

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 because we use administrative data for firms with an estimated measure of size below certain levels and because we request firms to provide separate reports for each industry in which they are engaged (see 2c. “Problems Requiring Specialized Reporting Arrangements” for additional info). The sample represents a universe of approximately three million establishments, based on the Census Bureau’s Business Register and economic census data.

2. Sampling Methodology and Estimation Procedures

a. Sampling Methodology

A new sample was introduced with the 2011 SAS. The new 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 the new sample. 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, Sampling Rates, and Allocation

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. We select sampling units expected to have a large effect on the precision of the estimates “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 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 Mater 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 2 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 Service Annual Survey sample from the October 2010 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 9 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 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 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 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.

Single-unit EINs selected into the sample with certainty are not dropped from canvass and tabulation if they are no longer BMF active. Rather, the firm that used the EIN is contacted, and if a successor EIN is found, it is added to the survey. For both inactive EINs and any previously inactive EINs that are now active, data are tabulated for only the 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. A detailed description of the benchmarking procedures for the introduction of the new SAS sample in 2011 is not yet available, but will be provided online at <http://www.census.gov/services/sas/sastechdoc.html>. Also, published estimates from the SAS are benchmarked using the latest results from the economic census as they become available as described on the website above.

3. Efforts to Maximize Response

The following actions have or will be taken to maximize response rates:

- Customized mailing arrangements for some large firms;
- Conducted outside consultations;
- Customized computer-imprinted instructions to clarify reporting criteria for selected industries;
- Planned follow-up actions to contact delinquent firms;
- Provide the option for firms to FAX their report;
- Provide a toll-free number that firms can call for assistance;
- Survey results available to survey respondents on Internet;
- Mandatory response;
- Provide Internet response option.

In instances when the survey coverage requires our obtaining data from various subsidiaries or operating units of the firm, specialized arrangements have been established to mail separate forms to each subsidiary or operating unit of the firm. Customized mailings of this type have proven to be effective in obtaining more timely response and thus reducing follow-up costs, minimizing errors in reporting that result from coverage problems and reducing respondent burden.

Through our contacts, we have confirmed that the data being requested were available from existing firm records or could be easily estimated, that reporting instructions were clear and helpful, and that terminology used on the questionnaires conformed to industry usage. Through these consultations, we were also able to establish an estimate of the number of hours necessary for a firm to complete the survey.

The follow-up actions listed below with approximate dates will be taken for delinquent firms in the SAS. (Dates assume an initial mailout in January).

- First reminder letter with form-- April
- Second reminder letter with form-- May
- Telephone follow-up-- May
- Third reminder letter (certified) with form-- August

Mandatory response provides an incentive for firms otherwise reluctant to provide the required information in the SAS.

4. Tests of Procedures or Methods

Since a considerable amount of information exists for previous canvasses of the SAS, no testing was conducted or is currently planned.

5. Contacts for Statistical Aspects and Data Collection

Direct questions regarding the sample design and statistical methodology used for this survey to William C. Davie Jr., Assistant Division Chief for Research and Methodology, Service Sector Statistics Division, U.S. Census Bureau, (301) 763-7182.

Direct questions regarding the planning and implementation of this survey to Ronald W. Farrar, Chief, Health Care and Consumer Services Branch, Service Sector Statistics Division, U.S. Census Bureau, (301) 763-6782.

List of Attachments

1. Survey Form Descriptions and Representative Selection of SAS Forms
2. Cover Letters

3. Total Respondent Burden Hours
4. Current Publishable NAICS Industries