



U.S. Department of Transportation
Office of the Secretary of Transportation

Bureau of Transportation Statistics



FREIGHT ANALYSIS FRAMEWORK VERSION 5 (FAF5)

Base Year 2017
Data Development Technical Report

About the Bureau of Transportation Statistics

Leadership

Pat Hu, *Director*

Publication Management

Bureau of Transportation Statistics

About This Report

Project Manager

Bureau of Transportation Statistics

Major Contributors

Ho-Ling Hwang (Oak Ridge National Laboratory), Hyeonsup Lim (Oak Ridge National Laboratory), Shih-Miao Chin (Oak Ridge National Laboratory), Majbah Uddin (Oak Ridge National Laboratory), Alec Biehl (Oak Ridge National Laboratory), Fei Xie (Oak Ridge National Laboratory), Stephanie Hargrove (BLD Group), Yuandong Liu (Oak Ridge National Laboratory), and Ross Wang (Oak Ridge National Laboratory)

Editor

Ross Wang (Oak Ridge National Laboratory)

Visualization Contributors

Ross Wang (Oak Ridge National Laboratory), Hyeonsup Lim (Oak Ridge National Laboratory), and Majbah Uddin (Oak Ridge National Laboratory)

Other Contributors

Federal Highway Administration

Report DOI

<https://doi.org/10.2172/1844893>

Title

Freight Analysis Framework
Version 5 (FAF5) Base Year 2017
Data Development Technical
Report

Key Words

Freight Analysis Framework; FAF5;
freight flows; origin-destination
database; freight modeling

Publication Date

December 2021

Abstract

This technical report documents the data sources and methodologies used to develop the Freight Analysis Framework version 5 (FAF5) base-year regional database, benchmarked to the 2017 Commodity Flow Survey. The FAF integrates data from multiple sources to provide a national picture of freight movements by commodity, mode, and geography within the United States. The FAF5 base-year database supports forecast development, annual estimates, and truck-flow assignment on the highway network.

Recommended Citation

United States Department of Transportation, Bureau of Transportation Statistics. *Freight Analysis Framework Version 5 (FAF5) Base Year 2017 Data Development Technical Report*. Washington, DC: 2021. <https://doi.org/10.2172/1844893>.

All material contained in this document is in the public domain and may be used and reprinted without special permission. Source citation is required.

BTS information service contact information:

Ask-A-Librarian <https://transportation.libanswers.com/>

Phone 202-366-DATA (3282)

Quality Assurance Statement

The Bureau of Transportation Statistics (BTS) provides high quality information to serve government, industry, and the public in a manner that promotes public understanding. Standards and policies are used to ensure and maximize the quality, objectivity, utility, and integrity of its information. BTS reviews quality issues on a regular basis and adjusts its programs and processes to ensure continuous quality improvement.

Notice

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The U.S. Government assumes no liability for its contents or use thereof.

Table of Contents

1. INTRODUCTION	7
1.1. The FAF Program	7
1.2. Development of FAF5 Base Year Data	7
1.3. ORGANIZATION OF THIS REPORT	9
2. CHANGES IN 2017 COMMODITY FLOW SURVEY THAT IMPACTED FAF5.....	11
2.1. GEOGRAPHICAL BOUNDARY	11
2.2. UNKNOWN COMMODITY	12
2.3. MODE ASSIGNMENT	12
3. ESTIMATION OF SUPPRESSED DOMESTIC CFS SHIPMENTS.....	14
3.1. OVERVIEW	14
3.2. ESTIMATION PROCESS FOR IMPUTING CFS SUPPRESSED DATA.....	14
3.2.1. Overview of Estimation Process	14
3.2.2. Estimation of Initial Seed Matrix	15
3.2.3. Iterative Proportional Fitting (IPF)	17
3.2.4. Data Quality Checks and Additional Refinements.....	17
4. ESTIMATION OF OOS FARM-BASED AGRICULTURAL SHIPMENTS ...	19
4.1. DATA SOURCES.....	19
4.1.1. 2017 Census of Agriculture	19
4.1.2. Agricultural Statistics 2018	19
4.1.3. USDA Statistical Bulletins.....	19
4.1.4. 2002 Vehicle Inventory and Use Survey.....	20
4.1.5. 2017 CFS Published Statistics	20
4.2. ESTIMATION PROCEDURES.....	20
4.2.1. Estimations of Farm-Based Agricultural Production by FAF Zone	20
4.2.2. Estimating Farm-Based Shipment OD Flows.....	22
5. ESTIMATION OF OOS MUNICIPAL SOLID WASTE AND CONSTRUCTION & DEMOLITION DEBRIS	25
5.1. MUNICIPAL SOLID WASTE (MSW)	25
5.1.1. Data Sources for MSW.....	26
5.1.2. Estimation Process for MSW.....	27
5.2. CONSTRUCTION AND DEMOLITION DEBRIS (C&D)	30
5.2.1. Data Sources for C&D.....	30
5.2.2. Estimation Process for C&D.....	31
6. ESTIMATION OF OOS RETAIL, SERVICES, AND HOUSEHOLD/BUSINESS MOVES	33
6.1. RETAIL.....	33
6.1.1. Data Sources for Retail Sector	33
6.1.2. Assumptions.....	34
6.1.3. Estimation Process for Retail	35
6.2. SERVICES	36
6.2.1. Data Sources for Services.....	36

6.2.2. Assumptions for Intra-zonal Shipments	38
6.2.3. Assumptions for Inter-zonal Shipments	38
6.2.4. Estimation Approach for Services	40
6.3. HOUSEHOLD/BUSINESS MOVES (HH&B)	41
6.3.1. Data Sources for Household/Business Moves	41
6.3.2. General Assumptions and Assignment of Commodities to SCTG	42
6.3.3. Estimation Process for Household/Business Moves	42
7. ESTIMATION OF OOS CRUDE PETROLEUM.....	44
7.1. DATA SOURCES FOR CRUDE PETROLEUM	44
7.1.1. EIA Data	44
7.1.2. Other Data Sources	47
7.2. ESTIMATION PROCESS	48
7.2.1. Flows of Domestic Crude Petroleum	48
7.2.2. Flows of Imported Crude Petroleum	48
7.2.3. Flows of Exported Crude Petroleum	49
8. ESTIMATION OF OOS NATURAL GAS	51
8.1. DATA SOURCES	51
8.1.1. EIA Data	51
8.1.2. Other Data Sources	52
8.2. ESTIMATION PROCEDURES FOR DOMESTIC NATURAL GAS FLOWS	53
8.3. ESTIMATION PROCEDURES FOR FLOWS OF NATURAL GAS IMPORTS & EXPORTS	55
8.3.1. Imports	55
8.3.2. Exports	55
8.3.3. Mode of Transportation on Imports & Exports	56
9. ESTIMATION OF OOS FOREIGN TRADE.....	57
9.1. DEFINITIONS	57
9.2. PRIMARY DATA SOURCES	58
9.2.1. Census Special Tabulation of 2017 USA Trade Online (USATO) Data	58
9.3. ESTIMATION PROCESSES	59
9.3.1. Step 1: Disaggregation of Commodity Groups to Associated 2-digit SCTG Codes	59
9.3.2. Step 2: Imputing Unknown States	60
9.3.3. Step 3: Unspecified Port Zones	60
9.3.4. Step 4: Unspecified Foreign Modes	61
9.3.5. Step 5: Adjustment of Port Zone Locations	61
9.3.6. Step 6: Disaggregation of State-Level Flows to FAF Regions	62
9.3.7. Step 7: Assignment of Domestic Mode	62
9.3.8. Step 8: Missing Weights and Values	62
9.3.9. Step 9: Foreign Mode is Pipeline or Other/Unknown	63
10. ESTIMATION OF OOS FISHERIES	64
10.1. DATA SOURCES	64
10.1.1. State Total and Process-at-Sea Total - 2017 Commercial Fishery Landings by State	64
10.1.2. Major Port Total - 2017 Commercial Fishery Landings by Top U.S. Ports	64
10.1.3. Historical Landings - 1981 – 2016 Landings by Top U.S. Ports	64
10.1.4. Species Data for Process-at-Sea	65

10.2. ESTIMATION PROCEDURE FOR COMMERCIAL FISHERIES SHIPMENTS.....	65
10.2.1. <i>Distribution of State Total into FAF Zones</i>	65
10.2.2. <i>Distribution of Process-at-Sea total into FAF Zones</i>	65
11. ESTIMATION OF OOS LOGGING.....	67
11.1. DATA SOURCES	67
11.1.1. <i>Forest Statistics by State – FIA DataMart</i>	67
11.1.2. <i>Timber Product Output (TPO) Data Download</i>	67
11.1.3. <i>Forest Industry Research Program, University of Montana</i>	68
11.1.4. <i>TimberMart-South Reports</i>	68
11.1.5. <i>State and Region Price Reports</i>	68
11.2. ESTIMATION METHODS	68
APPENDIX A. GLOSSARY OF TERMS, ABBREVIATIONS, AND	
 ACRONYMS	70
APPENDIX B. CFS AREA – FAF ZONE CROSSWALK TABLE	72
APPENDIX C. FAF5 COMMODITY CODES.....	75
APPENDIX D. AGRICULTURAL COMMODITY CATEGORIZED BY SCTG	
 CODE	76
APPENDIX E. ASSIGNMENT OF FOREIGN COUNTRIES TO FAF FOREIGN	
 ZONES	78

List of Figures

Figure 1-1. FAF5 Base Year Data Development Process	9
Figure 2-1. 2017 CFS Areas and Their Changes.....	11
Figure 3-1. Methodology for Imputing Suppressed Cells in CFS 2017	15
Figure 4-1. 2017 Example of Census of Agriculture Text File and Converted Spreadsheet Records	21
Figure 7-1. Definition of Petroleum Administration for Defense Districts (PADDs)	45

List of Tables

Table 1-1. FAF5 Base Year Data Composition	9
Table 2-1. Share of Rail vs. Truck-Rail under 2012 and 2017 CFS.....	12
Table 2-2. Mapping of Mode Definitions between 2017 CFS and FAF5.....	13
Table 3-1. Percentage of Suppressed Cells in the CFS Special Tables	15
Table 4-1. National Total for Farm-Based Agricultural Shipments in 2017.....	22
Table 4-2. Total Tonnages Originated from CFS Areas within a 50-mile Distance Range of FAF Zone 189 for SCTG 03.....	23
Table 4-3. Resulting OD Flows for the SCTG03 Shipment Example.....	24
Table 5-1. National Totals of MSW Generation (in million tons)	25
Table 5-2. Volume of C&D Generation (in million tons).....	30
Table 6-1. Percentage Assumptions of Truck Shipment for Retail Sales.....	34
Table 6-2. Assumed Relationship between NAICS and SCTG Codes for Retail Sales	35
Table 6-3. NAICS Codes Included under FAF5 OOS Services Component.....	37
Table 6-4. Assumptions for Intrazonal Services Shipments	38
Table 6-5. Assumptions for Interzonal Services Shipments	40
Table 7-1. Volume of Crude Petroleum (in million barrels).....	44
Table 9-1. FAF Foreign Regions.....	57
Table 9-2. Definition of SCTG Group in the Foreign Trade Data File	59
Table 9-3. Examples of Commodity Shares by Geographic Regions.....	60

1. Introduction

1.1. THE FAF PROGRAM

The Freight Analysis Framework (FAF) integrates data from a variety of sources to create a comprehensive national picture of freight movements among states and major metropolitan areas by all modes of transportation. The latest of this data series is FAF5, which is the fifth generation FAF and is benchmarked on Commodity Flow Survey (CFS) 2017. Except for FAF1 that provided estimates for truck, rail, and water tonnage for calendar year 1998, later generations of FAF (FAF2 through FAF5) were built based on their benchmark year CFS data, for 2002, 2007, 2012, and 2017 respectively. The FAF is produced under a partnership between Bureau of Transportation Statistics (BTS) and Federal Highway Administration (FHWA). The latest FAF information are available at the following websites:

- <https://www.bts.gov/faf/>
- https://ops.fhwa.dot.gov/freight/freight_analysis/faf/

As a major data product of the FAF program, the FAF regional database provides a national picture of freight flows to, from, and within the United States (among regions and states), by commodity and mode for the base year, as well as for forecasts up to 30 years into the future in a 5-year interval. Additional FAF data products also include FAF network flows database, where truck movements are routed onto the national highway network, estimates of annual projections, and synchronized historical data series.

This report is a technical document prepared to describe the data sources and methodologies applied in the process of building the FAF5 base-year 2017 regional database, released as FAF5.0 in February 2021. This report offers a description of the diverse data sources and modeling methods used in constructing the base year FAF5 regional database. The FAF5 base-year database is used as the base for development of forecasts and for assignment of truck flows on highway network. Similarly, the FAF5 base-year database will be used as the base to generate FAF5 annual estimates. In addition to this report, users are encouraged to refer to the FAF5 User's Guide¹, which provides basic information of the data, including definitions of the data attributes, information on how to access the data and tool, as well as detailed data dictionary and code tables.

1.2. DEVELOPMENT OF FAF5 BASE YEAR DATA

FAF5 2017 base-year Origin-Destination (OD) database provides estimates of freight volumes, in dollar values and tonnages by origin (O), destination (D), commodity class (C), and mode of transportation (M). Both domestic and foreign trade shipments are represented in the FAF5 flows.

The FAF is built on the CFS data. Similar to previous releases, the 2017 CFS captured shipments that accounted for approximately 71 percent of the FAF5-estimated total volumes by dollar value. Development of FAF5 required integrating additional data to estimate volumes of shipments from several industries that were not covered by the CFS. Since the CFS undercounts U.S. export and does not account for U.S. imports, only domestic shipments from

¹ FAF5 User's Guide, <https://faf.ornl.gov/faf5/data/FAF5%20User%20Guide.pdf>

the CFS were used for constructing FAF5 Origin-Destination commodity flow data. FAF5 freight flow data has the same definitions of the geographic areas² and commodity classes³ as used in the published 2017 CFS data.

In addition to domestic shipments from the 2017 CFS, the FAF5 includes shipments from establishments that were out-of-scope (OOS) to the 2017 CFS. The OOS components in FAF5 are:

- Farm-based Agriculture
- Municipal Solid Waste (MSW)
- Construction and Demolition Debris (C&D)
- Retail
- Services
- Household and Business Moves (HH&B)
- Crude Petroleum
- Natural Gas
- Foreign Trades (imports and exports)
- Fisheries
- Logging

Modes of transportation provided in FAF5 include truck, rail, water, air (include truck-air), multiple modes and mail, pipeline, and other/unknown. A *no domestic mode* was created under FAF to represent imported crude petroleum that is processed at the “dock” thus not involving any transport domestically. For foreign trades, foreign countries are grouped in 8 foreign regions in the FAF5. For more details on the FAF5 data element definitions, readers are referred to the FAF5 User’s Guide.

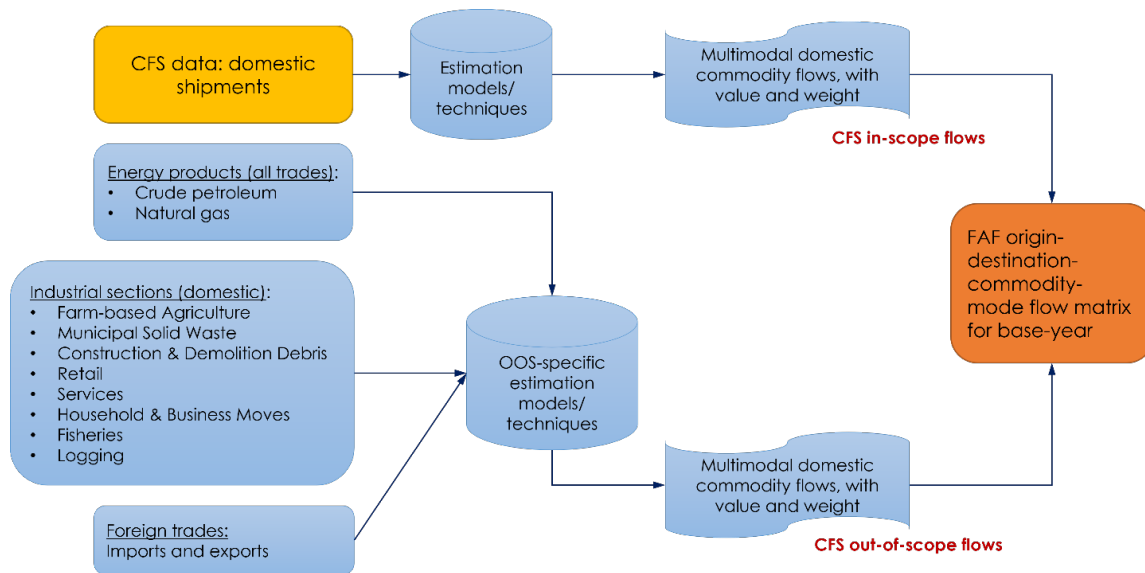
Estimates of each OOS-component for FAF5 were constructed based on various data sources, including administrative records (e.g., Census Foreign Trade Data) and other industry-based data. For missing or suppressed data elements — either due to disclosure concerns or geographic limitations of the source data files — models and imputation techniques (such as iterative proportional fitting and spatial interaction models) were employed to estimate those data components. Several improvements, over previous FAF process, were also applied in estimating flows of commodities for OOS sectors in FAF5.

As the overview processing flow shown in Figure 1-1, the estimated CFS out-of-scope movements were combined with CFS in-scope flows to form the final estimates of the FAF5 base year database.

² Commodity Flow Survey Geographies, <https://www.census.gov/programs-surveys/cfs/technical-documentation/geographies.html>

³ Commodity Flow Survey Commodity Codes, <https://www.census.gov/programs-surveys/cfs/technical-documentation/code-list.html>

Figure 1-1. FAF5 Base Year Data Development Process



The significance of OOS-component estimates can be seen in Table 1-1. Specifically, out of the total 19.7 billion tons and \$18.9 trillion of shipments estimated in FAF5 base year (2017), OOS components accounted for approximately 39.7 percent of the weight and 28.8 percent of the total value. Among the 11 OOS components, natural gas, foreign trades, farm-based shipments, and crude petroleum are four major components in terms of both their weights and values.

Table 1-1. FAF5 Base Year Data Composition

FAF5 Component	Total Weight (Million Tons)	Percent of Weight (%)	Total Value (Billion \$)	Percent of Value (%)
CFS (domestic)	11,886	60.28	13,444	71.21
Farm-based	1,217	6.17	377	2.00
Fishery	5	0.03	5	0.03
Logging	447	2.27	7	0.04
Crude Petroleum (domestic)	508	2.57	162	0.86
Crude Petroleum (foreign trade)	502	2.54	161	0.86
Natural Gas (domestic)	2,011	10.20	371	1.96
Natural Gas (foreign trade)	141	0.71	19	0.10
MSW	381	1.93	0	0
C&D	569	2.88	4	0.02
HH&B	58	0.29	276	1.46
Retail	357	1.81	281	1.49
Service	38	0.19	154	0.82
Foreign Trades ⁴	1,600	8.11	3,617	19.16
Total	19,719	100	18,879	100

1.3. ORGANIZATION OF THIS REPORT

This report details the data sources and methodologies applied to develop the base-year 2017 FAF5 database. Chapter 1 summarizes the document with an overview of the FAF program and the development of the FAF5 base year database. Effects on FAF5 from the changes in the

⁴ This foreign trades component do not include the foreign trade portions of crude petroleum and natural gas.

2017 CFS are highlighted in Chapter 2. Chapter 3 provides a general discussion on the process used in filling data gaps within the domestic CFS matrix; specifically, on the estimation of CFS suppressed cells. Chapters 4-11 describe the methodology and data sources used in estimating the CFS OOS components of FAF5. Several appendices are also provided at the end of this report to offer additional information.

2. CHANGES IN 2017 COMMODITY FLOW SURVEY THAT IMPACTED FAF5

2.1. GEOGRAPHICAL BOUNDARY

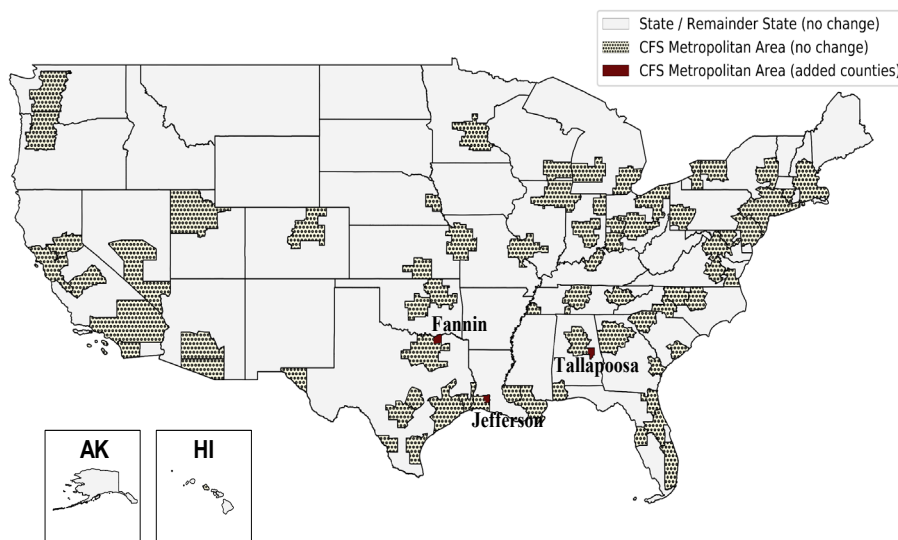
As in the 2012 CFS, the 2017 CFS contains 132 zones for U.S. domestic regions, which are directly carried over to the geography definitions in FAF5. These geographic areas (called “CFS areas”⁵ or “FAF zones”) can be classified as one of the following three types:

- *Metropolitan Area (MA):* The state part of a selected metropolitan statistical area (MSA) or combined statistical area (CSA).
- *The Remainder of State (ROS):* The portion of a state containing the counties that are not included in the MA type CFS Areas defined above.
- *Whole State:* An entire state where no MA type CFS Areas are defined within the state.

Although there is no change to the total number (132 CFS areas) of regions from 2012 CFS to 2017 CFS, boundaries of three 2017 CFS areas and three remainder of states have changed. Below listed counties are removed from the “Remainder of State” category and are included as a part of metropolitan area.

