The United States Census Bureau conducts the Quarterly Services Survey to provide national estimates of quarterly revenues for establishments classified in select service sectors. (See the coverage section for a complete description of the industries included in the Quarterly Services Survey.)

Questionnaires are mailed to a probability sample of firms selected from the larger Service Annual Survey sample. The sample includes firms of all sizes; however, firms without paid employees (nonemployers) are not included in the estimates.

Coverage

The estimates contained in this report are summarized by kind-of-business classification based on the 1997 North American Industry Classification System (NAICS). **NAICS** is a classification system that groups establishments into industries based on the activities in which they are primarily engaged. This industry classification system was developed as a result of a joint effort by statistical agencies in Canada, Mexico, and the United States so that common industry definitions would allow for comparability in statistics on business activity across North America.

Estimates in this report are presented for select industries in the following NAICS sectors:

- 51 Information
- 54 Professional, Scientific, and Technical Services
- 56 Administrative and Support and Waste Management and Remediation Services
- 622 Hospitals
- 623 Nursing and Residential Care Facilities

Dollar Values

All dollar values presented in this report are expressed in current dollars; that is, the estimates are not adjusted to a constant dollar series. Consequently, when comparing estimates, users also should consider price level changes.

Confidentiality

Title 13 of the United States Code authorizes the Census Bureau to conduct censuses and surveys. Section 9 of the same Title requires that any information collected from the public under the authority of Title 13 be maintained as confidential. Section 214 of Title 13 and Sections 3559 and 3571 of Title 18 of the United States Code provide for the imposition of penalties of up to five years in prison and up to $250,000 in fines for wrongful disclosure of confidential census information. In accordance with Title 13, no estimates are published that would disclose the operations of an individual firm.

The Census Bureau's internal Disclosure Review Board sets the confidentiality rules for all data releases. A checklist approach is used to ensure that all potential risks to the confidentiality of the data are considered and addressed.

Disclosure Limitation

A disclosure of data occurs when an individual can use published statistical information to identify either an individual or firm that has provided information under a pledge of

confidentiality. Disclosure limitation is the process used to protect the confidentiality of the survey data provided by an individual or firm. Using disclosure limitation procedures, the Census Bureau modifies or removes the characteristics that put confidential information at risk for disclosure. Although it may appear that a table shows information about a specific individual or business, the Census Bureau has taken steps to disguise or suppress the original data while making sure the results are still useful. The techniques used by the Census Bureau to protect confidentiality in tabulations vary, depending on the type of data.

SAMPLE DESIGN AND ESTIMATION PROCEDURES

Sampling Frame

The sampling frame for the Quarterly Services Survey sample is a subset of the Service Annual Survey sample and has the same types of sampling units as the Service Annual Survey frame—large, multiple-establishment firms and Employer Identification Numbers (EINs). The EIN is the identifier employer businesses use to report Social Security payroll withholdings to the federal government. Both sampling units represent clusters of one or more establishments owned or controlled by the same firm.

Stratification, Sampling Rates, and Allocation

The primary stratification of the Quarterly Services Survey frame is by industry group based on the detail required for the Quarterly Services Survey publication. We publish Quarterly Services Survey estimates at broader industry levels than the industry groupings used to publish Service Annual Survey estimates. Therefore, the industry stratification for the Quarterly Services Survey sample is broader than the industry stratification used for the Service Annual Survey sample.

Within industry group, we substratify the sampling units by a measure of size related to their annual revenue as reported in the Service Annual Survey. 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 one and a sampling weight of one). 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 1997 Economic Census data and data extracted from the Census Bureau's Business Register. We also use this analysis to determine the number of size substrata for each industry group and to set sampling rates needed to achieve specified sampling variability objectives on revenue estimates for different industry groups.

Sample Selection

We select the Quarterly Services Survey sample independently within each size substratum contained in an industry stratum. The actual selection procedure follows a systematic, probability proportional-to-size scheme. Because the Quarterly Services Survey sample is an independently selected subsample, it is possible that we select some units in the Service Annual Survey sample at a lower sampling rate than desired for the Quarterly Services Survey sample. We include such a unit in the Quarterly Services Survey sample and assign a sampling weight equal to the unit's Service Annual Survey sampling weight. The maximum sampling weight for an EIN selected for the Quarterly Services Survey sample is about 2,050.

