FAF^2: PROVISIONAL COMMODITY ORIGIN-DESTINATION ESTIMATES

REVISED METHODOLOGY REPORT

Submitted to

Office of Freight Management and Operations
Federal Highway Administration
Washington DC

Prepared by

Battelle
The Business of Innovation

In Association with

MacroSys Research and Technology

MACROSYS

University of Tennessee

May 6, 2009

TABLE OF CONTENTS

Page

Table of Contents (Continued)

Page

TABLE OF CONTENTS (CONTINUED)

Page

List of Appendices

TABLE OF CONTENTS (CONTINUED)

Page

List of Tables

List of Figures

LIST OF ACRONYMS

AAR	American Association of Railroads
ATA	American Trucking Association
BEA	Bureau of Economic Analysis
BLS	Bureau of Labor Statistics
BTS	Bureau of Transport Statistics
CBP	County Business Pattern Database
CFR	Code of Federal Regulations
CFS	Commodity Flow Survey
CIF	Customs Insurance and Freight
CIV	Customs-Import-Value
CPI	Consumer Price Index
CWS	Carload Waybill Sample
DOT	Department of Transportation
EIA	Energy Information Administration
FAF	Freight Analysis Framework
FAS	Free-alongside-ship
FedEx	Federal Express
FHWA	Federal Highway Administration
FTD	Foreign Trade Division
FTZ	Free Trade Zone
GDP	Gross Domestic Product
GSP	Gross State Product
HS	Harmonized Schedule
IPP	International Price Program
LPMS	Lock Performance Monitoring System
M3	Monthly Manufacturers' Shipments, Inventories, and Orders
MCF	Master Coordinates File
NAICS	North American Industry Classification System
NDC	National Data Center
OAI	Office of Airline Information
O-D	Origin-Destination
PAD	Petroleum Administration for Defense
POE/POD	Point of Entry/Point of Departure
PPI	Producer Price Index
PSA	Petroleum Supply Annual
RITA	Research and Innovative Technology Administration
SCTG	Standard Classification of Transported Goods
SMA	Simple Moving Average
SQL	Structured Query Language
STB	Surface Transportation Board
STCC	Standard Transportation Commodity Codes
TFD	Transborder Freight Database
USACE	United States Army Corps of Engineers
USDOT	United States Department of Transportation
YTD	Year-to-date

This page intentionally left blank.

1.0 INTRODUCTION

The Office of Freight Management and Operations of the Federal Highway Administration (FHWA) is responsible for the development of the Freight Analysis Framework (FAF) including the commodity origin-destination (O-D) database. FAF integrates data from a variety of sources to estimate commodity flows and related freight transportation activity among states, regions, and major international gateways. The commodity O-D database contains commodity flows between domestic origins and destinations, exports between domestic origins and foreign destinations and the port of exit, and imports between foreign origins and domestic destinations via a port of entry. Each record contains zone of origin, zone of destination, port of entry or exit (which applies only to export and import flows), type of commodity, mode of transportation for domestic portions of the flow, value in millions of dollars, and tons in thousands of short tons.

The FAF commodity O-D database lays the foundation for transportation infrastructure analysis. With regard to the first generation of the FAF, FHWA relied upon data provided by private and proprietary sources. However, this arrangement limited the data usage to United States Department of Transportation (USDOT) only. State DOTs and local metropolitan planning organizations (MPOs) do have access to the products from the usage of these data but not to the data itself.

To overcome the limitations of FAF's commodity data issue, FHWA developed the next generation of the FAF known as FAF^2. The commodity O-D data used in developing FAF^2 data are based on the 2002 Commodity Flow Survey (CFS) and a host of other public data sources. For data quality reasons, FAF^2 freight flow O-D coverage is limited to 131 freight analysis zones that include 114 CFS freight O-D zones and 17 major ports, border crossings, and freight ports.

In an attempt to make the FAF into a useful tool for measuring and analyzing the changing world of freight transportation, FHWA began developing annual provisional estimates of commodity movements including all modes of transportation starting with 2005. The goal is to provide practitioners in the area of economic development and transportation planning with the latest update – provisional data on goods movement. For example, the provisional estimate for calendar year 2006 was released early 2007; and the provisional estimate for 2007 was early 2008. The provisional estimates are developed based on publicly available freight data sources and methods that can be fully disclosed to the general public.

This report presents a description of the public data sources and methodologies for extracting freight information from yearly, quarterly, and monthly publicly available publications for the current year or past years and to generate provisional estimates of freight movement by mode for the current year.

1.1 Report Organization

The remainder of the report is organized as follows. Note that for each mode of transport, both domestic and international movements are included.

Chapter 2 presents the data sources and methodology used in developing the provisional estimates of volume and value of freight movement by highway. International movements by land border crossings include import and export between the U.S. and Mexico, and between the U.S. and Canada via land border crossings.

Chapter 3 presents the data sources and methodology used in developing the provisional estimates of volume and value of freight movement by air. International movements by air include international air cargo covering both import and export.

Chapter 4 presents the data sources and methodology used in developing the provisional estimates of volume and value of freight movement by rail. International movements by rail include all rail shipments to and from Canada, Mexico, and countries outside North America that sue rail for the domestic portion of the movement.

Chapter 5 presents the data sources and methodology used in developing the provisional estimates of volume and value of freight movement by water. International movements by water include import and export between the U.S. and the other seven international trade regions via seaports.

Chapter 6 presents the data sources and methodology used in developing the provisional estimates of volume and value of freight movement by pipeline. International movements by pipeline include import and export between the U.S. and the other seven international trade regions via seaports.

Chapter 7 describes the compilation of the provisional databases. Chapter 8 presents the conclusions of this methodology report, and Chapter 9 lists the works consulted and cited. An appendix compares selected data sources and forecasts, and contains supporting data in international air freight data.

2.0 FREIGHT MOVEMENT BY HIGHWAY

2.1 Introduction

This chapter presents the data sources and methodology used for preparing the 2008 FAF provisional freight tonnage and value estimates for the highway mode of transportation. It covers both domestic and transborder (international) highway freight transportation. The estimation methods are formulated based on the FAF^2 2002 benchmark estimates, and the latest publicly available and reliable information from different data sources. Improved approaches are applied by incorporating any new publicly available information since the last provisional estimates.

2.2 Principal Data Sources

The following are the main data sources used in developing the estimates for freight movement by highways.

Monthly Trucking Tonnage Report – Published by the American Trucking Associations (ATA) and provides up-to-date information on the trends of for-hire trucking activities. This monthly trucking tonnage index is based on an ongoing ATA survey of monthly tonnage by Class I and II general freight carriers. It includes both large and small truckload carriers, along with less-than-truckload carriers. The data are released with five weeks of time lag.

County Business Pattern (CBP) Database – Published by the U.S. Census Bureau on an annual basis and provides national, state, and county level data on payroll, employment, and number of establishments by detailed North American Industry Classification System (NAICS) industries. The series provides subnational economic data by industry and excludes data on self-employed individuals, employees of private households, railroad employees, agricultural production employees, and most government employees. The report is released with a two-year time lag. It can be accessed at http://www.census.gov/econ/cbp/index.html.

Gross State Product (GSP) – Prepared by the Bureau of Economic Analysis (BEA) of the U.S. Department of Commerce and provides data on GSP including components of GSP such as compensation of employees, operating surplus, taxes, etc. Gross domestic product (GDP) by state is the state counterpart of the nation's GDP and is derived as the sum of the GDP originating in all the industries in the state. The data are published with a one-year time lag. It can be accessed at http://www.bea.gov/regional/.

State Personal Income – Published by BEA of the U.S. Department of Commerce on a quarterly basis. Data on state personal income, employment, and compensation for NAICS industries are available from this source. Personal income is the income received by all persons from all sources. It is measured before the deduction of personal income taxes and other personal taxes and is reported in current dollars (no adjustment is made for price changes). Data are published with a three-month time lag. It can be accessed at http://www.bea.gov/regional/.

Monthly Manufacturers' Shipments, Inventories, and Orders (M3) Survey – Conducted by the Census Bureau, it provides broad-based monthly statistical data on the economic conditions in the domestic manufacturing sector. It measures current industrial activity and provides an indication of future production commitments. The value of shipments measures the value of goods delivered during the month by domestic manufacturers. The data are released with a two-month time lag. The survey results can be accessed at http://www.census.gov/indicator/www/m3/.

Monthly Wholesale Trade Survey – The Census Bureau provides monthly estimates of sales and inventories of wholesale trade industries. This provides statistics on sales and inventory/sales ratios along with standard errors. Data are both seasonally adjusted and unadjusted. The data are released six weeks after the close of the reference month. It can be accessed at http://www.census.gov/mwts/www/mwts.html.

Surface Transborder Freight Database (TFD) – Published by the Bureau of Transportation Statistics (BTS) and contains data on North American merchandise trade by commodity, surface mode (rail, truck, pipeline, mail, and other), and by port of entry and geographic detail for the U.S. trade to and from Canada and Mexico. This source provides the dollar value of both imports and exports, and tonnage of imports. The data are published with a three-month time lag. It can be accessed at http://www.bts.gov/programs/international/transborder/.

Producer Price Index (PPI) – Measures the average changes over time in the prices received by domestic producers of goods and services. This measures price changes from the point of view of the producer. The data are reported by detailed industry and detailed type of commodities. The Bureau of Labor Statistics (BLS) publishes these data on a monthly basis with a time lag of one month. The data can be obtained from http://www.bls.gov/ppi/home.htm.

Import/Export Price Indexes – The International Price Program (IPP) of the BLS produces import and export price indexes, which measure the change over time in the prices of goods or services purchased from abroad by U.S. residents (imports) or sold to foreign buyers by U.S. residents (exports). BLS publishes the Import/Export Price Indexes monthly with a time lag of two weeks. The data are available at http://www.bls.gov/mxp/home.htm.

Commodity Flow Survey (CFS) – The CFS is the primary source of national and state level data on domestic freight shipments by American businesses. It is a shipper-based survey that collects information on how U.S. establishments ship raw materials and finished goods; the types of commodities shipped by mode of transportation; the value, weight, origin, and destinations of shipments; and the distance shipped. The survey is conducted every five years. The latest survey covers 2007 and was released in December 2008.

2.3 Methodology for Domestic Highway Freight

The method used for preparing the annual provisional O-D freight flow estimates for domestic highway freight shipment involves the following steps:

1. Determine the annual growth of highway freight tonnage and value at the national level.
2. Estimate the growth factors for each O-D pair at the FAF region level.
3. Estimate the annual growths of each O-D pair by applying the respective O-D regional growth factors to the national annual growth.
4. Determine the provisional freight level (in terms of tonnage and value) of each O-D pair for 2008 by adding the growths to the freight level of the corresponding O-D pair in the FAF^2 benchmark year or the provisional estimate of the previous year.

This approach can be characterized as an *updating approach*. In comparison to producing provisional commodity O-D estimates entirely from updated input data, this approach, which produces provisional estimates by adding estimated growths (or changes) to the corresponding estimates in the benchmark year, has the following advantages:

- It fully utilizes all relevant new information including the most recent data, which become available after the benchmark year, to allow the provisional estimates to capture any changes that occurred after the benchmark year.
- It takes full advantage of the knowledge and detailed information embodied in the estimates of the benchmark year, but not available for the provisional estimates.

2.3.1 Determine Annual Growth of Highway Freight at the National Level

The following four definitions are important for the discussions presented in this document.

Estimates of highway freight – indicates the level of or volume of highway freight in a year with units of short ton and dollar. “Estimates of highway freight” and “highway freight” are used interchangeably in this discussion.

Growth – is defined as the change in highway freight either in terms of tons or dollar value between two years. Unless otherwise specified, it is calculated as the difference between highway freight for the current year and highway freight for the previous year. The units are tons and dollars.

Rate of growth – indicates the relative magnitude of growth when growth is compared to the level of the base year. Rate of growth is expressed in percent. “Rate of growth” and “growth rate” are used interchangeably in this report.

Current year – refers to the year for which provisional estimates are prepared.

2.3.1.1 Freight Tonnage

The monthly trucking tonnage index published by the ATA in the Monthly Truck Tonnage Report, the 2007 CFS preliminary results, and the truck tonnage reported in the FAF benchmark year are used in preparing the provisional domestic freight tonnage carried by trucks on the highway.

First, the growth rates of for-hire and private trucking freight shipments are calculated between 2002 and 2007 using data from CFS. This growth rate is applied to the 2002 FAF benchmark estimate to derive the total tons of freight for 2007. The result of this effort provides a more reasonable and robust estimate for 2007, based on which the 2008 provisional estimates are prepared. The formula for deriving the 2007 estimate is given by the following equation.

$$T_t = T_0 * \frac{T_{c2007}}{T_{c2002}}$$

Where
T_t = Total tons by truck for year t (t = 2007)
T_0 = Total tons by truck for FAF benchmark year, i.e., 2002
T_{c2007} = CFS truck (private and for-hire trucking) tonnage for 2007
T_{c2002} = CFS truck (private and for-hire trucking) tonnage for 2002.

This estimate provides the 2007 tons of freight shipment by truck. Then, the 2007 total tons of freight is disaggregated at two-digit Standard Classification of Transported Goods (SCTG) commodity level of detail by implementing the following procedure.

1. First, the 2007 value of output at the two-digit SCTG commodity level is multiplied by the corresponding ton-per-value of output ratios for the FAF benchmark year to derive the tons of freight by type of commodity for 2007. The information on the value of output by type of commodity is obtained from the Census Bureau's Monthly M3 Survey. The type of commodity from this data source is established based on each industry's primary product.
2. Second, commodity shares are calculated based on the commodity distribution of the above tons of freight estimates for 2007, and these shares are used to break down the growth in national highway tons of freight, estimated with CFS data, into growth in tons of freight by two-digit SCTG commodity.
3. Finally, the growth in tons by type of commodity is added to the 2002 estimates to obtain the national freight tonnage by type of commodity for 2007.

In this method, both highway freight weight/value ratio by commodity and highway shipment tonnage to output-value ratio by commodity in 2007 year are assumed to remain the same as in the FAF benchmark year. The advantage of this method is that it utilizes the latest and most reliable indicator on the growth of highway freight from the 2007 CFS.

The 2008 provisional freight flow estimates are prepared on the basis of the 2007 estimates and the monthly trucking tonnage index from ATA. The trucking tonnage index of 2008 divided by the trucking tonnage index of 2007 and multiplied by the 2007 tons of freight would provide the total tons of freight shipment by truck mode for 2008. The formula is given as follows:

$$T_t = T_0 * \frac{I_t}{I_0} * 100$$

Where
T_t = Total tons of freight by truck for year t (t = 2008)
T_0 = Total tons of freight by truck for year 0 (year 0 = 2007)
I_t = Trucking tonnage index for year t (t = 2008)
I_0 = Trucking tonnage index for year 0 (year 0 = 2007)

The total tons of freight shipment for the provisional year are disaggregated at the two-digit SCTG commodity level of detail using the following procedure:

1. First, the 2008 value of output by commodity is multiplied by the corresponding ton-per-value of output ratios for 2002 to derive tons by type of commodity for the provisional year.
2. Second, commodity shares are calculated based on the commodity distribution of the above tonnage estimates, and these shares are used to break down the current year growth in national highway freight tons into commodity details.
3. Finally, the current year growth in tons by type of commodity is added to the 2007 estimates to provide the national tons of freight by type of commodity for the provisional year.

In this method, both highway freight weight/value ratio by commodity and highway shipment tonnage to output-value ratio by commodity in the current year are assumed to remain the same as in the FAF benchmark year.

2.3.1.2 Highway Freight Value

Freight value is determined not only by its weight but also by its weight-value ratio. Weight-value ratio, in turn, changes over time due to changes in the commodity composition of freight and changes in their prices. The value of freight for the provisional year has been estimated based on data from the 2002 FAF benchmark database, the 2007 CFS, ATA's trucking tonnage index, and the value of output by industry.

The growth rate of the value of freight shipment in current dollars for the highway (truck) mode is calculated using the combined values of freight for private and for-hire trucking from the 2002 and 2007 CFS. This growth rate has been applied to the 2002 FAF benchmark estimate to obtain the national total value of freight in current dollars for 2007. Then, the growth in the 2007 total value of freight is disaggregated at two-digit SCTG commodity detail by implementing the following steps.

1. First, the 2007 value of output by type of commodity is multiplied by the corresponding ton-per-value of output ratios for 2002 to obtain freight tonnage by type of commodity.

2. Second, commodity shares are calculated based on the above tons of freight estimates, and these shares are used to disaggregate the growth in total tons of highway by commodity.
3. Third, multiplying the growth in tons at the two-digit SCTG commodity level with their corresponding value-per-ton ratios of the FAF benchmark year yields the growth of freight values (in 2007 dollars) by type of commodity for 2007. Note that the values utilized in calculating the 2002 value-per-ton ratios are adjusted for prices to make them in 2007 current dollars by applying changes in the PPI of commodities. The PPIs are obtained from the BLS.
4. Fourth, based on the value estimates above, value shares by type of commodity are calculated, and these shares are applied to distribute the growth in the total value of freight among two-digit SCTG commodities in 2007.
5. Fifth the 2002 FAF freight values at the two-digit commodity level of detail for the highway mode are valued at 2007 prices by applying the PPIs.
6. Sixth, the results from step four and five are combined and value shares by commodity are calculated, and these shares are applied to disaggregate the total current value of freight for 2007.
7. Finally, the growth in the current dollar value of freight are adjusted for inflation by applying changes in PPIs at two-digit SCTG commodity levels to obtain values in 2002 dollars. These are added to the FAF benchmark values to arrive at the 2007 values in 2002 dollars.

The 2008 freight values are prepared using the 2007 estimates, the value of output by industry, and the FAF benchmark value-per-ton ratios.

1. First, the current year output of commodities is multiplied by the corresponding 2008 ton-per-value of output ratios to arrive at the tons of freight by type of commodity for 2008.
2. Second, commodity shares are calculated based on the above tons, and these shares are used to disaggregate the growth in tons (i.e., total tons for 2008 minus total tons for 2007) into two-digit SCTG commodity details.
3. Third, multiplying the growth in tons by commodity with the corresponding value-per-ton ratios of 2002 provides the 2008 growth of freight values in 2002 dollars. These are added to the 2007 values (in 2002 dollars) to provide the provisional year value of freight in 2002 dollars.
4. Finally, the 2007 current dollar value of freight estimates, and the 2008 freight value growth (in 2002 dollars) are adjusted for inflation, and these two are added to obtain the current dollar value of freight for 2008.