- Tallapoosa County, AL added to Birmingham-Hoover-Talladega, AL - *FAF zone 011*
- Fannin County, TX added to Dallas-Fort Worth, TX-OK (TX Part) - *FAF zone 484*
- Jefferson Davis Parish, LA added to Lake Charles-Jennings, LA - *FAF zone 222*

Figure 2-1. 2017 CFS Areas and Their Changes



⁵ 2017 Commodity Flow Survey Methodology, <https://www2.census.gov/programs-surveys/cfs/technical-documentation/methodology/2017cfsmethodology.pdf>

2.2. UNKNOWN COMMODITY

During the 2017 CFS data processing, a more comprehensive imputation process has been implemented to assign Standard Classification of Transported Goods (SCTG)⁶ codes for those that were missing from records provided by the survey respondents. The 2017 CFS methodology document⁷ describes that a machine learning technique was utilized to impute commodity codes where the respondent provided a description of the product but not an SCTG code. According to the 2017 CFS methodology document, approximately 106,000 shipments (1.65 percent of all shipments with SCTGs) were assigned to certain SCTG codes with a high degree of confidence during the first phase of the 3-phase SCTG imputation process. With this SCTG imputation effort, the 2017 CFS, consequently the associated FAF5, does not contain “unknown” commodity shipments in their records.

2.3. MODE ASSIGNMENT

The location of transfer point for shipments involving rail or water is a key information while determining whether the shipments are to be assigned as single mode or multiple mode (e.g., truck-rail or truck-water). Under the 2017 CFS data processing, the selection of transfer point was determined based on commodity, volume, and distance. However, commodity type was not considered in estimating the transfer point during the 2012 CFS processing. Because of this methodological change, many previously would be classified as single mode of rail/water shipments were assigned to truck-rail/truck-water (i.e., multiple mode) shipments under the 2017 CFS.

To show the impact due to the aforementioned mode assignment change, Table 2-1 compares the split between rail and truck-rail using the 2012 and 2017 CFS published statistics. As shown in the table, rail as a single mode decreased by 16 and 26 percentage points, in tonnage and value, respectively, between the two years. Since FAF5 is built on the 2017 CFS data, impacts from CFS mode changes are carried directly into the FAF5. Although this change did not have any impact on the FAF5 base year estimates, it did cause comparability issues with prior FAF data. Therefore, adjustments were needed on historical FAF databases to allow for a consistent comparison over time. This historical trend synchronization process was conducted as a separate FAF activity, which is out of scope from this report.

Table 2-1. Share of Rail vs. Truck-Rail under 2012 and 2017 CFS

CFS	Tonnage				Value			
	Rail	Rail	Truck-Rail	Truck-Rail	Rail	Rail	Truck-Rail	Truck-Rail
	KTons	Percent	KTons	Percent	M\$	Percent	M\$	Percent
2012	1,628,537	88	213,814	12	473,070	68	224,833	32
2017	1,251,240	73	471,398	27	254,209	42	348,047	58
Difference	-377,297	-16	257,584	+16	-218,861	-26	123,214	+26

⁶ <https://www.bts.gov/sites/bts.dot.gov/files/docs/browse-statistical-products-and-data/surveys/commodity-flow-survey/210866/2017-cfs-commodity-code-sctg-manual.pdf>.

⁷ <https://www2.census.gov/programs-surveys/cfs/technical-ocumentation/methodology/2017cfsmethodology.pdf>

As in the past, FAF adopted the 2017 CFS definitions of transportation modes with some aggregations, mainly due to data limitations. These modes, as described under the FAF5 context, are listed below:

1. Truck – Includes company-owned and for-hire trucks
2. Rail – Includes any common carrier or private railroad
3. Water – Includes inland water, deep sea, Great Lakes shipments, and shipments operating over any combination of water modes
4. Air (includes truck-air) – Includes shipments that are moved by air or a combination of truck and air in a commercial or private aircraft; includes airfreight and air-express
5. Multiple Modes and Mail – Includes intermodal shipments, shipments by multiple modes, and shipments by parcel delivery services, U.S. Postal Service, and couriers; excludes shipments typically weighing more than 150 pounds that move by a combination of truck and air. This category is not limited to containerized or trailer-on-flatcar shipments
6. Pipeline – Includes flows from offshore wells to land
7. Unknown or Other – is defined as “other (single)” mode in the CFS, mostly conveyor belts

Furthermore, Table 2-2 provides a crosswalk of the mode definitions between 2017 CFS and FAF5. FAF5 data also includes “no domestic mode” (coded as “8”) which is created to represent imported crude petroleum that is processed at the “dock” (i.e., refinery located at the port) and thus does not involve further domestic transport. Commodity flow for crude petroleum is not captured in CFS and imported shipment is also out-of-scope for the CFS, therefore the “no domestic mode” is not a CFS attribute.

Table 2-2. Mapping of Mode Definitions between 2017 CFS and FAF5

2017 CFS				FAF5	
Most Detailed Mode Codes		1st Collapsing	2nd Collapsing		
04	For-hire truck	03 Truck	02 Single	1	Truck
05	Company-owned truck	03 Truck	02 Single	1	Truck
06	Rail	06 Rail	02 Single	2	Rail
08	Inland Water	07 Water	02 Single	3	Water
09	Great Lakes	07 Water	02 Single	3	Water
10	Deep Sea	07 Water	02 Single	3	Water
101	Multiple Waterways	07 Water	02 Single	3	Water
11	Air (including truck & air)	11 Air (including truck & air)	02 Single	4	Air (include truck-air)
12	Pipeline	12 Pipeline	02 Single	6	Pipeline
19	Other (single) mode	19 Other (single) mode	02 Single	7	Other and unknown
14	Parcel, USPS, or courier	14 Parcel, USPS, or courier	13 Multiple	5	Multiple modes & mail
15	Truck and rail	20 Non-parcel multiple	13 Multiple	5	Multiple modes & mail
16	Truck and water	20 Non-parcel multiple	13 Multiple	5	Multiple modes & mail
17	Rail and water	20 Non-parcel multiple	13 Multiple	5	Multiple modes & mail
18	Other multiple mode	20 Non-parcel multiple	13 Multiple	5	Multiple modes & mail
Not defined in the CFS (imports are not in-scope for the CFS)				8	No domestic (only applies to certain crude petroleum imports)

3. ESTIMATION OF SUPPRESSED DOMESTIC CFS SHIPMENTS

3.1. OVERVIEW

The CFS publishes origin, destination, commodity, and mode (ODCM) of freight transportation activity by business establishments in mining, manufacturing, wholesale, and selected retail and services trade industries. Over 90 percent of domestic movements (in terms of tonnage and value) are captured in unsuppressed cells but not all data are released. This is due to two major reasons. First, measured and expanded activity captured by the CFS survey may be suppressed from the published tables due to (a) protection of the confidentiality of identifiable shippers, and (b) sampling variability (namely coefficients of variation above 50 percent). Second, it is possible that certain infrequent shipment activities may not be captured under the CFS data collection due to survey sampling limitation, especially at the most detailed ODCM level, more cells are suppressed in the CFS matrix, likely due to confidentiality or data reliability reasons.

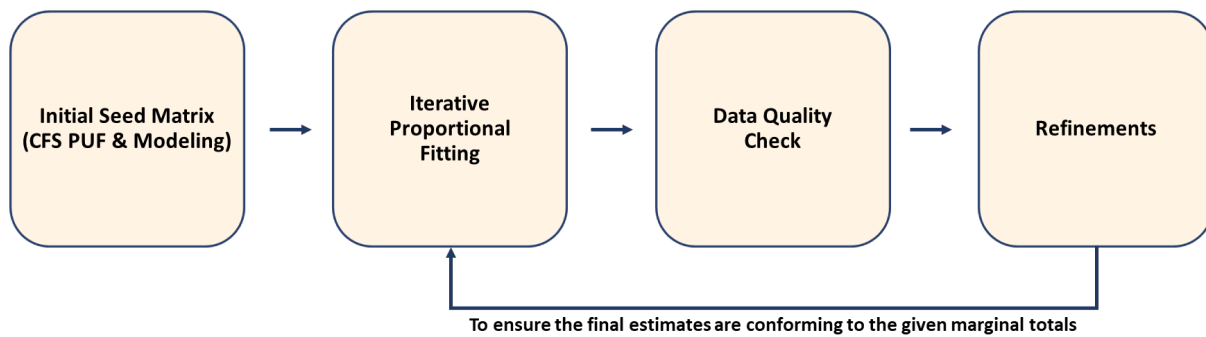
FAF5 suppressed flow estimation process utilized Census-provided special tabulation of domestic-only movements (i.e., excluding exports) as the input data. This special tabulation has fewer suppressions than the published tables. It includes count of shipments for each ODCM cell for both unsuppressed and suppressed cells. This additional information on suppressed cell enabled the FAF team to fill in data gaps in published version of the CFS and helped produce a more comprehensive ODCM matrix.

3.2. ESTIMATION PROCESS FOR IMPUTING CFS SUPPRESSED DATA

3.2.1. Overview of Estimation Process

Figure 3-1 illustrates an overview of the imputation process for estimating suppressed data using the Census-provided CFS special tables. First, the tonnage and value of those suppressed cells were estimated using the 2017 CFS Public Use File (PUF) data and a model-based imputation method. In this first step, the imputed tonnages and values did not align with respect to known marginal totals, thus the resulting matrix was referred to as an “initial seed matrix.” Next, these estimated tonnages and values (i.e., imputed suppressed cells) were iteratively adjusted to conform to all known marginal totals at lower dimensions (e.g., O, D, C, M, OD, OC, ..., DCM) using CFS published tables. This process is known as the iterative proportional fitting (IPF).

Figure 3-1. Methodology for Imputing Suppressed Cells in CFS 2017



Under the third step, a series of comprehensive data quality checks were conducted to examine reasonableness of resulting value-to-weight ratios, modes, and to identify any null/zero values, etc. Finally, if necessary, the estimated tonnages and values were updated/revised according to the finding from step 3. The process from steps 2 to 4 was repeated as needed until the desired level of data quality was reached for the final estimates. More detailed discussions on each of the processing steps are provided in the subsections below (Sections 3.2.3-3.2.4).

Table 3-1 shows the scale of suppressed cells in the CFS special tables. As seen in the table, over 40 percent of cells in the 2017 CFS domestic flow matrix are suppressed at the full ODCM level. These suppressed cells accounted for about 8 percent by total domestic tonnage in the 2017 CFS and 6 percent by value. Expectedly, the percent of suppressed cells decreases as the dimension of matrix is reduced, particularly when measured by tonnage and value. Also, there are no suppressions at one-dimensional marginal totals (i.e., O, D, C, M).

Table 3-1. Percentage of Suppressed Cells in the CFS Special Tables

Comb.	% of suppressed cells	% of total suppressed tonnage	% of total suppressed value	Comb.	% of suppressed cells	% of total suppressed tonnage	% of total suppressed value
OD	1.0	0.0	0.1	ODC	29.3	3.5	6.7
OC	3.7	0.1	0.1	ODM	16.9	2.4	2.4
OM	6.9	0.2	0.0	OCM	18.6	1.6	1.0
DC	0.9	0.0	0.0	DCM	16.2	1.7	1.0
DM	7.6	0.2	0.1	ODCM	40.2	7.7	6.2

3.2.2. Estimation of Initial Seed Matrix

As mentioned previously, the first step of imputing the suppressed cells was to generate the initial seed matrix, which was then fed into the IPF process. Under prior FAF process, this imputation process was implemented using a “log-linear” method, which considered all effects estimated from geometric means. An analysis to explore alternative methods that could potentially improve the estimation of the initial seed matrix was carried out for FAF5 prior to starting the imputation process. The alternative methods considered for FAF5 include maximum likelihood estimates (MLE), multiple imputation, sequentially applying MLE, and utilizing estimates obtained from the 2017 CFS PUF (i.e., shipment-level microdata). Note that the 2012

CFS Public Use Microdata (PUM) data, the first in the CFS data series, was not available at the time of constructing the FAF4 base-year matrix.

The use of CFS PUF data appear to yield the lowest error overall among the alternative methods considered. Therefore, in FAF5, tonnages and values by the ODCM combinations based on the 2017 CFS PUF data, when available, were used in the initial seed matrix. This covers 97 percent of the suppressed cells that need to be estimated in the 2017 CFS in-scope domestic flows. The remaining 3 percent of the suppressed cells were then estimated by the sequentially applied MLE, which outperformed the other model-based alternative methods. The basic idea of this “sequentially applied MLE” modeling approach is to estimate those combinations with only fewer suppressed cells first, while maximizing likelihood of the set of estimates given the marginal total.

Like the log-linear method used in the prior FAF process, the sequentially applied MLE method also utilized the effects of each possible combinations (e.g., O, D, C, M, OD, CM, ..., DCM), obtained from normalizing values by geometric means of all known cells, as exploratory variables in the modeling process. The major improvement with the current FAF5 method comes from two aspects. First, the model leverages the coefficient of variations to enable model parameters to be estimated, where the likelihood of each set of estimates are maximized within the given marginal total. Second, the process of MLE is applied “sequentially,” where the combinations with least number of suppressed cells are estimated in each iteration.

The underlying assumption of the “sequential” approach is to estimate the suppressed cells that are expected to have less uncertainty and gradually reduce the suppressed cells, so that it would eventually improve the overall imputation results. For each given marginal total, the final estimates were obtained by the following equations:

$$y^*_{\theta_{suppressed}} = \arg \max \prod \mathcal{L} \left(y_{\theta_{suppressed}} \mid y_{\theta_{suppressed}} \sim N(y'_{\theta_{suppressed}}, \sigma_{y'_{\theta_{suppressed}}}) \right)$$

$$\text{subject to } \sum y^*_{\theta_{suppressed}} = Y_{\theta} - \sum y_{\theta_{known}}$$

where

$\theta_{suppressed}$ is the set of combinations that are suppressed

θ_{known} is the set of combinations that are known or estimated

y' is the initial estimate for tonnage or value

$\sigma_{y'}$ is the initial estimate of standard deviation for tonnage or value

Y_{θ} is the marginal total of set of combinations θ

y^* is the final estimate of tonnage (or value)

The resulting initial estimates were then used in the refinement step to ensure that the final estimates were conforming to all known marginal totals in lower-dimensional CFS tables.

3.2.3. Iterative Proportional Fitting (IPF)

The IPF process is a procedure for adjusting a matrix of data cells so that they sum up to marginal totals at all given dimensions. This technique is commonly used to generate estimates at a disaggregate level, where many input data are available only at more aggregate levels.

Although the CFS special table provides marginal totals at 1 to 3 dimensional levels (e.g., O, D, C, M, OD, DC, ..., DCM) in most cases, some marginal totals also have suppressed tonnages and/or values (see Table 3-1). To apply the conventional IPF, which works well with a complete set of marginal totals, these suppressed marginal totals will need to be estimated. Doing so, however, the imputed marginal totals would potentially impose additional variations in the final estimates. Therefore, only known marginal totals should be used as the primary control totals. To overcome this limitation, a heuristic approach of IPF was developed to adjust the full ODCM combinations wherever a matrix of marginal totals includes suppressed tonnages and/or values.

The key idea of this heuristic IPF approach is to utilize all “known” marginal totals wherever applicable, while the remaining suppressed marginal totals are grouped dynamically during the iterations. Specifically, for any dimension of the marginal total matrix, the control totals were classified into two groups: 1) “known marginal totals” and 2) “suppressed marginal totals.” In this scheme, the suppressed cells in the known marginal total group were simply adjusted with the given totals. More specifically, for each marginal total, the tonnage and value of all the cells within the marginal total combination were summed and compared to the control totals. If different, all the suppressed cells were adjusted up or down by a constant factor to match the control. Note that no adjustments were made to tonnages and values of known cells within the CFS special table.

For those with suppressed marginal totals, the expected total was obtained by subtracting the sum of all known marginal totals from the national total. This expected total was then applied to adjust (rebalancing) all remaining suppressed cells to ensure that the combined total of all suppressed cells and all known cells matches to the national total. The heuristic IPF process described above was repeated until changes between subsequent iterations of estimated tonnages and values reached a preset threshold value.

3.2.4. Data Quality Checks and Additional Refinements

The processes of estimating initial seed matrix and IPF for estimating suppressed cells were performed separately for tonnage and value. Therefore, at the end of these processes, an examination of value-to-weight ratios on those imputed estimates was conducted to determine whether they were within a reasonable range for the given commodity and given mode. To mitigate issues found from this evaluation, the initial estimates were adjusted based on corresponding value-to-weight ratios obtained from the 2017 CFS PUF data. The IPF process was then repeated to ensure the revised estimates were conforming to all known marginal totals.

Furthermore, additional data checks were performed to ensure the quality of final estimates. This process included validating the reasonableness of modes, reviewing null/zero values, identifying any implausible or unlikely combinations, and resolving all issues found. For benchmarking, external data sources, such as Carload Waybill Sample from the Surface Transportation Board and Waterborne Commerce Statistics published by the U.S. Army Corps of Engineers were utilized. Due to definition differences, these benchmarking data sources were

used mainly to ensure FAF estimates are consistent with general trending, volume scale, and mode share as indicated by the reference datasets.

4. ESTIMATION OF OOS FARM-BASED AGRICULTURAL SHIPMENTS

CFS does not include survey of farmhouse or ranch and therefore it does not account for farm-based agricultural shipments from the field (i.e., farm) to grain elevator, distribution center, processing center, or slaughterhouse. Development of FAF5 base year data includes estimation of farm-based agricultural shipment and this chapter describes methods applied to estimate transport of commodities involved in these shipments. In most cases, farm-based agricultural shipments were assumed to be moved by truck for all states except for Hawaii. Water mode (vessel or barge) is assumed for shipments between islands in the State of Hawaii.

Commodities estimated under this OOS component are allocated to: SCTG 01 (live animal and fish), 02 (cereal grains), 03 (agricultural products), 04 (animal feed, eggs, honey, etc.), and 07 (other prepared foodstuffs, e.g., milk).

4.1. DATA SOURCES

The following is a list of primary data sources that were used in estimating farm-based agricultural shipments.

4.1.1. 2017 Census of Agriculture

The Census of Agriculture⁸ is a census conducted once every five years by the U.S. Department of Agriculture (USDA), with the same cycle as the Economic Census where the CFS is a part of. It provides a complete counts of U.S. farms and ranches with statistical information at the national, state, and county (or county equivalent) levels. All establishments involved in agricultural production (e.g., farms, ranches, nurseries, and greenhouses) are included in this USDA census.

4.1.2. Agricultural Statistics 2018⁹

The Agricultural Statistics is an annual publication prepared by the National Agricultural Statistics Service (NASS) of the USDA. It provides information on agricultural production, supplies, consumption, facilities, costs, and returns. Weights, measures, and conversion factors are also published in this reference book. Information from the 2018-issue of this publication (i.e., for 2017 statistics) was used for FAF5 estimation process.

4.1.3. USDA Statistical Bulletins¹⁰

The NASS also issues a series of bulletins that contain final estimates for agricultural data series based on the review of the 2017 Census of Agriculture data and other information. Specifically, these bulletins contain estimates on historic record and benchmark for current year,

⁸ Information available at: <https://www.nass.usda.gov/Publications/AgCensus/2017/index.php>

⁹ https://downloads.usda.library.cornell.edu/usda-esmis/files/j3860694x/2z10wz00c/9019s972q/Ag_Stats_2018.pdf

¹⁰ NASS statistical bulletins are available at: https://www.nass.usda.gov/Publications/Statistical_Bulletins/index.php

covering all major types of agricultural commodities (e.g., field crops, cattle, chicken and eggs, and citrus fruits).

4.1.4. 2002 Vehicle Inventory and Use Survey

As a part of the Economic Census, the Bureau of Census collects information on the physical and operational characteristics of the private and commercial truck population in the U.S. under the Vehicle Inventory and Use Survey (VIUS). The VIUS was discontinued prior to 2007, making the 2002 VIUS the latest available dataset in this series. Although dated, it is still the only available source for estimating the average travel distance for farm-based agricultural shipments nationwide.

4.1.5. 2017 CFS Published Statistics

Tonnage and value by origin information from the 2017 CFS PUF for shipments involving farm-based agricultural commodities were used in determining associated origin-destination flows in the FAF5 flow estimation process.

4.2. ESTIMATION PROCEDURES

There are three major procedures in generating estimates for farm-based OOS flows. They are (1) estimations of agricultural production by FAF zone, (2) identification of farm-based shipment destinations and their totals by zone, and (3) completion of farm-based agricultural shipment OD flows at the FAF zone level.

4.2.1. Estimations of Farm-Based Agricultural Production by FAF Zone

4.2.1.1. Data Extraction from the 2017 Census of Agriculture Dataset

The 2017 Census of Agriculture data contains both state and county levels product quantities by farm product type. Different from prior releases of the same data series, the 2017 Census of Agriculture data provides individual text files by state which includes both state-level and county-level statistics on different farm commodities (e.g., dollar value of milk, bushels of corn) and assets (e.g., acres of harvested farmland). As the first step, all relevant line items were extracted from the USDA text file and converted into a spreadsheet format for uses in later steps. As shown in Figure 4-1, each row (i.e., a line item) of the spreadsheet corresponds to one commodity statistic item with specific units (e.g., dollar value, acres of land, count, and weight measure), while the columns provide information for both states and counties within the state. The commodity statistics provided in the USDA data include dollar value (usually in \$1,000) and tonnage (measured in various metrics such as tons, acres, bushels).

Figure 4-1. 2017 Example of Census of Agriculture Text File and Converted Spreadsheet Records

Table 25. Field Crops: 2017 and 2012
[For meaning of abbreviations and symbols, see introductory text.]

