

**Generic Information Collection Request:  
Pilot Study and Respondent Debriefings for the Commodity Flow Survey**

**Request:** The Census Bureau plans to conduct additional research under the generic clearance for questionnaire pretesting research (OMB number 0607-0725 for the Commodity Flow Survey (CFS)). Staff from the Census Bureau's Economic Reimbursable Surveys Division (ERD) will be working with the Economic Statistical Methodology Division (ESMD) to conduct a pilot study and respondent debriefings for the Commodity Flow Survey.

The CFS is administered every 5 years through a partnership between the U.S. Census Bureau and the Bureau of Transportation Statistics (BTS). It provides data on the movement of goods in the United States including commodities shipped, their value, weight, and mode of transportation, as well as the origin and destination of shipments of commodities from manufacturing, mining, wholesale, and selected retail and services establishments. These data are used by policy makers and transportation planners to evaluate the demand for transportation facilities and services, energy use, and safety risk and environmental concerns. Additionally, the data collected from the CFS are also used to analyze 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.

Respondents to the CFS have traditionally been asked to provide a sample of their shipments four separate times during the reporting year, which occurred every 5 years. Some respondents to the CFS reported this process of creating a sample to be burdensome, and expressed interest in altering this method of data collection. Census Bureau staff created a secure depository (i.e., portal) for respondents to upload large amounts of shipment data from their records, thus eliminating the need for respondents to create a quarterly sample. Small scale participation in a trial of this new method of 'passive data collection' was found to be successful for some companies, and had promising implications for companies of similar attributes (i.e., size, employee count, resources) who could potentially participate as well.

In 2019, researchers conducted early stage scoping interviews to assess the feasibility of wide scale implementation of this data collection change, as well as its potential effects on respondent burden. Additionally, usability research is currently underway to ensure the functionality of the instrument. These qualitative studies were approved via separate requests under the Census Bureau's generic clearance, letters dated June 21, 2019, and April 13, 2020 respectively.

To further refine the CFS instrument and evaluate this new data collection strategy, the ERD will be conducting a pilot study and post-collection data analysis. Following the pilot, respondent debriefings will be conducted by ESMD staff.

The pilot test is designed to focus on the following topics:

- Functionality and usability of the instrument using live data
- The impact of a digitally native data submission upload option on data quality
- Error fix functionality and respondents' level of willingness and ability to correct errors
- Viability of the new process for medium to small companies

- Title/role of the individual who completes the questionnaire, and what other people/departments are involved in the response process
- The approximate burden for completing the survey

Further information regarding the Commodity Flow Survey can be found at this website: <https://www.census.gov/econ/cfs/> .

**Purpose:** The purpose of this pilot test is to determine the level of effort in submitting data to the Commodity Flow Survey via this new method, and furthermore to determine whether this new method was easier compared to the original response process. Debriefing interviews will illuminate generally who is tasked with filling out the questionnaire and what departments need involvement. Post hoc analysis will determine the quality of data submitted compared to previous submissions in 2017.

Results from the pilot will likely be summarized in a working paper. Following debriefing interviews, a report and supporting presentation will be developed that outlines the findings and recommendations for improvement.

**Population of Interest:** U.S. companies that ship goods, across a variety of sizes, industries, and locations.

**Timeline:** The pilot test will be conducted in October-November, 2020. Debriefing interviews are expected to take place in December, 2020.

**Language:** Testing will be conducted in English only.

**Method:** Approximately 100 CFS respondents will be recruited for the pilot primarily via email and phone calls. The utilization of paper mail is still to be determined as the Coronavirus pandemic means fewer people will be in their places of business to receive it.

Pilot participants will be asked to provide data regarding the shipment and movement of goods from their establishment. The primary change in this reporting method from the original reporting method is that respondents will no longer need to upload a sample of their shipments, and can instead upload all of their shipment records without needing to pare them down. Additionally, previous reporting required respondents to look up Census Bureau specific product codes, (i.e., SCTG codes) for each shipment record, which was a time consuming part of previous CFS completion. Now, machine learning will be implemented to apply product codes based on respondents' product text descriptions.