This approach takes advantage of the available up-to-date information on the growth of highway freight tonnage and value, as well as the changes in composition of commodities and prices over time. In this method, the 2008 tons-to-value of output ratios by commodity is assumed to remain the same as that in 2002.

2.3.2 Estimate Growth Factors for Each O-D Pair by FAF Regions

The purpose of preparing growth factors is to enable the annual provisional commodity O-D estimates to capture the impacts of differences in regional growths on freight shipments. A State-County-FAF Region approach was used in estimating the regional growth factors. There are three reasons for using this approach. 1) All the necessary economic data for estimating regional growth factors are currently available at the state level, not at the FAF regional level. 2) Most of the available county-level economic data that can be used for estimating regional growth factors are not readily available in a timely fashion for preparing provisional estimates. Such data are usually released with a time lag of more than one year, and hence could not be used as primary inputs for our purpose. 3) Counties are sub-regions to both states and FAF regions, and hence they provide a bridge for the crosswalk between states and FAF regions.

The approach for estimating growth factors for each O-D pair involves the following steps:

1. Determine annual state growth rates
2. Estimate county share of state growth
3. Estimate annual FAF regional growth
4. Estimate annual growth factors for each O-D pair at the FAF region level.

2.3.2.1 Determine Annual State Growth

The best indicator of the size and growth of a state's economy is its GSP. Similar to GDP at the national level, GSP measures the annual net output of a state's economy. Given the positive link between freight and output, freight grows as the economy grows, and hence GSP can serve as a reasonable indicator of freight growth.

GSP estimates are published by the BEA and can be used to directly calculate the annual growth rates of states. However, the GSP data are only available with a lag of one year. Currently, 2006 is the latest year for which GSP data are readily available. This creates a timeliness problem for FAF annual provisional commodity O-D estimates, whose annual updates for a year are scheduled to be completed at the end of the same year. In order to overcome this problem, the State Quarterly Personal Income statistics from BEA are used to calculate state annual growth rates for the current year.[1] Using the quarterly personal income statistics, the current year growth of GSP by state is estimated by the following relationship:

$$\Delta GSPs = SG_s * GSP_{s,t-1}$$

Where:
ΔGSP_s = Current year growth of GSP for state s ($)
SG_s = Current year GSP growth rate (approximated by the growth of personal income) for states (%)
$GSP_{s,t-1}$= GSP for previous year for state s ($).

[1] State personal income is the income that is received by the residents of that state. Personal income is the most significant component and the main driving force of the GSP of a state.

2.3.2.2 Estimate County Share of State Growth

In order to calculate the growth in a FAF region, state growth factors are allocated among counties of that state. Then the county's share of the state growth is estimated. Then county growths are summed up to yield FAF regional growth.[2] Current year growths of counties are estimated using the following formula:

$\Delta CG_{k,s} = \Delta GSP_s * CS_{k,s}$

Where:
$\Delta CG_{k,s}$ = Current year growth of county k in state s ($)
ΔGSP_s = Current year growth of GSP for state s ($)
$CS_{k,s}$= Share of county k in the GSP of state s ($).

The county share of state GSP is estimated with the most recent data on total payroll of a county, which is obtained from the Census Bureau's CBP. These data are released with a lag of two years.

2.3.2.3 Estimate FAF Regional Growth

Current year FAF regional growth is calculated by summing up current year growths of counties within a given FAF region.

$\Delta RG_j = \sum \Delta CG_{k,j}$

Where:
ΔRG_j = Growth for region j
$CG_{k,j}$ = Current year growth of county k in region j.

2.3.2.4 Estimate Annual Growth Factors for Each O-D Pair of FAF Regions

Estimates of current year growths for all FAF regions provide the basic input information necessary for estimating annual growth factors of FAF O-D pairs. Instead of attempting to estimate the economic-spatial relationship between each pair of FAF regions using geo-spatial interaction models, such as various gravity models, the approach uses an interregional flow modifier method[3] for deriving growth factors of FAF O-D pairs. The method involves the following basic steps.

A. Converting economic growth into pseudo-growth in highway freight

The conversion of economic growth into pseudo-growth in highway freight tonnage is given by the following relationship:

$$\Delta PGT_{i,j} = \frac{\Delta \mathrm{CEG}_{i,j}}{\Delta \mathrm{CGSP}_{i,j,t-1}} * T_{i,j,t-1}$$

---

[2] Note that some FAF regions and states are the same, which means that the state growth and FAF regional level growth will be the same.

[3] Developed MacroSys as part of multi-regional Input-Output modeling research.

Where:
$\Delta PGT_{i,j}$ = Pseudo-growth in highway freight tonnage between two regions (region i and region j)
$\Delta CEG_{i,j}$ = Combined economic growth of the two regions for the current year , (i.e., region i and region j)
$CGSP_{i,jt-1}$ = Combined economic size of the two regions (i.e., region i and region j) for previous year or t-1
$T_{i,j,t-1}$ = Highway freight tonnage between region i and region j for previous year.

The combined economic growth and the combined economic size of region i and j in the above formula are established based on the real dollar GSP of region i and j.

Similarly, the conversion of economic growth into pseudo-growth in highway freight value is accomplished using the following formulation:

$$\Delta PGV_{i,j} = \frac{\Delta \mathrm{CEG}_{i,j}}{\Delta \mathrm{CGSP}_{i,j,t-1}} * V_{i,j,t-1}$$

Where:
$\Delta PGV_{i,j}$ = Pseudo-growth in highway freight value between two regions (region i and region j)
$\Delta CEG_{i,j}$ = Combined economic growth of the two regions for the current year , (i.e., region i and region j)
$\Delta CGSP_{i,j,t-1}$ = Combined economic size of the two regions (i.e., region i and region j) for previous year or t-1
$V_{i,j,t-1}$ = Highway freight value between region i and region j for previous year.

The freight value is estimated using current dollar values of the combined economic growth and the combined economic size of regions i and j. The combined economic size and growth of the regions are estimated using GSP statistics.

B. Estimating annual growth factor for each FAF O-D pair

Let $\Delta TPGT$ be the sum of all pseudo-growths of all FAF O-D pairs in highway freight tonnage ($= \Sigma \Delta PGT_{i,j}$), and let $\Delta PGT_{i,j}$ be pseudo-growth in highway freight tonnage of each O-D pair. Then the annual freight tonnage growth factor for each FAF O-D pair, $GFT_{i,j}$, is given by:

$$GFT_{i,j} = \frac{\Delta PGT_{i,j}}{\sum \Delta PGT_{i,j}} \text{ or } \frac{\Delta PGT_{i,j}}{\Delta TPGT}$$

Let $\Delta TPGV$ be the sum of all pseudo-growths of all FAF O-D pairs in highway freight value ($= \Sigma \Delta PGV_{i,j}$), and let $\Delta PGV_{i,j}$ be the pseudo-growth in freight value of each O-D pairs. Then the annual freight value growth factor for each FAF O-D pair, $GFV_{i,j}$, is given by:

$$GFV_{i,j} = \frac{\Delta PGV_{i,j}}{\sum \Delta PGV_{i,j}} \text{ or } \frac{\Delta PGV_{i,j}}{\Delta TPGV}$$

The separation between tonnage growth factors and value growth factors recognizes the differences in commodity components and their prices among FAF O-D pairs. The main advantage of the interregional flow modifier method is that it captures the special economic-spatial relationships developed over time among FAF regions and at the same time recognizes recent changes in these relationships.

2.3.3 Estimate Growth of Highway Freight for Each FAF O-D Pair

Once the annual growth factors are established, the estimation of growth in highway freight for each FAF O-D pair is straightforward and is obtained through the following formula.

A. Let ΔGT be the annual growth of national highway freight tonnage,[4] let $GFT_{i,j}$ be the annual freight tonnage growth factor of an FAF O-D pair between region i and region j. The annual growth for the FAF O-D pair in highway freight tonnage, $G_{i,j}$, is given by the formula:

   $\Delta G_{i,j} = \Delta GT*GFT_{i,j}$

B. Let ΔGV be the annual growth of national highway freight value, let $GFV_{i,j}$ be the annual freight value growth factor of FAF O-D pair between region i and region j. The annual growth for the FAF O-D pair in highway freight value, $\Delta G_{i,j}$, is given by the formula:

   $\Delta G_{i,j} = \Delta GV*GFV_{i,j}$

2.3.4 Determine the Provisional Freight Flow Estimates for Each FAF O-D Pair

The provisional estimate of highway freight tonnage of a FAF O-D pair for the current year is calculated by adding its estimated annual tonnage growth to its freight tonnage in the 2002 FAF benchmark year (or the provisional estimate of the previous year if the current year is two or more years away from the benchmark year).

$FT_{i,j,t} = FT_{i,j,t-1} + \Delta GT_{i,j,t}$

Where:
$FT_{i,j,t}$ = Highway freight tonnage for O-D pair i and j for year t
$FT_{i,j,t-1}$ = Highway freight tonnage for O-D pair i and j for year t-1.
$\Delta GT_{i,j,t}$ = Estimated annual tonnage growth for O-D pair i and
j for year t.

Similarly, the provisional estimate of highway freight value of a FAF O-D pair is calculated in the updating year by adding its estimated annual growth of freight value to its freight value in the FAF^2 benchmark year (or the provisional estimate of the previous year if the updating year is two or more years away from the benchmark year).

---

[4] Let T_t be the current year national tonnage and T_{t-1} be the previous year national tonnage. Then the growth in national tonnage for the current year is equal to T_t-T_{t-1}.

$FV_{i,j,t} = FV_{i,j,t-1} + \Delta GV_{i,j,t}$

Where:
$FV_{i,j,t}$ = Highway freight value for O-D pair i and j for year t
$FV_{i,j,t-1}$ = Highway freight value for O-D pair i and j for year t-1
$\Delta GV_{i,j,t}$ = Estimated annual growth of value for O-D pair i and j for year t.

2.4 Methodology for International Highway Freight

U.S. international highway freight shipments include freight flows between the U.S. and Canada, and between the U.S. and Mexico. Statistics on imports to and exports from Canada and Mexico by surface mode (highway, rail, and pipeline) are available from the North American TFD. Three sets of data are reported in the TFD, namely, state imports and exports by type of commodity using the Harmonized Schedule (HS) commodity classification method, state imports and exports by port of exit or entry, and U.S. imports and exports by port and commodity. The data on U.S. imports and exports by port and commodity details is a new addition to the TFD, beginning in 2007. However, no details are available on imports and exports by type of commodity, port of exit or entry, and origin and destination states in this database or from any other data source.

The TFD provides values and tons of imports, and value of exports. Data are not available on the weight of exports from this or any other known data sources. Beginning from 2007, the transborder data are reported for the 50 U.S. states and the District of Columbia, and for an unidentified (unknown) U.S. state.

2.4.1 Methodology for International Highway Freight

Based on import and export statistics from the TFD, and using the newly available information, an enhanced methodology of estimation has been developed. The methodology for preparing the 2008 provisional estimates involves the following steps:

1. Prepare details of state imports and exports by type of commodity, and port of exit or entry.
2. Allocate imports and exports of the unknown state to the 50 states and the District of Columbia.
3. Estimate the weight of exports, which is not provided in the TFD.
4. Convert data from TFD port of exit or entry to FAF international gateways.
5. Disaggregate state imports and exports into FAF Regions.
6. Convert data from HS commodity classification to SCTG commodity classification.

2.4.1.1 Preparing State Imports and Exports by Commodity and Port Detail

As pointed out above, an improved approach has been formulated and applied to estimate state imports and exports at the commodity and port of exit or entry level of detail. This approach keeps the original state imports and exports by type of commodity, and state imports and exports by port of exit or entry unchanged. In other words, when state-level imports and exports by type of commodity, and state-level imports and exports by port of exit or entry are derived from detailed estimates of state imports and exports by FAF region, international gateways, and type of commodity estimates, the results would be consistent with the actual data from the TFD. The methodology for imports and exports is exactly the same. The methodology for import freight flows is presented below.

Let

m_c = U.S. total imports from country c (i.e., Canada or Mexico)

m_{ic} = U.S. total imports of commodity *i* from country c

m_c^s = Total imports of State *s* from country c

$m_c{}^p$ = U.S. total imports through Port *p* from country c

$m_c{}^{sp}$ = Total imports of State *s* through Port *p* from country c

m_{ic}^s = Imports of commodity *i* by State *s* from country c

m_{ic}^p = Imports of commodity *i* through Port *p* from country c

m_{ic}^{sp} = Imports of commodity *i* by State *s* through Port *p* from country c.

Among the above variables, m_{ic}^{sp} is unknown. The unknown variable can be estimated from available data. Assuming that the imports of a commodity by a state through a port is proportional to the share of that port in the U.S. total imports of that commodity, the imports of a commodity by a state through a port can be derived using the following equation.

$$m_{ic}^{sp} = m_{ic}^s \times (m_{ic}^p / m_{ic}) \qquad (1)$$

The estimates derived from this equation should have the following two important properties:

a) $\sum_{p=1}^{P} m_{ic}^{sp} = m_{ic}^s$, since $\sum_{p=1}^{P} m_{ic}^{sp} = \sum_{p=1}^{P} (m_{ic}^s \times m_{ic}^p / m_{ic}) = m_{ic}^s \times \sum_{p=1}^{P} (m_{ic}^p / m_{ic})$

where $\sum_{p=1}^{P} (m_{ic}^p / m_{ic}) = 1$.

b) $\sum_{s=1}^{S} m_{ic}^{sp} = m_{ic}^p$, since $\sum_{s=1}^{S} m_{ic}^{sp} = \sum_{s=1}^{S} (m_{ic}^s \times m_{ic}^p / m_{ic}) = m_{ic}^p \times \sum_{s=1}^{S} (m_{ic}^s / m_{ic})$

where $\sum_{s=1}^{S} (m_{ic}^s / m_{ic}) = 1$.

However, the problem with m_{ic}^{sp} derived from equation (1) is that $\sum_{i=1}^{I} m_{ic}^{sp} \neq m_c^{sp}$. This means that state imports by port when derived from m_{ic}^{sp} may not be consistent with the actual data from the Transborder Freight Database. In order to address this problem, we introduce a measure of state port propensity for imports (pp_{ic}^{sp}), which is given by the following equation:

$$pp_{ic}^{sp} = [(m_c^{sp} / m_c^s)/(m_c{}^p / m_c)] / \sum_{p=1}^{P} [(m_c^{sp} / m_c^s)/(m_c{}^p / m_c)].$$

A state port propensity defined this way has the following important property:

$$\sum_{p=1}^{P} pp_{ic}^{sp} = 1.$$

Different sets of port propensity are used for imports from Canada and imports from Mexico. The purpose of port propensity is to adjust the results of equation (1) or m_{ic}^{sp} in order to make them closer to reality by taking into account the special relationship between a port and a state. The implicit assumption for this adjustment is that all commodities of a state's imports have the same port propensity. The adjustment of state imports by type of commodity through port (M_{ic}^{sp}) can be done using the following relationship.

$$M_{ic1}^{sp} = m_{ic}^{sp} \times pp_{ic}^{sp} = (m_{ic}^s \times m_{ic}^p / m_{ic}) \times pp_{ic}^{sp} \qquad (2)$$

Where M_{ic1}^{sp} = adjusted M_{ic}^{sp}

Since the assumption that all commodities of a state's imports have the same port propensity is not true, the adjusted estimates of the imports of a commodity by a state through a port (M_i^{sp}) loses the two important properties that m_i^{sp} has, which means that $\sum_{p=1}^{P} M_{ic}^{sp} \neq m_{ic}^s$, and

$$\sum_{s=1}^{S} M_{ic}^{sp} \neq m_{ic}^p.$$

In order to bring these two important properties back to the estimates, further adjustments are required. The adjustments can be done through an iterative process. The number of rounds of iteration depends on the accuracy requirement for the estimates. The higher the accuracy, the more rounds of iteration will be needed.

The goal is to make $\sum_{p=1}^{P} M_{ic}^{sp} = m_{ic}^s$.

From equation 2, $\sum_{p=1}^{P} M_{ic}^{sp} = \sum_{p=1}^{P} (m_{ic}^s \times m_{ic}^p / m_{ic}) \times pp_{ic}^{sp} = m_{ic}^s \sum_{p=1}^{P} (m_{ic}^p / m_{ic}) \times pp_{ic}^{sp}$. One way to achieve the goal is to adjust the $\sum_{p=1}^{P} M_{ic}^{sp}$ by a factor ∂_{ic}^s, where $\partial_{ic}^s = 1 / \left(\sum_{p=1}^{P} (m_{ic}^p / m_{ic}) \times pp_{ic}^{sp} \right)$.

Since m_{ic}^p, m_{ic}, pp_{ic}^{sp} can be determined from available data, ∂_{ic}^s can be also calculated from the data. Once ∂_{ic}^s is determined, we can adjust $\sum_{p=1}^{P} M_{ic}^{sp}$ as follows:

$M_{ic2}^{sp} = \sum_{p=1}^{P} \partial_{ic}^{s} M_{ic1}^{sp} = \partial_{ic}^{s} \times m_{ic}^{s} \sum_{p=1}^{P} (m_{ic}^{p} / m_{ic}) \times pp_{ic}^{sp} = m_{ic}^{s}$, where M_{ic2}^{sp} is adjusted M_{ic1}^{sp}.

$$M_{ic2}^{sp} = \partial_{ic}^{s} M_{ic1}^{sp} = \partial_{ic}^{s} \times m_{ic}^{s} \times (m_{ic}^{p} / m_{ic}) \times pp_{ic}^{sp}.$$

Next, we try to achieve $\sum_{s=1}^{S} M_{ic1}^{sp} = m_{ic}^{p}$ by following the same approach.

Since $\sum_{s=1}^{S} M_{ic2}^{sp} = \sum_{s=1}^{S} \partial_{ic}^{s} \times m_{ic}^{s} \times (m_{ic}^{p} / m_{ic}) \times pp^{sp} = m_{ic}^{p} \sum_{s=1}^{S} \partial_{ic}^{s} \times (m_{ic}^{s} / m_{ic}) \times pp_{ic}^{sp}$,
we can adjust the $\sum_{s=1}^{S} M_{ic2}^{sp}$ by a factor β_{ic}^{p}, where $\beta_{ic}^{p} = 1 / \left(\sum_{s=1}^{S} \partial_{ic}^{s} \times (m_{ic}^{s} / m_{ic}) \times pp_{ic}^{sp} \right)$.