2017				
Harvested				
Irrigated				
Geographic area	Farms	Acres	Quantity	Farms
BARLEY FOR GRAIN (BUSHELS)				
State Total				
Massachusetts	8	46	1,606	-
Counties				
Berkshire	-	-		
Franklin	-	-		
Hampshire	5	28		
Middlesex	2	(D)		
Worcester	1	(D)		
BUCKWHEAT (BUSHELS)				
State Total				
Massachusetts	2	(D)		
Counties				

Processed data for FAF

VarName	Table	Column	Item	Additional	Alabama	Sumter	Dallas	Lee	DeKalb	Colbert	Ma
VarCol1	Table 14		4 State Total All Goats	17,133	-	-	81	225	564	203	-
VarCol2	Table 16		7 State Total Mohair CLI	1,008	-	-	-	-	-	-	-
VarCol3	Table 18		4 HORSES A -	6,996	-	12 (D)	-	34	151	79	-
VarCol4	Table 18		4 MULES, BU -	551	(D)	-	5 (D)	-	21	11 (D)	-
VarCol5	Table 20		4 CHUKARS -	78,000	-	-	-	-	-	-	-
VarCol6	Table 20		4 DUCKS -	23,330	-	-	(D)	(D)	-	30 -	-
VarCol7	Table 20		4 EMUS -	-	-	-	-	-	-	-	-
VarCol8	Table 20		4 GEESE -	355	-	-	-	-	-	-	-
VarCol9	Table 20		4 GUINEAS -	1,590	-	-	(D)	-	60 -	-	-
VarCol10	Table 20		4 HUNGARI -	-	-	-	-	-	-	-	-
VarCol11	Table 20		4 OSTRICHE -	-	-	-	-	-	-	-	-
VarCol12	Table 20		4 PEACOCKS -	589	-	-	-	-	-	-	-
VarCol13	Table 20		4 PHEASANT -	78,515	-	-	-	(D)	-	-	-
VarCol14	Table 20		4 PIGEONS -	525	-	-	-	-	82 -	-	-
VarCol15	Table 20		4 QUAIL -	3,001,615	(D)	(D)	(D)	-	7,884 (D)	(D)	-
VarCol16	Table 20		4 RHEAS -	-	-	-	-	-	-	-	-
VarCol17	Table 20		4 ROOSTERS -	737,325	(D)	-	(D)	-	146,973 (D)	-	-
VarCol18	Table 20		4 OTHER PO -	(D)	-	-	-	-	120 -	-	-

USDA data



4.2.1.2. Estimating Agricultural Production at State Level

The state-level dollar values for agricultural commodities covered under this farm-based OOS component can be obtained directly from the 2017 Census of Agriculture data. Unlike values, the tonnage estimates for farm-based agricultural commodities require additional unit conversions. This is because agricultural products are typically measured in various weight units depending on the types of products (e.g., pounds, bushels, hundredweight, barrels, tons, etc.). In many cases, the conversion factors (to tonnage) are different even though the “same” measuring unit was used. For example, the approximate net weight for a bushel of buckwheat is 48 pounds, while a bushel of corn is 70 pounds. All conversion factors used in this study are based on information published in the “Weights, Measures, and Conversion Factors” section of the *Agriculture Statistics 2018*.

Because USDA use a much more detailed commodity categorization, agricultural commodities were regrouped into the SCTG categories needed by FAF to the extent possible and reasonable. Under FAF5, several additional USDA commodities (not considered under previous FAFs) are also included. As a result, estimates for this farm-based OOS component is considered to be more inclusive than prior FAFs. New agricultural product items included under the FAF5 are:

- Bedding/garden plants, cut flowers and cut florist greens, foliage plants, potted flowering plants, and other floriculture and bedding crops
- Nursery stock crops
- Cut Christmas trees
- Greenhouse vegetables & herbs
- Mushrooms
- Sod
- Maple syrup

Note that all new items added above are all covered under SCTG 03 in FAF. Therefore, commodity categories of farm-based agricultural shipments covered under the FAF5 remain the same as in FAF4, which includes SCTGs 01 (live animal), 02 (cereal grains), 03 (agricultural products), 04 (animal feed, eggs, etc.), and 07 (other prepared foodstuffs).

Results from this farm-based OOS component shows that nearly 1.2 billion tons valued at \$377 billion of shipments in total were moved from farms across the country in 2017. Table 4-1 shows the total by SCTG at the national level.

Table 4-1. National Total for Farm-Based Agricultural Shipments in 2017

SCTG	Commodity Description	Weight (KTons)	Value (M\$)
01	Animal and fish (live)	79,035	158,564
02	Cereal grains	593,139	63,320
03	Agricultural products (include tobacco)	368,322	114,889
04	Animal feed, eggs, honey, and other animal products	72,064	3,455
07	Other prepared foodstuffs (milk)	104,509	36,994
Total:		1,217,070	377,222

4.2.1.3. Estimating Agricultural Production at FAF Zone Level (Origin of Farm-Based Shipments)

In addition to state-level statistics, the 2017 USDA county level statistics are used to estimate FAF zone level agricultural production tonnages and values by SCTG. When county level estimates were not available, the state total values presented in the USDA data were disaggregated, using the county-level commodity statistics as the share to distribute the state total into individual counties. Once county-level tonnage and dollar value per SCTG group were calculated for each county, County-level production estimates (tonnage and value) were summed to obtain the FAF zone level estimates.

4.2.2. Estimating Farm-Based Shipment OD Flows

4.2.2.1. Determining Destinations of Farm-Based Shipments

In order to estimate OD flows of farm-based agricultural shipments, destination locations of those shipments had to be determined first. The criteria used in selecting potential destinations of farm-based shipments were the same as those used in FAF4. Basically, it was assumed that farm-based commodities are being moved from farms (i.e., origin of the shipments) to grain elevators, distribution/process centers, slaughterhouses, or like kinds (i.e., destinations of farm-based shipments), where shipments beyond those points are covered under the CFS (i.e., in-scope establishments).

With this rationale, CFS shipment origins were utilized as potential destinations for the farm-based shipments within this OOS component. Depending on the commodities being shipped, two types of CFS origins were considered as potential ending locations of given shipments from a farm: (1) CFS areas that shipped out commodity coded in the same 2-digit SCTG and (2) CFS areas that shipped out products associated with the farm-based commodity (e.g., processing of the farm-based commodity into other products). For example, for farm-based shipment of live animals & fish (SCTG 01), the choice of potential destinations could be CFS origins that ship out

SCTG 01 (live animal) or SCTG 05 (meat & fish products). In addition, for farm-based shipment of SCTG 03 (agricultural products), its potential destinations could be CFS origins that ship out SCTG 03 (for non-tobacco products) or SCTG 09 (tobacco products) for tobacco leaves harvested at farms.

4.2.2.2. Estimating OD Flows of Farm-Based Shipments

The OD flow assignment process is identical to the methods used in FAF4. As in prior FAFs, VIUS 2002 data was used in estimating the distribution of average shipment distances. Specifically, VIUS data provides information on the typical “area of operation” of trucks carrying agricultural products. This information is given in categories such as: off-the-road; 50 miles or less; 51 to 100 miles; 101 to 200 miles; 201 to 500 miles; 501 miles or more; not reported; and not applicable (i.e., vehicle not in use). Because the primary interest of FAF is on commodity movements on the national transportation systems, off-road activities were not applicable to this study. Furthermore, since farm-based shipments were generally assumed to be relatively local (shorter trips), over 500 miles categories were eliminated from the estimation process. Using mid-points of the remaining range categories and the distribution of operating ranges, a distribution of shipment length was estimated for each of the associated SCTG commodity-carrying truck groups (i.e., principal product carried) by state.

When multiple destinations (CFS-based) were within the same distance range from a given “production” area (i.e., origin of the farm-based shipment), the estimated total of this shipment was divided proportionally among all involved destination regions, based on their shares of tonnages for the given commodity.

As an illustration of this estimation process, assuming there are a total of 100 thousand tons of agricultural product (SCTG 03) shipments¹¹ originating from the “Rest of IN” (FAF zone 189). Based on the VIUS data, 94 percent of Indiana-based trucks that moved agricultural shipments traveled within 50 miles. Thus, in this example, 94 thousand tons of these shipments will be moved within a 50-mile radius of the origin zone “189.”

With the aid of a Geographic Information System (GIS) tool, four CFS areas are identified as within the 50-mile range (measured from centroid to centroid of two zones) of zone “189.” These four zones along with their associated CFS tonnages (amount originating from each given zone) for the selected commodity are listed in Table 4-2 below.

Table 4-2. Total Tonnages Originated from CFS Areas within a 50-mile Distance Range of FAF Zone 189 for SCTG 03

CFS Area	Destination	2017 CFS total from the given zone for SCTG 03 (1,000 tons)	Zone Share (%)
211	Cincinnati (KY Part)	21.22	12.6
183	Fort Wayne, IN	6.04	3.6
182	Indianapolis, IN	12.44	7.4
189	Rest of IN	128.72	76.4

Using the shares obtained from Table 4-2 above, the 94 thousand tons of SCTG 03 originating from zone “189” are proportionally distributed to four destinations zones: 211, 183, 182, and 189. As a result, four OD flows are created for this example case, all with commodity SCTG 03

¹¹ For simplicity of this example, it is assumed no tobacco-related products are involved.

and domestic mode of truck. Table 4-3 below presents the assigned OD pairs and their shipment tonnages for the example case.

Table 4-3. Resulting OD Flows for the SCTG03 Shipment Example

Origin	Destination	Shipment Weight (1,000 tons)
189 - Rest of IN	211 - Cincinnati (KY Part)	11.84
189 - Rest of IN	183 - Fort Wayne IN	3.38
189 - Rest of IN	182 - Indianapolis IN	6.96
189 - Rest of IN	189 - Rest of IN	71.82
189 - Rest of IN Total		94.00

5. ESTIMATION OF OOS MUNICIPAL SOLID WASTE AND CONSTRUCTION & DEMOLITION DEBRIS

Both municipal solid waste (MSW) and construction and demolition debris (C&D) involve industries that were out of scope to the CFS survey frame, thus their transportation activities were not captured under the 2017 CFS. MSW commodities are allocated to SCTG 41-Waste/Scrap and C&D commodities are allocated to SCTGs 10 (building stone), 11 (natural sands), 12 (gravel), 13 (other non-metallic minerals), 23 (other chemical products), 26 (wood products), 31 (non-metallic mineral products), 32 (base metals), 33 (articles of base metal), and 41 (waste/scrap). This chapter presents the data sources and methodologies used in the FAF5 estimation process for shipments of MSW and C&D.

5.1. MUNICIPAL SOLID WASTE (MSW)

The MSW is mostly common trash or garbage that consists of everyday items people dispose. It is generally produced from homes, schools, hospitals, and businesses. It includes containers and packaging (e.g., soft drink bottles and cardboard boxes), durable goods (e.g., furniture and appliances), nondurable goods (e.g., newspapers, trash bags, and clothing), and other wastes (e.g., food scraps and yard trimmings). The majority of MSW is typically disposed of in landfills and some portion of it is processed in incinerators (i.e., combustion with energy recovery) and in resource recovery facilities.

Table 5-1 presents a summary of national totals on MSW generation, recycling and composting, and combustion with energy recovery for 2012, 2016, and 2017 as reported by the U.S. Environmental Protection Agency (EPA). According to the EPA report, Americans generated about 268 million tons of MSW in 2017. Of that total, about 94 million tons of MSW were recycled and composted (approximately 35 percent recycling rate) and about 34 million tons were combusted with energy recovery (approximately 12 percent incineration rate). The total generation of MSW increased by about 17 million tons from 2012 to 2017, while the amount of MSW disposed of in landfills increased less than 5 million tons in the same period.

Table 5-1. National Totals of MSW Generation (in million tons)

MSW	2012	2016	2017
Landfill	135.0	139.2	139.6
Recycling and Composting	86.6	93.7	94.2
Combustion with Energy Recovery	29.3	33.9	34.0
Total Generation	250.9	266.8	267.8

Data Source: Environmental Protection Agency (EPA)¹²

¹² <https://www.epa.gov/facts-and-figures-about-materials-waste-and-recycling/advancing-sustainable-materials-management>

5.1.1. Data Sources for MSW

Multiple data sources were utilized for the MSW movement estimation, and they are briefly described below.

5.1.1.1. State Solid Waste Management Reports

All states, except Alaska and Idaho, release annual statistics on their solid waste management facilities and activities. These reports and, in some cases, online databases provide information such as the volume of MSW and recycled MSW generation, movement of MSW across state borders, and allocation of MSW to landfills at the county and state levels. Some examples of these reports include Florida's *2017 Municipal Solid Waste Management Annual Report*¹³ and South Carolina's *Solid Waste Management Annual Report*¹⁴. Some examples of state agency online databases include Pennsylvania's *Solid Waste Disposal Information Tool*¹⁵ and Minnesota's *SCORE Program Report Tool*¹⁶.

5.1.1.2. BioCycle - State of Garbage in America

A 2011 survey conducted by the Columbia University Earth Engineering Center (Shin, 2014)¹⁷ on the MSW data generation was utilized. This survey served as the continuation of BioCycle's *State of Garbage in America* survey and was one of the primary data sources for FAF4. However, the Columbia University survey was discontinued in 2014. As pointed out above, state statistics for Alaska and Idaho are not available. Therefore, under FAF5, the MSW generation data for the two states (Alaska and Idaho) were obtained from the BioCycle survey. Note that BioCycle data included C&D within its reported total tonnage of MSW. For FAF5, the same methodology used in FAF4 was applied to estimate the MSW tonnages for the above two states.

5.1.1.3. EPA Advancing Sustainable Materials Management: 2017 Fact Sheet

The EPA report, *Advancing Sustainable Materials Management: 2017 Fact Sheet*¹⁸, contains statistics for MSW materials generation, recycling, composting, combustion with energy recovery and landfilling in the U.S. This report provided a high-level overview of how MSW is managed and the distributions among those management methods.

5.1.1.4. EPA Landfill Methane Outreach Program (LMOP) Database

The LMOP Landfill and Landfill Gas Energy Project Database¹⁹ contains information about physical address, latitude and longitude, owner/operator organization, operational status, design capacity, etc. for more than 2,600 MSW landfills. The locations of existing open MSW landfills for 2017 were used to identify destinations for landfilled MSW shipments.

¹³ <https://floridadep.gov/waste/waste-reduction/content/2017-municipal-solid-waste-management-annual-report>

¹⁴ <https://scdhec.gov/environment/land-management/solid-waste/sc-solid-waste-management-annual-report>

¹⁵ http://cedatareporting.pa.gov/reports/powerbi/Public/DEP/WM/PBI/Solid_Waste_Disposal_Information

¹⁶ <https://www.pca.state.mn.us/waste/report-2017-score-programs>

¹⁷ <https://secureservercdn.net/198.71.233.199/epm.300.myftpupload.com/wp-content/uploads/2020/10/DollyShinThesis.pdf>

¹⁸ Advancing Sustainable Materials Management: 2017 Fact Sheet, EPA, available at: https://www.epa.gov/sites/production/files/2019-11/documents/2017_facts_and_figures_fact_sheet_final.pdf

¹⁹ <https://www.epa.gov/lmop/landfill-technical-data>

5.1.1.5. Recycling Partnership Municipal Materials Recovery Facility Database

A comprehensive map of municipal materials recovery facilities (MRF) in the U.S. has been compiled and publicly released by the Recycling Partnership²⁰. This map contains location, latitude, and longitude, MRF name/company name, etc. for more than 350 residential MRFs. Locations of the MRFs were used in determining destinations for recycled MSW shipments.

5.1.1.6. Energy Recovery Council Waste-to-Energy Facility Database

The directory of waste-to-energy (WTE) facilities in the U.S. is available from Energy Recovery Council²¹. This directory provides information about county, owner/operator company, start year, MSW input capacity, etc. for all WTE facilities. In 2017, there were 76 operating WTE facilities in 21 states. This information was used to identify destination locations for WTE MSW shipments.

5.1.1.7. EIA Electricity Power Plant Level Database

The Energy Information Administration (EIA) collects detailed electric power data on electricity generation, fuel consumption, receipts at the power plant level using the EIA-923 survey form. This data contains the annual volume of total MSW consumption at the individual WTE plant level. These volumes were used in estimating flows of MSW that were converted to energy.

5.1.1.8. Carload Waybill Sample

Published by the Surface Transportation Board (STB), the Carload Waybill Sample²² database is a stratified sample of carload waybills for all U.S. rail traffic submitted by rail carriers that terminate 4,500 or more revenue carloads annually. The 2017 *Carload Waybill Sample* database captured over 670 thousand waybills reported electronically in 2017. This data source served as the benchmark for determining the pattern and volume of MSW shipments being moved by rail.

5.1.1.9. Waterborne Commerce Data

Waterway network link commodity data²³ from the Waterborne Commerce Statistics Center (WCSC), a part of the U.S. Army Corps of Engineers, contains tonnage for waste commodity by link and direction. Waterway links represent either actual shipping lanes (i.e., channels, Intracoastal waterways, rivers) or serve as representative paths in open water (where no defined shipping paths exist). This data was used in estimating MSW shipments by water.

5.1.2. Estimation Process for MSW

Basically, at the state-level, MSW generation data for landfilling, recycling, and WTE can be obtained and then sum up to estimate the total state-level MSW generation volumes. The methodology applied in estimating MSW flows at the FAF-zone level is described in the following subsections.

²⁰ <https://recyclingpartnership.org/residential-mrfs/>

²¹ <http://energyrecoverycouncil.org/wp-content/uploads/2019/10/ERC-2018-directory.pdf>

²² For access of Carload Waybill Sample data visit Surface Transportation Board website at: <https://prod.stb.gov/reports-data/waybill/> or access the 2017 documentation at: <https://prod.stb.gov/wp-content/uploads/2017-STB-Waybill-Reference-Guide.pdf>

²³ <https://usace.contentdm.oclc.org/digital/collection/p16021coll2/id/3004>

5.1.2.1. Obtain/Estimate State-Level MSW Generation Volumes by Weight

Forty-eight states and Washington, DC provided their total amounts of MSW generation that were landfilled in the state. For the remaining two states (Alaska and Idaho), data from BioCycle was used. When landfilled data was not available for the base year (2017), data for other years were used to generate estimates for 2017, based on population changes over time in the given state.

Thirty-nine states and Washington, DC provided their total amount of MSW that were recycled in the state. Similar to landfilled data, recycled volumes were generated for 2017 based on state population changes when the available data was for a different year. For states with no available data, a regression model was used to estimate their recycled amounts. The independent variables selected for this regression model were “In-state landfill amount” and “Recycling requirement” (i.e., whether a state has at least one mandatory recycling requirements), where the “Recycling requirement” is a binary variable and “In-state landfill amounts” is a continuous variable. States with mandatory recycling requirements were determined based on a report by the Northeast Recycling Council²⁴. The regression equation used for estimating recycled amounts for states without reported data is provided below.

$$\begin{aligned} \text{Recycled amount} \\ = \exp(6.0010 + 0.4941 \times \log(\text{In-state landfill amount}) + 0.7218 \times \text{Recycling requirement}) \end{aligned}$$

In the case of MSW for WTE, information from the EIA electricity power plant level database was obtained. This EIA data provided state-level generation amounts for all 21 states with WTE facility for 2017. Based on above discussed three datasets, it is estimated that a total of 381 million tons of MSW was landfilled, recycled, or incinerated in 2017, the base year of FAF5.

Note that, state reports also indicated that MSW volumes that was moved to landfills across state borders. Based on state reported information, out-of-state OD pairs and their associated tonnages of the MSW was determined. It was estimated that 28 million tons of MSW were transported across state borders in 2017. These out-of-state MSW movements accounted for about 7 percent of the total MSW moved in 2017. An evidence that most MSW shipments are moved within state (i.e., intra-state movements).

5.1.2.2. Obtain/Estimate FAF zone Level MSW Generation Volumes by Weight

For shipments of MSW that moved entirely within a state, when county-level data were available, they can be aggregated to produce the estimated volume at the FAF zone level. When county-level data is not readily available, the state-level total MSW tonnages were disaggregated using population shares to estimate FAF zone level generation tonnages. The 2017 population data published by the Census was used in this FAF5 process. The rationale here is that the amount of MSW generation is closely correlated to the population size of a region (FAF zone in this case).

²⁴ Disposal Bans and Mandatory Recycling in the United States, Northeast Recycling Council, available at: https://nerc.org/documents/disposal_bans_mandatory_recycling_united_states.pdf

5.1.2.3. Identify the Nearest MSW Processing Facility

The destination information of the MSW shipments was not given in majority of state reports and databases. In those cases, MSW is assumed to be transported from its origin county to the nearest facility (i.e., landfill for landfilled MSW, or MRF for recycled MSW). For incinerated MSW, WTE facility websites and online sources provide information on MSW source region (county). Therefore, origin county for incinerated MSW can be identified; this is in a reverse order when compared to identifying facilities for landfilled and recycled MSW.

As mentioned above, facility locations for landfilled and recycled MSW were identified considering the closest facility. A geospatial analysis method using the Voronoi diagram was applied to allocate the origin counties for each facility. Note that the Voronoi diagram is not spatially limited by state boundaries. For that, the closest open facility to a county centroid can be in a neighboring state, in which case the shipment would be counted toward the out-of-state shipment totals. Because of this, there is a need to examine the reasonableness of OD county pairs. Specifically, if the results produced from the geospatial analysis method indicated certain existence of out-of-state OD pairs but no evidence of such flows existed in the given state report, the destinations of these OD pairs are then reassigned to the nearest facility within the same state. Similarly, if the results show no out-of-state OD pairs but the state report suggests otherwise, then the nearest out-of-state facilities are assigned as the destinations instead of the Voronoi diagram identified locations.

5.1.2.4. Estimate FAF zone Level MSW Flows

For a few states (e.g., Indiana and Washington), county-level OD pairs MSW data were available and were used directly for FAF estimates. For other states, completed county-level OD pairs were established through the steps and processes discussed above. With these, the list of county-level OD pairs for all MSW shipments was compiled and flows for associated FAF zone level OD pairs was then generated. Note that MSW is categorized under SCTG code “41-Waste/scrap” in FAF and assumed to have no dollar value.

5.1.2.5. Assignment of MSW Flows by Mode of Transportation

- **Rail:** Information on MSW shipments by rail were obtained from state reports and databases as well as online sources. FAF team members also had a discussion with one representative from a major waste disposal company in the U.S., which further assisted with the refinement of OD flow matrix. Lastly, the flow matrix was benchmarked against the pattern and tonnage of MSW shipments from the 2017 Carload Waybill Samples.
- **Water:** Waterway network link commodity data from WCSC provided a flow map for waste materials. This data has waste tonnages by direction on the National Waterway Network. The flow map was spatially joined with the FAF-zone data layer. Then, by examining flow directions on links and how flow tonnages changed inside a FAF zone, a FAF zone level OD flow matrix was estimated. It should be noted that waste materials defined by WCSC include a few additional materials which are not MSW or C&D. Due to this reason, OD flows having destination link with nearby landfills were assumed to be carrying MSW or C&D. In addition, if online sources mentioned movements of MSW or C&D on certain waterway links, the above assumption was also made. To estimate MSW shipments by water for those OD flows, a factor estimated as the ratio of landfilled MSW over landfilled C&D at the state-level was applied.
- **Truck:** Besides the MSW shipments by rail and water, all remaining MSW shipments are assumed to be via truck. Thus, MSW shipments by truck were estimated by subtracting

shipments by rail and water from the total FAF zone level OD flows. In terms of tonnage, MSW flows by truck has the largest share in 2017 (approximately 98 percent).

