Supporting Statement B
U.S. Department of Commerce
U.S. Census Bureau
Service Annual Survey
OMB Control Number 0607-0422

B. Collections of Information Employing Statistical Methods

1. Description of Universe

The Service Annual Survey (SAS) is a sample survey of approximately 72,000 firms having one or more service establishments. The sample size differs from the number of reporting units, which is 82,240, because we request firms to provide separate reports for each industry in which they are engaged. The sample represents a universe of approximately 4.6 million establishments, based on the Census Bureau's Business Register and economic census data.

2. Sampling Methodology and Estimation Procedures

a. Sampling Methodology

The current sample was introduced with the 2011 SAS. This sample was designed to produce estimates based on the 2007 North American Industry Classification System (NAICS). This section describes the design, selection, and estimation procedures for this sample. Similarly, a new sample will be introduced with the 2016 SAS survey year based on the 2012 Economic Census and will produce estimates based on the 2012 North American Industry Classification System (NAICS). The design, selection, and estimation procedures for that new sample will be the same as those described below. For descriptions of prior samples, see the Service Annual Survey publications.

Sampling Frame

The Service Annual Survey sampling frame has two types of sampling units: Employer Identification Numbers (EINs) and large, multiple-establishment firms. Both sampling units represent clusters of one or more establishments owned or controlled by the same firm. The information used to create these sampling units was extracted from data collected as part of the 2007 Economic Census and from establishment records contained on the Census Bureau's Business Register as updated to October 2010. The sample is redesigned and reselected every 5 to 7 years to redistribute reporting burden and to introduce sampling and processing efficiencies.

To create the sampling frame, we extract the records for all employer establishments located in the United States and classified in service sectors as defined by the 2007 NAICS. For these establishments we extract revenue, payroll, employment, name and address information, as well as primary identifiers and, for establishments owned by multi-unit firms, associated EINs.

To create the sampling units for multi-unit firms, we aggregate the economic data of the establishments owned by these firms to an EIN level by tabulating the establishment data for all service establishments associated with the same EIN. Similarly, we aggregate the data to a multi-unit firm level by tabulating the establishment data for all service establishments associated with the same firm identifier. No aggregation is necessary to put single-unit establishment information on an EIN basis or a firm basis. Thus, the sampling units created for single-unit firms simultaneously represent establishment, EIN, and firm information.

Stratification

The SAS uses a stratified, one-stage design with primary strata defined by industry, tax status, and government ownership. There are 391 primary strata. We further stratify the sampling units within industry group by a measure of size (substratify) related to their annual receipts. Sampling units expected to have a large effect on the precision of the estimates are selected "with certainty." This means they are sure to be selected and will represent only themselves (i.e., have a selection probability of 1 and a sampling weight of 1). To identify the certainty units, we determine a substratum boundary (or cutoff) that divides the certainty units from the non-certainty units. We base these cutoffs on a statistical analysis of data from the 2007 Economic Census. We also use this analysis to determine the number and boundaries of non-certainty substrata for each industry stratum and to set sampling rates needed to achieve specified sampling variability constraints on revenue estimates for different industry groups. The size substrata and sampling rates are later updated through analysis of the sampling frame.

Sample Selection

Sample selection is a two-step process and begins by identifying the firms selected with certainty. If a firm's annual receipts are greater than the corresponding certainty cutoff, that firm is selected into the SAS sample with certainty.

In the second step, all firms not selected with certainty are subjected to sampling on an EIN basis. If a firm has more than one EIN, we treat each of its EINs as a separate sampling unit. To be eligible for the initial sampling an EIN had to have nonzero payroll in 2009. The EINs are stratified according to their major industry and their estimated receipts (on a 2007 basis). Within each non-certainty stratum, a simple random sample of EINs is selected without replacement.

Sample Maintenance

We update the sample to represent EINs issued since the initial sample selection. These new EINs, called births, are EINs, recently assigned by the IRS, that have an active payroll filing requirement on the IRS Business Master File (BMF). An active

payroll filing requirement indicates the EIN is required to file payroll for the next quarterly period. The Social Security Administration attempts to assign industry classification to each new EIN.

EINs with an active payroll filing requirement on the IRS Business Master File are said to be "BMF active" and EINs with an inactive payroll filing requirement are said to be "BMF inactive."

We sample EIN births on a quarterly basis using a two-phase selection procedure. To be eligible for selection, a birth must either have no industry classification or be classified in an industry within the scope of the Service Annual Survey, the Annual Wholesale Trade Survey, or the Annual Retail Trade Survey, and it must meet certain criteria regarding its quarterly payroll. In the first phase, we stratify births by broad industry groups and a measure of size based on quarterly payroll. A relatively large sample is drawn and canvassed to obtain a more reliable measure of size, consisting of receipts in two recent months and a new or more detailed industry classification code. Births that have not returned their questionnaire after 30 days are contacted by telephone.