Similar to ∂_{ic}^{s}, β_{ic}^{p} can be determined from available data. Once β_{ic}^{p} is determined, the adjustment of M_{ic2}^{sp} can be done as follows:

$\sum_{s=1}^{S} \beta_{ic}^{p} M_{ic2}^{sp} = \sum_{s=1}^{S} \beta_{ic}^{p} \partial_{ic}^{s} M_{ic1}^{sp} = \beta_{ic}^{p} \times m_{ic}^{p} \sum_{s=1}^{S} \partial_{ic}^{s} \times (m_{ic}^{s} / m_{ic}) \times pp_{ic}^{sp} = m_{ic}^{p}$, and

$$M_{ic1\beta}^{sp} = \beta_{ic}^{p} M_{ic2}^{sp} = \beta_{ic}^{p} \partial_{i}^{s} M_{ic1}^{sp} = \beta_{ic}^{p} \times \partial_{i}^{s} \times m_{i}^{s} \times (m_{i}^{p} / m_{i}) \times pp_{ic}^{sp}.$$

The adjustment of M_{ic2}^{sp} using β_{ic}^{p} will achieve the goal of $\sum_{s=1}^{S} m_{ic2}^{sp} = m_{ic}^{p}$. However, the adjustment changes the actual commodity composition of state imports and hence $\sum_{p=1}^{P} M_{ic1\beta}^{sp}$ may not be equal to m_{ic}^{s} ($\sum_{p=1}^{P} M_{ic1\beta}^{sp} \neq m_{ic}^{s}$). Therefore, we need to continue the adjustment process until the two important properties of detailed estimates hold true. At the 2nd round of the adjustment process, we will have:

$$\partial_{ic1}^{s} = 1 / \left(\sum_{p=1}^{P} \beta_{ic}^{p} \partial_{ic}^{s} (m_{ic}^{p} / m_{ic}) \times pp_{ic}^{sp} \right)$$

$$M_{ic3}^{sp} = \beta_{ic}^{p} \times \partial_{ic}^{s} \times \partial_{ic1}^{s} \times m_{ic}^{s} \times (m_{ic}^{p} / m_{ic}) \times pp_{ic}^{sp}$$

$$\beta_{ic1}^{p} = 1 / \left(\sum_{s=1}^{S} \beta_{ic}^{p} \partial_{ic}^{s} \partial_{ic1}^{s} (m_{ic}^{s} / m_{ic}) \times pp_{ic}^{sp} \right)$$

$$M_{ic2\beta}^{sp} = \beta_{ic}^{p} \times \beta_{ic1}^{p} \times \partial_{ic}^{s} \times \partial_{ic1}^{s} \times m_{ic}^{s} \times (m_{ic}^{p} / m_{ic}) \times pp_{ic}^{sp}$$

And at the n^{th} round of the iteration process, we will have:

$$\partial_{icn-1}^{s} = 1 / \left(\sum_{p=1}^{P} \beta_{ic}^{p} \beta_{ic1}^{p} \beta_{ic2}^{p} \cdots \beta_{icn-2}^{p} \partial_{ic}^{s} \partial_{ic1}^{s} \partial_{ic2}^{s} \cdots \partial_{icn-2}^{s} (m_{ic}^{p} / m_{ic}) \times pp_{ic}^{sp} \right)$$

$$M_{icn}^{sp} = \beta_{ic}^{p}\beta_{ic1}^{p}\beta_{ic2}^{p}\cdots\beta_{icn-2}^{p}\partial_{ic}^{s}\partial_{ic1}^{s}\partial_{ic2}^{s}\cdots\partial_{icn-1}^{s}\times m_{ic}^{s}\times(m_{ic}^{p}/m_{ic})\times pp_{ic}^{sp}$$

$$\beta_{icn-1}^{p} = 1/\left(\sum_{s=1}^{S}\beta_{ic}^{p}\beta_{ic1}^{p}\beta_{ic2}^{p}\cdots\beta_{icn-2}^{p}\partial_{ic}^{s}\partial_{ic1}^{s}\partial_{ic2}^{s}\cdots\partial_{icn-1}^{s}(m_{ic}^{s}/m_{ic})\times pp_{ic}^{sp}\right)$$

$$M_{icn\beta}^{sp} = \beta_{ic}^{p}\beta_{ic1}^{p}\beta_{ic2}^{p}\cdots\beta_{icn-1}^{p}\partial_{ic}^{s}\partial_{ic1}^{s}\partial_{ic2}^{s}\cdots\partial_{icn-1}^{s}\times m_{ic}^{s}\times(m_{ic}^{p}/m_{ic})\times pp_{ic}^{sp}$$

The iteration process continues until ∂_{icn}^{s} and β_{icn}^{p} are approaching 1, and the difference between $\sum_{p=1}^{P} m_{icn}^{sp}$ and m_{ic}^{s}, and between $\sum_{p=1}^{s} m_{icn}^{sp}$ and m_{ic}^{p} are approaching zero.

Although the above discussion concentrates on imports, the same approach has been used for estimating details of state exports at commodity and port level of detail.

2.4.1.2 Allocating Imports and Exports of Unknown State to the Fifty States and the District of Columbia

Starting from 2007, the TFD provides imports to and exports from an unknown U.S. state. The imports and exports of the unknown state should be allocated among the fifty states and the District of Columbia. For this purpose, two sets of data (imports and exports by port, and imports and exports by commodity) are obtained for the unknown state. First, similar to other states, details of state-level imports and exports by type of commodity and port of exit or entry are estimated using the methodology discussed above. Second, the share of each state's imports and exports by port and commodity detail in the total U.S. imports and exports by port and commodity detail are calculated. Finally, these shares are applied to the data on imports and exports of the unknown state to allocate the unknown state imports and exports to the 50 states and the District of Columbia. As a result of this adjustment, the final estimates of imports and exports for each state would be greater than the actual figure reported in the TFD. However, the total imports and exports, including the commodity composition or port of exit and entry data, would not be affected by this adjustment.

2.4.1.3 Estimating Export Tonnage

Data on the tonnage of U.S. exports to Canada and Mexico by truck mode are not available from the TFD or any other data source. In order to fill in this data gap, an imports weight-value ratio approach has been used. Two sets of weight-value ratios at the two-digit SCTG commodity level of detail are used for this purpose. One set of weight-value ratios is calculated based on imports statistics from Canada and the other set of ratios is computed using imports from Mexico. These ratios are country-specific and, therefore, recognize differences in the characteristics of U.S. trade with these two countries. The ratios have been smoothed using a simple moving average (SMA) method to reduce the impacts of over time and extreme variations. Multiplying the export values by the weight-value ratios of imports provides the tonnage of exports. This method assumes that the weight-value ratios of U.S. exports to Canada are the same as the weight-value ratios of U.S. imports from Canada, and the weight-value ratios of exports to Mexico are the same as the weight-value ratios of imports from Mexico.

2.4.1.4 Converting Data from Transborder Port to FAF International Gateways

The category of port of exit or entry in the TFD and the international gateways in the FAF database are different. We converted the data from the former to the latter. The conversion involves two steps: (1) establishing a crosswalk between Transborder's port of exit/entry and FAF international gateways; this requires identifying the exact location of the port of exit/entry of the TFD and assigning them into the proper FAF international gateways; and (2) applying the crosswalk to the freight data from the TFD.

2.4.1.5 Disaggregating State Imports and Exports into FAF Regions

The statistics on state-level imports and exports by type of commodity and international gateways are disaggregated at the FAF region O-D level of detail using the level of economic activity of each region during the year, which reflects the changes in the characteristics of freight shipment among regions. The economic activity of each FAF region has been approximated by employees' salaries and benefits received during the period. For this purpose, data on county-level payroll have been obtained from the Census Bureau's CBP. The county-level payroll is aggregated at the FAF region level of detail. The shares of FAF region payroll are calculated and applied to disaggregate state imports and exports by type of commodity and international gateways at the FAF region O-D level of detail.

2.4.1.6 Converting Data from Harmonized Schedule (HS) to Standard Classification of Transportation Goods (SCTG)

The commodity details in the TFD are reported using the HS commodity classification method. The data were converted into SCTG commodity classifications using a BTS working cross-walk between HS and SCTG. The cross-walk used for this purpose is based on the two-digit HS commodity since the data from the TFD are provided only at the two-digit HS commodity level of detail. For a more precise conversion of HS to SCTG, it would be advisable to use a four- or six-digit HS commodity data, and a cross-walk established based on this greater level of detail.

2.4.2 Value of Imports and Exports in Constant Dollars

The current dollar values of imports and exports are deflated by import and export price indexes, respectively, to obtain constant dollar freight values. The BLS publishes price indexes for imports and exports by selected HS commodity using 2000 as the base year. The price indexes are converted to the two-digit SCTG commodity level of detail by employing relative weights from the BLS. These price indexes are adjusted to reflect 2002 as the base year for deriving freight values in constant dollars.

3.0 FREIGHT MOVEMENT BY AIR

3.1 Introduction

Air cargo is a key part of the overall freight transported in terms of its dollar value, time sensitivity, and reliance on other shipment modes. This report outlines a method to integrate U.S. Census Bureau value data with the Department of Transportation (DOT), Office of Airline Information (OAI) weight data in order to develop two datasets containing data on commodity, commodity value, and weight of air shipments by O-D for domestic shipments; and data on commodity, value, weight, and origin-port of entry/exit-destination for international shipments.

The aviation component of the provisional commodity O-D estimates combines Office of Airline Information (OAI) data on the weight of shipments for the U.S. airline industry with Census/Customs (hereafter Census) data on commodity type, value and weight for imports and exports by air, and the FAF^2 domestic aviation value and weight data. The major reasons to use OAI data are the ability to estimate a port-of entry/exit and that it is considered the definitive source for tons of U.S. air freight shipped. While the Census data do provide port-of-entry/exit identities, these are based on the port in which a shipment clears customs rather than the first port after/before crossing the border. The main reasons for using the Census data are the availability of information on commodity type and value. The major contribution of the FAF^2 domestic aviation data is to capture commodity type and value differences between the international and domestic data. This report specifies the process for combining the OAI, Census, and FAF^2 data and the methodologies for estimating the port-of-entry/exit and for forecasting data for months that have not yet been reported.

3.2 Air Freight Data Sources

3.2.1 OAI Data

The Office of Airline Information (OAI-BTS/RITA), part of the U.S. DOT Research and Innovative Technology Administration, publishes the Form 41T-100 and T-100 (f) traffic data monthly on both a market and segment basis. The T-100 data contain information on the weight of air freight and mail by carrier, origin airport, and destination airport, as well as additional identifying and operational information. The OAI data are considered the definitive source of tons shipped for the U.S. airline industry. OAI shipments are defined differently than FAF^2 shipments in that OAI shipments use an airport basis (from airport origin to airport destination) rather than an establishment basis. In OAI market data, airport O-D refers to tons enplaned by a specific carrier at the origin airport and deplaned by the carrier at the destination airport. The T-100 market data will exclude the port-of-entry/exit whenever the port is an intermediate stop for the shipment. O-D for each record on the segment component of the T-100 data refers to a non-stop leg and reports tons transported rather than tons enplaned. The T-100 segment data will include the port of entry/exit for international shipments, but will exclude the ultimate O-D when a shipment has multiple stops. Combining the market and segment data to add ports-of-entry/exit is one of the main objectives of this project. The T-100 data covering freight shipments by U.S. carriers is publicly available approximately 60 days after the end-of-month and the T-100 (f) data covering foreign carriers is publicly available approximately six months after the end-of-month. The data can be found at http://www.transtats.bts.gov/ (the T-100 (f) data are included in the versions having all carriers).

Two other differences between the OAI data and FAF^2 are the lack of information on the value and commodity composition of shipments. In order to provide information for FAF^2 international air shipments, U.S. Customs data on commodity type and value are combined with the OAI data.

The coverage of the OAI data may be summarized with a few aggregate statistics. In 2003, freight data were recorded for almost 1,500 airports worldwide. About 600 of these were international airports where they were engaged in shipments between the U.S. and other countries. About 200 of these international airports were located in the U.S. and its territories. The OAI T-100 data cover large certificated U.S. commercial carriers; since 10/2002, commuter and small certificated carriers are covered as well, although these will account for only a negligible amount of international air shipments. The T-100 (f) covers foreign carriers serving the U.S. Included in these carriers are parcel, courier, and express carriers, which are treated as a separate mode in FAF^2. In 2003, the T-100 and T-100 (f) showed 244 air carriers shipping freight in the U.S., and 188 carriers shipping freight between the U.S. and other countries (119 of these were foreign carriers). Like FAF^2, the public version of the T-100 data excludes in-transit shipments from the market data and foreign-to-foreign shipments from the segment data; however, see the Further Notes on the Data below for a qualification. The T-100 data do not include private or illegal shipments of freight, and passenger baggage is not counted as freight.

3.2.2 Census Foreign Trade Data

The Census Bureau Foreign Trade Division (FTD) (http://www.census.gov/foreign-trade/reference/products/index.html) publishes two monthly paid subscription series that largely satisfy the need for International Air data. The data are collected by the U.S. Customs Service and published as: 1) U.S. Exports of Merchandise – Monthly – DVD ROM (information on the value, quantity, method of transportation, and shipping weights for 9,000 export commodities, 240 trading partners, and 45 Districts; 2) U.S. Imports of Merchandise – Monthly – DVD ROM (data on more than 17,000 commodities for 240 trading partners and 45 Districts. The data provide value, quantity, method of transportation, shipping weights, import charges, duties, and much more.) Shipments are for all merchandise between foreign countries and U.S. Customs Territories (50 states, District of Columbia, Puerto Rico, the U.S. Virgin Islands, and U.S. Foreign Trade Zones [FTZs]). The objective is to capture the physical movement of merchandise between foreign countries and the U.S. It includes government and non-government shipments and does not depend on the shipment being part of a commercial transaction.

A shipment's O-D on the Census data is based on Customs Districts and where the shipment is processed by the Customs Service. For FAF^2 purposes, a Customs Districts may include more than one state and a state may have more than one Customs District. The Export data satisfies the need for mode-destination-port of origin-tonnage-dollar value, but lacks port-of-exit data. The Import data satisfies the need for mode-origin-destination-tonnage-dollar value, but defines port-of-entry as the port in which the shipment clears customs rather than the first port after crossing the border. Commodities are reported using the 10-digit HS (Schedule B for exports), which can be translated to SCTG using a crosswalk provided by FHWA. Export values are reported free-alongside-ship (FAS). Import values are available both by customs-import-value (CIV.), which excludes duties, freight, insurance, and other costs of importation; or by customs-insurance-freight (CIF), which adds freight and insurance to the CIV. For FAF^2 the CIF values are used to better reflect the shipment's value at the border. The data are available approximately three months after the end-of-month. Export data are recorded in the month in which the shipment leaves the country, corresponding to the FAF^2 definition. However, import data are recorded in the month in which the shipment clears customs and may therefore not correspond to the month the shipment was transported into the country due to time spent in bonded warehouses or FTZs. Like FAF^2, the Census data exclude in-transit shipments. Although the Census Bureau data provide vital information for the FAF^2 project, there is also a substantial on-going cost to subscribe to the dataset, currently $2,700/year for both Imports and Exports. Therefore it may be useful to consider a related subset of data that is available on-line for $75 for a one-month subscription at http://www.usatradeonline.gov/usatrade.nsf?Open&mc=F9000 for future use. Appendix A compares the dimensions of the data sources used to produce the provisional estimates.

3.2.3 Further Notes on the Data

a) Although the T-100 market/segment data include information on the largest cargo carriers, it excludes information for some all-cargo carriers.

b) The methodology below can be applied to cargo and mail either separately or together as freight and mail combined. One concern with using cargo and mail separately is an on-going dispute between Federal Express (FedEx) and OAI as to how U.S. mail should be reported. Federal Express lumps mail with freight due to concerns about disclosing the size of its contract with the U.S. Postal Service. Cargo by itself will then tend to overstate actual cargo shipments. For purposes of this report, combined freight and mail is used while recognizing that FAF^2 treats parcels/mail as a separate mode.

c) Neither of the OAI datasets gives information about the initial origin at the manufacturer or the ultimate destination at the purchaser of the products. However, it is very unlikely that a given shipment of air freight originates outside of the FAF zone where the airport is located. Also it is very unlikely that a shipment of air cargo will be transported outside the FAF region where the airport is located. One more issue is that the data are carrier-based, so a shipment that involves more than one carrier will have misrepresented the initial origin and ultimate destination of the shipment. The methodology here assigns airport origins and destinations and will distribute to ultimate origins and destinations based on the methodology used for the 2002 FAF^2.

d) The OAI market data are reported by carriers and covers enplanements and deplanements of freight and mail. Although the public version of the dataset excludes in-transit shipments, it is likely that some in-transits are included, since a shipment that changed either carriers or planes would not be excluded.

e) There was a substantial expansion in coverage of the T-100 OAI data in October 2002 to include all-cargo carriers, small-certificated carriers, and commuter carriers. For estimation of growth rates across years involving 2002, carrier growth rates in revenue ton-miles of freight and mail (available from the T1 data) were used to backcast 2003 monthly data for individual routes for the largest all-cargo carriers. For example, the FedEx tons enplaned at Memphis and deplaned in Seattle in January 2003 would be decreased by the January 2002 to January 2003 growth rate in FedEx domestic revenue ton-miles of freight and mail to obtain an estimate of January 2002 tons shipped by FedEx between Memphis and Seattle.

3.3 Combining Census and OAI Data

3.3.1 Cross-Walks for Commodity and Geographic Information

Combining the OAI and Census data into a FAF^2 dataset requires reconciling the different levels of detail at which commodity and geographic identifying information are collected and stored. In the case of commodity types and values, the OAI data are at a more general level than is required by FAF^2 – a topic that is covered in the sections below on estimation. This section covers the cross-walks used to reconcile differences between the commodity types on the Census and FAF^2 datasets and the geographic information on all three datasets.

Several of cross-walks were already available from FHWA. Commodity cross-walks between the HS used in the Census Foreign Trade files to the SCTG codes used in FAF^2 are available at http://ops.fhwa.dot.gov/freight/freight_analysis/faf/faf2_tech_document.htm. Cross-walks between countries and foreign trade regions are available upon request from FHWA (contact Michael Sprung at Michael.Sprung@fhwa.dot.gov). A third cross-walk from U.S. counties to FAF regions was also provided by FHWA.

