

2012 CFS Survey Methodology

Revised: Feb 2016

Table of Contents

- Overview..... 1
- Objective..... 2
- Industry coverage..... 2
 - Advance survey..... 2
 - CFS industries..... 2
- Shipment coverage..... 5
- Sample design overview..... 5
- First stage – establishment selection..... 5
 - Stratification..... 6
 - Determining the sample sizes, stratifying by MOS size class, and sample selection..... 8
- Second stage – reporting week selection..... 9
- Third stage – Shipment selection..... 9
- Data collection..... 9
 - Commodity coding changes for 2012..... 10
- Imputation of Shipment Value or Weight..... 11
- Mileage calculation..... 12
 - Methodological changes to Mileage Calculation for the 2012 CFS..... 12
 - Air versus Parcel mode..... 12
 - Routing a shipment when Mode is Other, Unknown, or Missing..... 13
 - Private truck versus For-hire truck..... 13
- Estimation..... 13
- Reliability of Estimates..... 15
 - Unpublished estimates..... 15
- Sampling error..... 15
- Non-Sampling error..... 17
 - Nonresponse..... 17
 - Response rates..... 18
- Disclosure avoidance..... 19
- Comparability of Estimates..... 19
 - Geographic area changes..... 19
 - Industry changes..... 20
 - Mode changes..... 20
 - Routing software changes..... 21
 - Commodity coding changes..... 21
 - Application of noise infusion..... 21
 - Sampling variability and nonresponse..... 22

OVERVIEW

The Commodity Flow Survey (CFS) is a joint effort by the Bureau of Transportation Statistics (BTS) and the U.S. Census Bureau, U.S. Department of Commerce. The survey is the primary source of national and state-level data on domestic freight shipments by establishments in mining, manufacturing, wholesale, auxiliaries, and selected retail and services trade industries located in the 50 states and the District of Columbia. Data are provided on the type, origin and destination, value, weight, modes of transportation, distance shipped, and ton-miles of

commodities shipped. The CFS is conducted every five years as part of the Economic Census. It provides a modal picture of national freight flows, and represents the only publicly available source of commodity flow data for the highway mode. The CFS was conducted in 1993, 1997, 2002, 2007 and most recently in 2012.

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

OBJECTIVES

The primary objective for the 2012 CFS was to estimate shipping volumes (value, tons, and ton-miles) by commodity and mode of transportation at varying levels of geographic detail. A secondary objective was to estimate the volume of shipments moving from one geographic area to another (i.e., flows of commodities between states, regions, etc.) by mode and commodity. A detailed description of the survey coverage and sample design for the 2012 CFS is provided below.

INDUSTRY COVERAGE

The 2012 CFS covers business establishments with paid employees that are located in the United States and are classified using the 2007 North American Industry Classification System (NAICS) in mining, manufacturing, wholesale, and selected retail and services trade industries, namely, electronic shopping and mail-order houses, fuel dealers, and publishers. Additionally, the survey covers auxiliary establishments (i.e., warehouses and managing offices) of multi-establishments companies.

Advance Survey

For the 2012 CFS, a targeted advance survey was conducted in 2011 to improve the quality of the data on the frame for certain industries or types of establishments. The groups included in this advance survey were:

Advance Survey Group	Number of Establishments
Auxiliaries (NAICS 484. 4931. 551114)	34,985
Small electronic shopping mail order establishments (NAICS 4541)	13,431
Small publishers (NAICS 5111)	11,804
Large establishments	39,608
Total	99,828

For the first three groups, the purpose was to identify those establishments that actually conduct shipping activities. In these groups, surveyed establishments that reported that they did not conduct any shipping activity were excluded from the eventual CFS sample universe. For large establishments the objective was to obtain an accurate measure of their shipping activity.

CFS Industries

In-scope industries for the 2012 CFS were selected based on the 2007 NAICS. Industries included in the 2007 and 2002 CFS were selected based on the 2002 and 1997 versions of the NAICS, respectively. The industries in the

1997 CFS and the 1993 CFS were selected based on the 1987 Standard Industrial Classification System (SIC) and, although attempts were made to maintain similar coverage among the SIC based surveys (1993 and 1997) and the NAICS based surveys (2002, 2007, and 2012), there have been some changes in industry coverage due to the conversion from SIC to NAICS. Most notably, coverage of the logging industry changed from an in-scope Manufacturing (SIC 2411) to the out-of-scope sector of Agriculture, Forestry, Fishing, and Hunting under NAICS 1133. Also, publishers were reclassified from Manufacturing (SIC 2711, 2721, 2731, 2741, and part of 2771) to Information (NAICS 5111 and 51223) and were excluded in the 2002 CFS. The 2007 and 2012 CFS, however, includes publishers and retail fuel dealers.

The (2007) NAICS industries covered in the 2012 CFS are listed in the following table:

NAICS Code	Description
212	Mining (Except Oil and Gas)
311	Food Manufacturing
312	Beverage and Tobacco Product Manufacturing
313	Textile Mills
314	Textile Product Mills
315	Apparel Manufacturing
316	Leather and Allied Product Manufacturing
321	Wood Product Manufacturing
322	Paper Manufacturing
323 ¹	Printing and Related Support Activities (except 323122)
324	Petroleum and Coal Products Manufacturing
325	Chemical Manufacturing
326	Plastics and Rubber Products Manufacturing
327	Nonmetallic Mineral Product Manufacturing
331	Primary Metal Manufacturing
332	Fabricated Metal Product Manufacturing
333	Machinery Manufacturing
334	Computer and Electronic Product Manufacturing
335	Electrical Equipment, Appliance, and Component Manufacturing
336	Transportation Equipment Manufacturing
337	Furniture and Related Product Manufacturing

339	Miscellaneous Manufacturing
4231 ²	Motor vehicle and parts merchant wholesalers
4232 ²	Furniture and home furnishing merchant wholesalers
4233 ²	Lumber and other construction materials merchant wholesalers
4234 ²	Commercial equip. merchant wholesalers
4235 ²	Metal and mineral (except petroleum) merchant wholesalers
4236 ²	Electrical and electronic goods merchant wholesalers
4237 ²	Hardware and plumbing merchant wholesalers
4238 ²	Machinery, equipment, and supplies merchant wholesalers
4239 ²	Miscellaneous durable goods merchant wholesalers
4241 ²	Paper and paper product merchant wholesalers
4242 ²	Drugs and druggists' sundries merchant wholesalers
4243 ²	Apparel, piece goods, and notions merchant wholesalers
4244 ²	Grocery and related product merchant wholesalers
4245 ²	Farm product raw material merchant wholesalers
4246 ²	Chemical and allied products merchant wholesalers
4247 ²	Petroleum and petroleum products merchant wholesalers
4248 ²	Beer, wine, and distilled alcoholic beverage merchant wholesalers
4249 ²	Miscellaneous nondurable goods merchant wholesalers
4541	Electronic Shopping and Mail-Order Houses
45431	Fuel Dealers
4841 ³	General Freight Trucking
4842 ³	Specialized Freight Trucking
4931 ³	Warehousing and Storage
5111 ⁴	Newspaper, Periodical, Book, and Directory Publishers
551114 ⁵	Corporate, Subsidiary, and Regional Managing Offices

¹ Excludes Pre-Press Services (NAICS 323122).

