RD-02-14. Washington DC, 20590

20. Transport Canada. Canadian Vehicle Survey, 2001. (http://www.tc.gc.ca/pol/en/cvs/cvs.htm)

21. UK DfT (2002) *Transport of Goods Bulletin: Transport of Goods by Road in Great Britain 2001.* National Statistics SB(02)1 7. United Kingdom Department of Transport, Local Government and the Regions, London, England. (http://www.dft.gov.uk/stellent/groups/dft_transstats/documents/page/dft_transstats_506336.pdf)

22. Canadian National Roadside Intercept Survey - 1999. (http://www.tc.gc.ca/pol/en/Truck_Bus/NRSIndex_e.htm)

23. Battelle (1999) *Heavy-Duty Truck Activity Data.* Final Report. Prepared by Battelle memorial Institute for the Office of Highway Information Management, Federal Highway Administration, Washington, DC 20590. April, 1999.

24. Rizet, C., Guilbauly, M., van Merijeren, J.C., Bjister, E. and Houee, M. (2003) *Tracking Along The Transport Chain Via The Shippers Survey*. International Conference on Transport Survey Quality and Innovation. (http://www.its.usyd.edu.au/)

25. GAO (2003) *Customs Service Modernization. Automated Commercial Environment Progressing, but Further Acquisition Management Improvements Needed*. US Government Accounting Office GAO-03-406. February, 2003.

26. Southwest Technology Development Institute (2003) *Vehicle Detector Clearinghouse*. New Mexico State University, Las Cruces, New Mexico. July 14, 2003. (http://www.nmsu.edu/~traffic/ProductInformation.htm)

27. Skszek, S.L. (2001*) "State-of-the-Art" Report on Non-Traditional Traffic Counting Methods*. Report prepared for the Federal Highway Administration and Arizona Department of Transportation. FHWA-AZ-01-503. Washington D.C. 20590.

28. TRB (2001) Remote Sensing and Transportation. Products and Results: Foundations for the Future. Conference Proceedings 29. Transportation Research Board, Washington DC, 20001

29. DOT-NASA (2002) *Achievements of the DOT_NASA Joint Program on Remote Sensing and Spatial Information Technologies. Application to Multimodal Transportation.* US Department of Transportation/National Aeronautics and Space Administration. April, 2002.

30. FMCSA (2002) Commercial Vehicle Information Systems and Networks (CVISN). Recommendations for Primary Identifiers. Baseline Version V1.0. White Paper. Federal Motor Carrier Safety Administration. Washington DC 20590. October 2002.

(http://www.jhuapl.edu/cvisn/index.html)

31. RFID Journal Inc. (2003) *RFID Journal (on-line)* .
http://www.rfidjournal.com/

32. EU (2000) Sustainable Urban and Regional Freight Flows (SURFF). EU/DG XIII Transport Telematics Applications Program. TR 1053. European Union. (http://www.euroweb.net/surff/about.htm)

33. Furness, A. (2002) Present and Future Smart Active Label (SAL) Enabling Technologies – An Introductory Overview, Centre for Automatic Identification and Intelligent Systems ,Technology Innovation Center, Birmingham, England. (http://www.sal-c.org/3_resources/Overview%20on%20Current%20and %20Future%20SAL%20Technologies.pdf)

34. BTS. (2003) Measurement of Ton-Miles and Value-Miles of International Trade Traffic Carried by Highways for Each State. (Draft). International Trade Traffic Study (TEA-21, Section 5115). Bureau of Transportation Statistics, US Department of Transportation, Washington, DC 20590.

35. BTS (2003) *Disclosure Limitation Methods for Tabular Data.* Task 2 Final Report. Bureau of Transportation Statistics, US Department of Transportation, Washington, DC 20590.

36. BTS (2002) *Review of the State of the Disclosure Limitation Methods for Tabular Data and Research Proposal*. Task 1 Final Report. Bureau of Transportation Statistics, US Department of Transportation, Washington, DC 20590.

37. Russell, J.N., Kelly, J.P. & Glover, F. (2002). The Bureau of Transportation Statistics' Statistical Disclosure Limitation Method for Tabular Data: A Review. 2002 Proceedings of the American Statistical Association, Government Statistics Section [CD-ROM], Alexandria, VA: American Statistical Association.