5.2. CONSTRUCTION AND DEMOLITION DEBRIS (C&D)

C&D comprises a significant portion of the solid waste stream in the U.S. Shipments originating from activities in this OOS sector include companies or establishments engaged in the construction of residential and commercial buildings, utility systems, roadways and bridges, and from specialty trade contractors (e.g., Foundation, structure, and building exterior contractors) that are out of scope to the CFS. In general, there are two major methods employed in the U.S. for C&D management: landfilling and recycling. The majority of C&D is recycled; however, the tracking of tonnage data has been limited in the past.

According to EPA (see Table 5-2), a total of 569 million tons of C&D were generated in the nation in 2017. Of which, about 33 million tons were generated during construction (approximately 6 percent) and about 536 million tons were demolition debris (approximately 94 percent). The total generation of C&D increased by about 22 million tons from 2015 to 2017.

Table 5-2. Volume of C&D Generation (in million tons)

C&D Generation	2015	2017
Waste during construction	29.6	33.3
Demolition debris	518.2	536.1
Total C&D generation	547.8	569.4

Data source: Environmental Protection Agency (EPA)²⁵

The C&D flow estimation process for FAF5 OOS here is similar to the process used in estimating MSW flows. It is common for MSW and C&D to be disposed of in the same landfills. Note that only the domestic portion of the C&D was considered here because shipments involving foreign trades are covered under a separate OOS component and discussed in Chapter 9 of this report.

5.2.1. Data Sources for C&D

Data sources used to estimate C&D shipment flows are summarized below.

5.2.1.1. Construction & Demolition Recycling Association (CDRA) C&D Recycling Center Database

The CDRA provides a database²⁶ with the location information of the major C&D recycling centers across the U.S. These recycling centers were used in estimating destinations for recycled C&D.

²⁵ <https://www.epa.gov/facts-and-figures-about-materials-waste-and-recycling/advancing-sustainable-materials-management>

²⁶ <https://cdrecycling.org/directory/results/>

5.2.1.2. Census Building Permits Data

Census provides data on the number of new housing units authorized by building permits²⁷. The county-level new housing unit numbers were used in the disaggregation of state-level C&D data into county-level.

5.2.1.3. Data Sources Same as those Used for MSW

Several of the data sources used to estimate C&D movements were the same as those used for MSW (see Section 5.1.1), including

- State Solid Waste Management Reports
- EPA Advancing Sustainable Materials Management: 2017 Fact Sheet
- EPA LMOP Database
- Carload Waybill Sample
- Waterborne Commerce Data

5.2.2. Estimation Process for C&D

The methodology applied for estimating C&D flows at the FAF zone level is summarized in the following subsections.

5.2.2.1. Obtain/Estimate State-Level C&D Generation Volumes by Weight

The total state-level C&D generation can be calculated as the sum of landfilled and recycled amounts. Forty-one states and Washington, DC provided their total amounts of C&D generation that were landfilled in the state. When landfilled data was not available for the base year (2017), data for other years were used to generate estimates for 2017, based on changes in housing units authorized by building permits over time in the given state. For states with no available data providing the total amounts of C&D generation that were landfilled in the state, a regression model was used to estimate the landfilled amounts. The model, for estimating C&D generation that was landfilled in-state, is provided below, where “Housing unit” denotes the number of housing units authorized by building permits.

$$\text{Landfilled amount} = \exp(5.5553 + \log(\text{Housing unit}))$$

Once C&D generation amount for landfilling in all the states were determined, the generation total for recycling was estimated by subtracting total landfilled C&D amounts from the total C&D generation. Then state-level recycling amounts were estimated by distributing the national total using the share of landfilled amounts. Out-of-state OD pairs and associated tonnage of the C&D being moved were determined based on state reports. It was estimated that 7 million tons of C&D were transported across state borders in 2017. These out-of-state shipments accounted for about 1 percent of the total C&D moved.

5.2.2.2. Obtain/Estimate FAF zone Level C&D Generation Volumes by Weight

For shipments of C&D that moved entirely within a state, when county-level generation data were available, those were aggregated to produce the estimated generation volume at the associated FAF zone level. When county-level data are not readily available, the state-level

²⁷ <https://www.census.gov/construction/bps/>

C&D tonnages were disaggregated to FAF zones within the given state, using the shares of housing units from building permits.

5.2.2.3. Identify the Nearest C&D Processing Facility and Estimate FAF zone Level C&D Flows

The steps for identifying the nearest processing facility and estimating FAF zone level flows for C&D are the same as those used for MSW. Please refer to Section 5.1.2 for details.

5.2.2.4. Assignment of Flows by Mode

The assignment of C&D flows by mode (i.e., rail, water, and truck) is the same as the process done for MSW described in Section 5.1.2.

5.2.2.5. Assignment of Flows by Commodity

The landfilled C&D flows were assigned to SCTG commodity code “41-Waste/scrap” and were assumed to have no dollar value. The recycled C&D flows were assigned to nine SCTG commodity codes depending on the materials involved. The commodities are SCTG 10 (building stone), 11 (natural sands), 12 (gravel), 13 (nonmetallic minerals), 23 (chemical products), 26 (wood products), 31 (nonmetal mineral products), 32 (base metals), and 33 (articles-base metal). These C&D flows were distributed among impacted commodity codes based on their tonnage shares from the 2017 CFS. As for the value of these flows, dollar-per-ton estimates were obtained from CFS and applied. Note that a factor of 0.05 was used for the value estimates, to reflect the fact that recycled materials have much lower values as compared to “new” materials represented by the CFS data.

6. ESTIMATION OF OOS RETAIL, SERVICES, AND HOUSEHOLD/BUSINESS MOVES

As detailed in its accompanying documentation²⁸, the 2017 CFS only covers shipments from retail establishments under North American Industry Classification System (NAICS) 4541 (Electronic shopping and mail-order houses) and 45431 (Fuel dealers). Similarly, for the services sector, the 2017 CFS only covers one services group, which combines shipments under NAICS 5111 (Newspaper, periodical, book, and directory publishers) and 51223 (Music publishers). Therefore, to expand on the representation of freight movements that are OOS to CFS, activities involving other retail and services sectors were conjectured to include goods of notably high value and/or weight under FAF. The SCTG commodities captured by these sectors are detailed later in this chapter.

Similarly, transport of household and business (HH&B) goods by storage and moving industry for relocation are not captured in the 2017 CFS. The commodities covered under the HH&B OOS sector are as follows: SCTGs 35 (electronic equipment), 36 (motorcycles & bicycles), 37 (transportation equipment), 39 (furniture), 40 (sporting goods), and 43 (miscellaneous).

Data sources and estimation processes for shipments from OOS components of retail, services, and household/business moves are summarized in this chapter.

6.1. RETAIL

6.1.1. Data Sources for Retail Sector

6.1.1.1. Census Annual Retail Trade Survey

One of the main data sources used in generating retail sector (i.e., NAICS code 44-45) estimates was the Annual Retail Trade Survey (ARTS)²⁹, published by the U.S. Census Bureau. Namely, the 2017 revised sales estimates were extracted from the table “Estimated Annual U.S. Retail Trade Sales - Total and E-commerce: 1998-2018” by 3-digit NAICS code and used as national control totals for the industries associated with the retail sector. Total retail trade sales in 2017 were estimated at \$5,053 billion, including \$459 billion from e-commerce; reflecting 17.3 percent and 100.4 percent increases, respectively, as compared to their 2012 volumes. Businesses with or without paid employees are included in these Census estimates. The Census defines e-commerce as “transactions sold online whether over open networks such as the Internet or proprietary networks running systems such as Electronic Data Interchange (EDI).” Note that NAICS 441 (motor vehicle and parts dealers), 447 (gasoline stations), and 454 (non-store retailers) are not included in FAF OOS retail estimates.

6.1.1.2. 2017 Economic Census Receipts

Estimates of state-level business data including sales receipts by NAICS (specifically 44-45 for this OOS sector³⁰) are also published by the Census based on information collected under the 2017 Economic Census. Receipts are defined by the Census as “operating revenue for goods

²⁸ <https://www.census.gov/content/dam/Census/library/publications/2017/econ/ec17tcf-us.pdf>

²⁹ <https://www.census.gov/data/tables/2018/econ/arts/annual-report.html>

³⁰ <https://www.census.gov/data/tables/2017/econ/economic-census/naics-sector-44-45.html>

produced or distributed, or for services provided ... excludes local, state, and federal sales and other taxes collected from customers or clients.” Because the Economic Census is conducted by the Census in a five-year interval, receipts data are available for years ending in “2” and “7”. The 2017 Economic Census estimated a total of \$4,950 billion in receipts generated from the retail sector, which is 2 percent lower than the ARTS. It is important to note that receipt data at the 3-digit NAICS level was available for all U.S. states and Washington, DC; in the case of 2012, however, this disaggregated data was available only for 26 states and all but two of the retail sectors in the District of Columbia.

6.1.1.3. 2017 County Business Patterns and Population Estimates

Employment payroll data from the Census Bureau’s County Business Patterns (CBP)³¹ at the 3-digit NAICS level is utilized in distributing state-level estimates to the FAF areas. In addition, 2017 American Community Survey (ACS) county population estimates are used for determining the volume of trade between adjacent FAF areas. Further details are included later.

6.1.2. Assumptions

It is expected that most retail sales are made at store locations, where customers transport the purchased goods to their home. However, certain large items purchased from the retail stores, such as furniture or appliances, are likely delivered by truck to a customer’s home. Due to lack of data on actual share of sales associated with goods delivered versus carried out, a set of assumptions on percent of retail-sales receipts involving truck were made for the FAF5 process. This is in addition to removing e-commerce shipments to establish estimates for overall physical shipment totals. Table 6-1 details this percentage breakdown.

Table 6-1. Percentage Assumptions of Truck Shipment for Retail Sales

NAICS Code	Description	ARTS Total Revenue (billion \$)	ARTS E-Commerce (billion \$)	Assumed Delivery Truck Shipment (%)
442	Furniture and home furnishings stores	113.0	1.3	70
443	Electronics and appliance stores	98.6	2.1	20
444	Building materials, garden equip. & supplies dealers	365.7	2.6	30
445	Food and beverage stores	725.1	3.4	3
446	Health and personal care stores	333.3	0.7*	2
448	Clothing and clothing access. stores	260.6	9.4	5
451	Sporting goods, hobby, musical instrument, and bookstores	84.2	2.4	4
452	General merchandise stores	687.1	0.4*	4
453	Miscellaneous store retailers	124.5	4.0	10

*Imputed due to suppression in ARTS table.

³¹ <https://www.census.gov/programs-surveys/cbp.html>

Table 6-2. Assumed Relationship between NAICS and SCTG Codes for Retail Sales

NAICS Code	SCTG Codes ³²
442	30, 31, 39, 40
443	34, 35
444	24, 26, 31, 33, 34
445	03, 04, 05, 06, 07, 08, 27, 28, 29
446	20, 21, 23, 28, 38
448	30, 40
451	29, 30, 40
452	03, 04, 05, 06, 07, 08, 20, 21, 23, 27, 28, 29, 30, 31, 34, 35, 39, 40
453	03, 04, 09, 13, 28, 38, 40, 43

Furthermore, each NAICS code must be translated into a corresponding set of SCTG codes to be compatible with FAF; here, an equal distribution of total value is assumed across all selected commodities. This is summarized in the Table 6-2 above. Note that, for comparative purposes, this one-to-many correspondence slightly differs from the FAF4 process.

6.1.3. Estimation Process for Retail

6.1.3.1. Estimating Total Retail Values Shipped by States

As mentioned earlier, 2017 Economic Census receipts data at the 3-digit NAICS is available for all states. Therefore, total state-level retail receipts associated with industry sectors involving truck deliveries can be estimated directly using the shares and assumptions, shown in Table 6-1, based on the proportion of physical shipments as indicated by 2017 ARTS estimates. Next, for each NAICS code, total state-level shipment values are distributed among the associated SCTG commodities using a simple equal-share assumption.

6.1.3.2. Distribution of State Totals to FAF Zones

The state-level estimates of values are distributed to each FAF zone within the given state using their shares of total annual payroll amounts obtained from the 2017 CBP dataset. Note that the calculation of shares considered only payroll information associated with the 3-digit NAICS codes included in FAF. For simplicity, this FAF zone level distribution is conducted uniformly over all impacted SCTG commodity codes within a given NAICS code.

6.1.3.3. Estimating FAF Zone Level Shipment Weight

Subsequently, two value-to-weight ratios are calculated based on attribute data for domestic shipments from the 2017 CFS PUF at origin-commodity-mode and at commodity-mode. The shipment weights are then estimated using the most detailed ratio available. Based on the processes described above, it is estimated that a total of \$281 billion, weighing 357 million tons, of CFS OOS retail goods were transported by truck in 2017. Compared to 2012, these reflect a 36.4 percent increase in value and 59.4 percent increase in weight.

6.1.3.4. Estimating Retail Trade between FAF Zones

Under FAF4, OOS retail activities were assumed to be intrazonal, thus the origin and destination FAF zones were assumed the same. Under FAF5, this assumption was expanded to include truck deliveries made to adjacent FAF zones. The following steps are conducted to estimate this

³² Detailed SCTG codes and their corresponding commodities in FAF5 can be found in APPENDIX C.

inter-zonal retail trade for all FAF zone types (i.e., metro areas, rest of states, and full states). Note that Alaska and Hawaii are treated separately from the contiguous United States.

Step 1: Generate county neighbor matrix and assign 2017 population estimates. A total of 2,676 valid and distinct core-neighbor pairs are established after quality checks.

Step 2: Remove neighboring counties whose populations exceed the mean county population within a given core FAF zone. This is done to reflect the “intervening opportunities” concept in geographic theory.

Step 3: Aggregate county populations to the corresponding FAF zone. To reflect the concept of distance decay, the population of neighboring FAF zones is reduced by 2/3. Thus, not all counties in a FAF zone are included in this aggregation process.

Step 4: Values and weights are assigned to each OD pair (including flows that stay within the core FAF zone) based on the proportion of the destination area’s population among the combined population totals of the core FAF zones and its neighbors.

For Alaska, since it is a single FAF zone without domestic neighbors, all retail flows are assumed to stay within the state. For Hawaii, which is comprised of two FAF zones, inter-zonal retail trade based on populations of all counties associated within each FAF zone. Furthermore, the following corrections were applied to the initial estimates, which then resulted in a recalculation of corresponding shipment weights:

- In the case of metro-metro adjacency, for any core FAF zones whose intra-zonal retail estimates were identified as “low” based on expert knowledge, the fraction of retail goods for each core-neighbor pair was compared with its counterpart in the “mirrored” case (i.e., origin and destination are switched). If these fractions surpassed a specified difference threshold, indicating certain kind of “trade imbalance,” 2/3 of the original inter-zonal trade fraction was added to the intra-zonal trade fraction for the core FAF zone.
- Any trade flows whose values were less than one percent of the total within any given core-neighbor estimation set were merged into the intra-zonal trade estimate for the core FAF zone.

This completed the FAF-zone level estimation process for the OOS retail sector.

6.2. SERVICES

The NAICS codes involved in this OOS component are listed in Table 6-3.

6.2.1. Data Sources for Services

6.2.1.1. 2017 Service Annual Survey Data

The Census conducts the Service Annual Survey³³ (SAS) to provide national estimates of annual revenues and expenses of establishments classified in select service industries. In particular, the NAICS-based table “Estimated Revenue for Employer and Nonemployer Firms: 2013 Through 2017” was used to establish national control totals for 2017 revenue totals for industries associated with the services sector, spanning 3- to 5-digit NAICS codes. The 2017

³³ <https://www.census.gov/data/tables/2017/econ/services/sas-naics.html>

SAS estimated a total of \$13.8 trillion in services receipts across NAICS codes 51-81, which is a 24.3 percent increase from 2012. The three bolded codes in Table 6-3, i.e., NAICS 7111, 7112 and 712, rely on different data sources and consequently their treatment in FAF5 estimation differs considerably compared to FAF4. This will be elaborated upon later.

Table 6-3. NAICS Codes Included under FAF5 OOS Services Component

NAICS Code	Description
51912	Libraries and archives
5322	Consumer goods rental
5324	Commercial and industrial machinery and equipment rental and leasing
562	Waste management and remediation services
6216	Home health care services
7111	Performing arts companies
7112	Spectator sports
712	Museums and Similar Institutions
7223	Special food services
8123	Dry cleaning and laundry services

6.2.1.2. 2017 Economic Census

As with the retail sector, information collected under the 2017 Economic Census was also used in the estimation process for the services OOS component. Census estimates a total of \$13.9 trillion in receipts was generated from the services sector (NAICS 51-81) in 2017. Receipts data associated with the services sector in all states and Washington, DC are published by the Census at an aggregated 2-digit NAICS level. However, receipts data for services subsectors and industries at the 3-digit to 5-digit NAICS levels are not always released for each geographic area, therefore requiring imputation based on national control totals obtained from the SAS.

6.2.1.3. County Business Patterns

As in the retail OOS estimation process, the 2017 CBP data was used to disaggregate state-level estimates to the FAF zones. It was also used to impute suppressed cells in the Economic Census based on the number of firms within each FAF zone corresponding to the appropriate NAICS codes.

6.2.1.4. Data Sources for NAICS 7111, 7112, and 712

In FAF5 estimation, it was decided that movements of NAICS codes 7111, 7112, and 712 would be more accurately represented with alternative datasets that allowed for the construction of OD networks and research on the value and/or weights associated with shipments. Specifically,

- For NAICS 7111 (performing arts), a list of the Top 200 North American concert tours in 2017 based on earned revenue was downloaded from Pollstar³⁴. From this list, the top 100 tours for which the full set of cities could be extracted from a reliable web resource were used to create a full network for the shipment of touring equipment.
- For NAICS 7112 (spectator sports), the 2017 season schedules for baseball (MLB), men's basketball (NBA), women's basketball (WNBA), pro football (NFL), college football

³⁴ https://www.pollstar.com/Chart/2018/01/2017YearEndTop200NorthAmericanTours_622.pdf

(NCAA Division I), hockey (NHL), men's soccer (MLS), and women's soccer (NWSL) were extracted from Sports Reference³⁵. In addition, the racing schedules for IndyCar, NASCAR Cup, NASCAR Trucks, and NASCAR Infinity leagues were extracting from Racing Reference³⁶.

- For NAICS 712 (museums), datasets were downloaded from the Institute of Museum and Library Services (IMLS) website³⁷. Using expert judgment based on web searchers, a subset of “major institutions” corresponding to specific museum types that would likely host valuable exhibits was derived so that a simulated network of shipments could be constructed; this process is elaborated upon in the Estimation section.

6.2.2. Assumptions for Intra-zonal Shipments

Analogous to the retail sector, the table “Estimated Revenue from Electronic Sources for Employer Firms: 2015 Through 2017” from the *2017 Service Annual Survey* data was used to generate estimates for the proportion of physical shipments in the services sector, and a set of assumptions on percent of services receipts involving truck were made for the FAF5 process. It is important to note that this SAS table does not have values for each NAICS code level, so the more disaggregated one available is used for the reductions. A summary is provided in Table 6-4, including corresponding SCTG codes (these do not differ from FAF4).

6.2.3. Assumptions for Inter-zonal Shipments

The main challenge of this process is to determine the value and/or weight being moved between cities for each involved industry category. The methodologies are described as follows.

Table 6-4. Assumptions for Intrazonal Services Shipments

NAICS Code	SAS Total Revenue (billion \$)	SAS E-Commerce (billion \$)	Delivery Truck Shipment (%)	SCTG Codes
51912	178.0	126.7	50	23, 29, 35
5322	159.9	28.0	50	35, 39
5324	159.9	28.0	100	34, 35, 38, 39
562*	---	---	50	24
6216	2,541.3	8.6	5	38
7223	929.1	72.8	33	03, 04, 05, 06, 07, 08
8123	111.3	5.8	20	30

*It is assumed that there is no e-commerce for waste services, so no reduction required.

6.2.3.1. Performing Arts (NAICS 7111)

After web searches for mentioned estimates of a particular performance art's touring equipment (which is assumed to be static throughout the duration of a tour), it was determined that an appropriate approximation for monetary value assigned to each shipment is half of the average per-show revenue. Thus, cumulatively, each directed “edge” of the concert network will have a distinct weight, which could be used for modeling future movements (as opposed to repeating the tedious process of building the network from scratch each year). Weights were then derived

³⁵ <https://www.sports-reference.com/>

³⁶ <https://www.racing-reference.info/>

³⁷ <https://www.imls.gov/research-evaluation/data-collection/museum-data-files>

using the value-to-weight ratios from 2017 CFS PUF commodity-mode pairs, with the mode assumed to be truck except in implausible cases (for which the mode was assumed to be truck-air).

6.2.3.2. *Spectator Sports (NAICS 7112)*

Similarly, web searches on the composition and average weight of transported sports equipment for baseball teams were conducted to produce a basis for determining the weights corresponding to each OD pair. The “total” weight was scaled depending on how the allotted active team roster compared to the 25-person standard for baseball teams (i.e., football is more, basketball is less, and so on) since no information could be found regarding other sports. Values were then derived using the value-to-weight ratios from 2017 CFS PUF commodity-mode pairs. Only ground shipments to/from airports and the site of each sporting match were included in the estimates since air movements involve either commercial or private jets and therefore fall outside the scope of FAF. In most cases, this yield intra-zonal movements, but with a handful of exceptions in most leagues (e.g., the FAF zone containing Cincinnati’s airport is in Kentucky while the stadiums are in Ohio).

For the four racing leagues, web searches pointed to the approximate values and weights for each class of vehicle, the number of drivers competing in each race, as well as the average total weight of additional equipment transported between matches. Values for non-vehicle goods were then derived using the value-to-weight ratios from 2017 CFS PUF commodity-mode pairs, with mode assumed to be truck unless implausible (for which truck-air is the assigned mode).

6.2.3.3. *Museums (NAICS 712)*

Two distinct movement networks were established based on institution category: the first one consists of major art museums, while the second one merged major natural history & science museums, science & technology museums and planetariums, and children’s museums; in the latter case, the assumption is that there is suitable overlap in exhibit type to anticipate potential shipments between pairs of institutions.