The cross-walks to be developed are translations between different levels of specificity for geographic information between the OAI data and Census/FAF^2. The OAI geography is based on airports, the most specific level of detail, and is used as link between the other two. Each airport is assigned to both a Customs District and a FAF^2 region so that the relevant (dis)aggregation can be accomplished.

The first cross-walk developed for FAF^2 International Aviation is from U.S. airports to counties, which is used in combination with the existing cross-walk from counties to FAF^2 regions. The matching process requires two supplemental files: the Master Coordinates File (MCF) from OAI, available at http://www.transtats.bts.gov/Tables.asp?DB_ID=595&DB_Name=Aviation%20Support%20Tables&DB_Short_Name=Aviation%20Support%20Tables, and the county subdivision file from Census, available at http://www.census.gov/geo/www/gazetteer/places2k.html.

Two other sources also proved useful when the assignment of an airport to a county was unresolved from the first round of processing: Mapquest ® at http://www.mapquest.com/maps/ and the National Association of Counties website at http://www.naco.org/Template.cfm?Section=Data_and_Demographics&Template=/cffiles/counties/city_srch.cfm. Both the MCF and County Subdivision files have information on the state and on latitude and longitude. Within each state, the airports are matched to the two closest county subdivisions. Two subdivisions are matched because an airport may be near the border of its actual county and closer to the geographic center of another county. When the two closest subdivisions were in the same county, the airport was assigned to that county. When the two closest subdivisions were in different counties, the airport city name from the MCF was used to determine the county using either Mapquest ® or the National Association of Counties website.

The second cross-walk developed for FAF^2 International Aviation is from U.S. airports to U.S. Customs Districts. As noted above, the MCF provides information on airports in the form of airport name, state, city name, and latitude and longitude. In order to assign airports to Customs Districts a hierarchical matching method is used. Matching airports to Customs Districts is more complicated because Customs Districts are less uniform than counties, i.e., a state may have multiple Customs Districts, no named District or Sub-District, or a Customs District may span more than one state. While Customs Sub-Districts also consist of places in the usual geographic sense of cities or regions, they may also be airports and business places (e.g., FedEx processing centers).

Matching Customs Sub-Districts to Airports

1. For those Sub-Districts which are also airports, assign the airport to that Sub-District. A list of Customs Districts/Sub-Districts is available at http://www.census.gov/foreign-trade/schedules/d/dist.txt.
2. For the remaining Sub-Districts, match the Sub-District name to a Census Place Name (a list of census place names, that includes latitude and longitude, is also available at http://www.census.gov/geo/www/gazetteer/places2k.html). This process required much hand-editing and also the use of supplemental information from the Code of Federal Regulations (CFR), customs, Mapquest ®, and the National Association of Counties websites.
3. Use the latitude and longitude information available from both the Census Places file and the MCF to determine the closest Sub-Districts for each airport. Choose between the two based on the airport city name or additional information. Note that an airport may be, and often is, on the outskirts of its actual place, and therefore closer to a second place.

The matching process resulted in each U.S. airport being assigned to a Customs Sub-District.

3.3.2 Estimating Flows by Weight

Estimation for domestic and international data is substantially different, with the estimation for domestic data being straightforward. Estimation of domestic data consists of calculating growth rates from OAI domestic market data by FAF origin region between the CFS year and the provisional year required for FAF. These growth rates are then applied to the individual commodity weights from 2002 FAF by origin region to obtain the estimates.

The OAI market data for international shipments are missing the port-of-entry/exit, while the Census foreign trade data are missing the port-of-exit for exports. Also, the port-of-entry for imports does not necessarily correspond to the FAF definition of port-of-entry. This section outlines a procedure for reconciling the differences between the two datasets and assigning a port-of-entry/exit to the OAI market data based on the OAI segment data. The guiding philosophy behind the algorithm is to impose aggregate efficiency by minimizing the distance transported at each step. The specification of the algorithm is based on a port-of-exit. The extension to a port-of-entry is straightforward.

Notation:
Superscripts:
 1st Position: M = market data, S = segment data, F = FAF^2 results.
 2nd Position: t = time period. Time periods are annual.
Subscripts:
 1st Position: i = origin airport
 Last Position: j = destination airport
 Intermediate Position in the case of 3 nodes: k = port-of-entry/exit

T = tons shipped.

a) For market routes that match non-stop segment routes for both origin and destination and by carrier, assign the min($T^{Mt}_{ij,}$ $T^{St}_{ij,}$) to T^{Ft}_{ij} and reduce both T^{Mt}_{ij} and T^{St}_{ij} by T^{Ft}_{ij} to obtain residual tonnage for each market route and port.

b) Determine the remaining market and segment routes from the remainders from step a) and all market and segment routes that did not match in step a).

c) Create two-leg routes from the segment data in which the origin and destination match the origin and destination from the market routes in step b) with the intermediate stop restricted to be domestic: T^{St}_{ikj} = min($T^{St}_{ik,}$ T^{St}_{kj}) where {i, j} correspond to {i, j} from b) and k is domestic.

    i. For each carrier and each market route, find the best (based on shortest distance) intermediate stop. Let the distance for this route be given by dist(ik^1j).
    ii. For each carrier and each market route, find the second best intermediate stop with distance given by dist(ik^2j).
    iii. Calculate the cost savings for each route using the best intermediate stop = dist(ik^2j) - dist(ik^1j).
    iv. For each carrier, find the route which gives the greatest cost savings and denote this route (ikj)*. Then let T^{Ft}_{ikj*} = min($T^{Mt}_{ij,}$ T^{St}_{ikj*}).
    v. Bookkeeping: Reduce T^{Mt}_{ij}, T^{S}_{ik} and $T^{S}_{.kj}$ by T^{Ft}_{ikj*} for the carrier.
    vi. Repeat steps a-e until all market routes have been evaluated for all carriers.

d) Determine the remaining market and segment routes from the remainders from step b) and those routes for which no two-leg routes could be formed.

e) For each carrier, aggregate airports to their FAF region level. Recalculate distance as the ratio of ton-miles to tons transported rather than airport-to-airport distance.

f) Rerun step c) using the FAF region level rather than the airport level.

g) For the international routes that were unmatched in steps a, c, and f, assign the port-of-entry/exit to be the domestic destination/origin.

h) Create an international dataset based on international routes from steps a, c, and f, plus the international market routes from step g). (Note: Step h) is actually done in tandem with the assignment of commodity type and value. This aspect is excluded here for simplicity.

i) Aggregate the results over carriers and airports to FAF regions.

The result of this algorithm will be a dataset with shipment weights by origin-port of exit-destination. Matching is done at the carrier level to preserve the correspondence between market and segment data in the OAI datasets. Appendix B, Table B1, provides round-by-round results of the estimation. Round 1 corresponds to step a), round 2 to step c), round 3 to step f), and round 4 to step h). For imports (exports), rounds 2 and 3 assign a different port-of-entry (exit) than the domestic destination (origin). These two rounds accounted for about 17% of imports and 13% of exports. A large majority of the data, more than 70% for both imports and exports, is assigned its port-of-entry/exit in the first round where it is equal to the original port-of-entry (exit) for imports (exports).

3.3.3 Estimating Commodity Composition and Value

3.3.3.1 International Routes

Commodity composition and value are available from Census for exports at the domestic origin, and for imports at the domestic destination based on the Census geographic definitions. The Census information is used to estimate commodity composition and value for the OAI data. The corresponding domestic origin for exports and domestic destination for imports from the OAI data will be referred to as the matching ports.

The first step in the estimation process is to determine whether it is reasonable to combine the two data sources. Evidence that combing the data is reasonable is given by the high correlation for tonnage values at the matching ports between the two data sources (see the bottom of Table B2 in Appendix B). Two caveats to this estimation need to be noted. The first is that the OAI data are about 20% larger than the Census data for both imports and exports. Although there are several differences between the data sources, the strongest explanation is that the difference is due to the OAI data including more in-transit shipments than the Census data. The OAI data are based on carrier reporting with market routes defined by enplanement and deplanement of the cargo. An in-transit shipment that switches carriers in the U.S., or which is transferred from one plane to another by the same carrier, would appear as an import/export on the OAI data. However, the same shipment would be more likely to be designated as in-transit in Customs' reporting to Census. Additional evidence that differences are due to in-transit shipments is that the Customs Districts with the largest differences are also likely transshipment ports: New York, Miami, and Anchorage. This is the second caveat, because it affects whether it is reasonable to apply Census information to the OAI data on a district-by-district basis. The second caveat is addressed in the estimation process.

The estimation philosophy is to assign the Census commodity distribution and value-per-ton by commodity (prices) to the OAI data for each matching port while keeping the aggregate commodity distribution and prices equal to that for the Census data. Because of the differences in tonnage for some key ports, a straightforward port-by-port application would result in large differences at the aggregate level. The first step in the port-by-port estimation is to rescale Census exports/imports by the ratio of the respective OAI-to-Census aggregates. The approach taken here has two parts. For the share of a matching port's tons that is on both the OAI and rescaled Census data, the distribution and prices are taken directly from the Census data for that port. The remainder can be either excess Census tons, or excess OAI tons. The commodities and values for matching ports with excess Census tons are then aggregated to define residual commodity shares and prices. The residual commodity shares and prices are then applied to those Customs Districts with excess OAI tons. The result is an OAI-based dataset that reflects the Census commodity distribution and prices at the aggregate level and also captures a large share of Census port-level differences in commodities and prices.

More formally, the estimation algorithm can be written in terms of exports as follows:
Notation:
Superscripts:
1st Position: O = OAI data, C = Census data, , R = residual of OAI minus Census, F = FAF^2 results.
2nd Position: t = time period. Time periods are annual.
Subscripts:
1st Position: i = origin Customs District, i=1,…,I, I=41.
2nd Position: j = commodity j, j=1,…,J, J=33
When the capital of the subscript letter is used, it denotes the sum over all values of the subscript.

T = tons shipped.
V = value.
α = a 33x1 vector of commodity shares, $\alpha_{ij}^{Ct} = T_{ij}^{Ct} \Big/ \sum_{j=1}^{J} T_{ij}^{Ct} = T_{ij}^{Ct} \Big/ T_{iJ}^{Ct}$

p = a 33x1 vector of commodity prices, $p_{ij}^{Ct} = V_{ij}^{Ct} \Big/ T_{ij}^{Ct}$

σ = the export scale factor = $\sum_{i=1}^{I}\sum_{j=1}^{J} T_{iji=1}^{Ot} \Big/ \sum_{i=1}^{I}\sum_{j=1}^{J} T_{ij}^{Ct} = T_{IJ}^{Ot} \Big/ T_{IJ}^{Ct}$

i. Let $T_{iJ}^{Rt} = T_{iJ}^{Ot} - \sigma T_{iJ}^{Ct}$

ii. Let $A = \{i \mid T_{iJ}^{Rt} < 0\}.and.B = \{i \mid T_{iJ}^{Rt} > 0\}$

iii. Let $\alpha_{j}^{At} = \sum_{i \in A} T_{ij}^{Rt} \Big/ \sum_{i \in A} T_{iJ}^{Rt}$

iv. Let $p_{ij}^{At} = \sum_{i \in A} p_{ij}^{Ct} \cdot \alpha_{ij}^{Ct} \cdot T_{iJ}^{Rt} \Big/ \sum_{i \in A} \alpha_{ij}^{Ct} \cdot T_{ij}^{Rt}$

v. Let $M_{iJ}^{t} = Min(T_{iJ}^{Ot}, \sigma T_{iJ}^{Ct})$

vi. Then $T_{ij}^{Ft} = \alpha_{ij}^{Ct} M_{iJ}^{t} + \max(\alpha_{j}^{At} T_{iJ}^{Rt}, 0)$,

vii. and $\alpha_{ij}^{Ft} = T_{ij}^{Ft} \big/ T_{iJ}^{Ft}$,

viii. and $p_{ij}^{Ft} = \left[p_{ij}^{Ct} \cdot \alpha_{ij}^{Ct} \cdot M_{iJ}^{t} + \max(p_{ij}^{At} \cdot \alpha_{j}^{At} \cdot T_{iJ}^{Rt} \right] \big/ T_{ij}^{Ft}$

The resulting FAF commodity shares and prices are then applied at the airport level before aggregating to create tons and value at the FAF regional level.

3.3.3.2 Domestic Routes

The only information available on commodity distribution and prices for domestic routes is the 2002 CFS survey, which is also the basis for the 2002 FAF database, and this is used as the base for estimation[5]. In order to more accurately reflect the values for non-survey years, the commodity price information from the CFS is updated using the commodity price data from Census on exports. Exports are used for two reasons: exports more closely resemble domestic production than imports and, for 2002, commodity shares and prices of exports are more highly correlated with CFS commodity shares and prices. In particular, commodity prices are calculated at the national level for exports for the provisional estimation year, and then the ratio of the commodity price in the provisional year to the 2002 level is used to inflate/deflate domestic commodity prices obtained from 2002 FAF data.

Domestic air freight commodity shares are unchanged at the individual route level, or when aggregated by origin FAF region, since individual commodity weights are estimated to grow at the same rate as the weight of OAI shipments from the origin FAF region. Note, however, that commodity shares for destination regions can change since they receive shipments from more than one origin region.

3.4 Forecasting Air Freight Data for the Remainder of the Year

The first release of current year estimates are to be made available in December of the current year so forecasts are required for weight, value and the commodity distribution for unreported months[6]. For both Census data and domestic shipments by domestic carriers on the OAI data, the missing data consists of the fourth quarters of the most recent year. Foreign carriers on the OAI data (who have minimal domestic shipments) and domestic carriers' international shipments will require the third and fourth quarters to be forecast. The specific forecast techniques are selected based on historical evidence, but the basic approach is to first forecast tons shipped based on the OAI data, and then to forecast values and the commodity distribution based on the Census data. The forecasts use the most recent data on annual changes to update the available data from the fourth quarter of the previous year. Using the available data from the fourth quarter helps to retain the seasonal pattern for routes, commodity distribution, and relative prices

[5] An adjustment is made to the CFS data to account for observations that have been rounded to zero for either value or tonnage. National level prices are calculated for each commodity based only on observations for which both value and tonnage are greater than zero. The national level price is then used to calculate the missing tonnage (value) by dividing the non-zero value by the price (multiplying the non-zero tonnage by the price).

[6] For revised releases, the most recently available data may be used, avoiding the need to use forecasts.

for the fourth quarter of the current year. The results of the forecasts are then used to supplement the available data so that the methods described above for estimating air freight flows can be applied.

The specific technique will be part of the broad class called time-series techniques. The general alternatives to time-series techniques are model-based techniques, which hypothesize relations between variables and estimate a model based on those relations. The problem with model-based techniques for the FAF^2 is that the use of variables outside the database (e.g., fuel prices) restricts how the forecast data can be used (e.g., how does the price of fuel affect congestion) for independent study. In the case of using fuel prices to help forecast missing data, the effect of fuel prices would be pre-determined by the forecast model rather than reflecting actual conditions. Time-series techniques in contrast use only the past histories of the variables of concern to forecast the future.

3.4.1 Forecasting the OAI Data for Tons Shipped

Two significant events have changed the characteristics of the OAI data and limited the efficacy of using the history of the series prior to 2002. The first is the 9/11/01 terrorist attacks, which had a profound, direct effect on aviation. The second is the carrier coverage of the T-100 data, which expanded in 10/2002 to include small certificated, commuter, and all-cargo carriers. Carriers that began full reporting in 2002 will be referred to as new-reporters while those who fully reported prior to 2002 will be referred to as prior-reporters. The primary impact of this change is on domestic tons shipped, because international operations were already reported prior to 10/2002. Of particular significance, domestic operations of Federal Express were not publicly reported prior to 10/2002. Given these events, the historical period used as a base for forecasting is restricted to 2002 and later. The growth rates that are the basis of the forecasts are also restricted to depend only on information from the previous year to allow for an evolving trend following September 11. The limited availability of data reduces the number of parameters that can be estimated and the ability to apply standard statistical tests. For these reasons, simple techniques that depend on only one estimated parameter were considered.

The techniques examined consist of using data on annual growth rates between the previous and current calendar year and then applying these growth rates to missing quarter(s) from the previous calendar year. For example, one of the forecasts for domestic carriers uses the growth rate from the third quarter of the previous year to the third quarter of the current year. The forecast for the fourth quarter of the current year is obtained by applying this growth rate to the level of tons enplaned in the fourth quarter of the previous year. Annual growth rates are used to avoid seasonal effects, which may have also have changed since September 11. The forecasts considered differ along three dimensions: whether the time period used to calculate the growth rates is the year-to-date (YTD) or the most recent completed quarter (depending on availability) relative to the same period in the previous year, whether to forecast domestic and international routes separately or in combination, and whether to forecast prior- and new-reporting carriers separately or in combination.

Another potential problem with the most recent data may be carriers who are late in reporting their data to OAI. To correct for the missing carrier effect, an adjustment is made to the data for the most recent year. The adjustment is based on the assumption that late-reporting carriers grew at the same rate as those who reported on time. Adjusted growth rates are calculated for each month after January, with the growth rate for each month based on aggregate enplaned tons from the subset of carriers who reported in both the current and previous month. The adjusted growth rate is then consecutively applied to each month after January, subject to a constraint that the adjusted aggregate enplaned tons is greater than aggregate enplaned tons obtained directly from the data (since the adjustment is to account for late reporters).

3.4.1.1 Mathematical Specification of the Adjusted Tons and Forecasts[7]

Let $T_{t,i,j,k}^{m/q,h}$ = tons enplaned and $F_{t,i,j,k}^{m/q,h}$ = the forecast of tons enplaned as defined below.

Let m/q = m
Where m = current month i.e., the latest month for which data is available for the carrier type (in the current year (generally September for domestic carriers and June for foreign carriers)
q = current quarter i.e., the latest quarter for which data is available for the carrier type in the current year (generally the third quarter for domestic carriers and the second quarter for foreign carriers).

Let h = n, r, a, u index carrier subsets
where n = no restrictions on carriers included for the respective group
r = carriers are restricted to have reported in both the current and previous month
a = adjusted for missing carriers
u = unadjusted for missing carriers.

Let t = y indicates the current year.

Let i = b, d, s, c index route groupings
where b = international (border-crossing) routes
d = domestic routes
s = the sum of forecasts of international and domestic routes
c = the forecast based on the growth rate of combined domestic and international routes.