Sample Maintenance

Periodically, we update the Quarterly Services Survey sample to represent EINs issued since the initial sample selection. These new EINs, called births, are EINs recently assigned by the Internal Revenue Service (IRS) on the latest available IRS mailing list for FICA taxpayers and assigned an industry classification (if possible) by the Social Security Administration.

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 one of the following: the Annual Retail Trade Survey, the Annual Trade Survey, or Service Annual Survey, and it must meet certain criteria regarding its number of paid employees or quarterly payroll. In the first phase, we stratify births by industry and a measure of size based on expected employment or quarterly payroll. A relatively large sample is drawn and canvassed to obtain a more reliable measure of size, consisting of sales in two recent months, and a new or more detailed industry classification.

Using this more reliable information, 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 Annual Retail Trade Survey, Annual Trade Survey, and Service Annual Survey samples. 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 samples approximately nine months after they begin operation.

Updates to the Quarterly Services Survey sample occur in the same manner and at the same time as updates to the Service Annual Survey sample. The births selected for the Quarterly Services Survey sample are a subset of the births selected for the Service Annual Survey sample. These births are selected using sampling rates equivalent to those used in selecting the initial Quarterly Services Survey sample.

Estimation Method

The final Quarterly Services Survey estimate for the level of quarterly receipts is the product of a direct expansion estimator and a ratio estimate that corrects for additional sampling error between the Quarterly Services Survey sample and the Service Annual Survey sample. The direct expansion estimate of receipts for each quarter is the sum of the weighted quarterly receipts for each unit (reported or imputed). The assigned weight for each sample unit is the reciprocal of its probability of selection into the Quarterly Services Survey sample. The adjustment ratio is computed as the estimate of annual receipts from the Service Annual Survey divided by an estimate of annual receipts computed using annual data from the Service Annual Survey for units in the Quarterly Services Survey subsample weighted by their Quarterly Services Survey weights. This adjustment ensures that if the same data were used to tabulate the Quarterly Services Survey sample and the Service Annual Survey sample, the estimates would be identical. The resulting estimates may be referred to as ratio-adjusted. This ratio adjustment is done at detailed industry levels. Estimates for aggregate levels are derived by summing the ratio adjusted detail level estimates to preserve additivity.

The estimate of change in quarterly receipts is based on the ratio-adjusted levels for the appropriate quarters.

RELIABILITY OF THE ESTIMATES

The estimates in this publication may differ from the actual, but unknown, population values. For a particular estimate, statisticians define this difference as the total error of the estimate. When describing the accuracy of survey results, it is convenient to discuss total error as the sum of sampling error and nonsampling error. Sampling error is the error arising from the use of a sample, rather than a census, to estimate population values. Nonsampling error encompasses all other factors that contribute to the total error of a survey estimate. The sampling error of an estimate can usually be estimated from the sample; whereas, the nonsampling error of an estimate is difficult to measure and can rarely be estimated. Consequently, the actual error in an estimate exceeds the error that can be estimated. Data users should take into account the estimates of sampling error and the potential effects of nonsampling error when using the statistics in this publication.

Further descriptions of sampling error and nonsampling error are provided in the following sections.

Sampling Error

Because the estimates are based on a sample, exact agreement with results that would be obtained from a complete enumeration of the sampling frame using the same enumeration procedures is not expected. However, because each firm on the sampling frame has a known probability of being selected into the sample, it is possible to estimate the sampling variability of the survey estimates.