² Wholesale establishments exclude manufacturers sales offices and own brand importers.

³ Includes only captive warehouses that provide storage and shipping support to a single company. Warehouses offering their services to the general public and other businesses are excluded. NAICS 4841 and 4842 are new industries to the 2012 CFS. For tabulation and publication purposes, NAICS 484 is grouped with NAICS 4931.

⁴ In 2007, NAICS 51223 Music Publishers was tabulated and published in NAICS 5111. However, for the 2012 cycle, NAICS 51223 was not sampled.

⁵ Includes only those establishments in NAICS 551114 with shipping activity.

Notes:

Excluded industries: Foreign establishments, establishments classified in transportation, construction, and most retail and services industries are excluded. Other industry areas that are not covered, but may have significant shipping activity, include agriculture and government. For agriculture, specifically, the CFS does not cover shipments of agricultural products from the farm site to the processing centers or terminal elevators (most likely short-distance local movements), but does cover the shipments of these products from the initial processing centers or terminal elevators onward.

General exclusions: Data for government-operated establishments are excluded from the CFS. These include public utilities, publically-operated bus and subway systems, public libraries, and government-owned hospitals. The CFS also excludes establishments or firms with no paid employees.

SHIPMENT COVERAGE

The CFS captures data on shipments originating from select types of business establishments located in the 50 states and the District of Columbia. The CFS does not cover shipments originating from business establishments located in Puerto Rico and other U.S. possessions and territories.

Likewise, shipments traversing the United States from a foreign location to another foreign location (e.g., from Canada to Mexico) are not included, nor are shipments from a foreign location to an initial U.S. location. However, imported products are included in the CFS from the point that they leave the importer's initial U.S. location for shipment to another location. Shipments that are shipped through a foreign territory with both the origin and destination in the United States are included in the CFS data. The mileages calculated for these shipments exclude the foreign country segments (e.g., shipments from New York to Michigan through Canada do not include any mileages for Canada). Export shipments are included, with the domestic destination defined as the U.S. port, airport, or border crossing of exit from the United States. See the Mileage Calculation section for additional detail on how mileage estimates were developed.

SAMPLE DESIGN OVERVIEW

The sample for the 2012 CFS was selected using a three-stage design in which the first-stage sampling units were establishments, the second-stage sampling units were groups of four 1-week periods (reporting weeks) within the survey year, and the third-stage sampling units were shipments.

FIRST STAGE – ESTABLISHMENT SELECTION

To create the first-stage sampling frame, a subset of establishment records (as of July 2011) was extracted from the Census Bureau's Business Register. The Business Register is a database of all known establishments located in the United States or its territories. An establishment is a single physical location where business transactions take place or services are performed. Establishments located in the United States, having nonzero payroll in 2010, and classified in mining (except oil and gas extraction), manufacturing, wholesale, electronic shopping and mail order,

fuel dealers, and publishing industries, as defined by the 2007 NAICS, were included on the sampling frame. Certain manufacturers (Prepress services) and wholesalers (manufacturers' sales offices, agents and brokers, and certain importers) were excluded from the frame.

Auxiliary establishments (e.g. truck transportation facilities, warehouses, and central administrative offices) with shipping activity were also included on the sampling frame. Auxiliary establishments are establishments that are primarily involved in rendering support services to other establishments within the same company, instead of for the public, government, or other business firms. All other establishments included on the sampling frame are referred to as nonauxiliary establishments.

Establishments classified in forestry, fishing, utilities, construction, and all other transportation, retail, and services industries were not included on the sampling frame. Farms and government-owned entities (except government-owned liquor stores) were also excluded from the sampling frame. The resulting frame comprised approximately 716,000 establishments as shown in the table below.

Trade Area	Establishments on Frame	
	2012 CFS	2007 CFS
Mining	6,543	6,789
Manufacturing	305,805	327,826
Wholesale	345,511	356,477
Retail	27,697	25,190
Services	15,599	22,539
Auxiliaries	14,959	14,878
Total	716,114	753,699

For each establishment, sales, payroll, number of employees, a 6-digit NAICS code, name and address, and a primary identifier were extracted, and a measure of size was computed. The measure of size was designed to approximate an establishment's annual total value of shipments for the year 2009.

All of the establishments included on the sampling frame had state and county geographic codes. We used these codes to assign each establishment to one of the 83 CFS metropolitan areas (CFS Areas) defined as a state part of a metropolitan statistical area (MSA) or combined statistical area (CSA). Establishments not located in one of these specified metropolitan areas (MAs) were assigned to a Rest of State (ROS) CFS Area.

Stratification

The sampling frame was stratified by geography, industry, and measure-of-size (MOS) class (with some exceptions for auxiliary establishments and hazardous materials establishments, as described below). The geography by industry cells form the primary strata for the main part of the sample.

Geographic strata were defined by a combination of the 50 states, the District of Columbia, and specific metropolitan areas (called CFS areas) selected based on their population and importance as transportation gateways. These CFS Areas were defined using the 2009 Office of Management and Budget’s definitions. All other metropolitan areas were collapsed with the non-metropolitan areas within the state into Rest of State CFS Area strata. When a metropolitan area crossed state boundaries, we considered the size of each state part of the metropolitan area when determining whether or not to create strata in each state in which the MA was defined. For example, the Chicago CSA makes up two CFS Areas: the IL part and the IN part. The WI part of Chicago was too small to be a separate CFS Area and was combined into the Rest of Wisconsin CFS Area. The table below (second column) summarizes the number of CFS Areas used for sampling by type.