Let j = n, p, s, c, f index carrier groups
where n = new-reporters
p = prior-reporters
s = the aggregate of separate domestic carrier groups (the sum n and p forecast individually)
c = domestic carriers combined over both new- and prior-reporters
f = foreign carriers.

[7] Note: The specification is geared toward the usual situation where data are available through September for domestic carriers and June for foreign carriers. In the event the forecast is implemented when fewer or greater months are available, then modifications would be required. For growth rates, the latest quarter would refer to the latest available three months. For example, if data are available only through August for domestic carriers, then the latest quarter would be June through August, and growth rates would be calculated relative to the same period in the previous year. On the other hand, the base to which the growth rates are applied to generate forecasts consists of the unavailable months. So if data are available only through August, the growth rates are multiplied by tons shipped in the September to December period of the previous year.

Let k = 1, 2
where 1 indicates growth rates based on the most recent available quarter
2 indicates growth rates based on the most recent available year-to-date.

Calculation of Adjusted Growth Rates

Let $G_{y,i,j,.}^{m,r} = T_{y,i,j}^{m,r} \big/ T_{y,i,j}^{m-1,r}$, for m = 2,..,M, i = b,d,c and j = f,n,p,c.

Let $T_{y,i,j}^{1,a} = T_{y,i,j}^{1,u}$ for i = b,d,c and j = f,n,p,c.

Then $T_{y,i,j}^{m,a} = Max\left(T_{y,i,j}^{m,u}, G_{y,i,j}^{m,r} \cdot T_{y,i,j}^{m-1,a}\right)$, for m = 2, …, M, i = b,d,c and j = f,n,p,c.

Forecasts:
First define atomistic levels of the forecast variables in a general sense and then define specific forecasts for the general level of all carriers and all routes.

Let $G_{y,i,j,1}^{.,a} = T_{y,i,j}^{Q,a} \big/ T_{y-1,i,j}^{Q,u}$ and $G_{y,i,j,2}^{.,a} = \sum_{q=1}^{Q} T_{y,i,j}^{q,a} \Big/ \sum_{q=1}^{Q} T_{y-1,i,j}^{q,u}$

and

$F_{y,i,j,k}^{q,a} = G_{y,i,j,k}^{.,a} \cdot T_{y-1,i,j,.}^{q,u}$, where i=b,d,c and j=f,n,p,c, k=1,2 and q depends on j. If j = f then q = 3,4 and q = 4 otherwise.

General Level Forecasts for All Carriers and Regions
There are four forecasts for the third quarter of the current year:

Separate Regions (Domestic and International)
$F_{y,s,.,k}^{3,a} = T_{y,c,c,.}^{3,a} + F_{y,d,f,k.}^{3,a} + F_{y,b,f,k.}^{3,a}$ for k=1,2

Combined Regions
$F_{y,s,.,k}^{3,a} = T_{y,c,c,.}^{3,a} + F_{y,c,f,k.}^{3,a}$ for k=1,2

There are eight forecasts for the fourth quarter of the current year:

Separate Regions (Domestic and International) – Separate Carrier Groups
$F_{y,s,s,k}^{4,a} = F_{y,d,p,k}^{4,a} + F_{y,b,p,k}^{4,a} + F_{y,d,n,k.}^{4,a} + F_{y,b,n,k.}^{4,a} + F_{y,d,f,k.}^{4,a} + F_{y,b,f,k.}^{4,a}$ for k=1,2

Combined Regions – Separate Carrier Groups
$F_{y,c,s,k}^{4,a} = F_{y,c,p,k}^{4,a} + F_{y,c,n,k.}^{4,a} + F_{y,c,f,k.}^{4,a}$ for k=1,2

Separate Regions – Combined Carrier Groups
$F^{4,a}_{y,s,c,k} = F^{4,a}_{y,d,c,k} + F^{4,a}_{y,b,c,k.} + F^{4,a}_{y,d,f,k.} + F^{4,a}_{y,b,f,k.}$ for k=1,2

Combined Regions – Combined Carrier Groups
$F^{4,a}_{y,c,c,k} = F^{4,a}_{y,c,c,k} + F^{4,a}_{y,c,f,k.}$ for k=1,2

For purposes of illustration, the numerical results of the forecasts over the period 2002 to 2006 are given in Tables A-2 (third quarter forecasts for foreign carriers) and A-3 (fourth quarter forecasts for all carriers) in the Appendix A. Results from Table A-2 are presented for completeness but do not enter into the selection process. Three summary measures are given for each forecast in both levels and percentage terms: average error, standard deviation, and absolute error. The selection decision will be based on measures of percentage error because of the large change in levels with the addition of new-reporting carriers. Due to the small sample size, the forecast is selected based on a subjective evaluation of these measures rather than using formal statistical tests. As an aid to reading Table A-2, the best measure for each group of four forecasts varying by time-period and regional grouping is highlighted. The measures in Table A-2 clearly indicate using a forecast based on growth rates calculated using the latest available quarter rather than year-to-date, because all of the best measures under percentage error fall in this category. Selecting between forecasts based on separate or combined regional groups and separate or combined carrier groups is less clear. However, the evidence slightly favors basing the forecasts on growth rates calculated using separate regional groups and combined carrier groups. Therefore, the selected forecast is:

$$F^{4,a}_{y,s,c,1} = F^{4,a}_{y,d,c,1} + F^{4,a}_{y,b,c,1.} + F^{4,a}_{y,d,f,1.} + F^{4,a}_{y,b,f,1.}.$$

The growth rates for each group in this forecast will then be applied to the most recently available fourth quarter (and third for foreign carriers) segment and market data from OAI at the individual carrier and route level. Thus, for the aviation components of the 2006 provisional estimates, the missing fourth quarter for 2006 is obtained by applying: $G^{\cdot,a}_{2006,d,c,1}$ to all fourth quarter 2005 domestic routes flown by domestic carriers; $G^{\cdot,a}_{2006,b,c,1}$ to all fourth quarter 2005 international routes flown by domestic carriers; $G^{\cdot,a}_{2006,d,f,1}$ to all fourth quarter 2005 domestic routes flown by foreign carriers; and $G^{\cdot,a}_{2006,b,f,1}$ to all fourth quarter 2005 international routes by foreign carriers. Applying these values at the disaggregated level allows the above methods for calculating ports-of-entry, values, and prices to be applied without modification.

3.4.2 Forecasting the Commodity Distribution and Price

The commodity distribution and price are forecast based on historical data from Census. Exports and imports are forecast separately for international shipments, and the export distribution and price are then used to forecast domestic shipments. As above, due to the disruptions to the aviation industry, only data for 2002 and later are used as a basis for the forecasts.

3.4.2.1 International Shipments

Census data on imports and exports are the only timely source for the value and commodity distribution of air shipments. Historical data from Census are used to forecast the price (value divided by weight) and commodity shares for the fourth quarter of the most recent year. The forecasts are then combined with the aggregate weight forecast from the OAI data, and the techniques outlined above for estimating the value and commodity distribution are then applied.

Forecasts use available information to generate estimates of unavailable information. Evaluation of forecast techniques is based on applying the technique to generate "historical forecasts," which can then be used to calculate errors based on the known values and summarized based on the forecast criterion. The basic philosophy is that the historical forecasts should be generated using only information that would have been available to a forecaster under the same production conditions as future forecasts will be generated. For FAF^2 purposes, a forecast for the fourth quarter of 2007, for example, uses only information that would have been available in December 2007 (Census data up to the third quarter of 2007). The evaluation criterion used here is the average squared error for the forecasts of prices and the standard deviation of the forecast errors for the commodity distributions. The reason for the different criterion is that the price forecasts are not necessarily mean zero and the standard deviation would fail to incorporate undesired bias effects.

Three forecasts of prices are considered: (i) year-to-date (YTD) third quarter prices are used for the fourth quarter, (ii) annual increases in individual commodity prices based on YTD in the current year and the same period in the previous year, and (iii) the annual increase in the price of aggregated commodities based on YTD in the current and previous year is applied to all individual commodities. The second and third forecasts of the price increase are then applied to the individual fourth quarter prices from the previous year. The first forecast will not include seasonal effects on prices, while the second and third forecasts will include seasonal effects. The final forecasts of commodity prices may use different techniques for different commodities, since seasonal effects may be important for some commodities, but not others, and because forecasts for the individual commodities are independent.

3.4.2.2 Domestic Shipments

The export commodity distribution and prices used for the latest provisional year are based on the implied forecast of these values for exports given above.

3.4.3 Conclusion

The aviation portion of the provisional commodity O-D data requires estimating key components that are missing from the available data and forecasting a portion of the data to provide timely information for analysis. The techniques outlined above provide a reasonable approach to filling the data gaps that will provide useful information to users of the database.

4.0 FREIGHT MOVEMENT BY RAIL

4.1 Introduction

This chapter describes the data sources and methodology for generating the provisional estimates of volumes and value for freight moved by rail. This includes both domestic and international origins and destinations. Specific commentary on the application of the methodology described below in generating 2008 provisional estimates is provided at the end of this chapter.

The most efficient method for updating the railroad freight flows depicted within the FAF^2 would employ the confidential version of the Surface Transportation Board's annual Carload Waybill Sample (CWS). However, the scope of the current project explicitly states that data sources used in the updates must be publicly available. The public use CWS is of limited use for flow analysis given that origin and/or destination information is removed in many records to preserve confidentiality. Accordingly, an alternative method was developed that makes heavy use of the public use CWS, but which also relies on the existing FAF^2 flows, annual rail traffic data from the Association of American Railroads (AAR), the Surface Transborder Freight Dataset (TFD), and data available from the U.S. Census Bureau.

The railroad freight flows depicted within FAF^2 have two dimensions – commodity and geographic. It is essential to reflect variations across both dimensions. The proposed process begins by using the AAR annual car loading data to obtain traffic growth rates for the 19 AAR commodities. Next, the 19 AAR commodity groups were mapped into the 43 FAF^2 commodity groups, so that each FAF^2 commodity flow will have an associated overall commodity-based growth rate. These rates may be used to develop control flow totals for each commodity.

Next, the growth rates for each FAF^2 flow element were modified to reflect any volume variations that are attributable to O-D location. Adjustments may be based on any combination of three factors – demographic variations, variations in industry-specific economic activity, and variations in general economic conditions.[8]

The final methodological step combines the O-D- specific growth rates with existing FAF^2 flow data in an application of a traffic growth factor model such as that developed by Fratar.[9] These models have a number of attributes and limitations. Most notably, however, they do not require the incorporation of impedances, and so are relatively simple to implement.[10]

The methodology for estimating international flows employed the annual rail records of the STFD, which are complete for rail movements. These shipments were disaggregated to the FAF^2 region level. For the purpose of the FAF^2 update, railroad carload and rail/truck intermodal

[8] In the past, a similar method was used to allocate FAF flows to more disaggregated geographic units.

[9] T.J. Fratar, "Vehicular Trip Distribution by Successive Approximation," Traffic Quarterly, Vol. 8, pp. 53-64 (1954).

[10] While it would be possible to generate impedances based on rail distances between origins and destinations, it is our view that creating these values and incorporating them into the estimation process would add little to the robustness of the results.

shipments were treated as separate modes, even though many calculations were integrated across the two.

There are a number of challenges inherent in the processes described above. These include, but are likely not limited to:

- There are difficulties in mapping the 19 AAR commodity groups into the 43 FAF2 commodity groups.
- The AAR data are presented for each Class I carrier. Thus, it is possible for a particular shipment to be duplicated if it is interchanged between two carriers.
- There is no immediately available source for establishing changes in commodity values for domestic shipments.
- AAR data are expressed in carloads; FAF2 flows are expressed in tons. Thus, changes in car loading weights could distort the AAR-based growth rates.
- Time lag between data availability and development of provisional estimates

The proposed response to each of these challenges is provided in the descriptions of the methodologies presented in the following sections.

4.2 Principal Data Sources for Rail Freight

Weekly Railroad Traffic – This weekly publication contains information on carload and intermodal traffic for the U.S. Class I railroads, the two large Canadian railroads, a major Mexican railroad, and selected U.S. non-Class-I railroads. It includes carload information for 19 commodity groups and intermodal traffic, which is reported for trailers and containers.

Carload Waybill Sample (CWS) – A stratified sample of carload waybills for terminated shipments by railroad carriers. Waybill data are used to create a movement-specific Confidential Waybill File and a less detailed Public Use Waybill File. This is published by the Surface Transportation Board (STB) and can be accessed at: http://www.stb.dot.gov/stb/industry/econ_waybill.html

Surface Transborder Freight Database (TFD) – Published by the BTS, this database contains data on North American merchandise trade by commodity, surface mode (rail, truck, pipeline, mail, and other), and by port of entry and geographic detail for the U.S. trade to and from Canada and Mexico. This source provides the dollar value of both imports and exports, and tonnage of imports. The data are published with a three-month time lag. It can be accessed a: http://www.bts.gov/programs/international/transborder/

County Business Pattern (CBP) Database – Published by the U.S. Census Bureau on an annual basis, this resource provides national, state, and county level data on payroll, employment, and number of establishments by detailed NAICS industries. The series provides subnational economic data by industry and excludes data on self-employed individuals, employees of private households, railroad employees, agricultural production employees, and most government employees. The report is released with a two-year time lag. It can be accessed at: **http://www.census.gov/econ/cbp/index.html**

Producer Price Index (PPI) – Measures the average change over time in the prices received by domestic producers of goods and services. This measures price changes from the point of view of the producer. The data are reported by detailed industry and detailed type of commodities. The BLS publishes these data on a monthly basis with a time lag of one month. It can be accessed at: http://www.bls.gov/ppi/

4.3 Domestic Rail Flows

The primary data source for the FAF^2 domestic rail flow update was the STB Public Use version of the CWS. The waybill sample is a stratified sample of the population of rail movements that originate or terminate within the United States. The actual degree of sampling depends on the variability of shipment characteristics across commodities. In most cases, the sample represents between one and two percent of the overall shipment population.

The initial step involved gleaning tonnage variations between 2002 and 2008. CWS commodity definitions (based on Standard Transportation Commodity Codes) were bridged to the Standard Classification of Transported Goods (SCTG) definitions employed within the FAF^2. Once this was accomplished, national traffic growth factors were calculated based on variations in tonnage between 2002 and 2008.

Unfortunately, the O-D information in the Public Use CWS is left incomplete in order to protect the confidentiality of both shippers and rail carriers. Consequently, it was impossible to use the CWS to identify geographic variations in traffic volumes that might be relevant to inter-temporal variations in the FAF^2 flows. To remedy this problem, industry-specific employment values, derived from CBP data, were used to build indexes reflecting FAF region employment, population, and income trends for both origin and destination regions. The hypothesized relationship between employment, demographic values, and rail flows are summarized in Table 4.1.

Table 4-1. Relationship between Employment, Demographic, and Rail Flow

STCG	Origin Index Component	Destination Index Component
2	NAICS 111 Employment	NAICS 311 Employment
3	NAICS 111 Employment	NAICS 311 Employment
4	NAICS 311 Employment	NAICS 311 Employment
5	NAICS 311 Employment	NAICS 311 Employment
6	NAICS 311 Employment	NAICS 311 Employment
7	NAICS 311 Employment	NAICS 311 Employment
8	NAICS 312 Employment	NAICS 445 Employment
9	NAICS 312 Employment	NAICS 447 Employment
10	NAICS 327 Employment	NAICS 234 Employment
11	NAICS 327 Employment	NAICS 234 Employment
12	NAICS 327 Employment	NAICS 234 Employment
13	NAICS 327 Employment	NAICS 327 Employment
14	NAICS 212 Employment	NAICS 331 Employment
15	NAICS 212 Employment	NAICS 221 Employment
16	NAICS 211 Employment	NAICS 324 Employment
17	NAICS 324 Employment	NAICS 447 Employment
18	NAICS 324 Employment	NAICS 221 Employment
19	NAICS 324 Employment	FAF Region Total Employment
20	NAICS 325 Employment	FAF Region Total Employment
21	NAICS 325 Employment	FAF Region Total Employment
22	NAICS 325 Employment	NAICS 111 Employment
23	NAICS 325 Employment	FAF Region Total Employment
24	NAICS 326 Employment	FAF Region Total Employment
25	NAICS 113 Employment	NAICS 321 Employment
26	NAICS 321 Employment	NAICS 321 Employment
27	NAICS 322 Employment	FAF Region Population
28	NAICS 322 Employment	FAF Region Population
29	NAICS 323 Employment	FAF Region Population
30	NAICS 313 Employment	FAF Region Population
31	NAICS 327 Employment	FAF Region Population
32	NAICS 331 Employment	FAF Region Population
33	NAICS 332 Employment	NAICS 332 Employment
34	NAICS 333 Employment	FAF Region Population
35	NAICS 335 Employment	FAF Region Population
36	NAICS 336 Employment	FAF Region Population
37	NAICS 336 Employment	FAF Region Population
38	NAICS 334 Employment	FAF Region Population
39	NAICS 337 Employment	FAF Region Population
40	FAF Region Population	FAF Region Population
41	FAF Region Population	FAF Region Population
43	FAF Region Population	FAF Region Population

The formal derivation of the indexes is explained in the following equation. Once established, the indexes were applied to the national growth factors to simulate geographic variations in flows. Finally, the adjusted growth factors were applied to the 2002 FAF^2 values to yield 2008 estimates.[11]

$$1 + \frac{(OVAL_{05} - 0VAL_{02}) + (TVAL_{05} - TVAL_{02})}{OVAL_{02} + TVAL_{02}}$$

where OVAL and TVAL represent the appropriate O-D employment or population variables.

Revising the FAF^2 data also required estimating changes in the value of commodity flows. Changes in values are a function of both changed flow volumes and per-unit commodity value variations. Flow changes were based on tonnage volume flows as described above. Per-unit values were based on commodity-specific variations where possible, as captured by changes in the components of the PPI. The bridge between FAF commodity definitions and PPI values is provided in Table 4.2.

Table 4-2. FAF Commodity Definitions and PPI

Producer Price Index Component	2002 – 2005 Percentage Change
Industrial Commodities Less Fuels	0.106993
Farm Products	0.196970
Industrial Chemicals	0.480754
Lumber	0.164127
Pulp Paper And Allied Products	0.089833
Crude Petroleum	1.210604
Chemicals And Allied Products	0.263989
Iron And Steel	0.263989
Steel Mill Products	0.523855
Motor Vehicle Parts	0.001771
Plastics Material And Resins Manuf.	0.534587
Aluminum Plate, Sheet, And Foil Manuf.	0.112846
Automobile And Light Duty Vehicle Manuf.	0.001483

[11] Unlike estimates for other modes, the commodity-specific estimates for domestic rail movements did not necessarily sum to the tonnage total change observed between 2002 and 2005. Therefore, ultimately, commodity-specific FAF flows were adjusted downward by roughly 5 percent to conform with observed tonnage totals.

