

Precanvass Survey for the 2007 Commodity Flow Survey

Promod Chandhok¹

Bureau of Transportation Statistics, Research and Innovative Technology Administration, U.S. Department of Transportation, 1200 New Jersey Avenue SE
Washington, D.C. 20590

Keywords: auxiliary establishments, sample design, non-shipping auxiliaries, duplicates.

1. Introduction

In planning for the 2007 Commodity Flow Survey (CFS), the joint Bureau of Transportation (BTS) and Bureau of the Census CFS estimation review team analyzed the issues with the 2002 CFS and decided on undertaking a precavass survey to improve the efficiency of the sample design, delete non-shipping establishments from the sampling frame and update contact information.

The Commodity Flow Survey (CFS) is a mail-out mail-back survey of shipping establishments in the mining, manufacturing, wholesale, and selected retail and information industries. It has been conducted every five years through a partnership between the U.S. Department of Transportation, Research and Innovative Technology Administration, Bureau of Transportation Statistics (BTS), and the U.S. Department of Commerce, Bureau of the Census. The survey produces national and subnational estimates of movement of goods including commodities shipped, their value, weight, and mode of transportation, origin and destination of shipments, as well as ton-miles and average miles per shipment. The data from the CFS are used by public policy analysts for transportation planning and decision making to assess the demand for transportation facilities and services, energy use, safety risk and environmental concerns. Earlier Commodity Flow Surveys were carried out in 2002, 1997, and 1993.

The sample for the 2007 Commodity Flow Survey (CFS) was selected using a stratified three-stage design in which the first-stage sampling units were establishments, the second-stage sampling units were reporting weeks, and the third-stage sampling units were shipments. In the first stage the approximately 750,000 in-

scope establishments appearing on the Census Bureau's Business Register were divided into primary strata mainly by geography and industry. In addition, some primary strata were constructed to improve estimates of hazardous material shipments. Then, in each primary stratum, establishments were sub-stratified by measure of size (MOS) and a simple random sample was selected without replacement from each MOS stratum. We used a target coefficient of variation (CV) constraint of 1.7% for each primary stratum and 0.036% at the national level to attain an overall sample of approximately 100,000 establishments. Note that the CV constraints are only for the first stage of selection. The second-stage was a systematic sample of four 1-week periods (reporting weeks) within the calendar year 2007. In the third-stage the respondent constructs a list of all shipments made in the reporting week and reports information on all shipments if there are 40 or less in that week. If the number of shipments exceeds 40, then the respondent selects a systematic sample of shipments with a sampling interval based on the number of shipments in the reporting week and reports information only on the selected shipments. For every establishment we get information on up to 40 shipments for the reporting week.

The 2007 CFS questionnaire requests the respondent to provide the following information about the sample of their establishment's shipments for their one-week period:

- date on which the shipment was made,
- it's value and weight,
- description of the commodity shipped and commodity code,
- mode(s) of transportation,
- whether the shipment was shipped intermodally,
- whether the shipment was a hazardous material,
- domestic destination and, for exports, the port of exit and foreign destination.

For the first time we also ask establishments to report on their usage of third-party logistics providers. This will be done only in the fourth quarter. For hazardous materials shipments, we also ask for the corresponding United Nations/North American Hazardous Materials (UN/NA) Code. Each sampled establishment will be contacted four times throughout 2007 - once every quarter.

Section 2 of this paper discusses some of the issues in the 2002 CFS that we attempted to resolve with the precanvass survey. In Section 3 we discuss what was done in the precanvass, Section 4 discusses the findings and Section 5 the accomplishments of the precanvass.

2. Some issues with the 2002 CFS

Some issues with the 2002 CFS that we attempted to resolve by the precanvass survey were:

Inadequate information on auxiliaries In planning of the 2002 CFS there was inadequate information on auxiliaries. Auxiliaries are establishments of a multiunit company that provide support for the other establishments in the company. Examples of auxiliaries are warehouses and corporate, subsidiary, and regional managing offices. First, the measure of size information that could be used in their selection was not always accurate or reliable. Second, most auxiliaries don't ship, but for the ones that do, they either ship a little or ship a lot (e.g. warehouses). Collection of this information from auxiliaries would assist in their selection by probability proportional to size and in deletion from the frame of any that do not ship.

Wasted questionnaires The number of establishments that responded to at least one questionnaire was approximately 31,000 out of a sample of approximately 51,000 establishments and the response rate for the survey was approximately 74%. Among the 47,212 nonauxiliaries in the sample, 4,473 (9%) were ineligible because they were either nonshippers or out-of-scope, whereas among the 3,793 sampled auxiliaries, 1,184 (31%) were ineligible for the same reason. We saw that a higher proportion of the auxiliaries were ineligible compared to nonauxiliaries and decreasing the

number of ineligible establishments would result in more completed records in the CFS data.

