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 70,753 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 four 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 2005 SAS. The new sample was designed to produce estimates based on the 2002 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: large, multiunit firms and Employer Identification Numbers (EINs). 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 2002 Economic Census and from establishment records contained on the Census Bureau's Business Register as updated to December 2004. The sample is redrawn every 5 years to redistribute reporting burden and to introduce sampling and processing efficiencies.

To create the sampling frame, we extract the records for all establishments located in the United States and classified in service sectors as defined by the 2002 NAICS. For these establishments, we extract revenue, payroll, employment,

name and address information, as well as primary identifiers and, for establishments owned by multiunit firms, associated EINs.

To create the sampling units for multiunit 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 multiunit firm level by tabulating the establishment data for all service establishments associated with the same company 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 primary stratification of the sampling frame is by industry group based on the detail required for publication. We assign each sampling unit to the industry stratum corresponding to the industry that contributes the most to the unit's measure of size (estimated annual receipts).

Within each industry stratum, we substratify the sampling units by a measure of size 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 noncertainty units. We base these cutoffs on a statistical analysis of data from the Census Bureau's Business Register. We also use this analysis to determine the number and boundaries of noncertainty substrata for each industry stratum and to set sampling rates needed to achieve specified sampling variability constraints on revenue estimates for different industry groups.

Sample Selection

The first step in the sample selection identified firms selected with certainty. If a firm's annual receipts were greater than the corresponding certainty cutoff, that firm was selected into the sample with certainty.

All firms not selected with certainty were subjected to sampling on an EIN basis. If a firm had more than one EIN, we treated each of its EINs as a separate sampling unit. The EINs were stratified according to their major industry and their estimated receipts (on a 2002 basis). Within each noncertainty stratum, a simple random sample of EINs was selected without replacement.

Sample Maintenance

We update the sample to represent EINs issued since the intial sample selection. These new EINs, called births, are EINs recently assigned by the IRS on the latest available IRS mailing list for FICA taxpayers and have been assigned an industry classification code (if possible) by the Social Security Administration (SSA).

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 number of paid employees or 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.

Using this more reliable information, in the second phase we subject the selected births from the first phase to probability proportional-to-size sampling with overall probabilities equivalent to those used in drawing the initial Service Annual Survey sample. 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.

We include births that are selected in the quarterly birth-selection procedure in November of the survey year in the initial mailing of the Service Annual Survey questionnaires in January of the following year. To better represent all EIN births in the reference year, and specifically to account for the time it takes to identify and select new EINs, we add births to the Service Annual Survey sample that are selected in February, May, and August the year following the reference year. We mail survey forms to these births in June and August to supplement the initial survey mailing.

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 be on the latest available IRS mailing list for FICA taxpayers from the previous quarter.
- It must have been selected from the Business Register in either the initial sampling or during the quarterly birth-selection procedure.

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 single-unit firm was selected with certainty, 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.

EINs selected into the sample with certainty are not dropped from canvass and tabulation if they are no longer on the IRS mailing list. 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 and reactivated EINs, data are tabulated for only the portion of the reference year that these EINs reported payroll to the IRS.

b. Estimation Procedure

The current sample was introduced with the 2005 Service Annual Survey to compute estimates based on the 2002 North American Industry Classification System (NAICS). This new sample replaces one that was designed to produce estimates based on the 1997 NAICS.

Totals estimated from this sample survey are computed as the sum of weighted data (reported and imputed) for all selected sampling units that meet the tabulation criteria given in the Sample Maintenance section. The weight for a given sampling unit is the reciprocal of its probability of selection into the sample. The sample-based estimated totals are then adjusted to the 2002 Economic Census using the procedure described below.

For industries affected by the change from 1997 to 2002 NAICS, published census-adjusted revenue estimates for 1998 through 2004 from the prior sample are restated on a 2002 NAICS basis, using revenue distributions from the 2002 Economic Census that link the two sets of classification codes. Of particular note, the estimates for Sector 51 (Information) are revised due to the creation of new industries for Internet publishing and broadcasting and Web search portals. For industries not affected by the change from 1997 to 2002 NAICS, there is no need to restate the published census-adjusted revenue estimates from the prior sample.

Revenue estimates for 2005 and subsequent years from the current sample are adjusted to the 2002 Economic Census by linking these estimates to the published census-adjusted estimates from the prior sample, after historical corrections are made to data from the current sample for 2004 and subsequent years. The linking is performed by multiplying the sample-based revenue estimate for a given detailed industry, which is generally defined by a 6-digit NAICS code, by a ratio. The numerator and denominator of the ratio are as follows:

- The numerator is the 2004 published census-adjusted revenue estimate of the industry on a 2002 NAICS basis from the prior sample.
- The denominator is the 2004 revenue estimate of the industry on a 2002 NAICS basis from the current sample.

For data items other than revenue, such as total expenses and other revenue items, a similar method is used to adjust the estimates for 2004 and subsequent years from the current sample, by linking these estimates to those from the prior sample. First, the ratio described above is applied to the sample-based estimates for the given detailed industry for 2004 and subsequent years. Then, the difference between the 2004 adjusted estimate from the current sample and the 2004 published adjusted estimate from the prior sample is taken into account by applying a second ratio to the published adjusted estimates for 1998 through 2003 from the prior sample. The numerator and denominator of the second ratio are as follows:

- The numerator is the 2004 adjusted estimate for the industry on a 2002 NAICS basis from the current sample.
- The denominator is the 2004 published adjusted estimate for the industry on a 2002 NAICS basis from the prior sample.

Total estimates at 2-, 3-, 4-, and 5-digit NAICS levels are computed by summing the adjusted totals for the appropriate detailed industries comprising the aggregate. Year-to-year change estimates are computed using the appropriate adjusted totals for the industry and time period.

Note that estimates for the following Truck Transportation (NAICS 484) data items are produced directly from the sample, without adjustment:

- Inventories of Revenue Generating Equipment by Type of Carrier;
- Number of Truck Miles.

c. Problems Requiring Specialized Reporting Arrangements

Some multi-establishments firms are engaged in multiple service activities. These firms were subjected to sampling using their primary kind-of-business (that accounted for the largest proportion of their estimated service receipts) but with a measure of size for sampling that reflected their total receipts for all service establishments. We have identified selected firms of this type and, where these multiple activities are significant, we mail separate questionnaires to obtain data covering each of the firms' service activities.

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;
- Providing the option for firms to FAX their report;
- Providing a toll-free number that firms can call for assistance;
- Free copy of survey results available to survey respondents;
- Mandatory response.
- Providing Internet response option.

In instances when the survey coverage requires our obtaining data from various subsidiaries or operating units of the company, 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 company 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 company 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-- mid-April
- Second reminder letter with form-- mid- May
- Telephone follow-up-- May
- Third reminder letter (certified) with form-- mid- August

Mandatory response provides an incentive for firms otherwise reluctant to provide the required information in the SAS. Experience has shown that mandatory authority yields a report response of approximately 90 percent 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

Questions regarding the sample design and statistical methodology used for this survey should be directed to Ruth Detlefsen, Assistant Division Chief for Research and Methodology, Service Sector Statistics Division, U.S. Census Bureau, (301) 763-5171. Planning and implementation of this survey are under

the direction of Jeff Barnett, Chief, Information and Business Services Branch, Service Sector Statistics Division, U.S. Census Bureau (301) 763-2787 and Ron 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