Geographic Stratum (CFS Area) Type	Number of Sampled CFS Areas	Number of Published CFS Areas
Actual CSA or MSA (state part)	83	82
CFS area = state (DC, RI)	2	2
ROS = whole state (AK,AR,ID,IA,ME,MS,MT,NM,ND,SD,VT,WV,WY)	13	13
ROS < whole state	36	35
Total number of CFS areas	134	132

Between the time the CFS sample of establishments was selected and publication of the data, there were changes to the definitions of the metropolitan areas used by the CFS. For sampling purposes, the CFS Areas were defined using the 2009 OMB metropolitan area definitions. For tabulation and publication, the 2013 OMB definitions were used to define the CFS Areas. As a result, two CFS Areas used for sampling (Stockton, CA and Remainder of NJ) disappeared and, for many others, the counties making up the CFS Areas changed. The rightmost column of the table above shows the number of CFS Areas for which data were eventually published.

The industry strata were defined as follows. Within each of the geographic strata, we defined 48 industry groups based on the 2007 NAICS codes:

- Three mining (four-digit NAICS).
- Twenty-one manufacturing (three-digit NAICS).
- Eighteen wholesale (four-digit NAICS).
- Two retail (NAICS 4541 and 45431).
- One services (NAICS 5111).
- Three auxiliary (combinations of NAICS 484, 4931 and 551114).

For auxiliaries that responded to the Advance Survey and were found to be shippers, 134 primary strata were created, one in each geographic stratum, combining NAICS 484, 4931, and 551114. For auxiliary establishments that did not respond to the Advance Survey, two national strata were created as follows:

- One stratum for nonresponding truck transportation establishments and warehousing and storage establishments (NAICS 484 and NAICS 4931).
- One stratum for nonresponding corporate, subsidiary, and regional managing offices establishments (NAICS 551114).

In order to produce good estimates of shipments of hazardous materials (HAZMAT), twenty 6-digit NAICS industries with high amounts of HAZMAT shipments were identified and used to form primary strata. The 2007 CFS data were used to identify these industries and in general, these industries were chosen because:

- They had a large (weighted) total value or total tonnage of hazardous materials.
- A high percentage of their (unweighted) shipments were HAZMAT shipments.

Thirteen of the 20 industries were made certainty strata and the remaining seven industries were made into primary strata defined by state and the 6-digit NAICS code.

The table below shows the number and types of primary strata for the main, auxiliary, and HAZMAT parts of the sample. Note that we are counting the number of strata before they are further stratified by MOS size class.

Part of the sample	Number of Primary Strata
Main part of the sample	6,030 (134 CFS areas x 45 industries)
Auxiliary part of the sample	
Responders to the Advance Survey	134 (134 CFS areas x 1 industry)
Nonresponders to the Advance Survey	2 (2 industries)
HAZMAT part of the sample	
Certainty (take-all) strata	13 (13 6-digit NAICS codes)
Noncertainty strata	357 (51 states (incl. DC) x 7 6-digit NAICS codes)

Determining the sample sizes, stratifying by MOS size class, and sample selection

The total desired sample size for the first stage sample was approximately 100,000 establishments and was fixed due to budget constraints. Therefore, in addition to defining the strata, a sample size was determined for each primary stratum. This was performed as follows:

- A target coefficient of variation (CV) for estimated total MOS was assigned to each primary stratum (geography by industry cell).
- Within each primary stratum, substrata defined by MOS were developed to minimize the sample size needed to achieve the target CV. The establishments in the largest MOS size class were taken with certainty. For the noncertainty substrata, the sample was allocated according to the Neyman allocation, since the Neyman allocation minimizes the sample size needed to achieve a target CV.
- Once the minimum sample sizes for each primary stratum were determined, these were added together and compared to the desired total sample size of 100,000. If the total was not close enough to 100,000, we multiplied all of the target CVs by a fixed factor and repeated the process until the total sample size was close to 100,000.
- The establishments in the geography by industry by MOS size class substrata were selected by simple random sampling without replacement. The total sample size was 102,565 establishments of which 46,265 were selected with certainty (see the table below).

Primary Strata Type	2012 Frame		2012 Sample			
	Estabs	Total MOS (\$mil)	Total Sample		Certainty Component	
			Estabs	MOS of Sampled Estabs (\$mil)	Estabs	MOS of Certainty Estabs (\$mil)
Main	680,128	\$8,361,138	95,678	\$6,215,482	42,187	\$5,620,044
Auxiliary	14,959	1,330,769	2,433	1,186,608	1,121	1,087,152
HAZMAT	21,027	775,739	4,454	685,595	2,957	669,835
Total	716,114	\$10,467,646	102,565	\$8,087,685	46,265	\$7,377,031

SECOND STAGE – REPORTING WEEK SELECTION

The frame for the second stage of sampling consisted of the 52-weeks in 2012. Each establishment selected into the 2012 CFS sample was systematically assigned to report for four reporting weeks, one in each quarter of the reference year (2012). Each of the 4-weeks was in the same relative position in the quarter. For example, an establishment might have been requested to report data for the 5th, 18th, 31st, and 44th weeks of the reference year. In this instance, each reporting week corresponds to the 5th week of each quarter. Prior to assignment of weeks to establishments, we sorted the selected sample by primary stratum (geography x industry) and measure-of-size.

THIRD STAGE – SHIPMENT SELECTION

For each of the four reporting weeks in which an establishment was asked to report, the respondent was requested to construct a sampling frame consisting of all shipments made by the establishment in the reporting week. Each respondent was asked to count or estimate the total number of shipments comprising the sampling frame and to record this number on the questionnaire. For each assigned reporting week, if an establishment made more than 40 shipments during that week, we asked the respondent to select a systematic sample of the establishment's shipments and to provide us with information only about the selected shipments. By design, this systematic sample consisted of between 20 and 40 shipments. If an establishment made 40 or fewer shipments during that week, we asked the respondent to provide information about all of the establishment's shipments made during that week; i.e., no sampling was required.

DATA COLLECTION

Each establishment selected into the CFS sample was mailed a questionnaire for each of its four assigned reporting weeks, that is, an establishment was sent a questionnaire once every quarter of 2012. For a given establishment, the respondent was asked to provide the following information about each of the establishment's reported shipments:

- Shipment ID number.
- Shipment date (month, day).
- Shipment value.
- Shipment weight in pounds.
- Commodity code from Standard Classification of Transported Goods (SCTG) manual.
- Commodity description.