Establishment address may not be up-to-date

Another issue was that the establishment address may not be up-to-date and the questionnaire may not be timely in reaching the person who has the responsibility for completing the questionnaire. In addition, the person responsible may not be the person who has the information needed to complete the questionnaire. If these issues are resolved before the first 2007 CFS questionnaire is mailed out then response rate would likely improve and fewer callbacks would be needed. This would also save time during the conduct of 2007 CFS, as the questionnaire can be addressed to the appropriate person.

Usage of third-party logistics providers

Another issue was whether the shipments subcontracted to a third-party logistics provider (3PL) were being properly counted. Third-party logistics providers are increasingly meeting a larger proportion of freight transportation needs, and collection of this information would give us a better picture of the commodity transportation landscape.

In order to improve the quality of the frame data for auxiliaries and to obtain better contact and 3PL usage information, the precanvass, also called an advance survey, was mailed to all auxiliary and selected large establishments to permit us to improve contact while at the same time eliminating non-shipping establishments from the frame and collecting MOS information from the shipping auxiliaries.

3. What we did in the precanvass

In the precanvass we focused on the likely certainty establishments and all the auxiliaries. The likely certainties were the largest establishments by MOS in strata formed by geography and industry. These were targeted because they contribute most to the final estimates. We wanted updated address and contact information from them and also wanted to identify duplicate establishments. The auxiliaries were targeted because many don't ship and these can be deleted from the frame. For the shipping auxiliaries we don't have a good measure of their annual value of shipments.

Two questionnaires were used – one for the large establishments and one for the auxiliaries.

A short questionnaire consisting of two questions shown in Figure 1 was sent in February 2006 to 45,000 likely certainties to identify the appropriate person to whom the 2007 CFS questionnaire should be sent. Those who did not send in their completed questionnaires were sent

a follow-up questionnaire in March 2006. Some establishments that had been sent multiple forms called in and resolving this led to the identification of some of the duplicates among the large establishments. This procedure resulted in getting a good address and the name of the appropriate contact person and we were also able to identify some of the duplicates on the frame.

Figure 1: Questions Asked of Likely Certainty Establishments

1. Please print the information below for the person at your establishment (location) who should receive the 2007 Commodity Flow Survey forms and who can coordinate responses about all shipments made to or from this location, including those made by logistics providers.

Name
014

Title
015

Address
016

City
017

State
018

ZIP Code
019

Area Code
020

Phone Number
021 -

2. Please print your name and phone number, and return the form in the enclosed envelope. Thank you.

Name
022

Area Code
023

Phone Number
024 -

An alternative questionnaire (see Figure 2) was mailed to all the auxiliaries (about 40,000) on the frame. Those establishments not responding were sent a second copy in March 2006 and in addition the auxiliaries that did not respond to the two mail-outs were followed up by phone in May 2006. There were 4,845 auxiliary establishments in the North American Industry Classification System (NAICS) 4931, which is Warehousing & Storage, and 35,152 in NAICS 551114, which is Corporate, Subsidiary, and

Regional Managing Offices. The information collected from this questionnaire was used to identify nonshipping auxiliary establishments, which were then removed from the frame; for auxiliaries that did ship, we collected information on the total annual value of all shipments originating from that location. This information was used to construct a measure of size for each auxiliary, which was used in the sample design, and later for editing the 2007 CFS data.

Figure 2: Part of Precanvass Survey Questionnaire for Auxiliary Establishments

1. Does the establishment at this location do any of the following:

001 1 Yes 2 No
 a. Ship or deliver products to customers or clients.

002 1 Yes 2 No
 b. Ship or deliver products to other locations of this company.

003 1 Yes 2 No
 c. Provide for customer pick-up of products.

004 1 Yes 2 No
 d. Maintain records of shipments or deliveries leaving this location.

2a. Does the establishment at this location subcontract out all or a portion of its freight transportation (including inbound) to a third-party logistics provider (3PL)? 005 1 Yes 2 No

2b. Do other locations of this company subcontract out all or a portion of the company's freight transportation needs to a third-party logistics provider (3PL)? 006 1 Yes 2 No

(Note: If you answered "No" to all parts of Item 1, skip to item 5.)

3. Annual Value of Shipments - Please mark the box that best represents your estimate of the total annual value of all shipments originating from this location.

Mark (X) only one

007 Less than \$5 million

008 \$5 million or more but less than \$20 million

009 \$20 million or more but less than \$50 million

010 \$50 million or more but less than \$200 million

011 \$200 million or more but less than \$500 million

012 \$500 million or more but less than \$1 billion

013 \$1 billion or more but less than \$5 billion

014 \$5 billion or more

Table 1 summarizes the breakdown of the precavass establishments by industry sector. The 44,985 likely certainties are divided by five

industry sectors and the 39,987 auxiliaries are divided by the two types.