4.3.1 Rail Flows to Deep-Draft Ports

Somewhat inexplicably, the FAF^2 rail flows over U.S. deep-draft ports entirely neglect import flows and capture export flows only. Consequently, export flows alone were updated from 2002 values to reflect changed economic conditions in 2008. The basis for this update was export data obtained from the U.S. Department of Commerce. As in the case of domestic rail flows, the initial step involved reconciling export data commodity definitions with the commodity definitions used within the FAF^2.

Next, because the Department of Commerce export volumes are expressed in dollar values only (as opposed to tons), it was necessary to account for intertemporal changes in commodity values between 2002 and 2008. As in the case of domestic flows, the PPI was used for this purpose. Once price variations were accounted for, export data were used to scale 2002 FAF flows to reflect 2008 FAF^2 deep-draft port flows. Given that no geographic variation is reflected in the Department of Commerce data, the revised FAF^2 values assume that the distribution of rail export flows across US ports is unchanged between 2002 and 2008.[12]

The final step in the adjustment to the FAF data involved again using price index data in order to inflate the value field in the 2002 data. Where possible, industry or product specific values were used. In the absence of such data, the value corresponding to “Industrial Commodities Less Fuels” was used.

4.4 Methodology for International Rail Freight

4.4.1 Methodology for International Freight

FAF contains international rail freight shipments of two types: (1) all-rail shipments to/from Canada and Mexico and (2) shipments to/from countries outside of North America that use rail for the domestic portion of the movement. Different methodologies are used for addressing the two categories.

4.4.2 Transborder Rail Freight to and from Canada and Mexico by U.S. State and Port of Entry or Exit

The approach for estimating rail freight flows between the U.S. and Canada, and between the U.S. and Mexico, is as follows:

1. Determine state-level transborder rail freight to and from Canada and Mexico for the current year using information from BTS’ TFD;
2. Disaggregate state level transborder rail freight flow to FAF region-level based on FAF patterns from the base year; and

[12] In the case of lower-valued bulk commodities, this assumption is probably non-problematic. However, in the case of higher valued exports, the validity of this assumption is more suspect.

3. Allocate FAF region level flows to and from Canada/Mexico to ports of entry/exit (actually border crossing points) based upon FAF^2 patterns from the base year or data on port use from the current year TFD.

4.4.3 Determine State-Level Transborder Rail Freight to and from Canada and Mexico

BTS's TFD provides freight data on tons and value of exports and imports from Canada and Mexico to the United States by rail. The data are reported by O-D state, country, and type of commodity. A separate group of files provide data on total weight and value through ports of entry/exit by O-D state and county, without regard to specific commodities. The rail records are extracted from all of the annual TFD datasets for the target year.

The TFD uses the HS commodity classification, rather than the SCTG employed by FAF. Using a cross-walkmatching HS and SCTG codes, the TFD records are processed to add the appropriate SCTG. The port of entry/port of departure (POE/POD) in the TFD is described using Customs Port Codes. A translation table maps these codes to FAF regions and international gateways.

The following sections describe processing steps to handle import and export data.

4.4.3.1 Imports

The following TFD files contain import data at the commodity level:

- 09yyyy–imports from Mexico with state of destination and 2-digit commodity detail, where yyyy is the year of release, e.g. 2008; and
- 10yyyy–imports from Canada with state of destination and 2-digit commodity detail.

The import records are processed to tally the total weight and value by HS commodity. Separate totals are kept for imports from Mexico and Canada. This information is used in processing the export records as described in the next section.

For compatibility with FAF, the weights and values in each record are converted from kilograms to thousand short tons expressed as kilotons and from dollars to millions of dollars, respectively. The origin field is set to the appropriate FAF code for Mexico or Canada.

4.4.3.2 Exports

Export data at the commodity level is contained in the following TFD files:

- 3ayyyy–exports to Mexico with state of origin and 2-digit commodity detail; and
- 4ayyyy–exports to Canada with state of origin and 2-digit commodity detail.

Unlike the import records, the export records lack weight. Accordingly, the weight for each flow must be imputed. To estimate weight, value is multiplied by a weight/value ratio for the commodity with different ratios used for Canada and Mexico. These ratios are derived using the commodity tallies collected from the import records. To minimize variance, the tally is based at the HS level, a lower level of aggregation than the SCTG.

As with the import records, the weights and values in each record are converted from kilograms to kilotons and from dollars to millions of dollars, respectively. The destination field is set to the appropriate FAF code for Mexico or Canada.

4.4.3.3 Initial Aggregation

Following initial processing, the import records for Mexico and Canada are combined into a single file representing all import rail traffic for the target year. The separate export record files are also combined. As a result of the conversion from HS to SCTG, each combined file may contain several records for a given origin, destination, and SCTG. The files are processed to combine these records into a single record containing the totals for the origin, destination, and SCTG.

4.4.4 Disaggregate Estimates of Transborder Rail Freight by State to FAF-Level Estimates

The state-level transborder rail freight tonnage and value are disaggregated to FAF-level estimates using the existing patterns from the original FAF base year.

4.4.4.1 Imports

The target year estimate of rail freight import tonnage $W_{i,c,r,t}$ of commodity *i* from country *c* to FAF region *r* in state *s* for year *t*, is calculated as:

$$W_{i,c,r,t} = W_{i,c,s,t} * P_{i,c,s,r,t-1}$$

where:

- W = Rail freight tonnage of imports,
- P = share variable,
- i = commodity,
- c = country of origin (Canada or Mexico),
- s = destination state,
- r = destination FAF region (in state *s*),
- t = target year, and
- $t-1$ = base year.

The share variable *P* is based upon the weight of *i* destined from *c* to *r* in the base year as a portion of the total weight destined from *c* to *s*. Domestic FAF regions lie entirely within state boundaries; a crosswalk table allows state totals to be derived from FAF totals. The share variable is formally calculated as:

$$P_{i,c,s,r,t-1} = W_{i,c,r,t-1} \Big/ \sum_{j \in s} W_{i,c,j,t-1}.$$

The value of the import tonnage for the commodity is calculated in a similar fashion using the same share variable, *P*:

$$V_{i,c,r,t} = V_{i,c,s,t} * P_{i,c,s,r,t-1}$$

where *V* is the value and all subscripts have the same meaning as previously.

Where commodity flows did not exist in the base year, the flow is allocated in equal portions to each FAF region in the state. Future refinements may use county level indicators from the Census CBP Database to disaggregate the flows.

4.4.4.2 Exports

Disaggregation of exports from the state to the FAF region level follows a similar approach to that of imports. The target year estimate of rail freight export tonnage of commodity *i* from FAF region *r* in state *s* to country *c* for year *t*, is calculated as:

$$W_{i,r,c,t} = W_{i,s,c,t} * P_{i,s,r,c,t-1}$$

where all variables and subscripts remain as previously defined. The share portion for *r* is calculated as follows:

$$P_{i,s,r,c,t-1} = W_{i,r,c,t-1} \Big/ \sum_{j \in s} W_{i,j,c,t-1}.$$

This share is also used to apportion value for the flow.

4.4.5 Allocate FAF-Level Flows to Ports

The final processing step is to allocate FAF region level imports and exports across ports. FAF defines a number of international gateways, and these correspond well to the limited number of international rail border crossings in North America. The term "port" is used here to generally include rail border crossings. However, a significant portion of the FAF rail transborder records do not presently use the designated gateways, instead having ports identified as regular FAF regions. No effort was made to improve allocation of shipments to the defined rail border crossings.

4.4.5.1 Imports

For flows occurring in both the base year and the target year, import tonnage of a given commodity from a foreign source to a FAF region is allocated among ports of entry using the following formula:

$$W_{i,c,p,r,t} = W_{i,c,r,t} * P_{i,c,p,r,t-1}$$

where:

W = Rail freight tonnage of imports,
P = share variable,
i = commodity,
c = country of origin (Canada or Mexico),
p = port,
r = destination FAF region (in state *s*),
t = target year, and
$t-1$ = base year.

The value of the import tonnage for the commodity is calculated in a similar fashion using the same share variable, P:

$$V_{i,c,p,r,t} = V_{i,c,r,t} * P_{i,c,p,r,t-1}$$

where V is the value and all subscripts have the same meaning as previously.
The port share is again based upon the weight of i destined from c to r via p in the base year as a portion of the total weight of i destined from c to r via all ports involved in the trade. Mathematically, this share is expressed as:

$$P_{i,c,p,r,t-1} = W_{i,c,p,r,t-1} \Big/ \sum_{x \in X} W_{i,c,x,r,t-1}.$$

The set X contains all ports handling commodity i between country s and region r during the base year.

In the case where a commodity was not handled by the FAF region in the base year, the allocation is based on the each port's share of total trade, by value, from the origin country to the FAF region's state during the target year. This information is found in the port level TFD, which provides total weight and value for all freight moved between a country and state via a port.

4.4.5.2 Exports

Allocation of export flows between a FAF region and foreign country via ports follows a similar approach to that of imports. The target year estimate of rail freight export tonnage of commodity i from FAF region r in state s to country c for year t, is calculated as:

$$W_{i,r,p,c,t} = W_{i,r,c,t} * P_{i,r,p,c,t-1}$$

with all variables and subscripts as previously defined. The share portion for r becomes:

$$P_{i,r,p,c,t-1} = W_{i,r,p,c,t-1} \Big/ \sum_{x \in X} W_{i,r,x,c,t-1}$$

with X being the set of ports handling commodity i between region r and country s during the base year. Again, the share is also used to apportion value for the flow. Flows not existing during the base year are apportioned to ports based upon the target year TFD port level export data for the state of origin.

4.4.6 Combine Import and Export Data

The previous steps result in two output files for transborder rail freight:

- Import flows (weight and value) by commodity, country, port, and destination domestic FAF region; and
- Export flows (weight and value) by commodity, domestic origin FAF region, port, and country.

The final processing step is to combine the two files and eliminate any records for which the weight is less than 0.01 kilotons and the value less than 0.01 million dollars. The current FAF limits these values to two decimal points, so values less than these will appear as zeros. In practice, a single rail carload carries 0.07 to 0.10 kilotons, so this step will have minimal impact on overall flows.

4.5 2008 Provisional Rail Estimates – Commentary on Methodology

The methodology described in the previous sections was used in updating the 2007 rail files so that they reflect 2008's economic changes. AAR data on commodity-specific car loadings were used to estimate traffic changes. Elements from the PPI were used to calculate associated changes in commodity values. The resulting adjustment factors were applied uniformly to the 2007 data. A summary of adjustment factors is provided in Table 4.3. The repetition in some parameter values reflects the difficulty in bridging across three very different sets of commodity definitions.

While this methodology yields a reasonable approximation of the changes to actual rail freight flows, it has two notable limitations. First, under such a methodology, existing traffic can never fully disappear over a network lane and, similarly, new flows cannot emerge. Unfortunately, this limitation cannot be mitigated without the use of primary data, which are unavailable for update purposes. The second limitation is linked to the geographic implication of the FAF regions over which rail freight flows are defined. Given the current methodology, the pair-wise flow volumes observed in the updated data are proportional to the pair-wise volumes evident in the preceding year's data. Thus, what appears to be dynamic is actually static.

Prior to developing the 2008 estimates, the study team investigated potential methodological changes that might, to some degree, improve the geographic distribution of rail freight flow adjustments, so that the resulting estimates more accurately reflect regional shifts in economic activity. This consideration, however, was undertaken with understanding that both temporal and fiscal constraints could not be substantially relaxed. The findings of this investigation are summarized below.

Table 4.4 summarizes estimated domestic rail tonnages by commodity for 2008. Clearly, both the value of and ability to improve the geographic precision of annual updates is confined to a few commodity groups – specifically, coal and related products, agricultural inputs and outputs, and chemicals.

While intermodal rail shipments briefly eclipsed coal as a source of rail revenue, coal continues to dominate rail tonnage. Therefore, any attempt to adapt current methodologies to better reflect economic trends should center on the production and consumption locations of this commodity.

Coal is mined where it is found. Nonetheless, coal characteristics (Btu, sulfur, ash content, etc.), resulting mine-mouth prices, and environmental constraints can easily affect coal sourcing decisions. Similarly, these same elements, combined with demand conditions, can affect electric utility dispatching practices. Finally, international market conditions routinely alter coal export volumes. Any consideration of alternative FAF update methods must be mindful of these three aspects.

Table 4-3. 2008 Adjustment Factors

STCG	*Commodity Definition*	*AAR Adjustment Factor*	*PPI Adjustment Factor*
2	Cereal grains	0.000	0.126
3	Other agricultural products	-0.129	0.126
4	Animal feed and products of animal origin, not elsewhere classified (n.e.c.)	-0.004	0.126
5	Meat, fish, seafood, and their preparations	-0.004	0.126
6	Milled grain products and preparations, and bakery products	-0.027	0.126
7	Other prepared foodstuffs and fats and oils	-0.004	0.126
8	Alcoholic beverages	-0.129	0.126
9	Tobacco products	-0.129	0.126
10	Monumental or building stone	-0.083	0.048
11	Natural sands	-0.083	0.048
12	Gravel and crushed stone	-0.083	0.048
13	Nonmetallic minerals n.e.c.	-0.083	0.048
14	Metallic ores and concentrates	-0.057	0.048
15	Coal	-0.009	0.048
16	Crude Petroleum	0.044	0.306
17	Gasoline and aviation turbine fuel	0.044	0.306
18	Fuel oils	0.044	0.306
19	Coal and petroleum products, n.e.c.	0.044	0.048
20	Basic chemicals	0.033	0.044
21	Pharmaceutical products	-0.020	0.044
22	Fertilizers	0.033	0.044
23	Chemical products and preparations, n.e.c.	-0.020	0.048
24	Plastics and rubber	-0.020	0.048
25	Logs and other wood in the rough	-0.119	-0.011
26	Wood products	-0.119	-0.011
27	Pulp, newsprint, paper, and paperboard	-0.078	0.019
28	Paper or paperboard articles	-0.078	0.019
29	Printed products	-0.020	0.019
30	Textiles, leather, and articles of textiles or leather	-0.020	0.048
31	Nonmetallic mineral products	-0.015	0.048
32	Base metal in primary or semi-finished forms and in finished basic shapes	-0.078	0.038
33	Articles of base metal	-0.078	0.038
34	Machinery	-0.020	-0.008
35	Electronic and other electrical equipment components and office equipment	-0.020	0.048
36	Motorized and other vehicles (including parts)	-0.053	0.000
37	Transportation equipment, n.e.c.	-0.053	0.000
38	Precision instruments and apparatus	-0.020	0.048
39	Furniture, mattresses and mattress supports, lamps, lighting fittings	-0.020	0.048
40	Miscellaneous manufactured products	-0.020	0.048
41	Waste and scrap	0.000	0.048
43	Mixed freight	-0.020	0.048

Sources: Association of American Railroads / US Department of Commerce

Table 4-4. 2008 Tonnage Estimates by Commodity – Domestic File

STCG	*Estimated 2008 Tons (X 1K)*	*Percentage of Total*
15	893,656	47.1%
2	185,126	9.8%
19	115,009	6.1%
22	90,285	4.8%
41	77,193	4.1%
20	72,468	3.8%
12	70,174	3.7%
7	47,194	2.5%
32	37,725	2.0%
26	32,806	1.7%
14	32,617	1.7%
31	32,375	1.7%
24	32,051	1.7%
13	25,243	1.3%
4	24,873	1.3%
27	24,390	1.3%
3	22,629	1.2%
11	16,512	0.9%
6	11,807	0.6%
36	11,095	0.6%
17	8,367	0.4%
18	6,579	0.3%
8	5,657	0.3%
33	5,420	0.3%
37	4,896	0.3%
34	2,143	0.1%
43	1,489	0.1%
28	1,269	0.1%
25	894	0.0%
35	786	0.0%
30	705	0.0%
23	610	0.0%
5	555	0.0%
29	544	0.0%
40	508	0.0%
39	362	0.0%

Source: 2008 FAF DOMESTIC and SEAPORT updates

Table 4.5 provides state-specific coal production totals for 2006 and 2007. Again, there are three important points. First, while volumes changed to a small degree, the relative shares of the leading coal producing states did not change in any sort of meaningful way. Second, the production activity, as identified for FAF regions within producing states almost certainly did not change given that the FAF regions are largely defined on a metro / non-metro basis. Finally, these data do not include values for 2008, the period for current FAF data estimates.

A similar accounting for annual changes in facility dispatching and corresponding coal consumption leads to a similar conclusion – if the base FAF estimates adequately capture the regional distribution of domestic coal flows, then increased precision in annual updates is both impossible and unnecessary.

Table 4-5. Coal Production by State ('000 tons)

State	2007 Production	2006 Production	Percent Change
Wyoming	453,568	446,742	-1.5%
West Virginia	153,480	152,374	-0.7%
Kentucky	115,280	120,848	4.6%
Pennsylvania	65,048	66,029	1.5%
Montana	43,390	41,823	-3.7%
Texas	41,948	45,548	7.9%
Colorado	36,384	36,322	-0.2%
Indiana	35,003	35,119	0.3%
Illinois	32,445	32,729	0.9%
North Dakota	29,606	30,411	2.6%
Virginia	25,346	29,740	14.8%
New Mexico	24,451	25,913	5.6%
Utah	24,307	26,018	6.6%
Ohio	22,575	22,722	0.6%
Alabama	19,327	18,830	-2.6%
Arizona	7,983	8,216	2.8%
Mississippi	3,545	3,797	6.6%
Louisiana	3,127	4,114	24.0%
Tennessee	2,654	2,804	5.3%
Maryland	2,301	5,054	54.5%
Oklahoma	1,648	1,998	17.5%
Alaska	1,324	1,425	7.1%
Kansas	420	426	1.4%
Missouri	236	394	40.1%
Arkansas	83	23	-260.9%
Washington	0	2,580	100.0%
TOTAL	1,145,479	1,161,999	1.4%

Source: US Department of Energy, Energy Information Administration

The volume of coal moved domestically for import and export is considerably more variable. Table 4.6 provides total import and export volumes for the period between 2002 and 2008. These volumes are even more erratic when the data are disaggregated to distinguish between trading partner and steam versus metallurgical coal.