- An indication of whether the shipment was temperature controlled.
- United Nations or North American (UN/NA) number for hazardous material shipments.
- U.S. destination (city, state, zip code)—or gateway for export shipment.
- Modes of transport.
- An indication of whether the shipment was an export.
- City and country of destination for exports.
- Export mode.

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

In addition, establishments were asked to provide information about the use and extent of use of rush delivery services.

Commodity Coding Changes for 2012

The definitions of several SCTG codes changed or new codes added between 2007 and 2012. These are:

SCTG	Type of Chg	Description
07-R	Definition	Prior to 2012 CFS, Fats and oils were all classified under Commodity Code 07. For CFS 2012 CFS, oils and fats treated for use as biodiesel moved to Commodity Code 18 under Fuel Oils.
074-R	Definition	Prior to the 2012 CFS, fats and oils intended for use as biodiesel were not specifically identified, but were included in Commodity Code 074. In the 2012 CFS, fats and oils intended for use as biodiesel were specified and classified in under Commodity Code 182 (biodiesel and blends of biodiesel).
0743-R	Definition	Prior to the 2012 CFS, fats and oils intended for use as biodiesel were not specifically identified, but were included in Commodity Code 0743. In the 2012 CFS, fats and oils treated for use as biodiesel were specified and classified under Commodity Code 182.
08-R	Definition	Prior to the 2012 CFS, alcohols intended for use as fuel were not specifically identified, and were included under SCTG 08. In the 2012 CFS, ethanol for fuel moved to SCTG 17. Additionally, beverages and denatured alcohol were more clearly identified.
083-R	Definition	Prior to the 2012 CFS, denatured alcohol of more than 80% alcohol by volume was included in Commodity Code 083. In the 2012 CFS, denatured alcohol of more than 80% by volume was moved to Commodity Code 084, and ethanol for use as biofuel was moved to Commodity Codes 175 and 176.
0831-R	Definition	Prior to the 2012 CFS, both Denatured ethyl alcohol, and undenatured ethyl alcohol of more than 80% alcohol by volume were included in Commodity Code 0831. In the 2012 CFS, denatured alcohol of more than 80% by volume was moved to Commodity Code 0841, and ethanol for use as biofuel was specified and moved to Commodity Codes 175 and 176.
084	New	Denatured ethyl alcohol, not for ingestion or use as biofuel
17-R	Definition	Prior to 2012 CFS, Denatured ethyl alcohol, and undenatured ethyl alcohol were all classified under SCTG 08. For CFS 2012 CFS, ethanol that is used for fuel was identified and removed from SCTG 08 to SCTG 17 under fuel alcohols. Also, kerosene, which prior to 2012 CFS, was included in Commodity Code 19, was moved under Commodity Code 17.
171-R	Definition	Prior to the 2012 CFS, Commodity Code 171 only included gasoline, and blend of gasoline and ethanol were not identified. In the 2012 CFS, Commodity Code 171 includes gasoline, and mixtures of up to 10% ethanol and gasoline.
172-R	Definition	Prior to the 2012 CFS, kerosene was included in Commodity Code 192, and type A jet fuel was classified under Commodity Code 172.. In the 2012 CFS, all kerosene are classified under Commodity Code 172.
1720-R	Definition	Prior to the 2012 CFS, kerosene was included in Commodity Code 192, and type A jet fuel was classified under Commodity Code 1720. In the 2012 CFS, all kerosene is classified under Commodity Code 1720.
175	New	Ethanol, ethanol blends of more than 10 percent ethanol, and other fuel alcohols
176	New	Ethanol, for use as biofuels
18-R	Definition	Prior to the 2012 CFS, fats and oils intended for use as fuel were not identified as such, and were included in Commodity Code 07. In the 2012 CFS, such fats and oils were identified as biodiesel and were moved under Commodity Code 18.

SCTG	Type of Chg	Description
181	New	Fuel oils including diesel, distillate heating oil, and bunker c (excludes biodiesel)
182	New	Blends of fuel oils including 5% or less biodiesel by volume (b5, or less)
1821	New	Blends of fuel oils with more than 5% biodiesel by volume, except b100
1822	New	Biodiesel derived from vegetable oils or animal fats, b100 (excludes mixtures of biodiesel and diesel fuel)

IMPUTATION OF SHIPMENT VALUE OR WEIGHT

To correct for nonresponse or an unacceptable value in either the value or weight item for a given shipment, the missing or unacceptable value is replaced by a predicted value obtained from a donor imputation model. Such a shipment is considered a “recipient” if its commodity code is valid and one of the two data items (either shipment value or shipment weight) is reported, greater than zero, and the shipment is otherwise useable. The recipient’s missing or unacceptable data item is imputed as follows:

First a donor shipment for a given recipient with the same 5-digit SCTG is selected at random from a pool of potential donor shipments (those with valid SCTGs and with reported and usable shipment value and weight). The donor pools are summarized below in order of preference (the lowest numbered donor pool with a matching shipment is used).

Donor Pool	Description of Donor Pool Shipments
1	From same establishment and in the same detailed shipment size class
2	From same company and in the same detailed shipment size class
3	From same geographic area and in the same detailed shipment size class
4	From same establishment and in the same broad shipment size class
5	From same company and in the same broad shipment size class
6	From same geographic area and in the same broad shipment size class
7	From same establishment (<i>no restriction on shipment size</i>)
8	From same company (<i>no restriction on shipment size</i>)
9	From same geographic area (<i>no restriction on shipment size</i>)

Then, the donor’s value and weight data are used to calculate a ratio, which is applied to the recipient’s reported item, to impute the item that is missing or failed edit. If a donor cannot be found in one of the nine donor pools then the recipient’s item is imputed using the median value-to-weight ratio computed using all shipments in the same SCTG as that of the recipient.

Approximately three percent of shipment values are imputed and similarly, approximately three percent of shipment weights are imputed.

MILEAGE CALCULATION

The CFS does not ask respondents to report the distance traveled for each shipment. Therefore, shipment mileages were calculated using GeoMiler, which is a routing tool developed by BTS specifically for CFS mileage calculations. GeoMiler used current Geographic Information System (GIS) technology and spatial multimodal network databases and integrated map-visualization features with route solvers to handle many alternative multimodal combinations. The software used algorithms that found the quickest path over spatial representations of the U.S. highway, railway, waterway, and airway networks. For waterborne export shipments, GeoMiler used a waterborne commerce database from the U.S. Army Corps of Engineers to route freight originating in the U.S. via the deep sea (ocean). For airborne export shipments, GeoMiler used a newly developed air export network from the BTS Office of Airline Information.