ERD staff will employ different methods to evaluate the data that are returned from the pilot test. We are particularly interested to learn what discrepancies arise leveraging this mode of collection, more akin to a "data dump" than previous collection modes, and what opportunities and challenges present themselves. As a check for new/increased coverage, analysts will simulate the sampling process that existed during 2017 data collection, by taking a sample of the shipment data provided via the new upload method. Secondly, the number of shipment records submitted to the new instrument will be compared to those submitted in 2017. In addition, staff

will check data quality along a number of dimensions: (a) comparing the percentages of shipment records that fail the survey's rule-based edit checks to 2017 responses, (b) comparing the differences in outlier and other substantively incorrect values, (c) evaluating our ability to automatically code free-text product descriptions that respondents submit via this mode, and (d) evaluating the amount of correction and data cleaning required to use the response data provided in estimates. Finally, the pilot will also involve a small number of follow-up respondent debriefing telephone interviews for select cases that ERD staff may want to probe further regarding the response process.

**Sample:** The pilot study will involve recruiting mainly larger establishments, which will be most likely to use the new method, followed by a smaller sample of medium and small sized establishments. (Measure of size of company is determined by annual revenue.) This sample will yield a suitable, broad representation of U.S. businesses for the pilot test. All respondents to the pilot test will be instructed to upload a sample of their shipment data.

The sampling frame for these interviews comes from a list of companies that were sampled to participate in the 2017 CFS. Participating companies will most likely have submitted responses to the CFS survey, but not necessarily. Participants will be informed that the Commodity Flow Survey pilot test is voluntary, as are the telephone debriefing interviews.

Follow-up telephone debriefing interviews will be conducted with approximately 30 respondents. The sample size necessary for this test was determined by qualitative research experience. We plan to conduct interviews with a variety of sizes and types (i.e., industries) of business establishments, with an emphasis on large and medium companies (measure of size determined by revenue). This is because in production, the largest companies will be the most likely establishments to utilize the new reporting method.

**Recruitment:** Respondents will be recruited via email and phone. Respondents will be informed that their participation is voluntary.

**Protocol:** The protocol for the respondent debriefing portion of the study is enclosed (see Attachment A). We anticipate that each interview will take 45 minutes to complete.

**Use of Incentive:** Monetary incentives for participation will not be offered.

Below is a list of materials to be used in the current study:

1. **Attachment A:** Protocol used to outline how the respondent debriefing portion of the research study will be conducted
2. **Attachment B:** Consent form to obtain participant consent for participation and recording of the respondent debriefing session
3. **Attachment C:** Screenshots of the pilot instrument
4. **Attachment D:** 2017 CFS survey to illustrate the survey being evaluated

**Length of interview:** For the pilot, we expect that each response will take no more than 60 minutes (100 cases x 60 minutes per case = 100 hours). Additionally, to recruit pilot participants, we expect to make up to 5 phone calls per completed case. The recruiting calls are

expected to last on average 3 minutes per call (5 attempted phone calls per completed case x 100 cases x 3 minute per case = 25 hours). Thus, the estimated burden for the pilot is 125 hours (100 hours for response + 25 hours for recruiting).

For the debriefing interviews, we expect that each interview will last no more than 45 minutes (30 cases x 45 minutes per case = 22.5 hours). Additionally, to recruit respondents we expect to make up to 5 phone contacts per completed case. The recruiting calls are expected to last on average 3 minutes per call (5 attempted phone calls per completed case x 30 cases x 3 minute per case = 7.5 hours). Thus, the estimated burden for the respondent debriefing interviews is 30 hours (22.5 hours for interviews + 7.5 hours for recruiting).

Thus, the total burden for the pilot test and the respondent debriefings is expected to be 155 hours.

The contact person for questions regarding data collection and statistical aspects of the design of this research is listed below:

Rebecca Keegan  
Data Collection Methodology & Research Branch  
Economic Statistics and Methodology Division  
U.S. Census Bureau  
Washington, D.C. 20233  
(301) 763-6003  
Rebecca.Keegan@census.gov

Cc:

Nick Orsini	(ADEP) with enclosure
Carol Caldwell	(ESMD) with enclosure
Diane Willimack	(ESMD) with enclosure
Amy Anderson Riemer	(ESMD) with enclosure
Rebecca Keegan	(ESMD) with enclosure
Kristin Stettler	(ESMD) with enclosure
Demetria Hanna	(ESMD) with enclosure
Christian Moscardi	(ERD) with enclosure
Kevin Deardorff	(ERD) with enclosure
Carla Medalia	(ERD) with enclosure
Jennifer Whitaker	(ERD) with enclosure
Berin Linfors	(ERD) with enclosure
Julie Parker	(BTS) with enclosure
Jennifer Hunter Childs	(ADRM) with enclosure
Jasmine Luck	(ADRM) with enclosure
Danielle Norman	(PCO) with enclosure
Mary Lenaiyasa	(PCO) with enclosure