Industry Sector and NAICS codes	Establishments
Mining -212	2,258
Manufacturing - 31, 32 and 33 except for 323122	20,704
Wholesale - 42, except 425	18,189
Retail - 4541 and 45431	2,814
Information - 5111 and 51223	1,010
Auxiliaries - 4931	4,845
Auxiliaries - 551114	35,152
Total	84,972

4. Findings

In Table 2 we classify the establishments in the precavass file into two groups: auxiliary and likely certainty (LC) establishments. Data were received from 87% of the auxiliaries and 75% of the large establishments. A higher response rate

from auxiliaries was expected because we did a telephone follow-up of the auxiliary nonrespondents. In all, we received data from 81% of the mailed questionnaires.

Table 2: Summary of Precanvass Processing

	Auxiliary	LC	Total
Initial Precanvass Mailout	39,997	44,975	84,972
Final Precanvass File	34,936	33,836	68,772
% of Initial Mailout	87%	75%	81%

To see the effectiveness of follow-up mailing and the telephone follow-up in the precavass we look at Table 3, which provides the source of the

precavass records. We note that most of the analyst additions resulted from calls from establishments getting multiple forms.

Table 3: Source of Precanvass Records

Source	Auxiliary	LC	Total
Initial Mailing	19,037	25,488	44,525
Follow-up Mailing	5,609	7,477	13,086
Telephone Follow-up	5,421	127	5,548
Analyst Addition	4,869	744	5,613
Total	34,936	33,836	68,772

Table 4 shows the shipper status of the responding establishments. We see that 69% of the responding auxiliary establishments were nonshippers. These were deleted from the frame. Though the questionnaire sent to the certainty establishments did not specifically ask whether

they shipped but comments such as those indicating that they were out of business were used to classify them as out-of-scope and deleted as such from the frame. About 5% of the large establishments were removed by this exercise.

Table 4: Shipper Status of Precanvass Records

Shipper Status	Auxiliary	LC	Total
Shipper	10,716	32,275	42,991
Non-Shipper	24,220	1,561	25,781
Total	34,936	33,836	68,772

Auxiliaries were also asked to report the range of their total annual value of all shipments

originating from that location. Table 5 provides the collected information.

Table 5: Measure of Size of Shipping Auxiliaries

MOS Class Midpoint (\$ 000)	Counts	Total MOS (\$ mil)
-	1,130	0
2,500	4,451	11,128
12,500	1,652	20,650
35,000	1,019	35,665
50,000	1	50
100,000	1	100
125,000	1,221	152,625
150,000	3	450
250,000	3	750
275,000	2	550
350,000	548	191,800
400,000	13	5,200
725,000	1	725
750,000	321	240,750
3,000,000	304	912,000
5,000,000	46	230,000
Total	10,716	1,802,443

More than 70% of the shipping auxiliaries have annual value of shipments less than \$50 million.

An estimate of the total annual value of shipments made from auxiliaries is:

$$\$1,802,443 * 10^6 * \left(\frac{10,716}{9,586} \right) \left(\frac{1}{0.87} \right) = \$2,315,995 * 10^6$$

where \$1.8 trillion is the total annual value of shipments of the 9,586 responding shipping auxiliaries out of 10,716 shipping auxiliaries and 87% is the response rate for auxiliaries (see Table 2). The calculated response rate for auxiliaries makes no adjustment for ineligible establishments. From the 2007 CFS frame we estimate the total annual value of all shipments from CFS establishments to be \$10.5 trillion (internal computation by Census Bureau). We note that auxiliaries contribute approximately 20% to that and this is by far the biggest group. Auxiliary warehouses (NAICS 4931) are primarily designed to ship and, as such, despite their small number, contribute a large amount to the total.

5. Accomplishments

The precanvass survey was carried out to improve the efficiency of the sample design and improve the quality of the frame. The main accomplishments of the precanvass were:

- 1) Obtained more accurate estimates of the annual value of shipments of auxiliaries that ship, which was used to develop a measure of size for their selection in the sample
- 2) Eliminated more than 24,000 non-shipping auxiliaries
- 3) Eliminated more than 1,000 duplicate addresses
- 4) Developed procedures for resolving undeliverable as addressed cases. These procedures were later refined for 2007 CFS use
- 5) Updated contact information of precanvass establishments

- 6) Collected information on the usage of third-party logistics providers

This resulted in a much-improved frame, and therefore a more efficient sample design for the 2007 CFS.

6. References

Dahl, S. and W. Davie (2007), *Designing the 2007 Commodity Flow Survey*, paper presented at the 2007 International Conference for Establishment Surveys, Montreal, Canada.

U.S. Census Bureau (2006), *Weighting Specifications for the 2007 CFS*, internal memorandum, Washington, DC.

¹ The author thanks RITA/ BTS and Bureau of the Census reviewers for their comments. However, the views expressed in this paper are attributable only to the author and are not necessarily those of BTS or of the Bureau of the Census.