For a domestic shipment, the mileage was calculated between the centroid (center of a geographic area) of the origin ZIP Code and the centroid of the destination ZIP Code. For shipments where the origin and destination were within the same ZIP code (Intra-ZIP shipments), the square root of the total ZIP code area in square miles was used as an estimate for the distance shipped.

For multimodal shipments (shipments involving more than one mode, such as truck-rail shipments), spatial joins (intermodal transfer links) were used to connect the individual modal networks together for routing purposes. An intermodal terminals database and a number of terminal transfer models were developed at BTS to identify likely transfer points. An algorithm was used to find the minimum impedance path between a shipment's origin ZIP Code to the transfer point and then from the transfer point to the destination ZIP Code. Thus, for multimodal shipments, the cumulative length of the spatial joins, plus links on the path, was used for estimating distances.

The mileage for an export shipment was calculated between the centroid of the origin ZIP Code and the border crossing on the path of minimum impedance to the foreign destination country (foreign city in the case of Canada and Mexico). Only the portion of the mileage measured within U.S. borders was included as domestic mileage in the CFS estimates.

Methodological Changes to Mileage Calculation for the 2012 CFS

With a valid origin and destination zip code, GeoMiler will calculate the distance traveled (in miles) by mode for each shipment reported in the CFS. The following types of methodological changes to mileage processing were incorporated in 2012:

- A shipment with a respondent-provided mode of Parcel must weigh 150 pounds or less; in addition, a shipment with a respondent-provided mode of Air was not given a weight restriction;
- A mode of transportation was imputed whenever a respondent provided a mode of Other, or Unknown, or otherwise failed to provide a modal response (missing mode) for a shipment;
- Private truck is considered a "short-haul" mode; hence Private truck shipments were not routed more than 500 miles during shipment routing.

Air versus Parcel Mode

According to the 2007 CFS Instruction Guide, an air shipment was defined as a shipment that weighed 100 pounds or more. During mileage processing for the 2007 CFS, an Air shipment was manually converted to Parcel if the weight of the shipment was less than 100 pounds.

However, airlines do not necessarily have minimum weight restrictions when transporting cargo. Hence, for the 2012 CFS, the definition of an Air shipment was changed. As a result, an Air shipment was acceptable as provided by the respondent, regardless of weight.

Furthermore, for the 2012 CFS, Parcel shipments conformed to the definition used by the parcel industry that a parcel is a shipment of 150 pounds or less. For shipments submitted by the respondent with mode of Parcel and a weight above 150 pounds, GeoMiler changed the mode to For-Hire Truck during mileage processing.

Routing a Shipment When Mode Is Other, Unknown, or Missing

On the survey form, respondents were given the following choices for mode of transport: Air, Highway (Private truck or For-hire truck), Rail, Waterway (Inland water or Deep sea), Parcel, Pipeline, Other Mode (meaning none of the above), or Unknown.

During the 2007 CFS mileage processing, 2.4% of shipments had a respondent-provided mode of Unknown or Other, and an additional 2.1% had no reported mode at all. In these situations, the mode of transport was imputed. For 2012 CFS mileage processing, if the shipment weighed less than 80,000 pounds, it was routed via Highway mode as For-hire truck; if the shipment weighed 80,000 pounds or more, it was routed via Rail mode.

Private Truck versus For-Hire Truck

Shipments via Private truck are generally "short-haul" in nature. Because of the number of shipments exceeding this norm in the 2007 CFS, Census Bureau analysts researched the Private truck shipments at or above 500 miles. In almost all cases, the mode should have been reported as For-hire truck instead of Private truck.

Consequentially, for 2012 CFS GeoMiler mileage processing, Private truck was converted to For-hire truck if the shipment mileage was equal to or greater than 500 miles, regardless of the commodity being transported. The 2012 CFS preliminary data shows a decrease from 2007 in average miles per shipment for Private truck with an average of 46 miles per shipment.

ESTIMATION

Estimated totals (e.g., value of shipments, tons, ton-miles) are produced as the sum of weighted shipment data (reported or imputed). Percent change and percent-of-total estimates are derived using the appropriate estimated totals. Estimates of average miles per shipment are computed by dividing an estimate of the total miles traveled by the estimated number of shipments.

Each shipment has associated with it a single tabulation weight, which was used in computing all estimates to which the shipment contributes. The tabulation weight is a product of seven different component weights. A description of each component weight follows.

CFS respondents provided data for a sample of shipments made by their respective establishments in the survey year. For each establishment, we produced an estimate of that establishment's total value of shipments for the entire survey year. To do this, we used four different weights: the shipment weight, the shipment nonresponse weight, the quarter weight, and the quarter nonresponse weight. Three additional weights are then applied to produce estimates representative of the entire universe. These are the establishment-level adjustment weight, the establishment (or first-stage sample) weight, and the nonresponse post-stratification adjustment weight.

Like establishments, we identified shipments as either certainty or noncertainty. (See the Nonsampling Error section below for a description of how certainty shipments were identified.) For noncertainty shipments, the **shipment weight** was defined as the ratio of the total number of shipments (as reported by the respondent) made by an establishment in a reporting week to the number of sampled shipments the respondent reported on the questionnaire for the same week. This weight uses data from the sampled shipments to represent all the establishment's shipments made in the reporting week. However, a respondent may have failed to provide sufficient information about a particular sampled shipment. For example, a respondent may not have been able to provide value, weight, or a destination for one of the sampled shipments. If this data item could not be imputed or otherwise obtained, then this shipment did not contribute to tabulations and was deemed unusable. (A usable shipment is one that has valid entries for value, weight, and origin and destination ZIP Codes.) To account for these unusable shipments, we applied the **shipment nonresponse weight**. For noncertainty shipments from a particular establishment's reporting week, this weight is equal to the ratio of the number of sampled shipments for the reporting week to the number of usable shipments for the same week. The shipment weight for certainty shipments from a particular establishment's reporting week is equal to one.

The **quarter weight** inflates an establishment's estimate for a particular reporting week to an estimate for the corresponding quarter. For noncertainty shipments, the quarter weight is equal to 13. The quarter weight for most certainty shipments is also equal to 13. However, if a respondent was able to provide information about all large (or certainty) shipments made in the quarter containing the reporting week, then the quarter weight for each of these shipments was set to one. For each establishment, the quarterly estimates were added to produce an estimate of the establishment's value of shipments for the entire survey year. Whenever an establishment did not provide the Census Bureau with a response for each of its four reporting weeks, we computed a quarter nonresponse weight. The **quarter nonresponse weight** for a particular establishment is defined as the ratio of the number of quarters for which the establishment was in business in the survey year (usually four) to the total number of quarters (reporting weeks) for which we received usable shipment data from the establishment.