The particular sample used in this survey is one of a large number of samples of the same size that could have been selected using the same design. If all possible samples had been surveyed under the same conditions, an estimate of a population parameter of interest could have been obtained from each sample. Estimates derived from the different samples would, in general, differ from each other. Common statistical measures of the variability among these estimates are the sampling variance, the standard error, and the coefficient of variation (CV). The sampling variance is defined as the squared difference, averaged over all possible samples of the same size and design, between the estimator and its average value. The standard error is the square root of the sampling variance. The CV expresses the standard error as a percentage of the estimate to which it refers. For example, an estimate of 200 units that has an estimated standard error of 10 units has an estimated CV of 5 percent. The sampling variance, standard error, and CV of an estimate can be estimated from the selected sample because the sample was selected using probability sampling. Note that measures of sampling variability, such as the standard error or coefficient of variation, are estimated from the sample and are also subject to sampling variability. (Technically, we should refer to the *estimated* standard error or the *estimated* coefficient of variation of an estimator. However, for the sake of brevity we have omitted this detail.) It is important to note that the standard error and coefficient of variation only measure sampling variability. They do not measure any systematic biases in the estimates.

The estimate from a particular sample and its associated standard error can be used to construct a confidence interval. A *confidence interval* is a range about a given estimator that has a specified probability of containing the average of the estimates derived from all possible samples. Associated with each interval is a percentage of confidence, which is interpreted as follows. If, for each possible sample, an estimate of a population parameter and its approximate standard error were obtained, then:

1. For approximately 90 percent of the possible samples, the interval from 1.65 standard errors below to 1.65 standard errors above the estimate would include the average of the estimates derived from all possible samples.

2. For approximately 95 percent of the possible samples, the interval from 1.96 standard errors below to 1.96 standard errors above the estimate would include the average of the estimates derived from all possible samples.

To illustrate the computation of a confidence interval for an estimate of total revenue, assume that an estimate of total revenue is $10,750 million and the coefficient of variation for this estimate is 1.8 percent, or 0.018. First obtain the standard error of the estimate by multiplying the total revenue estimate by its coefficient of variation. For this example, multiply $10,750 million by 0.018. This yields a standard error of $193.5 million. The upper and lower bounds of the 90-percent confidence interval are computed as $10,750 million plus or minus 1.645 times $193.5 million. Consequently, the 90-percent confidence interval is $10,432 million to $11,068 million. If corresponding confidence intervals were constructed for all possible samples of the same size and design, approximately 9 out of 10 (90 percent) of these intervals would contain the average of the estimates derived from all possible samples.

Nonsampling Error

Nonsampling error encompasses all other factors, other than sampling error, that contribute to the total error of a sample survey estimate and may also occur in censuses. It is often helpful to think of nonsampling error as arising from deficiencies or mistakes at some point in the survey process. Nonsampling errors are difficult to measure and can be attributed to many sources: inadequacies in the questionnaire, treatment of non-response, inaccurate reporting by respondents, errors in the application of survey procedures, incorrect recording of answers; differences in the interpretation of the survey questions; and other errors of collection, response, coverage, and processing. Additional nonsampling error may have been introduced by the method used to adjust the survey estimates using results of the 2002 Economic Census. Although no direct measurement of the effect of nonsampling error on the survey estimates has been obtained, the Census Bureau employs quality control procedures in all phases of the collection, processing, and tabulation of the data in an effort to minimize its influence.

A potential source of bias in the estimates is nonresponse. Nonresponse is defined as the inability to obtain all the intended measurements or responses about all selected units. Two types of nonresponse are often distinguished. *Unit nonresponse* is used to describe the inability to obtain any of the substantive measurements about a sampled unit. In most cases of unit nonresponse, the questionnaire was never returned to the Census Bureau, after several attempts to elicit a response. *Item nonresponse* occurs either when a question is unanswered or the response to the question fails computer or analyst edits.

For both unit and item nonresponse, a missing value is replaced by a predicted value obtained from an appropriate model for nonresponse. This procedure is called *imputation* and uses survey data and administrative data as input. For NAICS 51 (Information sector), imputed revenue amounts to about 10 percent of the total revenue estimate; for NAICS 54 (Professional, scientific, and technical services) imputed revenue amounts to about 25 percent; and for NAICS 56 (Administrative and support and waste management and remediation services) imputed revenue amounts to about 30 percent of total revenue.