There is no comprehensive database detailing the composition of traveling exhibits. Hence some assumptions were made to create a simulated network of shipments that is transferable across years. After sifting through revenue and institutional information in the IMLS database and conducting supplementary web searches, a total of 313 art museums comprises the first network and 285 “other” museums comprise the second network. The weighted edges are then estimated as follows. First, it was assumed that the network will represent seasonal movements, so each museum was given a weight of four in a random sampling procedure that assigns destinations for each originating museum. Each shipment was then assigned a value that was drawn from the product of two triangular distributions: the first distribution was generated from the revenue information in the database in which the left endpoint is the 15th percentile, the right endpoint is the 85th percentile, and the mathematical mode is the median revenue value; the second distribution then acted as a random scaling factor to shrink the shipment value to a more reasonable number based on a small proportion of total museum assets that is anticipated to be non-stationary over the calendar year. Tonnages of the shipments involved were then calculated using the value-to-weight ratios obtained from 2017 CFS PUF commodity-mode pairs, with mode of transportation assumed to be truck unless implausible (for which truck-air is the assigned mode), and analogously scaled using an appropriate triangular distribution.

6.2.4. Estimation Approach for Services

6.2.4.1. SCTG Code Assignment

The impacted SCTG codes for each NAICS under this OOS are listed in Table 6-5, which also includes how the allocations were made across the commodities. Note that, for sporting leagues, a specific value was designed for SCTG 30 based on statistics found for baseball and scaled independently from the other impacted commodities using the ratio of team size across sports. To reflect expected heterogeneity across museum exhibits in the “other” category, two SCTG codes were randomly sampled with equal probability from the assumed set of possible goods, followed by a random fraction drawn from a triangular distribution that was assigned to be the value of the first sampled commodity, with the remaining value assigned to the second commodity.

Table 6-5. Assumptions for Interzonal Services Shipments

NAICS Code	SCTG Codes	Value/Weight Allocation
7111 Performing arts	33, 39, 40	Uniformly distributed
7112: Sports	30	—
	06, 07, 35, 38, 40, 43	0.025, 0.025, 0.250, 0.250, 0.400, 0.050
7112: Racing	24, 33, 34, 35, 36	0.09, 0.17, 0.17, 0.17, 0.40
712: Art museums	40	—
712: Other museums	29, 31, 32, 33, 34, 35, 36, 37, 39, 40	Random selection & allocation

6.2.4.2. Intra-zonal shipments

For the seven NAICS codes assumed to stay within the originating FAF zone, the estimation processes mirror that for retail:

- Finalize total values for each state;
- Apply the same simple even-share approach to allocate shipment values among SCTG commodities within a given NAICS code;
- Distribute to FAF zones in the given state based on 2017 CBP data for the relevant NAICS codes; and
- Impute shipment weights using 2017 CFS PUF data with the same hierarchy of shipment attribute categories.

This process results in an estimated \$144 billion with 37.5 million tons in shipment weights.

6.2.4.3. Inter-zonal shipments

The estimation procedure follows logically from the assumptions described above, but two challenges within this process should be mentioned. The first was accounting for “breaks” in schedules for NAICS 7111 and 7112, which would be respectively due to international touring legs and the structure of sporting league schedules that are not contained within a specific calendar year. This includes the addition of “home base” movements at the beginning or end of each season, especially in the case of racing. The second was adjusting point-of-entry/exit regions for transborder ground movements with Canadian origins or destinations. Thus, these processes yield \$1.45 billion in value of goods being moved for performing arts, \$5.77 billion for spectator sports, and \$3.62 billion for museums.

6.3. HOUSEHOLD/BUSINESS MOVES (HH&B)

Trucking services provided by the household and business (HH&B) moving industry is under the NAICS code “484210”, Used Household and Office Goods Moving. Note that NAICS code “4842” is covered under the 2017 CFS survey frame. However, as a shipper-based survey, shipments of household and business goods in the moving industry are not captured by the CFS. This is because businesses that engaged in these moving activities do not typically produce freight or warehousing services. Commodities covered under this HH&B OOS sector include SCTGs 35 (electronic equipment), 36 (motorcycles & bicycles), 37 (transportation equipment), 39 (furniture), 40 (sporting goods), and 43 (miscellaneous).

6.3.1. Data Sources for Household/Business Moves

The primary data sources for the HH&B OOS component are the ACS County-to-County Migration Files published by the Census, the Consumer Durable Goods Current-Cost Net Stock from the Bureau of Economic Analysis (BEA), and information obtained from various publications of the American Moving and Storage Association (AMSA).

6.3.1.1. 2013-2017 ACS County-to-County Migration Files

The ACS program combines consecutive yearly datasets to increase the sample size and provide reliable estimates for smaller geographic areas (e.g., county). For this HH&B OOS component, the 2013-2017 release of ACS county-level migration data³⁸ was used. The data provided in the ACS county-to-county (including intra-county) flow files include county of current residence, county of residence 1 year ago, and the number of movers between the two years. County-level total population and total number of housing units in 2017 from the ACS were used in FAF5 to estimate average household size (i.e., number of persons in a household) for each FAF zone.

6.3.1.2. BEA Consumer Durable Goods, Current-Cost Net Stock

The BEA publishes national statistics on both fixed assets and consumer durables³⁹. While fixed assets cover items that are used continuously in processes of production for an extended period, consumer durable goods are generally defined as tangible products that can be stored or inventoried that have an average life of three or more years. The BEA's Current-Cost Net Stock of consumer durable goods⁴⁰ statistics for 2017 was used to identify commodities associated with moves and the value of goods being moved in 2017.

6.3.1.3. AMSA Published Statistics

Statistics released online by AMSA⁴¹ were used to adjust estimates produced from ACS in order to capture business moves. According to the available 2018 statistics, AMSA stated that

³⁸ 2013-2017 ACS County-to-County Migration Files Documentation, U. S. Census Bureau, accessed March 2020, <https://www.census.gov/topics/population/migration/guidance/county-to-county-migration-flows/2013-2017.html>

³⁹ “Detailed Data for Fixed Assets and Consumer Durable Goods,” National Economic Accounts, Bureau of Economic Analysis, U.S. Department of Commerce, see <https://www.bea.gov/national/FA2004/Details/Index.html>.

⁴⁰ Table 8.1 Current-Cost Net Stock of Consumer Durable Goods, Fixed Assets Accounts Tables, National Data, <https://www.bea.gov/iTable/iTable.cfm?ReqID=10&step=1#reqid=10&step=1&isuri=1>, Bureau of Economic Analysis, accessed March 2020.

⁴¹ American Moving and Storage Association, About Our Industry, About AMSA, <https://www.moving.org/newsroom/data-research/about-our-industry/>, accessed March 2020.

shipments from “corporate and other federal government” moves accounted for about 39 percent of total - goods movements that occurred in that year. Furthermore, AMSA reported that about 47 percent of the interstate household goods moves were carried out by consumer themselves (i.e., not by professional movers or by use of rental truck). These percentages were used to adjust ACS-based estimates to account for business moves as well as to remove unassisted moves.

6.3.1.4. CFS Value-to-Weight Factors

To estimate commodity weights for the HH&B goods, value-to-weight ratios calculated from the 2017 CFS PUF data, considering domestic shipments by commodity, were applied. Since most HH&B goods moved are typically “used” items, as compared to “new” CFS goods, a simple depreciation rate of 30 percent (as used in FAF4) was applied to adjust the CFS-based value-to-weight factors for estimating the associated weights of HH&B goods.

6.3.2. General Assumptions and Assignment of Commodities to SCTG

In addition to the depreciation rate of value-to-weight factors discussed above, assumptions were applied in assigning commodity codes to the associated durable goods and in determining the share of specific commodities being moved by truck (versus items carried by the household members or left with the “old” house). It is assumed that only households with 3 or more bedrooms, as obtained from the ACS county-to-county migration data, fall under the HH&B intra-county moves. Moves of households with 2 or fewer bedrooms are assumed to be done by consumers themselves unassisted.

The assignment of the commodity code was done by examining the types of consumer durable goods specified in the BEA Current-Cost Net Stock table. This process assigned consumer durable goods items into the five SCTG codes that were involved in this OOS segment, including codes 35 (electronic equipment), 36 (motorcycles & bicycles), 37 (transportation equipment), 39 (furniture), 40 (sporting goods), and 43 (miscellaneous).

6.3.3. Estimation Process for Household/Business Moves

The estimation process associated with flows for HH&B moves is straightforward. A brief description of the process is given below.

6.3.3.1. Estimating Total Number of Moves

The ACS county-to-county migration flows (including intra-county and inter-county moves), measured in population, are aggregated to the FAF zone level and converted to number of households by applying the average household size factors generated using Census population and housing unit data for each zone. It is assumed that each household made one move as a whole. Under this assumption, the number of migrated households would be treated as equivalent to the number of “moves” or “OD flows”. The number of FAF zone-level moves is adjusted “upward” to include moves associated with “corporate and other federal government” and “downward” to exclude consumer non-assisted moves. The national averages taken from AMSA, as discussed above, were applied to all regions.

6.3.3.2. Estimating Total Value per Move by Commodity Code

The national total reported in the 2017 BEA Current-Cost Net Stock can be divided by the total number of households to derive a per-household value for each of the 6 impacted commodity codes (at the national level) after adjustment to eliminate items not likely to be transported in a moving truck. These per-household values were then multiplied by the total number of mover-households from each region to obtain values of SCTGs associated with the HH&B OOS. Finally, the CFS-based value-to-weight factors described above were applied to the values to estimate shipment weights by SCTG.

National averages of per-household value by SCTG were applied to all moves regardless of the possible existence of regional differences. This is a current data availability limitation that could be further examined and improved upon in the future, if new data sources became available. Based on the FAF5 process, it was estimated that truck shipments worth over \$276 million, with about 58 million tons in weights, were generated from the HH&B OOS sector in 2017. Approximately two-thirds (67.0 percent) of the total value of these shipments are for common household items of SCTG 39 (furniture) and 35 (electronic).

7. ESTIMATION OF OOS CRUDE PETROLEUM

The transportation of crude petroleum is an OOS movement for the CFS since NAICS 211: Oil and Gas Extraction is not included under its survey frame. In this OOS segment, only one commodity code is impacted, which is SCTG 16-Crude petroleum. This OOS includes shipments from field or marine terminals and international pipelines to the refineries or long-term storage facilities. Foreign trade crude petroleum shipments (i.e., imports and exports), in addition to domestic crude petroleum movements, are also addressed in this chapter.

Crude petroleum is one of the major OOS components for the CFS. Table 7-1 presents volumes for domestic production and foreign traded crude petroleum in 2012, 2016, and 2017. Domestic production of crude petroleum has increased since 2012, with a 43 percent increase in volume from 2012 to 2017. Conversely, the imported volume of crude petroleum has decreased since 2012, with a 7 percent decrease in volume from 2012 to 2017. The exported volume of crude petroleum has increased significantly in recent years. This is due to the enactment of the Consolidated Appropriations Act⁴² in December 2015, which authorized the export of U.S. crude petroleum without a license. However, exports to embargoed or sanctioned countries still require authorization. Before December 2015, crude petroleum exports were restricted to (1) crude petroleum derived from fields under the State waters of Cook Inlet of Alaska; (2) Alaskan North Slope crude petroleum; (3) certain domestically produced crude petroleum destined for Canada; (4) shipments to U.S. territories; and (5) California crude petroleum to Pacific Rim countries.

Table 7-1. Volume of Crude Petroleum (in million barrels)

Type	2012	2016	2017
Domestic production	2,387	3,240	3,421
Imports	3,121	2,873	2,909
Exports	25	216	423

Data source: Energy Information Administration (EIA)⁴³

7.1. DATA SOURCES FOR CRUDE PETROLEUM

As in FAF4, basic information on crude petroleum production, imports, exports, and related activities at refineries was estimated using data collected by the EIA. In addition to EIA-published data, several other data sources were utilized in estimating crude petroleum movements. These data sources are discussed briefly below. Note that geographic regions for the EIA data are typically in the Petroleum Administration for Defense Districts (PADDs). Figure 7-1 shows a map of the PADDs as defined by the EIA.

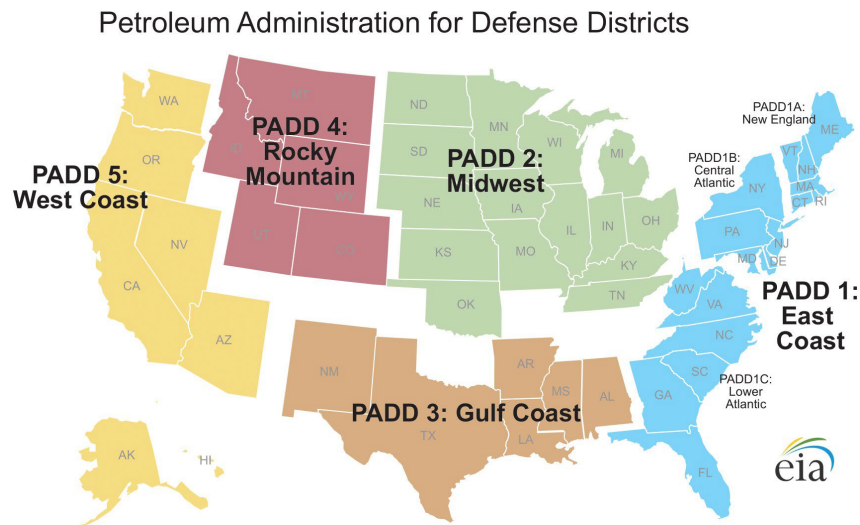
7.1.1. EIA Data

Several EIA-published tables provide the fundamental statistics needed in estimating 2017 crude petroleum volume and their flows.

⁴² Consolidated Appropriations Act, 2016: <https://www.congress.gov/bill/114th-congress/house-bill/2029/text>

⁴³ <https://www.eia.gov/petroleum/data.php>

Figure 7-1. Definition of Petroleum Administration for Defense Districts (PADDs)⁴⁴



7.1.1.1. Crude Petroleum Production

The *Crude Oil Production* table contains the annual production of crude petroleum, in thousands of barrels by state and PADD⁴⁵. The statistics published in this EIA table are based on amounts reported from states, which includes production estimates by state as well as summarized at the PADD and national levels. This table is used in the FAF5 processing as the primary reference for state-level production of crude petroleum. According to this EIA table, a total of 32 states produced crude petroleum in 2017.

7.1.1.2. Movements by Mode between PADDs

Annual volumes of crude petroleum movements, measured in thousand barrels, by transportation mode between PADDs, are published by the EIA⁴⁶. These modes include pipeline, tanker, barge, and rail. Statistics published in this table are based on information collected from state-reported EIA-813 (*Monthly Crude Oil Report*) form. This information is used as the control total for PADD-to-PADD movements during the estimation of FAF zone level flows.

7.1.1.3. Movements by Rail

The *Movements of Crude Oil by Rail* table contains rail crude petroleum movements as well as crude movement regions⁴⁷. This table provides detailed movements among PADDs as well as trade between PADDs and Canada. It is used as control totals for estimations of domestic flows, U.S. exports to Canada, and U.S. imports from Canada by rail.

⁴⁴ EIA, PADD map: <https://www.eia.gov/petroleum/marketing/monthly/pdf/paddmap.pdf>

⁴⁵ Table "Crude Oil Production", EIA: https://www.eia.gov/dnav/pet/pet_crd_crpdn_adc_mbb1_a.htm

⁴⁶ Table "Movements by Pipeline, Tanker, Barge and Rail between PAD Districts", EIA: https://www.eia.gov/dnav/pet/pet_move_ptb_a_EPC0_TNR_mbb1_a.htm

⁴⁷ Table "Movements of Crude Oil and Selected Products by Rail", EIA: https://www.eia.gov/dnav/pet/pet_move_railNA_a_EPC0_RAIL_mbb1_a.htm

7.1.1.4. Refinery Net Input

The EIA also publishes data on refinery net input for crude petroleum by PADDs and refining regions (sub-PADD level)⁴⁸. These statistics represent total crude petroleum input to crude petroleum distillation units and other refinery processing units in each region.

7.1.1.5. Refinery Capacity

Capacities of operable petroleum refineries can be found in EIA's Refinery Capacity Data⁴⁹. Individual refinery's processing capacity is provided in this dataset. These capacities are used to determine the shares of crude petroleum flow at the FAF zone level.

7.1.1.6. Refinery Receipts by Mode

The *Refinery Receipts of Crude Oil by Method of Transportation* table contains annual volumes of both domestic and foreign crude petroleum receipts by the refineries, measured in thousand barrels, by transportation mode and PADD⁵⁰. These modes include pipeline, tanker, barge, rail, and truck. Information obtained from this table is used to estimate the control total of intra-PADD (i.e., within PADD) movements.

7.1.1.7. Domestic Crude Oil First Purchase Prices by Area

The *Domestic Crude Oil First Purchase Prices by Area* table contains average annual domestic crude oil first purchase prices (dollars per barrel) by PADD⁵¹. These average prices are used to determine the values of domestic and exported crude petroleum.

7.1.1.8. Landed Costs of Imported Crude by Area

The *Landed Costs of Imported Crude by Area* table contains the average annual dollar per barrel price of imported crude petroleum at the port of discharge⁵². The average dollar per barrel price is used to determine the values of imported crude petroleum.

7.1.1.9. Company-Level Imports

This EIA database⁵³ contains monthly statistics regarding imports of crude petroleum and petroleum products at the company level. Specific information provided in this database includes importing company name and country, product name, port of entry location (city, state, PADD), import quantity (in thousand barrels), and other information.

⁴⁸ Table "Refinery Net Input", EIA, available at:

https://www.eia.gov/dnav/pet/pet_pnp_inpt2_a_epc0_YIY_mbbl_a.htm

⁴⁹ Refinery Capacity Data by individual refinery as of January 1, 2017, EIA:

<https://www.eia.gov/petroleum/refinerycapacity/archive/2017/refcap17.xls>

⁵⁰ Refinery Receipts of Crude Oil by Method of Transportation, EIA:

[https://www.eia.gov/dnav/pet/pet_pnp_caprec_a_\(na\)_8r00_Mbbl_a.htm](https://www.eia.gov/dnav/pet/pet_pnp_caprec_a_(na)_8r00_Mbbl_a.htm)

⁵¹ Table "Domestic Crude Oil First Purchase Prices by Area", EIA:

https://www.eia.gov/dnav/pet/pet_pri_dfp1_k_a.htm

⁵² Table "Landed Costs of Imported Crude by Area", EIA: https://www.eia.gov/dnav/pet/pet_pri_land1_k_a.htm

⁵³ Company Level Imports 2017, EIA:

<https://www.eia.gov/petroleum/imports/companylevel/archive/2017/data/impa17d.xlsx>

7.1.1.10. Exports by Destination Country and PADD

The *PAD District Exports by Destination* table contains annual volumes of exported crude petroleum, measured in thousand barrels, from each PADD by destination country⁵⁴. Information from this table is used as the control total for exported crude petroleum.

7.1.2. Other Data Sources

7.1.2.1. State Crude Petroleum Production Data

Several states reported annual production of crude petroleum at the county level for the year 2017. Some examples of these agencies include Mississippi's *State Oil and Gas Board*⁵⁵ and Nebraska's *Oil and Gas Conservation Commission*⁵⁶. In addition to state reports, three online databases provide some of the crude petroleum production data that were used to disaggregate state-level production to the FAF zone level. These databases are ShaleXP⁵⁷, DrillingEdge⁵⁸, and UpstreamDB⁵⁹, and each of them reports county level crude petroleum production data of several (but not all) states in different years.

7.1.2.2. Carload Waybill Sample

The Surface Transportation Board (STB) publishes and maintains the Carload Waybill Sample⁶⁰ database. This database is a stratified sample of carload waybills for all U.S. rail traffic submitted by those rail carriers terminating 4,500 or more revenue carloads annually. The 2017 *Carload Waybill Sample* database captured over 670 thousand waybills reported electronically in 2017. To protect sensitive shipping and revenue information of rail companies reporting the Waybill, STB provides a public-use version of the Waybill data that contains aggregated data. Additionally, STB provides confidential Waybill data that contains more detailed information but requires a special STB approval on its access.

7.1.2.3. Waterborne Commerce Data

The Waterborne Commerce Statistics Center (WCSC), a part of the U.S. Army Corps of Engineers, provides state to state water movement in tonnages for crude petroleum⁶¹. The WCSC also provides a dataset for foreign cargo trips⁶²; this dataset contains information on exported and imported crude tonnages moved by water along with their ports and country of import or export.

⁵⁴ Table "PAD District Exports by Destination", EIA:

https://www.eia.gov/dnav/pet/pet_move_expcp_a1_Z00_EPC0_EEX_mbb1_a.htm

⁵⁵ <https://www.ogb.state.ms.us/MSOGBDocuments/ProductionBooks/2017/Annual/2017AnnualProdReport.pdf>

⁵⁶ <http://www.nogcc.ne.gov/Publications/ANNSUM2017.pdf>

⁵⁷ <https://www.shalexp.com>

⁵⁸ <http://www.drillingedge.com>

⁵⁹ <https://www.upstreamdb.com>

⁶⁰ For access of Carload Waybill Sample data visit Surface Transportation Board website at:

<https://prod.stb.gov/reports-data/waybill/> or access the 2017 documentation at: <https://prod.stb.gov/wp-content/uploads/2017-STB-Waybill-Reference-Guide.pdf>

⁶¹ State to State and Region to Region tonnages for 14 major commodity groups-2017, WCSC, available at:

<https://publibrary.planusace.us/document/b1a0e3e1-61ec-4da2-b207-22846c14ed4a>

⁶² Foreign Cargo Inbound and Outbound Calendar Year 2017, WCSC, available at:

<https://publibrary.planusace.us/#/document/7022fdb0-12ab-462a-fff7-c7aabe7979c1>

7.1.2.4. Census Special Tabulation of 2017 USA Trade Online (USATO) Data

At the request of BTS, Census provided special tabulations that have more details for 2017 foreign trade shipments than what is published publicly. Please refer to Chapter 9 of this report for more details regarding these foreign trade datasets.