Using these four component weights and the reported (or imputed) shipment values, we computed an estimate of each establishment's value of shipments for the entire survey year. We then multiplied this estimate by a factor that adjusts this estimated value to a measure of the establishment's value of shipments or receipts obtained from the 2012 Economic Census. This weight, the **establishment-level adjustment weight**, attempts to correct for any sampling or nonsampling errors caused by the selection of specific reporting weeks or that occur during the sampling of shipments by the respondent.

The adjusted value of shipments estimate for an establishment was then weighted by the **establishment weight**. This weight is equal to the reciprocal of the establishment's probability of being selected into the first stage sample (see Sample Design below).

A final adjustment, the **nonresponse post-stratification adjustment weight**, adjusts the weighted shipment value (using all prior weighting factors) to the levels of tabulated revenue data from the 2012 Economic Census for specified post-stratification cells. This accounts for:

- Establishments which did not respond to the survey or from which we did not receive any usable shipment data.
- Changes in the universe of establishments between the time the first-stage sampling frame was constructed (2011) and the year in which the data were collected (2012).

For the preliminary 2012 CFS estimates, the nonresponse post-stratification cells were defined by industry categories, typically by 3-digit NAICS codes (for Manufacturing) or 4-digit NAICS codes (all other industries). There were approximately 45 nonresponse post-stratification cells.

For the final 2012 CFS estimates, the nonresponse post-stratification cells were defined by state by industry categories. The industry categories were the same as those described above for the preliminary estimates. There were approximately 2300 state by industry nonresponse post-stratification cells.

RELIABILITY OF THE ESTIMATES

The estimates presented by the 2012 CFS may differ from the actual, unknown population values. The difference between the estimate and the population value is known as the total error of the estimate. When describing the accuracy of survey results, it is convenient to discuss total error as the sum of sampling error and nonsampling error. Sampling error is the average difference between the estimate and the result that would be obtained from a complete enumeration of the sampling frame conducted under the same survey conditions. Nonsampling error encompasses all other factors that contribute to the total error of a sample survey estimate.

The sampling error of the estimates in this publication can be estimated from the selected sample because the sample was selected using probability sampling. Common measures related to sampling error are the sampling variance, the standard error, and the coefficient of variation (CV). The sampling variance is the squared difference, averaged over all possible samples of the same size and design, between the estimator and its average value. The standard error is the square root of the sampling variance. The CV expresses the standard error as a percentage of the estimate to which it refers. For percentage estimates, such as percentage change or percentage of a total, the standard error of the estimate is provided.

Nonsampling errors are difficult to measure and can be introduced through inadequacies in the questionnaire, nonresponse, inaccurate reporting by respondents, errors in the application of survey procedures, incorrect recording of answers, and errors in data entry and processing. In conducting the 2012 CFS, every effort has been made to minimize the effect of nonsampling errors on the estimates. Data users should take into account both the measures of sampling error and the potential effects of nonsampling error when using these estimates.

Unpublished Estimates

Estimates that had high sampling variability or poor response quality were not published. Some of these unpublished estimates can be derived directly from the CFS tables by subtracting published estimates from their respective totals. However, the (unpublished) estimates obtained by such subtraction would be subject to poor response, high sampling variability, or other factors that may make them potentially misleading. Estimates derived in this manner should not be attributed to the Census Bureau.

Individuals who use estimates in these tables to create new estimates should cite the Census Bureau as the source of only the original estimates.

More detailed descriptions of sampling and nonsampling errors for the 2012 CFS are provided in the following sections.

SAMPLING ERROR

Because the estimates are based on a sample, exact agreement with results that would be obtained from a complete enumeration of all shipments made in 2012 from all establishments included on the sampling frame using the same enumeration procedures is not expected. However, because probability sampling was used at each stage of selection, it is possible to estimate the sampling variability of the survey estimates. For CFS estimates, sampling variability arises from each of the three stages of sampling.

The particular sample of shipments used in this survey is one of a large number of samples of the same size that could have been selected using the same design. If all possible samples had been surveyed under the same conditions, an estimate of a population parameter of interest could have been obtained from each sample. These samples give rise to a distribution of estimates for the population parameter. A statistical measure of the variability among these estimates is the standard error, which can be estimated from any one sample. The standard error is defined as the square root of the variance. The coefficient of variation (or relative standard error) of an estimator is the standard error of the estimator divided by the estimator. For the CFS, the coefficient of variation also incorporates the effect of the noise infusion disclosure avoidance method (see Disclosure Avoidance below). Note that measures of sampling variability, such as the standard error and coefficient of variation, are estimated from the sample and are also subject to sampling variability and technically, we should refer to the estimated standard error or the estimated coefficient of variation of an estimator. However, for the sake of brevity, we have omitted this detail. It is important to note that the standard error only measures sampling variability. It does not measure systematic biases of the sample. The Census Bureau recommends that individuals using estimates contained in this report incorporate this information into their analyses, as sampling error could affect the conclusions drawn from these estimates.

An estimate from a particular sample and the standard error associated with the estimate can be used to construct a confidence interval. A confidence interval is a range about a given estimator that has a specified probability of containing the result of a complete enumeration of the sampling frame conducted under the same survey conditions. Associated with each interval is a percentage of confidence, which is interpreted as follows. If, for each possible sample, an estimate of a population parameter and its approximate standard error were obtained, then:

1. For approximately 90 percent of the possible samples, the interval from 1.833 standard errors below to 1.833 standard errors above the estimate would include the result as obtained from a complete enumeration of the sampling frame conducted under the same survey conditions.
2. For approximately 95 percent of the possible samples, the interval from 2.262 standard errors below to 2.262 standard errors above the estimate would include the result as obtained from a complete enumeration of the sampling frame conducted under the same survey conditions. The 1.833 and 2.262 values, used to compute the 90% and 95% confidence intervals, are taken from the t-distribution with nine degrees of freedom. This takes into account the uncertainty in the estimates of the CVs produced using the random group method with ten random groups.