Table 4-6. U.S. Coal Imports and Exports

(Values in thousands of tons)

Year	Exports	Imports
2002	39,601	16,875
2003	43,014	25,044
2004	47,998	27,280
2005	49,942	30,460
2006	49,647	36,246
2007	59,163	36,347
2008	59,191	25,107

Source: US Department of Energy, Energy Information Administration

Unlike domestic coal volumes, 2008 data are available, so that the available information is consistent with the FAF update period. Hence improved update methods might be possible. However, the rewards attributable to additional work are likely to be imperceptible. First, overall import and export volumes represent less than 10 percent of total U.S. production. And second production and port locations are highly invariant, so that even if volumes do change significantly, changes in *relative* regional activity are unlikely.

Because of coal's tonnage dominance, it is important to consider the possible avenues to and rewards of more advanced methods for annual FAF updates in some detail. However, a less rigorous consideration of other prominent railed commodities leads to similar conclusions. Agricultural inputs and outputs are the second most dominant commodity groupings – grains and fertilizers. Fertilizers are produced where inputs are available and shipped to where land is farmed. Farmland is where it is. Thus, while flow volumes may change noticeably based on both domestic and international market conditions, relative volumes in flow lanes are unlikely to change from one year to the next in a way that can be captured. The same is true for chemicals, aggregates, lumber, or other heavily railed commodities.

This page intentionally left blank.

5.0 FREIGHT MOVEMENT BY WATER

This chapter presents the data sources and methodology for estimating freight movement by water, both domestic and international. The data sources and their limitations to estimate freight tonnage and value for the estimation year are also described.

5.1 Domestic Waterborne Freight

5.1.1 Data Sources

One of the main data sources for use in developing the methodology and providing provisional estimates is the Internal U.S. Waterway Monthly Indicators data (http://www.iwr.usace.army.mil/ndc/wcsc/wcmthind.htm) published by the U.S. Army Corps of Engineers (USACE), Navigation Data Center (NDC). The USACE provides data on U.S. waterborne commerce, including the transport of goods by inland barge and ship over the nation's navigable rivers, across the Great Lakes, and within the U.S. Intra-Coastal Waterway. Domestic O-D movements are created by USACE from its Vessel Operating Reports, as well as from its Lock Performance Monitoring System (LPMS) database. Data are in theory reported by all vessels and provide estimates of annual tons moved by 5-digit commodity code for all commodities transported on U.S. waterways on a dock-to-dock basis. The NDC indicator is seasonally adjusted in terms of millions of short tons. Estimates for future years based on these historical data would represent the annual totals of domestic freight movement by waterway.

The NDC also publishes data on waterborne tonnage by state and ports. These reports list tonnages for Domestic, Foreign, Imports, Exports, and Intra-State waterborne movements (http://www.iwr.usace.army.mil/ndc/wcsc/statenm00.htm).

The following are some comments and inconsistencies noted in the data

- The public domain commodity codes employed in the data sources used were converted to STCG codes. Of the 46 STCG commodity codes, less than 12 are moved by water.
- In some states, certain commodities were inconsistently moved from or to that state from year to year. As such, some data points are missing for some commodities, which makes it necessary to base estimates for the current year on educated assumptions.
- The data sources do not contain information on the value of domestic waterborne commodity movements. For the estimation year, value can be estimated based on weight/value from 2002 benchmark data and corrected for the effect of inflation using the Consumer Price Index (CPI) published by the BLS.

5.1.2 Methodology

Provisional estimates of the volumes domestic waterborne freight were developed by calculating the total tonnages originating and destined for each state and then broken down by commodity between origin state and destination states. The state estimates are further disaggregated into FAF zones. The equivalent dollar values of these volumes were then estimated based on information on value per unit weight developed from 2002 FAF benchmark data. The steps involved in the methodologies are described in the following subsections.

5.1.2.1 Weight Estimates

The growth factor approach was used in developing the estimates of tonnage of freight movements with the FAF 2002 data as benchmark. This involved the following steps:

1. Estimate the growth factor for each state of origin and state of destination pair between 2002 and the estimation year. These growth factors are derived from total tonnages of commodity movements between each O-D pair.
2. Apply the growth factors in step 1 to the 2002 state O-D tonnages to obtain state totals for the provisional year of interest.
3. Estimate each commodity group's share of the state total tonnages originating from and destined to that state based on the previous year's provisional estimates. Then distribute the estimated state total tonnages among the commodity groups originating from or destined to that state.
4. Determine each FAF zone's share of the state's total tonnage of each commodity originating from and destined to that state based on the previous year's provisional estimate. Different factors are derived for freight originating from and destined to each FAF zone.
5. Expand the state-to-state estimates by commodity that were developed in the previous steps among the FAF zones in each state based on distribution of the 2002 FAF^2 data among the FAF zones. This approach assumes that a commodity's percent share of total tonnage in a FAF zone in the estimation year will be the same as in the base year.

5.1.2.2 Value Estimates

Because no information is available on commodity dollar values for domestic movements, weight/value ratios derived from the 2002 FAF^2 benchmark data were used together with the estimated volumes of the freight to estimate the dollar values. These rates were used as the basis in estimating the value of commodities moved by water in provisional estimation years.

5.2 International Waterborne Freight

5.2.1 Data Sources

The USACE generates the monthly *Waterborne Databank* that contains statistics on U.S. waterborne imported and exported freight. This databank is published in the Preliminary Foreign Waterborne Cargo Summary report that is available on the USACE web site at http://www.iwr.usace.army.mil/ndc/usforeign/pcsfiles.htm. The summary report contains information on weight and value information by type of service on U.S. waterborne imports and exports, along with year-to-date figures. Inbound and outbound in-transit data are not included.

The NDC of the USACE also publishes data on commodity movements. This database shows waterborne tonnage for principal U.S. ports and all 50 states and U.S. territories (http://www.iwr.usace.army.mil/ndc/wcsc/statenm00.htm). These reports list tonnages for Domestic, Foreign, Imports, Exports, and Intra-State waterborne traffic. This database presents a breakdown of freight tonnage by state and by port of entry/exit.

5.2.2 Methodology

Provisional estimates were developed for the volume and value of commodity movements from/to the ports of exit/entry and the seven destinations/origins outside the United States (i.e., Canada, Mexico, Europe, Latin America, Middle East, Asia, and Rest of World). The estimation process involves the following steps.

1. The first step is to group the foreign origins and destinations (i.e., countries) into the seven foreign destinations/origins outside the United States prior to calculating the growth rates.
2. Calculate the growth rates in freight tonnage between 2002 and the provisional estimation year in waterborne imports and exports through each of the U.S. ports of entry/exit and FAF foreign destinations and origins.
3. Apply the growth rates calculated above to the 2002 FAF benchmark freight tonnage to estimate the volume of freight by FAF O-D foreign region for the provisional year.
4. The growth rates in the tonnage of freight are applied to the 2002 tonnage of freight to derive the freight tonnage for the provisional year. This is done at the FAF O-D-port of exit or entry-commodity level of detail.
5. Apply the tonnage growth rates to the 2002 value/weight ratios to derive the growth in freight value for the estimation years. The value/weight ratios are calculated at the FAF O-D-port of exit/entry-commodity level of detail. The result of this effort provides the growth in the freight value in 2002 dollars for the provisional years.
6. Derive the growths in the current dollar value of freight by adjusting the growth in freight value in 2002 dollars by the changes in prices. The commodity export and import price indexes are used for this purpose. The price indexes are obtained from the BLS. The export price indexes by type of commodity are used to estimate the current dollar value of export freight, and the import price indexes are used to estimate the value of import freight. The BLS reports the price indexes in HS commodity classification method.
7. Add the growth in the current and constant dollar (i.e., 2002) value of freight to the 2002 freight values to derive the estimation year’s value of freight in current and constant dollars, respectively.

This page intentionally left blank.

6.0 FREIGHT MOVEMENT BY PIPELINE

6.1 Introduction

Movements of crude oil and petroleum products by pipeline between Petroleum Administration for Defense (PAD) Districts include trunk pipeline companies (operators of interstate, intrastate, and intra-company pipelines). PAD Districts are geographic aggregations of the 50 states and the District of Columbia into five districts, established by the PAD in 1950. These districts were originally defined during World War II for purposes of administering oil allocation. Figure 6.1 shows the PAD districts.[13]

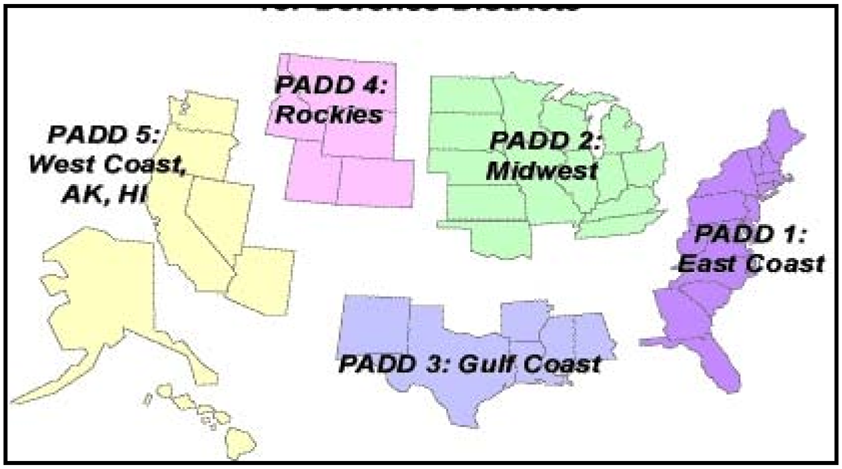

Figure 6-1. Petroleum Administration for Defense Districts (PADDs)

6.2 Data Sources

The primary source of data on freight movement by pipeline is the Petroleum Supply Annual (PSA) published by the U.S. Department of Energy, Energy Information Administration (EIA). http://tonto.eia.doe.gov/dnav/pet/pet_move_pipe_dc_R20-R10_mbbl_m.htm. This publication contains information on the supply and disposition of crude oil and petroleum products. Movements of crude oil and petroleum products are presented by pipeline, tanker, and barge between PAD Districts. This analysis focuses on movements by pipeline.

6.3 Estimation Methodology

Crude oil and petroleum movement data are aggregated at monthly and annual levels. Data are available for a limited number of inter-PAD movements by pipeline as shown in Table 6.1. No data are available for intra-PAD movements by pipeline. Furthermore, variations of the volume of crude oil and petroleum movements between origins and destinations (i.e., among PAD districts) from year to year are erratic. Based on this observation, the growth rate approach is

[13] http://www.eia.doe.gov/pub/oil_gas/petroleum/analysis_publications/oil_market_basics/paddmap.htm

used where it is assumed that trend between the two years preceding the estimating year will be maintained for the subsequent year.

Table 6-1. Data Availability for Freight Movement by Pipeline

From PAD (Origin)	To PAD District (Destination)				
	I	II	III	IV	V
I	0	X	X	0	0
II	X	0	X	X	0
III	X	X	0	X	X
IV	0	X	X	0	X
V	0	0	X*	X**	0

X – data available; 0 – no movements between this pair; X* - no shipments after February 2000
X** - no data available

6.3.1 Domestic Movements by Pipeline

The approach to developing the provisional estimates of domestic movements is based on the following assumptions:

- In distributing the PAD volumes among the states in each PAD district, it is assumed that the state's share of the total volume originating from a given PAD district is directly proportional to the volume of crude oil production in that state.
- A state's share of the total volume destined to a given PAD district is assumed to be proportional to the supply and sales volume reported in that state. The origin state-to-destination state total volumes for the estimation year are based on these shares.
- In expanding state-to-state volumes for FAF zone to FAF zone volumes, assume that the percent share among FAF zones in a given state in the estimation year would be the same as for the previous year.

The following steps are used in developing the provisional estimates:

1. Calculate the growth rates between the 2002 FAF^2 benchmark year and the estimation year for each of the total volumes of commodities moved by pipeline.
2. Apply the growth rates to the 2002 FAF^2 data to obtain the total volume for the estimation year.
3. Expand the total estimates to the FAF zones based on 2002 FAF^2 distribution among the FAF zones for the pipeline mode.
4. In estimating the dollar values of the volumes, calculate the weight/value ratios from the 2002 FAF^2 benchmark data.

5. Apply these ratios to the estimated volumes to derive the values in constant 2002 dollars for the estimation year. In order to derive the current dollar values, adjust the 2002 dollar value estimates by price indices obtained from BLS to account for changes in prices.

6.3.2 International Movements by Pipeline

A similar approach applies to developing the provisional estimates for international freight movements by pipeline, as outlined below.

1. Group countries that import crude oil and petroleum products from or export such products to the United States in the 7 foreign origins/destinations (i.e., Canada, Mexico, Europe, Latin America, Middle East, Asia, and Rest of World).
2. Calculate the total annual volumes of imports and exports of crude oil and petroleum products for the estimation year and base year 2002 from the data sources described above.
3. Calculate the growth rates for total volumes between 2002 and the estimation year.
4. Apply the growth rates in the previous step to the commodity O-D pairs for 2002 FAF^2 data to generate the corresponding estimation year volumes.
5. Repeat the above steps to generate the provisional estimates of the dollar values of commodities moved by pipeline to and from foreign destinations. This approach expresses the estimates in 2002 dollar values. To derive current dollar values, apply commodity export and import price indexes to account for changes in price and inflation.

This page intentionally left blank.

7.0 COMPILATION OF PROVISIONAL DATABASES

7.1 Introduction

The overall goal of the provisional estimates is to provide the latest updates of commodity – O-D movements that would serve as a source of data for practitioners in economic development, transportation planning, and transportation infrastructure analysis. This chapter describes the steps involved in compiling the databases from the individual modal estimates described in the preceding chapters.

7.2 Provisional Databases

Having developed the estimates for each mode, the next step was to compile databases as depicted in Figure 7-1 and described below. These databases were compiled using Structured Query Language (SQL) queries developed specifically for this purpose. The database structure is consistent with the 2002 FAF database. Each record contains FAF zone of origin, FAF zone of destination, port of entry or exit (which applies only to export and import flows), type of commodity, mode of transportation, value in millions of dollars, and tons in thousands of short ton.

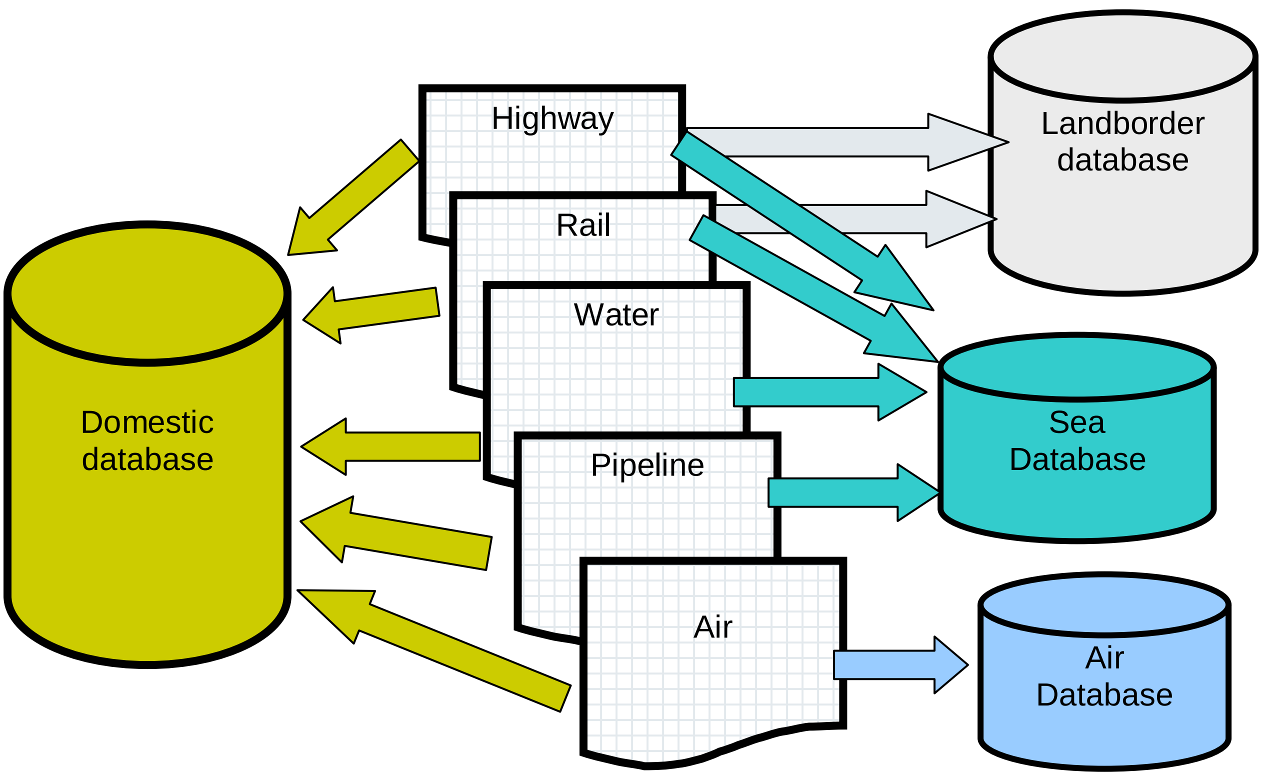

Figure 7-1. Provisional O-D Databases

The databases are:

1. Domestic, comprising all movements with origins and destinations within the U.S. by all modes – highway, rail, water, pipeline, and air.
2. Land border, including international movements, i.e., import and export via land border crossings between the U.S. and Mexico, and between the U.S. and Canada. This includes movements by truck and rail only.
3. Sea, including all international movements, i.e., imports and exports via the seaports between U.S. and other countries. This includes multimodal movements that include truck, rail, pipeline, and water.
4. Air, exclusively for international air cargo covering import and export via airports.

In addition to the databases, state and national summaries were prepared. These summaries present the total volumes and values of commodities originating from or destined to each state by mode as well as the percentages by mode of the total shipments for the given year for that state.

8.0 CONCLUSIONS

The provisional estimates of commodity movements between FAF zone origins and destinations provide annual updates to the FAF^2 database. These estimates are derived from public domain data sources. This report describes the data sources and the methodology used in developing the estimates for each mode of freight movement.