7.2. ESTIMATION PROCESS

7.2.1. Flows of Domestic Crude Petroleum

The EIA data provided the control totals for estimating domestic movements of crude petroleum at the PADD level. These PADD-level crude petroleum movements need to be disaggregated to their associated FAF zone level. The PADD-to-PADD crude petroleum flows were constructed in different matrices depending on their modes of transportation (i.e., pipeline, rail, water, and truck). This is because crude petroleum flows have different mode-specific OD patterns. The basic domestic flow estimation process is described in the following steps:

- Step 1: Estimate PADD-level movements by mode. First, obtain movements by mode between PADDs (inter-PADD movements) using information the EIA “Movements between PAD Districts” table. Then, estimate the intra-PADD movements by subtracting the inter-PADD movements from the crude petroleum amounts of refinery receipts for each transportation mode.
- Step 2: Estimate crude petroleum production at the FAF zone level. First, adjust county-level production data, obtained from online databases and state agencies, to match the state production control totals from EIA, if necessary. Then, aggregate county-level production to generate FAF zone level crude petroleum production estimates.
- Step 3: Estimate refinery input (i.e., attraction) at FAF zone level. The refinery input data provided by the EIA is at the state level. To estimate the FAF zone level attraction, the state level attraction was distributed proportionally to the operational capacity of each crude petroleum refinery and then aggregated to the corresponding FAF zone level. The assignment of refineries to the associated FAF zone was done using the spatial join feature of a GIS tool.
- Step 4: Apply the IPF procedure to estimate flows at the FAF zone level. This procedure used PADD-to-PADD movements as the control totals. An initial OD flow matrix at the corresponding FAF zone level was generated by considering the possibility of flows between the FAF zones. Specifically, when there is a possibility to have a flow between origin and destination FAF zones, the corresponding cell in the initial flow matrix is given a value of 1. Otherwise, the cell is assigned with a value of 0. For a PADD-PADD pair, using the initial flow matrix and their corresponding production and attraction marginal totals, a IPF procedure was applied. This process was then repeated for all PADD-PADD pairs by mode to obtain final estimates.

7.2.2. Flows of Imported Crude Petroleum

The EIA company-level imports data file provided fundamental details for estimating imported crude petroleum flows. However, this dataset does not have domestic mode or foreign mode information. The procedure for estimating these modes, and other shipment information, is provided below.

7.2.2.1. Imports by Rail (Trans-Border Only)

The 2017 STB Waybill data was used to derive patterns for imported crude petroleum movements by rail. According to the Waybill data, crude petroleum imported by rail occurred only from Canada in 2017. The foreign origin, domestic origin, and domestic destination for each imported movement were determined using this data. The derived waybill-based patterns by tonnage were used to distribute EIA-based control totals. Specifically, shares estimated from Waybill for each Canada-PADD pair are applied to EIA control totals to get estimated tonnages at the FAF zone level.

7.2.2.2. Imports by Other Modes

To estimate flows of imported crude petroleum by other modes (i.e., pipeline, water, and truck), amount on imports by rail were subtracted from EIA company-level imports data first. Then, the assignment of foreign-mode and domestic-mode for remaining imports was done. The foreign mode was determined by reviewing foreign country and port location for countries other than Canada. For these countries, the mode used for imported crude petroleum was assumed to be water. For imports from Canada, the shipments can be transported via pipeline, water, or truck, depending on port location's access to the corresponding transportation network. The domestic mode for an import is the mode that can serve both domestic origin (port of entry) and its domestic destination (e.g., facility location). After the mode assignment, the company-level import totals were aggregated into FAF zones. Lastly, the value of imported crude petroleum was estimated using the unit landed costs of imported crude petroleum obtained from EIA.

7.2.3. Flows of Exported Crude Petroleum

The EIA data, on PADD level exports by destination foreign region, provided the control totals for estimating exported crude petroleum for FAF5. However, EIA data does not provide any information on the foreign mode, domestic origin, port of exit (i.e., domestic destination) at the FAF zone level, or the mode used for the domestic move of exports. Therefore, Census provided Special Tabulation of 2017 USATO data and the WCSC data were utilized in this FAF5 process. The basic procedure is described below.

7.2.3.1. Exports by Water

Water movements for exported crude petroleum from WCSC data provided information on domestic destination (port of exit) and tonnage by foreign region using water as the foreign mode. The WCSC-based patterns by tonnage were used in FAF5 to distribute EIA-based control totals. The domestic origins of these exported flows were estimated based on the USATO data (methodological details are described in below subsection).

7.2.3.2. Exports by Other Modes

The USATO data provided information for foreign mode, commodity, domestic state of origin, and domestic destination (port of exit) at FAF zone. Since it provides domestic origins of exports at the state level, exports from FAF zone level geography needed to be estimated. In the cases where an entire state represents a FAF-zone, the assignment of domestic FAF-zone is straightforward. In the cases where there are several FAF zones in a state, the domestic origin is assumed to be in the top two FAF zones in terms of crude petroleum production. The domestic mode was selected as the mode that can serve both domestic origin and domestic destination. Then the EIA-based control totals at the PADD level were distributed based on

shares of shipment value (in \$) from the USATO to get estimated shipment tonnages at the FAF-zone level. Lastly, the value of exported crude petroleum was calculated using crude oil first purchase price data from EIA.

Note that both domestic and foreign trade crude petroleum movements, as produced from this mostly EIA-based estimation process, were included in the final FAF5 database. To avoid double counting, SCTG 16 estimates generated from the foreign trade OOS process (see Chapter 9) were removed before combining estimates from all OOS components with the domestic CFS matrix to form the final FAF5 ODCM flow matrix.

8. ESTIMATION OF OOS NATURAL GAS

Similar to crude petroleum, because the CFS does not cover shipments from industries in the NAICS 211: *Oil and Gas Extraction*, movement of natural gas is an OOS area for the CFS. Only SCTG 19 – *Other Coal and Petroleum Products, not elsewhere classified* is impacted under this OOS component. Note that CFS does capture some SCTG 19 shipments, which are not natural gas. Specifically, this OOS sector includes natural gas movements between processing plants, refineries, and terminals. This chapter summarizes the estimation methodology for domestic & foreign trade natural gas movements at the FAF zone level and describes the assumptions and data sources used in the FAF5 estimation process.

8.1. DATA SOURCES

As in FAF4, basic information on domestic, imports, and exports of natural gas flows was estimated using data collected by the EIA. In addition to EIA-published data, several other data sources were utilized in estimating natural gas movements. Since FAF5 has a base year of 2017, all data used in this OOS study were for the year 2017. Data sources used under this FAF5 estimation process are briefly discussed below.

8.1.1. EIA Data

8.1.1.1. EIA Interstate Natural Gas Flows

The EIA publication, *2017 Natural Gas Annual*, provides information on interstate flows of natural gas⁶³. The state-level natural gas flows are disaggregated to FAF zone level movements based on methodologies summarized in Section 8.2.

8.1.1.2. EIA Natural Gas Production Data (Domestic)

The EIA publication, *Natural Gas Annual*, provides information on the supply and disposition of natural gas. For estimating FAF zone level natural gas flows, specific information from Table 12⁶⁴ of the *Natural Gas Annual* report, entitled *Interstate movements and movements across U.S. borders of natural gas by state, 2017* is used. This information, along with the disposition by state data, is used to derive control totals for intrastate natural gas movements. In addition, the information on dry production and withdrawals from underground storage by individual states is also available from this report.

8.1.1.3. EIA Natural Gas Consumption by End-Use

This EIA data series provides annual statistics⁶⁵ on total “volumes delivered to consumers” by the state for end-use sectors of residential, commercial, industrial, vehicle fuel, and electric power plants. In 2017, among the five end-use sectors, natural gas being consumed for vehicle fuel accounted for only 0.1 percent. Due to the small contribution by this end-use sector, and the limitation of available data sources for FAF zone level estimation, only residential, commercial, industrial, and electric power sectors were considered under the FAF5 estimation process.

⁶³ Interstate natural gas flows in 2017: https://www.eia.gov/naturalgas/annual/archive/2017/csv/t2017_12.csv

⁶⁴ Natural Gas Annual: http://www.eia.gov/naturalgas/annual/archive/2017/pdf/table_012.pdf.

⁶⁵ Natural Gas Consumption by End Use: http://www.eia.gov/dnav/ng/ng_cons_sum_a_EPG0_VC0_mmcf_a.htm.

8.1.1.4. EIA Imports by Point of Entry

The major data source used in estimating imported natural gas flows is the *U.S. Natural Gas Imports by Point of Entry*, which is available from the EIA website⁶⁶, this information can also be obtained from Table 9 “Summary of U.S. natural gas imports by point of entry, 2008 – 2017” of the *Natural Gas Annual* report published by the EIA. This data series provides volume (in million cubic feet) and price for natural gas, liquefied natural gas (LNG), and compressed natural gas (CNG) transported by pipeline. In 2017, 94.4 percent of the U.S. natural gas imports were from Canada and Mexico by pipeline; the imports from other countries were transported by LNG tankers.

8.1.1.5. EIA Exports by Point of Exit

A data source for estimating exported natural gas flows was obtained from the U.S. Natural Gas Exports and Re-Exports by Point of Exit website⁶⁷, as well as available from Table 11 of the *Natural Gas Annual* report, entitled “Summary of U.S. Natural Gas Exports by Point of Exit, 2008-2017.”

8.1.2. Other Data Sources

8.1.2.1. ShaleXP County Production Data

ShaleXP⁶⁸ is a public website that provides county-level natural gas production data for 20 states in 2017. State production totals derived from ShaleXP data were compared with EIA state production totals and for states with comparable production totals between the two sources (± 6 percent), the county-level production shares were used to disaggregate EIA state production totals to FAF zone level.

8.1.2.2. State Reports

Several states publish natural gas production data on their websites. Examples of these state reports include Colorado⁶⁹, Idaho⁷⁰, and Kansas⁷¹. Similar to ShaleXP, County-level production shares from these state reports were used to disaggregate EIA production totals to FAF zone level estimates.

8.1.2.3. HIFLD Natural Gas Receipt and Delivery Point

The Homeland Infrastructure Foundation-Level Data (HIFLD) publishes the *Natural Gas Receipt Delivery Points*⁷² database online. This database contains geographic information representing locations on interstate natural gas pipelines where natural gas gathering systems connect (receipt) to the pipeline, or where natural gas local distribution systems and other end-users (delivery) connect to the pipeline within North America. This dataset also includes the annual

⁶⁶ U.S. Natural Gas Imports by Point of Entry, https://www.eia.gov/naturalgas/annual/archive/2017/csv/t2017_09.csv

⁶⁷ U.S. Natural Gas Exports and Re-exports by Point of Exit, http://www.eia.gov/dnav/ng/ng_move_poe2_a_EPG0_ENP_Mmcf_a.htm.

⁶⁸ ShaleXP county production data: <https://www.shalexp.com/>

⁶⁹ Colorado state report: <https://cogcc.state.co.us/COGCCReports/production.aspx?id=MonthlyGasProdByCounty>

⁷⁰ Idaho state report: https://ogcc.idaho.gov/wp-content/uploads/sites/88/2017/12/ar-2017-ogcc_pages.pdf

⁷¹ Kansas state reports: <http://www.kgs.ku.edu/PRS/petro/interactive.html> & https://ogcc.idaho.gov/wp-content/uploads/sites/88/2018/12/ar-2018-ogcc_pages.pdf

⁷² HIFLD Natural Gas Receipt Delivery Points: https://hifld-geopatform.opendata.arcgis.com/datasets/6e01edc178ea4b7e9cec874e206248a2_0

average of daily scheduled capacity of natural gas production, which was used to disaggregate state-level natural gas productions to FAF zone level for states that do not have data available either from state reports or ShaleXP.

8.1.2.4. Census County Business Patterns

The Census published CBP⁷³ is an annual series that provides county-level economic data by industry. It contains information on the number of establishments, employment size, and annual payroll. This supplemental dataset is mainly utilized in disaggregation models to distribute natural gas consumption estimates from state level to FAF zone level. Other data sources considered in the disaggregation process include U.S. Census County-level population, and EIA data on electric generating units (EIA-923)⁷⁴, which provides information on fuel consumption at the power plant.

8.2. ESTIMATION PROCEDURES FOR DOMESTIC NATURAL GAS FLOWS

EIA provides state level domestic natural gas shipments⁷⁵. Modeling approaches were developed to disaggregate the state level movements to the FAF zone level. To achieve this objective, natural gas production and consumption (attraction) data were disaggregated to county level movements then aggregated to produce FAF zone level estimates. The process for estimating domestic FAF zone level movement follows the three steps listed below:

Step 1. Disaggregate state level production to FAF zone level production.

As stated earlier, EIA publishes state level total production data⁷⁶. The state-level total production volume was disaggregated to county production and then aggregated to FAF zone level production using three different data sources: state reports, ShaleXP, and HIFLD Natural Gas Receipt and Delivery Point.

- For states that publish county productions, the EIA state total production was proportionally assigned to counties within the state based on state reported county productions.
- If state report was not available, county production obtained from ShaleXP was used to proportionally distributed the EIA state total production to each county within a state.
- For the remaining states, state production totals were proportionally distributed to counties based on the scheduled capacity obtained from HIFLD Natural Gas Receipt and Delivery Point.

The county production was then aggregated to corresponding FAF zones to obtain FAF zone level production.

Step 2. Disaggregate state level consumption to FAF zone level consumption.

The disaggregation of natural gas consumptions was conducted in a similar manner as production disaggregation. However, the proportional distributions at FAF zone level were obtained for four different end-user sectors: residential (RS), commercial

⁷³ U.S. Census County Business Patterns: <https://www.census.gov/data/datasets/2017/econ/cbp/2017-cbp.html>.

⁷⁴ EIA-923 detailed data can be found at: <https://www.eia.gov/electricity/data/eia923/>

⁷⁵ Interstate natural gas flows in 2017: https://www.eia.gov/naturalgas/annual/archive/2017/csv/t2017_12.csv

⁷⁶ State level production in 2017: https://www.eia.gov/naturalgas/annual/archive/2017/csv/t2017_03.csv

(CS), industrial (IN), and electric power (EU). The state-level natural gas consumptions by end-user types can be directly obtained from the EIA website⁷⁷.

- For residential (RS) use type, Census county population⁷⁸ was used to disaggregate the natural gas consumption by state into county level.
- For commercial (CS) and industrial (IN) end-uses, payroll information from relevant businesses in the 2017 CBP⁷⁹ data were used to disaggregate the natural gas consumption for the given state into its associated county consumptions.
- For electric power (EU) end use, the natural gas consumption quantity in sheet *Page 5 Fuel Receipts and Costs* from *EIA-923 Monthly Generation and Fuel Consumption Time Series File* data⁸⁰ was used for the disaggregation to FAF zone. Similarly, the total consumption is disaggregated based on the total consumption in all power plants within each county. The power plant location data is given in EIA – layer map⁸¹.

The county level consumptions by each end-user sector were then aggregated to obtain FAF zone level consumption.

Step 3. Apply IPF process to estimate FAF zone to FAF zone movement.

The IPF procedure is used for adjusting data cells so that they sum up to known marginal totals on both column wise and row wise. Details of the IPF process can be found in Section 3.2.3.

The natural gas production, consumption and movement data reported by EIA use million cubic feet as the unit. To obtain the shipments in tons, the unit of million cubic feet were converted based on the following factors:

- 1 million metric tons LNG = 48.7 billion cubic feet NG⁸²
- 1 metric ton = 1.10231 short tons⁸³
- 1 million cubic feet NG = $(1 / 48.7) \times 1.10231 = 0.022635$ thousand short tons

EIA's table *Average City Gate Price of Natural Gas in the United States, 2008-2017*⁸⁴ provides the average natural gas prices at the state level and it was used to further estimate the dollar values for these natural gas movements.

Domestic shipment modes were identified based on Figure 14 in EIA's *Natural Gas Annual* (NGA)⁸⁵ report. As indicated in this figure, over 99.9 percent of domestic natural gas movements are transported by pipeline. Therefore, all domestic natural gas shipments were assumed be moved by pipeline.

⁷⁷ The 2017 state-level natural gas consumption by end-use can be found at:

https://www.eia.gov/naturalgas/annual/archive/2017/csv/t2017_15.csv

⁷⁸ <https://www.census.gov/data/tables/time-series/demo/popest/2010s-counties-total.html>

⁷⁹ County Business Pattern: <https://www.census.gov/data/datasets/2017/econ/cbp/2017-cbp.html>

⁸⁰ Electric power plant https://www.eia.gov/electricity/data/eia923/archive/xls/f923_2017.zip

⁸¹ https://www.eia.gov/maps/map_data/PowerPlants_US_EIA.zip

⁸² The information of unit conversion from cubic feet to metric tons can be found at:

<https://www.extension.iastate.edu/agdm/wholefarm/pdf/c6-89.pdf>

⁸³ The information of unit conversion from metric tons to short tons can be found from the EIA website at:

<https://www.eia.gov/tools/faqs/faq.php?id=7&t=2>

⁸⁴ The EIA table *Average City Gate Price of Natural Gas in the United States, 2008-2017* can be obtained from the link, https://www.eia.gov/naturalgas/annual/archive/2017/csv/t2017_23.csv

⁸⁵ 2017 EIA Natural Gas Annual: <https://www.eia.gov/naturalgas/annual/pdf/nga17.pdf>

8.3. ESTIMATION PROCEDURES FOR FLOWS OF NATURAL GAS IMPORTS & EXPORTS

8.3.1. Imports

For the imports, data obtained from EIA's *Natural Gas Imports by Point of Entry*⁸⁶ table were used. The table provides natural gas import flows between foreign origins and point of entries with associated natural gas forms, from which the foreign mode can be determined as described in Section 8.3.3. The information on domestic mode and domestic destination are not given in the table. Two processing steps were performed to estimate natural gas import destination. These steps are briefly described below.

- Step 1.** Locations of entry points as obtained from EIA's *Natural Gas Imports by Point of Entry* data were assigned to their associated FAF zones.
- Step 2.** Determine whether domestic flows from the FAF zone level port of entry were in the set produced under Section 8.2.
 - a) If yes, then the amount of imported natural gas was distributed based on the share of flows from the port of entry to each possible destination FAF zone.
 - b) If no, then three destination counties were selected based on distances from the port of entry to potential destination counties that have natural gas consumptions. The amount of imported natural gas from the given port of entry was distributed to the selected counties based on their consumption shares generated during Step 2 in Section 8.2 and then aggregated to FAF zone level.

At this point, the process of estimating volume of imported natural gas flows between port of entries and destination FAF zones is completed. The process to estimate foreign and domestic mode is summarized in Section 8.3.3 below.

8.3.2. Exports

For the exports, data obtained from EIA's *Natural Gas Exports by Point of Exit*⁸⁷ table were used. The table provides natural gas movements between port of exits and foreign destinations with associated natural gas forms, from which the foreign mode was determined as indicated in Section 8.3.3. Domestic origin, foreign mode and domestic mode are not given in the table. The process for estimating exported natural gas origin is similar to the import process discussed above.

- Step 1.** The locations of all ports of exit, obtained from the abovementioned EIA table, were assigned to their corresponding FAF zones.
- Step 2.** Determine whether domestic flows to the FAF zone level port of exit were in the set produced under Section 8.2.
 - a) If yes, then the amount of exported natural gas was distributed based on the shares of domestic flows between potential origin FAF zones and the port of exit.
 - b) If no, then three origin counties that have natural gas productions were selected based on their distances to the port of exit. The amount of exported natural gas

⁸⁶ U.S. Natural Gas Imports by Point of Entry, https://www.eia.gov/naturalgas/annual/archive/2017/csv/t2017_09.csv

⁸⁷ U.S. Natural Gas Exports by Point of Exit, https://www.eia.gov/naturalgas/annual/archive/2017/csv/t2017_11.csv

being sent to the given port of exit was distributed to the selected origin by their production shares generated during Step 1 in Section 8.2 and then aggregated to FAF zone level.

At this point, the process of estimating volume of exported natural gas flows between origin FAF zones and port of exit is completed. The process to estimate foreign and domestic mode is summarized in Section 8.3.3 below.

8.3.3. Mode of Transportation on Imports & Exports

Import and export foreign modes were identified based on information found in Figure 14 and Tables 8 to 11 in EIA's *Natural Gas Annual* (NGA)⁸⁸ report. As indicated in Tables 8 to 11 of the NGA report, imported/exported natural gas is in three forms: pipeline, LNG (transported by either truck or vessel) and CNG (transported by either truck or vessel). For LNG and CNG, NGA's Figure 14 further provides the exact foreign mode (truck or vessel) between foreign countries and point of exits/entries. Combining the information given in both the figure and tables, the foreign mode was determined. The domestic mode used for transporting foreign trade natural gas was assumed to be pipeline, since Figure 14 of the NGA report indicates that over 99.9 percent of domestic shipments of natural gas is transported by pipeline.

⁸⁸ 2017 EIA Natural Gas Annual: <https://www.eia.gov/naturalgas/annual/pdf/nga17.pdf>

9. ESTIMATION OF OOS FOREIGN TRADE

The estimation process of imported and exported goods movement for FAF5 base year (2017) was primarily based on Census provided special tabulation of *U.S.A. Trade Online* statistics. Because Census foreign trade data only provide information at the State level (origin state for exports or destination state for imports), the main challenges for this OOS component were on determining FAF zone level origins (for exports) and destinations (for imports). Furthermore, Census foreign trade data do not collect information on mode uses by these foreign trades within the U.S., therefore assigning domestic transport modes used between origins/destinations and associate ports of exit/entry were also necessary. Additional imputation procedures were also required for estimating missing (e.g., unknown state) or suppressed attributes in some cases. This chapter describes the data sources used and methodologies applied in estimating the foreign trade flows under the FAF5 estimation process, which includes all commodities.

9.1. DEFINITIONS

Imports are defined as shipments originating in one of the eight foreign regions (see Table 9-1) and terminating inside the U.S. in one of the 132 domestic FAF zones. Upon entering the U.S., these imports pass through a designated port of entry, which function as the transfer points between foreign and domestic legs. Due to data limitations, imports from Canada and Mexico were generally assumed to have the same foreign and domestic mode unless that assumption produced an unfeasible result (e.g., to Hawaii via rail). Obviously, imports from other countries could only enter the U.S. via sea or air modes.