To illustrate the computation of a confidence interval for an estimate of total value of shipments, assume that an estimate of total value is \$10,750 million and the coefficient of variation for this estimate is 1.8 percent, or 0.018. First obtain the standard error of the estimate by multiplying the value of shipments estimate by its coefficient of variation. For this example, multiply \$10,750 million by 0.018. This yields a standard error of \$193.5 million. The upper and lower bounds of the 90-percent confidence interval are computed as \$10,750 million plus or minus 1.833 times \$193.5 million or \$354.7 million. Consequently, the 90-percent confidence interval is \$10,395 million to \$11,105 million. If corresponding confidence intervals were constructed for all possible samples of the same size

and design, approximately 9 out of 10 (90 percent) of these intervals would contain the result obtained from a complete enumeration.

NONSAMPLING ERROR

Nonsampling error encompasses all other factors that contribute to the total error of a sample survey estimate and may also occur in censuses. It is often helpful to think of nonsampling error as arising from deficiencies or mistakes in the survey process. In the CFS, nonsampling error can be attributed to many sources:

- Inability to obtain information about all units in the sample.
- Response errors.
- Differences in the interpretation of the questions.
- Mistakes in coding or keying the data obtained.
- Other errors of collection, response, coverage, and processing.

Although no direct measurement of the potential biases due to nonsampling error has been obtained, precautionary steps were taken in all phases of the collection, processing, and tabulation of the data in an effort to minimize their influence. The Census Bureau recommends that individuals using estimates in this report incorporate this information into their analyses, as nonsampling error could affect the conclusions drawn from these estimates.

Some possible sources of bias that are attributed to respondent-conducted sampling include:

- Misunderstanding the definition of a shipment.
- Constructing an incomplete frame of shipments from which to sample.
- Ordering the shipment sampling frame by selected shipment characteristics.
- Selecting shipment records by a method other than the one specified in the questionnaire's instructions.

The respondents who had reported a shipment with unusually large value or weight when compared to the rest of their reported shipments were often contacted for verification. In such cases, if we were able to collect information on all of the large shipments a respondent had made either for a particular reporting week or for the entire quarter, we then identified those large shipments as certainty shipments.

Nonresponse

A potential source of bias in the estimates is nonresponse. Nonresponse is defined as the inability to obtain all the intended measurements or responses from all units in the sample. Four levels of nonresponse can occur in the CFS:

- Item.
- Shipment.
- Quarter (reporting week).
- Establishment.

Item nonresponse occurs either when a particular shipment data item is unanswered or the response to the question fails computer or analyst edits. Nonresponse to the shipment value or weight items is corrected by imputation, which is the procedure by which a missing value is replaced by a predicted value obtained from an appropriate model. (See above for a description of the imputation procedure.)

Shipment, quarter, and establishment nonresponse describe the inability to obtain any of the substantive measurements about a sampled shipment, quarter, or establishment, respectively. Shipment and quarter nonresponse are corrected by reweighting (see the descriptions of the shipment and quarter nonresponse weights in the Estimation section above). Reweighting allocates characteristics to the nonrespondents in proportion to the characteristics observed for the respondents. The amount of bias introduced by this nonresponse adjustment procedure depends on the extent to which the nonrespondents differ, characteristically, from the respondents.

Establishment nonresponse is corrected during the estimation procedure by the nonresponse post-stratification adjustment weight. In most cases of establishment nonresponse, none of the four questionnaires have been returned to the Census Bureau after several attempts to elicit a response.

Response Rates

The CFS produces four different response rates: a participation response rate, a unit response rate, a weighted unit response rate, and a total quantity (item) response rate. The first three are based on the responses of the establishments selected into the survey. These unit response rates are shown in Table 1 below.

Table 1: 2012 CFS Unit Response Rates

Type of Response Rate	PRR, URR, WRR
Participation	57.0%
Unit	66.1%
Weighted Unit	76.7%

Participation Response Rate (PRR) - The Participation Response Rate is the total number of unweighted establishments that provided usable¹ data divided by the total number of establishments in the sample (102,565) (expressed as a percentage).

Unit Response Rate (URR) - The Unit Response Rate is defined as the ratio (expressed as a percentage) of the total unweighted number of establishments that provided usable data to the total number of establishments that were eligible (or potentially eligible) for data collection. URRs are indicators of the performance of the data collection process in obtaining usable responses.

Weighted Unit Response Rate (WRR) - The Weighted Unit Response Rate is defined as the percentage of the total weighted 2012 Economic Census adjusted receipts of establishments that provided usable data to the total weighted Economic Census adjusted receipts of establishments that were eligible (or potentially eligible) for data collection. This incorporates the size of the establishment as well as its establishment (first-stage sample) weight into the measure of response.

The fourth rate is based on the quality of the individual shipment data reported by the responding establishments. These total quantity response rates for the 2012 CFS are shown in Table 2 below.

Table 2: 2012 CFS Total Quantity Response Rates

CFS Variable	TQRR
VALUE	51.9%
TONS	50.9%
TON-MILES ⁶	63.2%

¹ "Usable data" means that an establishment provided at least one shipment that was used in the tabulation of published estimates.

⁶ For ton-miles (which is the product of shipment weight and distance shipped) the distance shipped component is derived from the respondent-reported destination zip code (see the Mileage Calculation section above). The respondent is not asked for the actual distance. This calculated distance is treated as equivalent-to-reported data for purposes of computing the TQRR for Ton-miles.

Total Quantity Response Rate (TQRR) - The Total Quantity Response Rate is defined as the percentage of the estimated (weighted) total of a given data item (VALUE, TONS, or TON-MILES) that is based on reported shipment data or from sources determined to be of equivalent-quality-to-reported data. The TQRR is an item-level indicator of the “quality” of each estimate. In contrast to the URR, these weighted response rates are computed for individual data items, so CFS produces several TQRRs.

The TQRR is the weighted proportion of the key estimates reported by responding establishments or obtained from equivalent quality sources. This measure incorporates the value of the individual shipment data items and the associated sampling and weighting factors.

DISCLOSURE AVOIDANCE

Disclosure is the release of data that have been deemed confidential. It generally reveals information about a specific individual or establishment or permits deduction of sensitive information about a particular individual or establishment. Disclosure avoidance is the process used to protect the confidentiality of the survey data provided by an individual or firm.

Using disclosure avoidance procedures, the Census Bureau modifies or removes the characteristics that put confidential information at risk of disclosure. Although it may appear that a table shows information about a specific individual or business, the Census Bureau has taken steps to disguise or suppress the original data while making sure the results are still useful. The techniques used by the Census Bureau to protect confidentiality in tabulations vary, depending on the type of data.