Using this more reliable information, in the second phase we subject the selected births from the first phase to probability proportional-to-size sampling so that the overall probabilities of selection are equivalent to those used in selecting the initial SAS sample from the Business Register. Because of the time it takes for a new employer firm to acquire an EIN from the IRS, and because of the time needed to accomplish the two-phase birth-selection procedure, we add births to the sample approximately nine months after they begin operation.

If a firm was selected with certainty and had more than one establishment at the time of sampling, any new service establishments that the firm acquires, even if under new or different EINs, are included in the sample with certainty.

However, if a firm was selected with certainty and had only one establishment (and thus, only one in-scope EIN) at the time of sampling, only future establishments associated with that firm's originally-selected EIN are included in the sample with certainty; any new EINs that might later be associated with that firm are subjected to sampling through the quarterly birth-selection procedure.

To be eligible for the sample canvass and tabulation, an EIN selected in the noncertainty sampling operations must meet both of the following requirements:

- It must have an active payroll filing requirement on the IRS Business Master File.
- It must have been selected from the Business Register in either the initial sampling or during the quarterly birth-selection procedure.

Each quarter, we check against the current Business Register to determine if any EINs on the survey have become BMF inactive. Typically, we do not canvass BMF inactive EINs during the reference year. Likewise, if any EIN on the survey was BMF inactive in a previous reference year but is now BMF active on the current Business Register, we again include these EINs in the canvass. In both cases, we only tabulate data for that portion of the reference year that these EINs reported payroll to the IRS.

Non-Response

Data are imputed using survey data and administrative data as input for unit non-response, item non-response, and for responses that fail computer or analyst edits.

b. Estimation Procedure

Totals are computed as the sum of weighted data (reported and imputed) for all selected sampling units that meet the sample canvass and tabulation criteria given above. The weight for a given sampling unit is the reciprocal of its probability of selection into the sample. The estimates are adjusted using the results of the most recent economic census. Variances are estimated using the method of random groups.

Benchmarking

There are two situations when benchmarking is performed for the SAS. Benchmarking is used when a new sample is introduced to link the new and prior samples and maintain the consistency of the time series. Also, published estimates from the SAS are benchmarked using the latest results from the economic census as they become available as described on
http://www.census.gov/services/sas/sastechdoc.html

3. Methods to Maximize Response

Letters (Attachment 2) are typically mailed one month after the data-year ends. The letter explains the necessity and use of the data and states the respondents' username and password where the respondent can report online and access the current year worksheet (Attachment 1). In an effort to promote electronic reporting, we eliminated paper forms from the initial and follow-up mailings and strongly encourage respondents to report data online. Worksheets can be returned via fax and mail.

SAS utilizes two follow-up mailings for delinquent cases. The first follow-up is sent approximately two months after the initial mailout. The second follow-up is sent approximately one month after the first follow-up. There is an option for a third follow-up mailing approximately three months after the second follow-up. Delinquent cases are also followed up by telephone approximately two weeks after the second follow-up mailing.

Firms are given at least 30 business days to respond to the initial mailing and are given extension dates upon request. The Census Bureau also provides a telephone number for assistance with any questions or concerns about the survey. The unit response rates for the most recent completed surveys are:

Year	Unit Response Rate
2013	72.20%
2012	74.74%
2011	74.90%

Unit and dollar volume response rates are analyzed prior to each follow-up mailout and survey closeout to ensure publication requirements are met. If response rates are low and fail to meet publication requirements, additional follow-up with delinquent companies may be performed including certified mailout and/or worksheet mailout.

4. Tests of Procedures or Methods

Procedures in every phase of SAS are tested – from mailout and data capture to editing and publication.

5. Contacts for Statistical Aspects and Data Collection

Direct questions regarding the planning and implementation of this survey to Anne Russell, U.S. Census Bureau, (301) 763-5173 or via email at anne.sigda.russell@census.gov . Questions regarding survey methodology should be directed to Katrina Washington, U.S. Census Bureau, (301) 763-7212 or via email at katrina.t.washington@census.gov.

List of Attachments

1. Survey Worksheet Descriptions and Representative Selection of SAS Worksheets
2. Initial, Follow-up and Due Date Reminder Letters
3. Screen Shots from the Electronic Instrument
4. Total Respondent Burden Hours
5. Letter of support from the Bureau of Economic Analysis
6. Letter of support from the Centers for Medicare and Medicaid Services.