The quality of the provisional estimates is determined primarily by the data from the various sources used. No revisions to the provisional estimates for a given year are expected. However, annual estimates will address any issues arising from the previous year's estimates. The provisional estimates are not intended to compete with private industry data sources. Rather the provisional estimates provide a single comprehensive source of data on all modes of freight movement for use in understanding trends and analyses related to economic development, transportation planning, and transportation infrastructure performance.

This page intentionally left blank.

9.0 REFERENCES

OAI

Code of Federal Regulations, vol. 14, part 241, National Archives, http://www.access.gpo.gov/nara/cfr/cfr-table-search.html#page1

Master Coordinates File, OAI, http://www.transtats.bts.gov/Tables.asp?DB_ID=595&DB_Name=Aviation%20Support%20Tables&DB_Short_Name=Aviation%20Support%20Tables.

Transtats T-100 data, OAI, http://www.transtats.bts.gov/

FHWA/FAF

(Report 1) An Overview of the 2002 Commodity Origin-Destination Database: Methodology and Data, FHWA, http://ops.fhwa.dot.gov/freight/freight_analysis/faf/faf2_tech_document.htm

(Report 6) Freight Analysis Framework Data Dictionary, FHWA, http://ops.fhwa.dot.gov/freight/freight_analysis/faf/faf2_tech_document.htm

FAF^2 Technical Documentation, FHWA, http://ops.fhwa.dot.gov/freight/freight_analysis/faf/faf2_tech_document.htm.

Harmonized System to Standard Classification of Transported Goods Cross-walk, at http://ops.fhwa.dot.gov/freight/freight_analysis/faf/faf2_tech_document.htm

Census/Customs

County, County Sub-District, and Places files, Census, http://www.census.gov/geo/www/gazetteer/places2k.html

Guide to Foreign Trade Statistics, Census, http://www.census.gov/foreign-trade/guide/

U.S. Exports of Merchandise – Monthly DVD ROM, Census Bureau Foreign Trade Division (FTD) (http://www.census.gov/foreign-trade/reference/products/index.html

U.S. Imports of Merchandise – Monthly DVD ROM, Census Bureau Foreign Trade Division (FTD) (http://www.census.gov/foreign-trade/reference/products/index.html

Schedule D - District/Port List, Census, http://www.census.gov/foreign-trade/schedules/d/dist.txt.

County Business Pattern Database, U.S. Census Bureau, at http://www.census.gov/econ/cbp/index.html

Monthly Manufacturers' Shipments, Inventories, and Orders (M3) Survey, U.S. Census Bureau, at http://www.census.gov/indicator/www/m3/

Monthly Wholesale Trade Survey, U.S. Census Bureau,
at http://www.census.gov/mwts/www/mwts.html

USA Trade Online, at http://www.usatradeonline.gov/usatrade.nsf?Open&mc=F9000

Other Sources

Carload Waybill Sample, Confidential Waybill File, Public Use Waybill File, U.S. Department of Transportation, Surface Transportation Board,
at http://www.stb.dot.gov/stb/industry/econ_waybill.html

Gross State Product, U.S. Department of Commerce, Bureau of Economic Analysis,
at http://www.bea.gov/regional/

Import/Export Price Indexes, Bureau of Labor Statistics, at http://www.bls.gov/mxp/home.htm

Internal U.S. Waterway Monthly Indicators, U.S. Army Corps of Engineers, at
http://www.iwr.usace.army.mil/ndc/wcsc/wcmthind.htm

Mapquest, at http://www.mapquest.com/maps/

National Association of Counties,
at http://www.naco.org/Template.cfm?Section=Data_and_Demographics&Template=/cffiles/counties/city_srch.cfm

Navigation Data Center, Waterborne tonnage for principal U.S. ports, U.S. Army Corps of Engineers, at http://www.iwr.usace.army.mil/ndc/wcsc/statenm00.htm

Petroleum Administration for Defense Districts Map, U.S. Department of Energy,
at http://www.eia.doe.gov/pub/oil_gas/petroleum/analysis_publications/oil_market_basics/paddmap.htm

Petroleum Supply Annual, U.S. Department of Energy, Energy Information Administration,
at http://tonto.eia.doe.gov/dnav/pet/pet_move_pipe_dc_R20-R10_mbbl_m.htm

Producer Price Index, Bureau of Labor Statistics, at http://www.bls.gov/ppi/home.htm

Surface Transborder Freight Database, Bureau of Transportation Statistics,
at http://www.bts.gov/programs/international/transborder/

Waterborne Databank, Preliminary Foreign Waterborne Cargo Summary, U.S. Army Corps of Engineers, at http://www.iwr.usace.army.mil/ndc/usforeign/pcsfiles.htm

APPENDIX A

COMPARISON OF DATA SOURCES AND FORECASTS

This page intentionally left blank.

A.1) A Comparison of Data Sources

Table A-1. A Comparison of the OAI and Customs Data with FAF

Dimensions	FAF/CFS	OAI	Census/Customs	Impact
Who Ships	U.S. business establishments in mining, manufacturing, wholesale trade, and select retail industries. Excludes household and government shipments in addition to business sectors not listed above.	Anyone shipping by commercial air carriers in or to/from the U.S., or by U.S. carriers abroad. Military are excluded from the public version of the T-100 data. Carriers designated as All-Cargo carriers, representing a small fraction of air cargo, are also excluded.	Anyone shipping to/from the U.S. except certain intra-governmental shipments.	The three industries excluded from FAF that are likely to have an impact on the OAI/Customs data are Services, Retail, and Transportation.
What is Shipped	All shipments by establishments listed above by weight, value, and commodity type that are directly related to the primary business of the establishment. Excludes by-products, imports, some commodity-types, and shipments by out-of-scope establishments. Shipments weighing less than 100 lbs are typically included in parcel delivery/courier/USPS	All property excluding passenger baggage excepting carriers' own shipments. Although mail is reported separately, and is a separate CFS category, it is questionable to separate mail because FedEx often reports mail as a freight to avoid revealing the size of their contract with the U.S.P.S. OAI is pursuing a remedy to this reporting difference.	Imports: > $2,000, Exports: >$2500 (Canada is an exception), Low-Value are estimated. Mail and parcels are not included in air exports. Shipments that are returned/to-be-repaired under warranty are excluded.	OAI data will cover the weight of shipments of by-products, non-covered commodities, and out-of-scope industries, and small shipments. Customs data will similarly pick up the weight of these shipments as well as their commodity type and value with the caveat that small shipments may be poorly estimated. OAI includes mail/parcel shipments while these are excluded from the air summary of the Customs data and reported under a separate mode for CFS data.

Table A-1. A Comparison of the OAI and Customs Data with FAF (Continued)

Dimensions	FAF/CFS	OAI	Census/Customs	Impact
Where Shipped	Within the U.S. - includes shipments which both start and end in the U.S., but pass through another country, and imports after they leave the importer's U.S. location, and exports to the border. In-transit shipments are excluded as well as shipments to households.	No foreign-to-foreign or in-transit on the publicly available market data. Purely foreign segments are excluded from the segment data.	In-transit is not necessarily included. Shipments to/from Puerto Rico will be attributed to Puerto Rico and therefore excluded from U.S. exports/ imports for the 50 states and the District of Columbia. Imports from the U.S. Virgin Islands are collected, but not exports to the U.S.V.I.	OAI data are likely to pick up a large portion of in-transit shipments that either change planes or carriers in the U.S., or are off-loaded to an FTZ or bonded warehouse.
How Shipped	All modes except pipeline.	Only air	While all modes are reported, only water and air are broken out as separate modes.	OAI data contain both the air-truck and parcel modes used by CFS.
How is it Counted	Through a survey of the above-referenced business establishments	As reported by carriers	Based on documents/electronic filings given to customs detailing what is shipped, standard weights applied. Imports from Canada are provided to Census by the Canadian government	
When is it reported	The survey is taken every five years with the last survey occurring in 2002.	Sixty days after the end of the month for data pertaining to U.S. carriers. Six months after the data month for foreign carriers.	Exports: when the shipment leaves the country. Imports: when the shipment is released to the importer. Note that a shipment may be further processed in a foreign-trade zone before it is released.	The Customs data for imports are likely to lag both the OAI and CFS data.

Table A-1. A Comparison of the OAI and Customs Data with FAF (Continued)

Dimensions	FAF/CFS	OAI	Census/Customs	Impact
How are values counted	Free-on-board at the plant	Not Available	Exports: Free-alongside-Ship (excludes freight and insurance). Imports: a) Customs Import Value – excludes duties, freight, insurance and other costs associated with exportation and assistance from the importer, b) Customs Insurance and Freight (CIF) – CIV + import charges (excludes duties)	The Customs and CFS data are comparable for exports. For imports, the CIF. Customs data best reflects the cost at the border.
Values Measure	Million Dollars	Not Available	Dollars	
Commodity Types	2-digit Standard Classification of Transported Goods	Not Available	10 digit Harmonized Tariff Schedule (Imports)/Schedule B (Exports)	Crosswalk provided by FHWA
Weight Measure	Kilotons (2,000,000 lbs)	Pounds	Kilograms	

Table A-1. A Comparison of the OAI and Customs Data with FAF (Continued)

Dimensions	FAF/CFS	OAI	Census/Customs	Impact
Geography:	CFS – county level and is aggregated to FAF region level for FAF	Airports with latitude and longitude included. Market data: origin airport is where the shipment was initially loaded by the carrier and destination is the airport where the carrier unloaded the shipment. Segment data: origin and destination refer to a non-stop segment whether or not the shipment was unloaded.	Customs District (Sub-Districts reported for some data products and available to help with matching) which may include multiple states. Exports are organized by Port of Exit (generally where loaded on transportation used to cross border) and Destination (foreign country). Imports are organized by Origin (foreign country) - Port of Entry (where it passes through customs) and Port of Unlading (where it is unloaded from the plane).	A crosswalk from counties to FAF regions was provided by FHWA. Supplemental files from Census were used: Counties, County Sub-Districts, and Places.
Other		There is a six-month delay in the release of foreign carrier data. U.S. carriers' foreign-to-foreign shipments are withheld for a period of three years.	Privacy/Disclosure Restrictions may undercount data at District and individual commodity level.	

Sources: CFS/FAF – FAF^2 Technical Documentation, http://ops.fhwa.dot.gov/freight/freight_analysis/faf/faf2_tech_document.htm.
OAI – Code of Federal Regulations, vol. 14, part 241, http://www.access.gpo.gov/nara/cfr/cfr-table-search.html#page1.
Customs – Guide to Foreign Trade Statistics, http://www.census.gov/foreign-trade/guide/

A.2) Forecasts

Table A-2. Third Quarter Forecasts of Tons Enplaned by Foreign Carriers (in 1,000s)

Year	Dom./ Intl.	Actual	Forecasts		Error		Percentage Error	
			YTD Second Quarter	Second Quarter	YTD Second Quarter	Second Quarter	YTD Second Quarter	Second Quarter
Carrier Groups Combined before Forecast								
2002	Combined	1,019	902	955	-117	-64	-11.5	-6.3
	Separate		902	955	-117	-64	-11.5	-6.3
2003	Combined	1,037	1,084	1,045	47	8	4.5	0.7
	Separate		1,084	1,045	47	8	4.5	0.7
2004	Combined	1,186	1,164	1,192	-22	6	-1.8	0.5
	Separate		1,164	1,192	-22	7	-1.8	0.6
2005	Combined	1,211	1,259	1,239	48	28	4.0	2.3
	Separate		1,259	1,239	48	29	4.0	2.4
2006	Combined		1,240	1,226				
	Separate		1,318	1,303				
Average	Combined				-11	-5	-1.2	-0.7
	Separate				-11	-5	-1.2	-0.7
Std Dev	Combined				78	40	7.4	3.8
	Separate				78	40	7.4	3.8
Avg Abs	Combined				58	27	5.4	2.5
	Separate				58	27	5.5	2.5

Notes: Domestic operations by foreign carriers are legally restricted and represent only a small portion of combined operations. The separate forecasts are reported here for consistency with the general framework.

Table A-3. Fourth Quarter Forecasts of Tons Enplaned by All Carriers (in 1,000s)

Year	Dom./ Intl.	Actual	Forecasts		Error		Percentage Error	
			YTD Latest Quarter	Latest Quarter	YTD Latest Quarter	Latest Quarter	YTD Latest Quarter	Latest Quarter
Carrier Groups Combined before Forecast								
2002	Combined	3,567	2,758	3,007	-809	-560	-22.7	-15.7
	Separate		2,766	3,017	-801	-550	-22.5	-15.4
2003	Combined	3,668	4,191	4,173	522	505	14.2	13.8
	Separate		4,204	4,194	536	526	14.6	14.3
2004	Combined	5,709	5,522	5,550	-187	-159	-3.3	-2.8
	Separate		5,685	5,675	-24	-34	-0.4	-0.6
2005	Combined	5,640	5,782	5,763	143	123	2.5	2.2
	Separate		5,747	5,747	107	107	1.9	1.9
2006	Combined		5,668	5,655				
	Separate		5,720	5,682				
Average	Combined				-83	-23	-2.3	-0.6
	Separate				-46	12	-1.6	0.1
Std Dev	Combined				564	450	15.4	12.2
	Separate				557	444	15.4	12.2
Avg Abs	Combined				415	337	10.7	8.6
	Separate				367	305	9.8	8.1
Groups Forecast Separately								
2004	Combined	5,709	5,830	5,821	121	112	2.1	2.0
	Separate		5,833	5,827	124	118	2.2	2.1
2005	Combined	5,640	5,771	5,753	131	114	2.3	2.0
	Separate		5,770	5,753	131	114	2.3	2.0
2006	Combined		5,475	4,913				
	Separate		5,545	4,958				
Average	Combined				-9	43	-1.0	0.5
	Separate				-3	52	-0.8	0.8
Std Dev	Combined				566	442	15.5	12.1
	Separate				566	445	15.5	12.2
Avg Abs	Combined				396	323	10.3	8.4
	Separate				398	327	10.4	8.5

Dom. = Domestic, Intl. = International, YTD = Year-to-Date

Groups: OAI classifies carriers into groups based on operating revenues and the type of certificate the carrier holds. The relevant groups above are carriers that reported fully prior to 10/2002 (majors, nationals, medium and large regionals) and those that began reporting in 10/2002: Federal Express, commuter, small-certificated and all-cargo carriers. Federal Express did report international operations prior to 10/2002, but not domestic. All-cargo carriers include only ABX (formerly Airborne Express).

Data for 2006 was adjusted to account for late reporting carriers as described in the main body of the text.

Notes: Based on OAI T-100 Market data downloaded between October 2006 and January 2007. The earlier years will not reflect the effects of data revisions that will be experienced under normal production conditions going forward. The data is affected by two significant events: the 9/11/2001 terrorist attacks and the 10/2002 expansion of the carrier coverage of the T-100 data to include small-certificated and commuter carriers, and domestic operations of Federal Express.

APPENDIX B

SAMPLE ESTIMATES FOR OAI INTERNATIONAL AVIATION DATA

This page intentionally left blank.

Table B-1. Round-by-Round Results for Estimating Domestic Ports of Entry/Exit for OAI International Aviation Data (2002)

Round	Inbound		Outbound	
	Tons	Share	Tons	Share
All Rounds	4,795,502		3,077,011	
1	3,500,541	73.0	2,300,522	74.8
2	712,928	14.9	366,357	11.9
3	97,311	2.0	44,210	1.4
4	484,722	10.1	365,923	11.9
Segment Data - All	5,081,136		3,387,022	
In-Transit	285,634	5.6	310,011	9.2

Note: Segment data include In-Transit routes that are excluded from the Market data. Additionally, U.S. Territories, Puerto Rico, and the U.S. Virgin Islands trade with each other is excluded from the In-Transit estimate.

Table B-2. A Comparison between Census and OAI Data of Tonnage Shipped by Air (2002)

Customs District		Exports ('000 tons)			Imports ('000 tons)		
Code	Name	Census	OAI	Difference	Census	OAI	Difference
0	All Districts	2,524	3,077	553	3,900	4,796	896
1	Portland	0	0	0	0	1	1
2	St Albans	0	0	0	1	0	-1
4	Boston	49	38	-11	57	62	5
5	Providence	0	0	0	0	0	0
7	Ogdensburg	6	0	-6	1	0	-1
9	Buffalo	12	2	-10	2	2	0
10	New York City	440	581	141	911	980	69
11	Philadelphia	66	90	24	74	107	33
13	Baltimore	2	1	-1	3	3	0
14	Norfolk	1	3	3	0	2	1
15	Charlotte	6	10	5	9	14	5
16	Charleston	1	0	0	1	0	-1
17	Savannah	114	117	3	207	192	-16
18	Tampa	10	15	5	15	21	6
19	Mobile	16	42	26	20	41	21
20	New Orleans	161	126	-35	130	163	34
21	Port Arthur	0	0	0	0	0	0
23	Laredo	4	1	-4	2	1	-1
24	El Paso	3	3	0	3	2	0
25	San Diego	1	3	2	2	3	1
26	Nogales	4	4	0	8	6	-2
27	Los Angeles	321	344	23	524	619	95
28	San Francisco	174	176	1	198	221	23
29	Columbia-Snake	12	9	-3	5	8	3
30	Seattle	56	54	-2	36	41	5
31	Anchorage	49	132	83	178	404	226
32	Honolulu	10	48	38	35	71	36
33	Great Falls	9	5	-4	8	8	-1
34	Pembina	3	0	-3	0	2	2
35	Minneapolis	16	12	-4	12	14	2

continued on next page

Table B2. A Comparison between Census and OAI Data of Kilo-Tons Shipped by Air (2002) contd.

Customs District		Exports ('000 tons)			Imports ('000 tons)		
Code	**Name**	**Census**	**OAI**	**Difference**	**Census**	**OAI**	**Difference**
36	Duluth	1	0	-1	0	0	0
37	Milwaukee	2	0	-2	0	0	0
38	Detroit	53	19	-34	45	45	0
39	Chicago	334	329	-6	520	500	-20
41	Cleveland	158	137	-21	118	129	10
45	St Louis	5	6	1	4	5	0
52	Miami	251	565	314	544	847	303
53	Houston-Galveston	71	68	-3	48	69	21
54	Washington	26	51	25	70	90	20
55	Dallas/Fort Worth	79	86	7	107	123	16
Ratio			1.219			1.230	
Correlation Coefficient			0.935			0.975	

Note: The Customs District for Exports is the District in which the shipment was loaded and the Customs District for Imports is the District in which the shipment was unloaded.

This page intentionally left blank.