Table 9-1. FAF Foreign Regions

Code	Foreign Region
801	Canada
802	Mexico
803	Rest of Americas
804	Europe
805	Africa
806	SW & Central Asia
807	Eastern Asia
808	SE Asia & Oceania

Similarly, exports were defined as shipments originating from one of the FAF zones, passing through a U.S. port of exit, and terminating in a foreign region. There were domestic and foreign legs for exported shipments as well, and as was the case with imports, domestic modes of exported shipments terminating in Canada and Mexico were generally assumed the same as their foreign modes, except in instances that yielded an unrealistic outcome. Similar to imports, exports to other countries could only exit the U.S. via water or air modes. Note that, as in previous FAF years, Puerto Rico and U.S. Virgin Islands were included in zone 803 under FAF5.

9.2. PRIMARY DATA SOURCES

9.2.1. Census Special Tabulation of 2017 USA Trade Online (USATO) Data

At the request of the BTS, the Census Foreign Trade Division began to provide more details regarding foreign-trade shipments since FAF4. The data includes the state-level origination/destination, commodity (in 2-digit SCTG), and port of entry/exit at the FAF-zone level. While there is still some missing/suppressed information, the availability of such special datasets has consequently streamlined the estimation of foreign trade flows between the state, port, and the foreign zones involved.

Two types of datasets were made available by the Census: (a) trade shipments moved between a foreign country and the U.S. and (b) “transshipment” data involving Canada or Mexico as the intermediary between U.S. and other countries. Under FAF, these “transshipped” goods were treated as imports or exports between the U.S. and Canada/Mexico. Since the transport mode reported in this specific foreign trade dataset reflects mode of movement between Canada/Mexico and other foreign countries, i.e., not the mode used for entering or existing the U.S., “multiple mode” was thus assumed under the FAF to reflect the ambiguity surrounding the North American component for these movements.

To ensure the consistency with other parts of FAF5 data and to reduce potential disclosure concerns, Census used a crosswalk conversion table prepared by the FAF team to convert the USATO data from its original Harmonized System class to the 2-digit SCTG commodity code. Similarly, Census assigned points of entry/exit locations and foreign countries to FAF5-defined zones based on a lookup table supplied by the FAF team.

The FAF5 has benefited from special tabulations provided by the Census, which offered more detail information for domestic segments of foreign-trade shipments for trade with all countries, not just Canada/Mexico. For most trade records, this USATO data included the state-level U.S. origination/destination, commodity in 2-digit SCTG, and port of entry/exit at the FAF-zone level. It included more specific information than what could be obtained from public foreign trade data, therefore circumventing the need for complex modeling efforts underpinning the processes prior to FAF4.

9.2.1.1. 2017 CFS Data on Exports

As pointed out previously, the USATO data did not provide details on the domestic segments of imported and exported shipments. Specifically, no information was available on domestic mode of foreign traded shipments and only state level geography of U.S. origin (on exports) or destination (on imports) was given. Since CFS did capture exported movements (as shipper reported), information from the 2017 CFS export shipments was used as a source for assigning domestic modes and identifying origins at the FAF5 zone (i.e., CFS area).

9.2.1.2. County Business Patterns

As for other OOS sectors discussed thus far in this report, payroll information from the 2017 CBP database was used to disaggregate state-level estimates to corresponding FAF5 regions within the given state, when appropriate.

9.3. ESTIMATION PROCESSES

To avoid disclosure issues, Census aggregated some commodities into less-detailed commodity groups, instead of the 2-digit SCTG as needed in FAF5. In addition, some states of origin/destination were not revealed in the Census file, either because that data were missing/unknown or to avoid a potential disclosure situation. Moreover, detailed information on the location of points of entry/exit (shown as a FAF zone) might not be available or non-specified for certain trade flows (e.g., vessel by its own power). Therefore, FAF5 imports and exports flow matrix includes application of models and assumptions grounded in expert knowledge. The applied estimation procedures are discussed in the following subsections.

9.3.1. Step 1: Disaggregation of Commodity Groups to Associated 2-digit SCTG Codes

The definition of an aggregated commodity group (i.e., SCTG group) as applied by Census to the USATO data (for FAF5 uses) is presented in Table 9-2. For example, commodity group “1G” covers five SCTG codes (i.e., 01-05), while a code of “5G” could mean any commodities that fall between SCTG 20 and SCTG 24. The very first step in preparing the USATO data for FAF5 flow estimation processes, was hence to disaggregate these SCTG-groups into their associated 2-digit SCTG codes.

Table 9-2. Definition of SCTG Group in the Foreign Trade Data File

SCTG Group Code	SCTG 2-digit Covered	Description
1G	01-05	Agriculture products and fish
2G	06-09	Grains, alcohol, and tobacco products
3G	10-14	Stones, non-metallic minerals, and metallic ores
4G	15-19	Coal and petroleum products
5G	20-24	Pharmaceutical and chemical products
6G	25-30	Logs, wood products, and textile and leather
7G	31-34	base metal and machinery
8G	35-38	Electronic, motorized vehicles, and precision instruments
9G	39-43	Furniture, mixed freight, and misc. manufactured products

A straightforward approach was employed to derive shares of SCTGs within each SCTG group based on known USATO statistics (in \$). To account for regional variations in commodities being shipped, commodity shares were classified by trade type, foreign zone, and the U.S. state involved. Note that, for simplicity, foreign zones outside Canada and Mexico were grouped together (i.e., FAF foreign zones of 803 through 808). Table 9-3 gives a few examples of the commodity shares used for FAF5 processing.

Table 9-3. Examples of Commodity Shares by Geographic Regions

Trade Type	Foreign Zone	U.S. State	SCTG Group	SCTG 2-digit	2017 USA Trade Online Value (\$K)	% Share
Imports	801	OH	3G	11	122,214	1.9
				12	4,996,425	75.7
				13	1,478,602	22.4
	803-808	OH	3G	11	7,654	0.5
				12	6,312	0.4
				13	1,627,081	99.1
Exports	802	IL	9G	40	76,145	97.1
				41	2,292	2.9
		TN	9G	39	6,323	0.6
				40	219,043	20.0
				41	869,468	79.4

9.3.2. Step 2: Imputing Unknown States

A relatively straightforward approach was used to account for shipment volumes from “unknown state” trades. Fundamentally, volumes from known USATO records with unknown states are proportionally allocated to other trade records of the same trade type (imports vs. exports) that have an attribute set match in the remaining data. Ideally, the “best” match would align based on the following four features: (1) commodity, (2) port of entry/exit, (3) foreign region, and (4) foreign mode. In the case where this matching yields an “empty” set, seven additional attribute combinations were tested: the four distinct tri-variate subsets of these attributes and three bi-variate subsets corresponding to the set with attributes (1)-(3) above. The match set yielding the minimum number of candidate states was selected; this algorithm was sufficiently exhaustive as to prevent the generation of all nonempty sets.

To demonstrate the basic principle underlying this process, assume an “unknown state” shipment of \$1,000 is matched to two “known state” records (State A and State B valued at, say, \$2,000 and \$500, respectively). The amount of \$1,000 from this unknown-state record would be divided between State A and State B with an 80-20 percent split. Using that ratio, the \$1,000 from the unknown state is split to State A and State B to produce an increase to \$2,800 from \$2,000 for State A and to \$700 from \$500 for State B.

In the transshipments case, an analogous method was applied, with slight variation in the attributes comprising the successive matching conditions. That is, the additional feature of “Canada or Mexico” as the North American port of entry/exit, as well as import/export status, was used to generate the “best” potential match in addition to the four attributes mentioned previously. If an empty set was produced from the remaining data, the six distinct five-attribute subsets were used to produce sets of candidate states, of which the smallest was selected for imputation.

Those records with assigned commodity groups that now have an origin or destination U.S. state were then disaggregated via the process described above.

9.3.3. Step 3: Unspecified Port Zones

Census used several special codes for ports, in place of FAF-zone codes, on shipments that met certain conditions in the UASTO data. Because of that, port (FAF-zone level) information for these shipments was not provided in the data file, therefore requiring estimation.

Port Zone Code 998 designates low-value shipments and mail and were thus assigned to “multiple modes and mail” for foreign mode in FAF. For imports, there were eight records, one for each foreign zone, with this designation, all of which had SCTG code 43 (“mixed freight”) and unknown state. Accordingly, these records were distributed proportionally across all 50 states and Washington, DC, based on value market shares for each originating foreign zone in the base dataset. The case for exports was somewhat trickier. While eight records mirrored the case with imports and were thus processed in the same way, the remaining 1,074 export records involved imputing an unknown state or a new port of exit, following a similar procedure to that described in Step 2.

Port Zone Code 997 designated “Vessel under its own power” in both imports and exports and were thus assigned “water” for foreign mode. Due to the nature of these vessels having to be large enough to travel across countries, it was assumed that the ports involved with these shipments would be in proximity to their origin/destination states. 21 import records and 28 export records existed with this designation in the USATO data; all were manually assigned ports of entry/exit.

Port Zone Code 991 was used for certain 32 exported coal shipments by vessels out of one of three ports: Norfolk, Mobile, or Charleston. However, a specific port was not identified for these records in the original FTD data file. A simple assignment by geographic location of originating state was applied to reassign FAF area codes for these shipments.

Port Zone Codes 981 and 982 represented shipments involving Puerto Rico and the U.S. Virgin Islands, respectively. For imports, the assumption was that most goods will remain in these islands, which was supported when comparing the total value of goods to local economic data⁸⁹. Thus, for the 907 records corresponding to Puerto Rico and 185 records corresponding to the Virgin Islands, those with an “unknown state” were removed from consideration in FAF, while the other records were assigned foreign zone 803 (“Rest of Americas”) and imputed port of entries. Meanwhile, for exports, there were 518 records with designation 981 and 67 records with designation 982. Again, unknown state and port of exit records underwent imputation.

9.3.4. Step 4: Unspecified Foreign Modes

There were records in both the USATO data files that did not have a foreign mode assignment. These were all imports and fall under the Foreign-Trade Zone designation. The imputation procedure proceeded as follows:

- Find candidates using commodity-port of entry-destination state attribute set. If there are none, assign the foreign mode to be “other.”
- Else, select the foreign mode that is the mode across identified candidates. In case of a tie, distribute proportionally according to the corresponding value share.
- Apply estimation Steps 2-4 as appropriate for “direct” shipments.

9.3.5. Step 5: Adjustment of Port Zone Locations

Ports of entry/exit as indicated in the original data were not necessarily located along the U.S. borders or coasts. As directed by the BTS, shipments with non-border and non-coastal port

⁸⁹ <https://tradingeconomics.com/puerto-rico/imports#:~:text=Imports%20in%20Puerto%20Rico%20averaged,Million%20in%20September%20of%202017>

zones were reassigned to geographically logical border or coastal ports based on foreign zone and mode type. The six general cases comprising this process are summarized below.

- Canada | Land
- Mexico | Land
- Canada/Mexico | Water/Multiple
- Rest of Americas | Water/Multiple
- Europe/Africa/SCW Asia | Water/Multiple
- E Asia/SE Asia and Oceania | Water/Multiple

In addition, two major shifts in air shipment ports were adopted to reflect the home location of major private parcel companies.

9.3.6. Step 6: Disaggregation of State-Level Flows to FAF Regions

As in the processing of other OOS areas, CBP payroll data was utilized to disaggregate state-level flows to associated FAF regions, as the origins for exports and destinations for imports.

9.3.7. Step 7: Assignment of Domestic Mode

Transshipments were assumed to use “multiple modes” for their foreign and domestic movements, thus the assignment of domestic mode was only needed for records included under the main USATO dataset. This process required several rounds of data processing and adjustments after the initial imputation effort. Importantly, for FAF5, the utilization of the 2017 CFS data on exports provided the basis for domestic mode assignment. Furthermore, records with foreign mode as pipeline or other/unknown were imputed in separate processes.

Records matching to the 2017 CFS export file were used for domestic mode imputation. The matching criteria considered various combination of variables, began with the highest number of combination and gradually reducing the dimension as needed for matching purposes. If no matches in the 2017 CFS exports data file were identified, the default assignment for the domestic mode was assumed to be the same as its foreign mode. This initial imputation stage either yielded one or multiple possible modes; in the latter case, the original shipment was split into two or more records proportionally based on the share of the mode across all identified attribute set matches.

A series of checks and corresponding adjustments were applied to the initial foreign trade flow matrix. These were based on knowledge accumulated from previous FAF versions as well as benchmarking to other data sources such as Waybill data and US Army Corps of Engineers (USACE) data for rail and water shipments, respectively.

9.3.8. Step 8: Missing Weights and Values

Numerous import and export records did not have an associated shipping weight, while only a small handful of records did not have an associated monetary value. Value-to-weight ratios based on 2017 CFS data were applied to estimate this missing information.

9.3.9. Step 9: Foreign Mode is Pipeline or Other/Unknown

For foreign trade entering/exiting the U.S. by pipeline, if the domestic origin/destination was different from the port of exit/entry, the domestic mode was assigned as truck, else we follow the standard transborder shipment convention. The shipping weights for the export records were imputed, if necessary, as described above.

For foreign trade entering/exiting the U.S. by other/unknown mode, the default was to assume the same for its domestic mode. However, when possible, the domestic mode was imputed using information in other foreign trade records as well as expert knowledge. Missing weights and small-value shipments were treated the same as before for all records.

10. ESTIMATION OF OOS FISHERIES

Fishery shipments, that are OOS for the CFS, are those occurred prior to the first point of processing, or before arrival at a distribution center. Once the fishery shipments reach these points, they become an in-scope commodity for the CFS. The commodity coverage for this OOS shipments is SCTG 01 (live animal and fish). It is assumed, under this OOS component, fishery production and distribution occur at the same locations. Although some fishery commodities could be directly shipped to locations other than the production sites, those shipment are usually within the scope of the CFS, and thus not included for the OOS estimation.

10.1. DATA SOURCES

Commercial fisheries landing information released by the National Oceanic and Atmospheric Administration (NOAA)⁹⁰ is used as the data source for the fishery OOS estimation process. Specifically, four sets of summary statistics generated from the NOAA data are used for FAF5. They are:

10.1.1. State Total and Process-at-Sea Total - 2017 Commercial Fishery Landings by State

This summary table contains the list of states that have commercial fishery landings. The table provides weight (pounds) and value (dollars) for each state's total commercial fishery landings. There are 28 states listed with an additional entry of "Process-at-Sea" (where fish are processed at sea) in the table.

10.1.2. Major Port Total - 2017 Commercial Fishery Landings by Top U.S. Ports

This table summarizes total commercial fisheries in weight (millions of pounds) and value (millions of dollars) by top U.S. ports (major port). There are 132 major ports listed in the 2017 table.

10.1.3. Historical Landings - 1981 – 2016 Landings by Top U.S. Ports

In addition to the 2017 major port table mentioned above, historical statistics are also obtained for this OOS estimation effort. Each year's report includes a list of major U.S. ports for commercial fishery landing in the given year and their weights and values. The combinations of these ports over all historical years yield a more comprehensive list of identifiable ports for commercial fishery landings in the U.S.

⁹⁰ Commercial Fisheries Landings, National Oceanic and Atmospheric Administration, <https://www.fisheries.noaa.gov/national/sustainable-fisheries/commercial-fisheries-landings>

10.1.4. Species Data for Process-at-Sea

For Process-at-Sea, NOAA data on detailed species were extracted so that the Process-at-Sea production quantity can be more accurately redistributed by geography. This data table includes total weight in pounds and value in dollars by species type.

10.2. ESTIMATION PROCEDURE FOR COMMERCIAL FISHERIES SHIPMENTS

As mentioned above, state total commercial fisheries (in pounds and dollars) for 28 states can be directly obtained from NOAA. Therefore, the FAF5 process for this OOS component mainly involves two procedures: (1) to disaggregate state total volumes to FAF zones within the given state and (2) to distribute the Process-at-Sea total to associated FAF zones. These procedures are briefly discussed below.

10.2.1. Distribution of State Total into FAF Zones

Both state level and major port level totals were utilized in this process. First, all major ports were assigned to their corresponding FAF zones, based on their locations to the extent possible. Since major ports published by NOAA include those top producers of commercial fisheries only, it does not include a complete list of all commercial fishery landing ports in the U.S. Therefore, the next step is to conduct a comparison between state total and the sum of volumes from all major ports-based FAF-zones within that given state. Any residual portion, as expressed in the equation below, will then be distributed to other identifiable ports in the state; or redistributed among major ports within the given state, if no additional ports could be identified.

$$\text{Residual portion for a state} = \text{State total} - \text{Sum of all major ports within the given state}$$

To achieve this, NOAA commercial fishery landing data at major ports over the historical period of 1981 to 2017 was used. Specifically, the average annual port production tonnages were applied to proportionally distribute the residual portion from each state to the identified “additional” ports (those not in the 2017 major ports list) within the given state. Similar procedures to assign ports to their corresponding FAF zones were then applied, thus updating all associated FAF zone level statistics (weight and value). For states where no “additional” ports were found, their residual portion of volumes, if exist, were distributed proportionally among their corresponding FAF zones, using the port volumes as the share for each given state.

10.2.2. Distribution of Process-at-Sea total into FAF Zones

NOAA detailed species data for the Process-at-Sea is examined under this process to identify major types of commercial fisheries being included here. Majority of commercial fishery being processed-at-sea is found to be Pacific hake, or pacific whiting, which accounted for over 99 percent of all production quantity in tonnage from those classified as processed-at-sea. Based on this finding, commercial fishery processed at sea was proportionally distributed to ports in Alaska, Washington, Oregon, and California, based on volume shares of their state total.

As in the prior FAFs, OOS fishery shipments were assumed to be local (i.e., staying around the docks). Thus, these OOS shipments were considered as intra-zone movements, i.e., within the same FAF zone. Moreover, movements for all shipments from this OOS component were assumed via truck.

11. ESTIMATION OF OOS LOGGING

This OOS component covers shipments from field (forests) to processing facilities that were captured under the CFS (timber cutting and/or transporting industries). The commodity involved in these logging industry shipments falls under the 2-digit SCTG 25 (logs and other wood in the rough). Under the FAF5 processing, OOS logging shipments were assumed to be transported in relatively short distances between its harvest location to the nearest processing facilities.

11.1. DATA SOURCES

The USDA Forest Service conducts regular surveys of mills to estimate industrial uses of roundwood. It also surveys logging operations to determine the logging sites and their types of operations. These data are collected and published under the Forest Inventory and Analysis (FIA) National Program through their web site⁹¹. Several reports and statistics produced under the FIA National Program were used for estimating the shipments involving this OOS component under the FAF5 process. These data sources are briefly summarized below.

11.1.1. Forest Statistics by State – FIA DataMart

The *FIA DataMart*⁹², previously released under *Forestry Inventory Data Online* (FIDO), is an annual report prepared by the USDA Forest Service under the FIA National Program. It contains the most recent state inventories, including statistics on status and trends in forest area and locations, and in total tree growth, mortality, and removals by harvest. Quantity (in board feet) of the harvest removals was collected by location and species type to determine the weight of the logs heading to processing facilities.

11.1.2. Timber Product Output (TPO) Data Download

The *Timber Product Output Data Download*⁹³ is a toolkit provided by the FIA that enables users to download TPO data at the state and county level. Historical year information from the TPO Reports, dated back to 1997, are published under this data download website. Note that TPO estimates are provided for reported states only within each given year. Therefore, not all state estimates are available from the same year. For 2017 (FAF5 base year), statistics for 13 states are available from the TPO report, including AL, AR, FL, GA, KY, LA, MS, NC, OK, SC, TN, TX, and VA.

Note that public release of the 2017 TPO data was delayed at the time of FAF5 OOS processing. Under the gracious assistance of an analyst at the USDA Forest Services, the ORNL team was able to obtain a spreadsheet file directly for the FAF5 use. The file being provided was originally created by the FIA analyst for use in their 2017 Resource Planning Act (RPA) assessments. This file includes the most current TPO estimates (at the time) for all states (except for Hawaii and DC). Note that the state of Hawaii does not have enough of an industry to be included in the TPO reporting. Also, DC does not have a TPO reporting either.

⁹¹ <http://www.fia.fs.fed.us/program-features/tpo/>

⁹² <https://apps.fs.usda.gov/fia/datamart/datamart.html>

⁹³ <https://usfs-public.app.box.com/s/y4ziirdb9v7zardus0cuajh7ziy9b2id>.

11.1.3. Forest Industry Research Program, University of Montana

The *Bureau of Business and Economic Research (BBER)* at the University of Montana⁹⁴ provides timber harvest data by county for 5 western states: MT, ID, WA, OR, and CA.

11.1.4. TimberMart-South Reports

*TimberMart-South*⁹⁵ is a non-for-profit organization that provide data for stumpage, delivered wood prices, and other forestry business information for the Southeastern U.S. Specifically, its coverage area includes Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, and Virginia.

11.1.5. State and Region Price Reports

To determine the value of the logging OOS, prices of soft and hard woods were obtained from State or Region Price Reports, when available. Additional information sources on timber prices were gathered from the *TimberMart-South*, *BBER*⁹⁶, and the Texas Forest Service's *Stumpage Prices Trends*⁹⁷.

11.2. ESTIMATION METHODS

The FAF5 estimation process for this logging OOS was straightforward, although rather tedious. Generally, estimates from this FIA-provided file, at county and state levels, were used directly for estimating the OOS logging volumes. Specifically, timber volumes in the 2017 RPA report provided estimates on timber removals, as coming from the growing-stock portion of sawtimber trees, were extracted for this OOS study. These include roundwood product harvested classified as either softwood or hardwood. For this FAF5 processing, the 2017 RPA reported roundwood volumes on sawlogs, veneer logs, and post/poles/pilings were included under the "softwood" category. Timber volumes on pulpwood/fiber byproducts, composite products, fuelwood/fuel-byproducts, and miscellaneous products/byproducts were considered as "hardwood."

The reported volumes of county-level totals of board feet by weight (in cubic feet) were taken directly from the source tables (spreadsheets) and summed to the FAF5 zone level, by type of timber (softwood and hardwood) and by state. Because these weights were measured in cubic feet, conversion factors were needed to determine green tonnage of logged shipments. Note that different factors are required for the conversions, depending on the forestry location and type of wood (i.e., softwood or hardwood).

Due to data limitation, available information obtained from the *TimberMart-South Notes*⁹⁸ was applied to convert logging volumes to tonnages for all states within the U.S. In addition, weight (tons) to dollar conversions for different product groups were also done based on average prices information available from the *TimberMart-South* website. Note that, the weights used for this FAF5 OOS component is in green tons, which is better aligned with raw woods covered

⁹⁴ <http://www.bber.umt.edu/FIR/default.asp>.

⁹⁵ <http://www.timbermart-south.com/>.