38. TRB (2003) Cybersecurity of Freight Information Systems. Special Report 274. Transportation Research Board, Washington, DC 20001. (http://gulliver.trb.org/news/blurb_detail.asp?ID=1498)

Appendix 1. Example Motor Vehicle Activity Surveys

Data Collection Method	Example Surveys	Sampling and Data Collection Methods	Objectives/Target Statistics	Data Elements Sampled	Pros and Cons
Vehicle (Registration) Based Sampling	United Kingdom Continuous Survey of Read Goods Transport (annual since 1994) [ref 21]	Mandatory annual mail out/mail back survey of motor carriers, sample drawn from population of vehicle licenses. Over 360 vehicles sampled per week in a "self- weighting" sample spread evenly over the year (annual sample of 18,000- 20,000 trucks); a survey of goods vehicles over 3.5 tonnes gross weight; reporting is via a weekly diary of all trips tken by a sampled truck. Stratiferd random sampling within 8 GVW groups. 95% of surveys returned in 2001, 87% useable.	The only national survey of road haulage in the UK; a key source of data for traffic forecasts and studies of vehicle regulation	*Vehicles: g*ross vehcile weight, carrying capacity, axles, trailer gross weight and body type, abnormal loads, private/ for-hire carriage, fuel, reason not used. *Trips:* origin, destination (nearest town, county), commodity class (19 types),distance laoded, distance empty, weight of freight carrierd, delivered and collected, if wieght or volume limited, if hazardous, if to port dock or rail terminal, nture of inter-modal transfers; and number of collection and delivery stops and trailer axles for trucks with > 5 stops.	*Pros:* detailed data on vehicle classes and types of operation. Well established sample frames exist from which to derive representative samples by vehicle class. *Cons:* Limited commodity detail as carrier often doesn't know exactly what's being hauled. Important to establish who in organizatrion fills in the form (affects data quality)
	Canadian Vehicle Survey (quarterly, since 4th quarter of 1999) [ref. 20]	Vehicles are selected randomly in each province and territory; a 7-day trip diary is mailed to the owners to record how these vehicles are used over the survey time period. Reporting is voluntary. Trucks over 4.5 tonnes surveyed as part of broad passenger plus freight vehicle activity survey; limited sample size; sample stratified by vehicle type; CATI initial interview followed by mail out/mail back week-long trip log: annual sample size of 29,211 vehicles in 2001.	Annual estimates of the amount of road travel: vehcile counts, vehicle -kilometers by Province/Territory, broken down by types of vehicle, type of driver, purpose, time of day and season. national fuel use estimates.	*Trips:* truck configuration*;* type of carriage (private, for-hire);trip purpose, start and end times; driver age and sex; road type (if < or > 50 mph speed limit); intra- or inter-province, transborder or US trip; if dangerous commodity transported; odometer-based mileages; fuel type used and quantity purchased.	*Pros:* Well established sample frames exist from which to derive representative samples by vehcile class. *Cons:* Low response (truck-day) response rates associated with voluntary survey:on the order of 27% and 43% for fully completed trip logs; large coefficients of variation for some cross -tabulations.
Vehicle/Driver Intercept Survey	Canadian National Roadside Intercept Survey (1991,95,99) [ref. 22]	Voluntary driver survey, 7 day continuous truck count for a representative week, random intercept sampling, yielding 65,000 interviews of randomly selected trucks at the 238 sites on road network; 88% response rate from drivers. Data is recorded on hand-held computers by roadside interviewers, specialized software is used to infer route	Number of trips, weight of cargo carried, weight of trucks, distance traveled, transport output (cargo weight x distance), road impact (truck weight x distance).	*Trucks* (size, weight, axles,cab-type, optional equipment; *Drivers (*age experience, home location, employment type);*Trips* (linehaul vs multi-stop "peddle" trip, origin, destination, facility by types visited, if trip interlined, start and end times); *Routes* (border ports, other stops en route, distance); *Carriers* (private, for-hire, leased for-hire, if dispatching center used).	*Pros:* high response rates possible; detailed data on vehicle classes and types of operation; captures the link betweeen vehicle use (including partial and empty loads), and commodity class; useful data on routes and times taken and stops made; actual trips sampled; driver data a bonus. *Cons:* commodity details may be difficult to ascertain as the driver often doesn't know exactly what's being hauled; data on initial origin or final destination for goods may not be known by driver; delay for driver/carrier during working hours; general traffic disruptions also possible, though can usually be minimized. Spatially based sampling design can be difficulty to implement when seeking generally represenative data.
Vehicle Tracking Survey	Heavy Duty Truck Tracking Study in California (1997-8) [ref 23]	Voluntary "(opportunity") sample of 24 (out of an initial 236) trucking firms; continuous vehicle tracking for 4 months, then for a further 8 months. Not necessarily a representative sample for general purposes. Data is recorded using GPS technology fitted to trucks, (140 useable recording devices from 167 installs), data downloaded weekly via hand-held computers.	Describes truck activity patterns in urban and rural areas for congestion studies and for on-road emissions forecasting.	Truck routes and stops en route, timing of vehicle starts and stops, travel times and speeds, truck idling and soak times	*Pros:* ability to follow trucks through complete trip activity circuits and driving cycles, and to capture detailed geography of stops, travel speeds and distances; easy to link data to GIS software for mapping purposes; non-intrusive data collction *Cons:* limited commodity detail as the driver often doesn't know exactly what's beig hauled. Costs to implement currently quite high (but falling); effort to implement implies voluntary survey, and may be unpopular with drivers and/or carriers: difficulty to obtain statistically represenative data; commodity classes as well as vehicle loads hard to capture using only tracking devices.