

PART B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS

1. Universe and Respondent Selection

Sampling Frame: The sampling frame used for the ARTS 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 in the Census Bureau's Business Register as updated to October 2010.

To create the sampling frame we extract the records for all employer establishments located in the United States and classified in the Retail Trade or Accommodations and Food Services sectors as defined by the 2007 NAICS. For these establishments we extract sales, 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 retail establishments associated with the same EIN. Similarly, we aggregate the data to a multi-unit firm level by tabulating the establishment data for all retail 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 ARTS uses a stratified, one-stage design with primary strata defined by industry. There are 84 primary strata. We further stratify the sampling units within industry group by a measure of size (substratify) related to their annual sales. 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). Within each industry stratum 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 2007 Economic Census. Accordingly, these values are on a 2007 sales basis. We also used this analysis to determine the number of size substrata for each industry stratum and to set preliminary sampling rates needed to achieve specified sampling variability constraints on sales 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 sales or end-of-year inventories are greater than the

corresponding certainty cutoff, that firm is selected into the ARTS 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 sales (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 that 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 Retail 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 selected 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 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 ARTS 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 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 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.

Estimation procedures: 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 ARTS 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 ARTS. 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 ARTS sample in 2011 is not yet available, but will be provided online at http://www.census.gov/retail/how_surveys_are_collected.html. Also, published estimates from the ARTS are benchmarked using the latest results from the Economic Census as they become available as described on the website above.

2. **Procedures for Collecting Information**

There were approximately 21,775 firms selected in the latest ARTS sample. However, the mail count will vary for the 2011, 2012, and 2013 survey years. It will increase due to the quarterly birth process (see “Sample Maintenance” in Part B, Section 1). It may also increase if we require, or if firms request, that reporting responsibility be broken out by different locations or type of activity. It will decrease as we do not canvass companies that are out-of-business, out-of-scope, or merged with other firms. Typically, the mail count decreases each year of the sample. For example, the mail counts for the most recent completed surveys decreased over time as shown below:

<u>Year</u>	<u>Mail Count</u>	<u>% Decline</u>
2009	21,644	-2.2%
2008	22,124	-1.4%
2007	22,443	-0.2%

Report forms (Attachment 2) are typically mailed a few weeks after the data year ends. Forms can be returned via the Internet, fax, or mail. A cover letter (Attachment 3) is included that explains the necessity and uses of the data, as well as stating the due date and mandatory status of the survey.

Analysis of the data received is performed which might require an analyst to contact the firm to discuss. Examples of analysis conducted include a comparison of the current annual data to prior years, current annual data to monthly estimates produced by the Monthly Retail Trade Survey, and comparison of annual ratios such as inventory-to-sales.

3. Methods to Maximize Response

Two follow-ups, for delinquent cases, are sent via regular (1st follow-up) and certified (2nd follow-up) mail approximately two and three months, respectively, after the initial mail. Any remaining delinquent cases are followed up by telephone approximately one month after the first certified mailing. A second optional certified mailing is also available, if necessary, after the first calling period. In addition, a second optional calling period is available, if necessary, after required and optional mailings have been completed.

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:

<u>Year</u>	<u>Unit Response Rate</u>
2009	81.6%
2008	81.7%
2007	81.3%

Unit and dollar volume response rates are analyzed prior to closeout to ensure publication requirements are met. If any fail to meet those publication requirements additional follow-up with delinquent companies may be performed.

4. Testing of Procedures

Procedures in every phase of the ARTS are tested - from mailout and data capture to editing and publication.

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 for this survey to Aneta Erdie, Chief of the Annual Retail Branch, Service Sector Statistics Division, U.S. Census Bureau, 301-763-4841.

Attachments

1. BEA Support Letter
2. Copy of Report Forms
3. Copy of Cover Letter