⁹⁶ http://www.bber.umt.edu/FIR/F_LogPrice.asp.

⁹⁷

https://tfsweb.tamu.edu/uploadedFiles/TFSMain/Manage_Forest_and_Land/Landowner_Assistance/Timber_Harvesting/Timber_Price_Trends/Prices_2011-2015.pdf.

⁹⁸ <http://timbermart-south.com/pdf/TMSNotes2019.pdf>

under this logging OOS sector. This information, however, was not available in prior FAFs, thus dry tons was used then instead.

Similar to prior FAF, movement of OOS logging shipments was assumed to be local activity (i.e., travel from forests to local processing locations) under FAF5. Thus, OOS logging shipments are assumed to be intra-zone movements. Furthermore, these OOS logging shipments were assumed to be transported by truck and all movements were considered under commodity code SCTG 25. For FAF5, shipments involving this OOS logging sector totaled at 447,337 thousand tons, value at \$7,215 million were moved in 2017.

Appendix A. Glossary of Terms, Abbreviations, and Acronyms

ACS	American Community Survey
AMSA	American Moving and Storage Association
ARTS	Annual Retail Trade Survey
BBER	Bureau of Business and Economic Research
BEA	Bureau of Economic Analysis
BTS	Bureau of Transportation Statistics
C&D	Construction and Demolition Debris
CBP	County Business Patterns
CDRA	Construction & Demolition Recycling Association
CFS	Commodity Flow Survey
CNG	Compressed Natural Gas
CSA	Combined Statistical Area
EDI	Electronic Data Interchange
EIA	Energy Information Administration
EPA	U.S. Environmental Protection Agency
FAF	Freight Analysis Framework
FHWA	Federal Highway Administration
FIA	Forest Inventory and Analysis
FIDO	Forestry Inventory Data Online
HH&B	Household and business
HIFLD	Homeland Infrastructure Foundation-Level Data
IPF	Iterative Proportional Fitting
LMOP	Landfill Methane Outreach Program
LNG	Liquefied Natural Gas
MA	Metropolitan Area
MLE	Maximum Likelihood Estimates
MRF	Materials Recovery Facilities
MSA	Metropolitan Statistical Area
MSW	Municipal Solid Waste
NAICS	North American Industry Classification System
NASS	National Agricultural Statistics Service
NG	Natural Gas
NGA	Natural Gas Annual
NOAA	National Oceanic and Atmospheric Administration
OD	Origin-Destination
ODCM	Origin, Destination, Commodity, Mode
OOS	Out-of-Scope of the 2017 CFS
ORNL	Oak Ridge National Laboratory
PADDs	Petroleum Administration for Defense Districts
PUF	Public Use File
PUM	Public Use Microdata
ROS	Remainder of State
RPA	Resource Planning Act
SCTG	Standard Classification of Transported Goods
STB	Surface Transportation Board
TPO	Timber Product Output

USACE	US Army Corps of Engineers
USATO	USA Trade Online
USDA	U.S. Department of Agriculture
VIUS	Vehicle Inventory and Use Survey
WCSC	Waterborne Commerce Statistics Center
WTE	Waste-to-energy

Appendix B. CFS Area – FAF Zone Crosswalk Table

CFS Area	FAF Zone	Description
01-142	011	Birmingham-Hoover-Talladega, AL CFS Area
01-380	012	Mobile-Daphne-Fairhope, AL CFS Area
01-99999	019	Remainder of Alabama
02-99999	020	Alaska
04-38060	041	Phoenix-Mesa-Glendale, AZ CFS Area
04-536	042	Tucson-Nogales, AZ CFS Area
04-99999	049	Remainder of Arizona
05-99999	050	Arkansas
06-348	061	Los Angeles-Long Beach, CA CFS Area
06-472	062	Sacramento-Roseville, CA CFS Area
06-41740	063	San Diego-Carlsbad-San Marcos, CA CFS Area
06-488	064	San Jose-San Francisco-Oakland, CA CFS Area
06-260	065	Fresno-Madera, CA CFS Area
06-99999	069	Remainder of California
08-216	081	Denver-Aurora, CO CFS Area
08-99999	089	Remainder of Colorado
09-25540	091	Hartford-West Hartford-East Hartford, CT CFS Area
09-408	092	New York-Newark, NY-NJ-CT-PA CFS Area (CT Part)
09-99999	099	Remainder of Connecticut
10-428	101	Philadelphia-Reading-Camden, PA-NJ-DE-MD CFS Area (DE Part)
10-99999	109	Remainder of Delaware
11-47900	111	Washington-Arlington-Alexandria, DC-VA-MD-WV CFS Area (DC Part)
12-300	121	Jacksonville-St. Marys-Palatka, FL-GA CFS Area (FL Part)
12-370	122	Miami-Fort Lauderdale-Port St. Lucie, FL CFS Area
12-422	123	Orlando-Deltona-Daytona Beach, FL CFS Area
12-45300	124	Tampa-St. Petersburg-Clearwater, FL CFS Area
12-99999	129	Remainder of Florida
13-122	131	Atlanta-Athens-Clarke County-Sandy Springs, GA CFS Area
13-496	132	Savannah-Hinesville-Statesboro, GA CFS Area
13-99999	139	Remainder of Georgia
15-46520	151	Urban Honolulu, HI CFS Area
15-99999	159	Remainder of Hawaii
16-99999	160	Idaho
17-176	171	Chicago-Naperville, IL-IN-WI CFS Area (IL Part)
17-476	172	St. Louis-St. Charles-Farmington, MO-IL CFS Area (IL Part)
17-99999	179	Remainder of Illinois
18-176	181	Chicago-Naperville, IL-IN-WI CFS Area (IN Part)
18-294	182	Indianapolis-Carmel-Muncie, IN CFS Area
18-258	183	Fort Wayne-Huntington-Auburn, IN CFS Area
18-99999	189	Remainder of Indiana
19-99999	190	Iowa
20-312	201	Kansas City-Overland Park-Kansas City, MO-KS CFS Area (KS Part)
20-556	202	Wichita-Arkansas City-Winfield, KS CFS Area
20-99999	209	Remainder of Kansas
21-178	211	Cincinnati-Wilmington-Maysville, OH-KY-IN CFS Area (KY Part)
21-350	212	Louisville/Jefferson County-Elizabethtown-Madison, KY-IN CFS Area (KY Part)

CFS Area	FAF Zone	Description
21-99999	219	Remainder of Kentucky
22-12940	221	Baton Rouge, LA CFS Area
22-324	222	Lake Charles, LA CFS Area
22-406	223	New Orleans-Metairie-Hammond, LA-MS CFS Area (LA Part)
22-99999	229	Remainder of Louisiana
23-99999	230	Maine
24-12580	241	Baltimore-Columbia-Towson, MD CFS Area
24-47900	242	Washington-Arlington-Alexandria, DC-VA-MD-WV CFS Area (MD Part)
24-99999	249	Remainder of Maryland
25-148	251	Boston-Worcester-Providence, MA-RI-NH-CT CFS Area (MA Part)
25-99999	259	Remainder of Massachusetts
26-220	261	Detroit-Warren-Ann Arbor, MI CFS Area
26-266	262	Grand Rapids-Wyoming-Muskegon, MI CFS Area
26-99999	269	Remainder of Michigan
27-378	271	Minneapolis-St. Paul, MN-WI CFS Area (MN Part)
27-99999	279	Remainder of Minnesota
28-99999	280	Mississippi
29-312	291	Kansas City-Overland Park-Kansas City, MO-KS CFS Area (MO Part)
29-476	292	St. Louis-St. Charles-Farmington, MO-IL CFS Area (MO Part)
29-99999	299	Remainder of Missouri
30-99999	300	Montana
31-420	311	Omaha-Council Bluffs-Fremont, NE-IA CFS Area (NE Part)
31-99999	319	Remainder of Nebraska
32-332	321	Las Vegas-Henderson, NV-AZ CFS Area (NV Part)
32-99999	329	Remainder of Nevada
33-148	331	Boston-Worcester-Providence, MA-RI-NH-CT CFS Area (NH Part)
33-99999	339	Remainder of New Hampshire
34-408	341	New York-Newark, NY-NJ-CT-PA CFS Area (NJ Part)
34-428	342	Philadelphia-Reading-Camden, PA-NJ-DE-MD CFS Area (NJ Part)
35-99999	350	New Mexico
36-104	361	Albany-Schenectady, NY CFS Area
36-160	362	Buffalo-Cheektowaga, NY CFS Area
36-408	363	New York-Newark, NY-NJ-CT-PA CFS Area (NY Part)
36-464	364	Rochester-Batavia-Seneca Falls, NY CFS Area
36-99999	369	Remainder of New York
37-172	371	Charlotte-Concord, NC-SC CFS Area (NC Part)
37-268	372	Greensboro--Winston-Salem--High Point, NC CFS Area
37-450	373	Raleigh-Durham-Chapel Hill, NC CFS Area
37-99999	379	Remainder of North Carolina
38-99999	380	North Dakota
39-178	391	Cincinnati-Wilmington-Maysville, OH-KY-IN CFS Area (OH Part)
39-184	392	Cleveland-Akron-Canton, OH CFS Area
39-198	393	Columbus-Marion-Zanesville, OH CFS Area
39-212	394	Dayton-Springfield-Sidney, OH CFS Area
39-99999	399	Remainder of Ohio
40-416	401	Oklahoma City-Shawnee, OK CFS Area
40-538	402	Tulsa-Muskogee-Bartlesville, OK CFS Area
40-99999	409	Remainder of Oklahoma
41-440	411	Portland-Vancouver-Salem, OR-WA CFS Area (OR Part)
41-99999	419	Remainder of Oregon
42-428	421	Philadelphia-Reading-Camden, PA-NJ-DE-MD CFS Area (PA Part)
42-430	422	Pittsburgh-New Castle-Weirton, PA-OH-WV CFS Area (PA Part)

CFS Area	FAF Zone	Description
42-408	423	New York-Newark, NY-NJ-CT-PA CFS Area (PA Part)
42-99999	429	Remainder of Pennsylvania
44-148	441	Boston-Worcester-Providence, MA-RI-NH-CT CFS Area (RI Part)
45-16700	451	Charleston-North Charleston-Summerville, SC CFS Area
45-273	452	Greenville-Spartanburg-Anderson, SC CFS Area
45-99999	459	Remainder of South Carolina
46-99999	460	South Dakota
47-368	471	Memphis, TN-MS-AR CFS Area (TN Part)
47-400	472	Nashville-Davidson--Murfreesboro, TN CFS Area
47-314	473	Knoxville-Morristown-Sevierville, TN CFS Area
47-99999	479	Remainder of Tennessee
48-12420	481	Austin-Round Rock, TX CFS Area
48-13140	482	Beaumont-Port Arthur, TX CFS Area
48-204	483	Corpus Christi-Kingsville-Alice, TX CFS Area
48-206	484	Dallas-Fort Worth, TX-OK CFS Area (TX Part)
48-238	485	El Paso-Las Cruces, TX-NM CFS Area (TX Part)
48-288	486	Houston-The Woodlands, TX CFS Area
48-29700	487	Laredo, TX CFS Area
48-41700	488	San Antonio-New Braunfels, TX CFS Area
48-99999	489	Remainder of Texas
49-482	491	Salt Lake City-Provo-Orem, UT CFS Area
49-99999	499	Remainder of Utah
50-99999	500	Vermont
51-40060	511	Richmond, VA CFS Area
51-545	512	Virginia Beach-Norfolk, VA-NC CFS Area (VA Part)
51-47900	513	Washington-Arlington-Alexandria, DC-VA-MD-WV CFS Area (VA Part)
51-99999	519	Remainder of Virginia
53-500	531	Seattle-Tacoma, WA CFS Area
53-440	532	Portland-Vancouver-Salem, OR-WA CFS Area (WA Part)
53-99999	539	Remainder of Washington
54-99999	540	West Virginia
55-376	551	Milwaukee-Racine-Waukesha, WI CFS Area
55-99999	559	Remainder of Wisconsin
56-99999	560	Wyoming

Appendix C. FAF5 Commodity Codes

Code	Commodity Description
01	Animals and Fish (live)
02	Cereal Grains (includes seed)
03	Agricultural Products (excludes Animal Feed, Cereal Grains, and Forage Products)
04	Animal Feed, Eggs, Honey, and Other Products of Animal Origin
05	Meat, Poultry, Fish, Seafood, and Their Preparations
06	Milled Grain Products and Preparations, and Bakery Products
07	Other Prepared Foodstuffs, Fats and Oils
08	Alcoholic Beverages and Denatured Alcohol
09	Tobacco Products
10	Monumental or Building Stone
11	Natural Sands
12	Gravel and Crushed Stone (excludes Dolomite and Slate)
13	Other Non-Metallic Minerals not elsewhere classified
14	Metallic Ores and Concentrates
15	Coal
16	Crude Petroleum
17	Gasoline, Aviation Turbine Fuel, and Ethanol (includes Kerosene, and Fuel Alcohols)
18	Fuel Oils (includes Diesel, Bunker C, and Biodiesel)
19	Other Coal and Petroleum Products, not elsewhere classified
20	Basic Chemicals
21	Pharmaceutical Products
22	Fertilizers
23	Other Chemical Products and Preparations
24	Plastics and Rubber
25	Logs and Other Wood in the Rough
26	Wood Products
27	Pulp, Newsprint, Paper, and Paperboard
28	Paper or Paperboard Articles
29	Printed Products
30	Textiles, Leather, and Articles of Textiles or Leather
31	Non-Metallic Mineral Products
32	Base Metal in Primary or Semi-Finished Forms and in Finished Basic Shapes
33	Articles of Base Metal
34	Machinery
35	Electronic and Other Electrical Equipment and Components, and Office Equipment
36	Motorized and Other Vehicles (includes parts)
37	Transportation Equipment, not elsewhere classified
38	Precision Instruments and Apparatus
39	Furniture, Mattresses and Mattress Supports, Lamps, Lighting Fittings, and Illuminated Signs
40	Miscellaneous Manufactured Products
41	Waste and Scrap (excludes of agriculture or food)
43	Mixed Freight

Appendix D. Agricultural Commodity Categorized By SCTG Code

SCTG 01: Live Animal and Fish

Calves sold (number)
Cattle sold (number)
Hogs and pigs sold (number)
Any poultry sold, layers 20 weeks old and older (number)
Any poultry sold, pullet chicks (number)
Any poultry sold, broilers (number)
Any poultry sold, turkeys sold (number)
Sheep and lambs sold (number)
Horses and ponies, sales (number)
Miscellaneous livestock (number)
Mink and their pelts, sales (number)
Ducks, sales (number)
Geese, sales (number)
Pigeons or squab, sales (number)
Pheasants, sales (number)
Quail, sales (number)
Emu and ostrich, sales (number)
Miscellaneous poultry, sales (number)
Other poultry, sales (number)
Poultry hatched, sales (number)
Mules, burros, and donkeys - sales (number)
Goats, total sales (number)
Rabbits and their pelts -sales (number)
Catfish, pounds (1,000)
Trout, pounds (1,000)
Hybrid Striped Bass, pounds (1,000)
Other fish, pounds (1,000)
Crawfish, clam, mussels, oysters, snails, pounds (1,000)
Other aquaculture products, pounds (1,000)

SCTG 02: Cereal Grains

Corn for grain or seed (bushels), harvested	Rye for grain (bushels), harvested
Sorghum for grain or seed (bushels), harvested	Wild rice (cwt), harvested
Wheat for grain, total (bushels), harvested	Popcorn (pounds, shelled), harvested
Barley for grain (bushels), harvested	Proso millet (bushels), harvested
Buckwheat (bushels), harvested	Safflower (pounds), harvested
Emmer and spelt (bushels), harvested	Triticale (bushels), harvested
Oats for grain (bushels), harvested	Corn for grain or seed (bushels), harvested
Rice (hundredweight), harvested	

SCTG 03: Agricultural Products

Canola and other rapeseed (pounds), harvested	Lespedeza seed (pounds), harvested
Canola (pounds), harvested	Orchardgrass seed (pounds), harvested
Other rapeseed (pounds), harvested	Red clover seed (pounds), harvested
Flaxseed (bushels), harvested	Ryegrass seed (pounds), harvested
Mustard seed (pounds), harvested	Sudangrass seed (pounds), harvested
Sunflower seed (pounds), harvested	Timothy seed (pounds), harvested
Cotton (bales), harvested	Vetch seed (pounds), harvested
Tobacco (pounds), harvested	Wheatgrass seed (pounds), harvested
Soybeans for beans (bushels), harvested	White clover seed (pounds), harvested
Dry edible beans, excluding dry limas (hundredweight)	Other seeds (pounds), harvested
Dry limas beans (hundredweight), harvested	Vegetables harvested, harvested (acres)
Dry edible peas (hundredweight), harvested	Fruits Total Production in 1,000 tons
Dry cowpeas and dry southern peas (bushels), harvested	Dill for oil (pounds), harvested
Lentils (hundredweight), harvested	Ginger root (pounds), harvested
Potatoes, excluding sweet potatoes (hundredweight), harvested	Ginseng (pounds), harvested
Sweet potatoes (hundredweight), harvested	Guar (pounds), harvested
Sugar beets for seed (pounds), harvested	Sesame (pounds)
Sugar beets for sugar (tons), harvested	Herbs, dried (pounds), harvested
Sugarcane for seed (tons), harvested	Hops (pounds), harvested
Sugarcane for sugar (tons), harvested	Jojoba harvested (pounds), harvested
Peanuts for nuts (pounds), harvested	Mint for oil (pounds of oil), harvested
Alfalfa seed (pounds), harvested	Mint for tea
Austrian winter peas (hundredweight), harvested	Pineapples harvested (tons), harvested
Bahia grass seed (pounds), harvested	Sorghum for syrup (pounds), harvested
Bentgrass seed (pounds), harvested	Sweet corn for seed (pounds), harvested
Bermuda grass seed (pounds), harvested	Taro (pounds), harvested
Birdsfoot trefoil seed (pounds), harvested	Switchgrass
Bromegrass seed (pounds), harvested	Miscanthus
Crimson clover seed (pounds), harvested	Camelia
Fescue seed (pounds), harvested	Maple Syrup
Kentucky Bluegrass seed (pounds), harvested	Mushrooms
Ladino clover seed (pounds), harvested	

SCTG 04: Animal Feed, Eggs, Honey, and Other Products of Animal Origin

Hay-alfal, other tame, small grain, wild, grass silage grass (tons)
Haylage/Grass Silage/Greenchop (tons)
Corn for silage or green chop (tons, green), harvested
Sorghum for silage or green chop (tons, green), harvested
Salt hay (tons), harvested
Sheep and lambs shorn (pounds of wool)
Honey, sales (pounds)
Mohair, sales (pounds)

SCTG 07: Other Prepared Foodstuffs, Fats and Oils

Milk and milk fat (million pounds)

Appendix E. Assignment of Foreign Countries To FAF Foreign Zones

FAF ZONE	Foreign Region	Countries
801	Canada	Canada
802	Mexico	Mexico
803	Rest of Americas	Anguilla, Antigua and Barbuda, Argentina, Aruba, Bahamas, Barbados, Belize, Bermuda, Bolivia, Bonaire, Sint Eustatius and Saba, Brazil, British Virgin Islands, Cayman Islands, Chile, Colombia, Costa Rica, Cuba, Curacao, Dominica, Dominican Republic, Ecuador, El Salvador, Falkland Islands, French Guiana, Greenland, Grenada, Guadeloupe, Guatemala, Guyana, Haiti, Honduras, Jamaica, Martinique, Montserrat, Nicaragua, Panama, Paraguay, Peru, Puerto Rico, St. Barthelemy, St. Kitts and Nevis, St. Lucia, St. Martin, St. Pierre and Miquelon, St. Vincent and the Grenadines, Sint Maarten, Suriname, Trinidad and Tobago, Turks and Caicos Islands, U.S. Virgin Islands, Uruguay, and Venezuela
804	Europe	Aland Islands, Albania, Andorra, Austria, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Channel Islands (Guernsey, Jersey, and Sark), Croatia, Czech Republic, Denmark, Estonia, Faroe Islands, Finland, France, Germany, Gibraltar, Greece, Hungary, Iceland, Ireland, Isle of Man, Italy, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Moldova, Monaco, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Romania, Russia, San Marino, Serbia, Slovakia, Slovenia, Spain, Svalbard and Jan Mayen Islands, Sweden, Switzerland, Ukraine, United Kingdom, and Vatican City
805	Africa	Algeria, Angola, Benin, Botswana, British Indian Ocean Territory, Burkina Faso, Burundi, Cameroon, Cape Verde, Central African Republic, Chad, Comoros, Congo (Brazzaville), Congo (Kinshasa), Djibouti, Egypt, Equatorial Guinea, Eritrea, Eswatini, Ethiopia, French Southern and Antarctic Lands, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Ivory Coast, Kenya, Lesotho, Liberia, Libya, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mayotte, Morocco, Mozambique, Namibia, Niger, Nigeria, Reunion, Rwanda, St. Helena, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, Somalia, South Africa, South Sudan, Sudan, Tanzania, Togo, Tunisia, Uganda, Western Sahara, Zambia, and Zimbabwe
806	Southern, Central, and Western Asia	Afghanistan, Armenia, Azerbaijan, Bahrain, Bangladesh, Bhutan, Cyprus, Georgia, India, Iran, Iraq, Israel, Jordan, Kazakhstan, Kuwait, Kyrgyzstan, Lebanon, Maldives, Nepal, Oman, Palestine, Pakistan, Qatar, Saudi Arabia, Sri Lanka, Syria, Tajikistan, Turkey, Turkmenistan, United Arab Emirates, Uzbekistan, and Yemen
807	Eastern Asia	China, Hong Kong, Japan, Macao, Mongolia, North Korea, South Korea, and Taiwan
808	Southeastern Asia and Oceania	American Samoa, Australia, Brunei, Cambodia, Christmas Island, Cocos Islands, Cook Islands, East Timor, Federated States of Micronesia, Fiji, French Polynesia, Guam, Heard and McDonald Islands, Indonesia, Kiribati, Laos, Malaysia, Marshall Islands, Myanmar, Nauru, New Caledonia, New Zealand, Niue, Norfolk Island, Northern Mariana Islands, Palau, Papua New Guinea, Philippines, Pitcairn Islands, Samoa, Singapore, Solomon Islands, Thailand, Tokelau, Tonga, Tuvalu, Vanuatu, Vietnam, and Wallis and Futuna