For the CFS the primary method of disclosure avoidance is Noise Infusion: Noise infusion is a method of disclosure avoidance in which the weighted values for each shipment are perturbed prior to tabulation by applying a random noise multiplier to shipment value and weight. Disclosure protection is accomplished in a manner that causes the vast majority of cell values to be perturbed by at most a few percentage points. For sample-based tabulations, such as CFS, the estimated relative standard error for a published cell includes both the estimated sampling error and the amount of perturbation in the estimated cell value due to noise. Other cells in the table may be suppressed because the quality of the data does not meet publication standards. By far, the most common reason for suppressing a cell is a high coefficient of variation (greater than 50 percent). These suppressed cells are shown with an “S” in the tables.

COMPARABILITY OF ESTIMATES

This section summarizes the coding and processing differences between the 2007 and 2012 surveys that limit the comparability of the published statistics or estimates across the survey years. Data users should exercise caution when comparing CFS data across any survey years.

Geographic Area Changes

Many CFS Areas, while similar in name from one survey to the next, are actually made up of slightly different sets of counties. For example, in 2007 the Chicago-Naperville-Michigan City, IL-IN-WI CFS Area (IL Part) consisted of 10 counties in Illinois. In 2012 the Chicago-Naperville, IL-IN-WI CFS Area (IL Part) was made up of 13 counties. Consequently, as a result of the change to the Chicago CFS Area, the number of counties included in the

Remainder of Illinois CFS Area was reduced. The addition of ten new CFS Areas in 2012 also reduced the number of component counties making up the remainders of the affected states. The table below provides a summary of the changes to CFS Areas from 2007 to 2012.

Summary of CFS Area Changes: 2007 to 2012

Type of CFS Area		Example	2007	2012	
			Total	Total	Number w/ Definition Changes (1)
Metropolitan Area	In 2007 and 2012	Portland (OR part)	74	74	46
	New for 2012	Portland (WA part)	-	10	N/A
Remainder of State	Full state	Idaho (= <i>Remainder of ID</i>)	15	13	31
	Less than full state	Remainder of OR	34	35	
Total			123	132	

(1) A Definition change means that there were changes to counties that made up the CFS Area.

Industry Changes

Industry coverage has changed slightly from survey year to survey year. The details of CFS industry coverage are described in the **Industry Coverage** section of the survey methodology. The most significant recent changes are:

- NAICS 484 was included as an in-scope auxiliary industry in 2012 but not 2007.
- NAICS 51223 was included as an in-scope publishing industry in 2007 but not in 2012.

The 2007 estimates were based on the industry classification of the sample establishments at the time those estimates were produced (Dec 2009). The 2007 or earlier estimates are never revised to account for subsequent industry classification changes to the sample establishments.

Mode Changes

There were changes to the detailed mode of transportation codes associated with water-borne shipments. The table below lists the water modes in 2007 and 2012. In addition to a new water mode (Multiple Waterways) in 2012, there were slight changes to the definitions of modes 08 and 10 that may have affected the respondent's choice of answer. See the 2007 and 2012 questionnaires and instruction guides for descriptions of the modes.

CFS Water Mode Codes

2007		2012	
Code	Meaning	Code	Meaning
07	Water	07	Water
08	Shallow Draft	08	Inland Water
09	Great Lakes	09	Great Lakes
10	Deep Draft	10	Deep Sea
		101	Multiple Waterways

The following methodological changes to mileage processing also affected mode assignment (and the shipment distance calculations).

- The maximum weight of a parcel shipment was limited to 150 pounds. In 2007 the limit was 1000 pounds. Shipments with weights above the maximum were re-assigned to a non-Parcel mode, usually a truck mode.
- For 2012, there was no minimum restriction on the weight of an air shipment. In 2007 air shipments with a weight of less than 100 pounds were reclassified as Parcel.
- Private truck shipments were not routed more than 500 miles during 2012 mileage calculation. In 2007 there was no mileage limit.
- In 2012 there was a major effort to re-code shipments, where a respondent provided a mode of Other or Unknown, to one of the more descriptive codes. For these type shipments in 2007, “Other” and “Unknown” modes were generally acceptable. At the conclusion of 2012 CFS mileage calculation operation, a review of these “Other mode” shipments was conducted. This analysis showed there to be a few truly “Other mode” shipments. Such shipments were often transported via conveyor belts. The table below compares the value and tonnage estimates for the Other-type modes in 2007 and 2012.

“Other” Modes of Transportation

Mode		2007		2012	
Code	Meaning	Value (\$mil)	Tons (000)	Value (\$mil)	Tons (000)
18	Other multiple modes	45,320	113,841	668	2,452
19	Other modes	279,113	271,567	1,026	36,844

More details about mileage calculation and related processing can be found in the **Mileage Calculation** section of the survey methodology.

Routing Software Changes

The underlying transportation network software used to model shipment distances was updated to reflect changes to the transportation infrastructure between 2007 and 2012. In particular:

- Due to the vendor change, the highway network used for routing was changed from 2007 to 2012. While the formula used to determine the most likely route for highway shipments did not change, the network was more expansive in 2012.
- Five airports were added to the air network in 2012. Other components of the 2007 domestic and international air networks were updated with the 2012 versions.
- The 2007 zip codes were replaced with 2012 zip codes.

Commodity Coding Changes

Several commodities in SCTGs 07, 08, 17, and 18 were redefined for 2012. See the *Commodity Coding Changes for 2012* table in the **Data Collection** section of this survey methodology for the details of these changes.

Application of Noise Infusion

For establishments that were in the survey in both 2007 and 2012, no effort was made to coordinate the direction or magnitude of the noise factor applied to these establishments from one survey to the next. For such an

establishment, the random noise multiplier may have been greater than 1.0 in 2007 but less than 1.0 in 2012 or vice versa. See the **Disclosure Avoidance** section of the survey methodology for more details.

Sampling Variability and Nonresponse

Through its sample design, the CFS tries to ensure the sample will include shipments originating from establishments in each CFS Area. However, estimates of other shipment characteristics, such as destination, commodity, and mode, depend entirely on the sample of shipments reported by responding establishments. See the sample design sections of the survey methodology for further information.

A particular combination of origin, destination, commodity, and mode (for example) may be common one year but rare or non-existent in the next survey. While this may reflect true changes in economic activity, it may also result from:

- Failing to include in the CFS sample, the establishments making these shipments, or
- If included, the sampled establishments failing to respond, or
- If responding, failing to include shipments with this particular combination of characteristics in the sample of shipments provided to the Census Bureau.