

**SUPPORTING STATEMENT FOR
ANNUAL REFILEING SURVEY**

OMB CONTROL NO. 1220-0032

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

- 1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection methods to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.**

1a. Universe

The universe for the ARS survey includes all private sector employers subject to State UI laws in all 50 states; Washington, D.C.; Puerto Rico; and the Virgin Islands. It also includes Federal agencies whose employees are covered by UCFE. This will result in a universe of approximately 10.6 million private sector business establishments. The most current Quarterly Census of Employment and Wages (QCEW) files for each State serve as the ARS sampling frame.

1b. Sample Size

The survey was designed so that roughly one-third of private sector reporting units would be sampled each year. However, due to the Consolidated Appropriations Act 2014, the BLS, in coordination with the Department of Labor (DOL) and OMB, curtailed the QCEW program through a permanent budget reduction of \$4.8 million. This curtailment was accomplished, in part, by reducing the scope of and changing the approach to the ARS. To develop a reduction strategy with minimal impact on QCEW data quality, an analysis considered how frequently establishments changed their industry, location, and ownership. The analysis looked at these changes over three ARS cycles (nine years) and showed that there were industries that experienced low rates of change. Thus, certain NAICS industries exist that traditionally have low volatility based on factors such as infrastructure, intellectual property, or advanced degrees that make them unlikely candidates for transitioning from one industry to another. Rather than surveying one-third of businesses in these low-change industries every year, the QCEW program will survey one-sixth of these slow-changing businesses each year, putting them on a six-year cycle. The six-year cycle preserves the ability to identify trends in these industries and to maintain the accuracy of industry and geographic codes while reducing costs and workload.

- 2. Describe the procedures for the collection of information including:**
- **Statistical methodology for stratification and sample selection,**
 - **Estimation procedure,**
 - **Degree of accuracy needed for the purpose described in the justification,**
 - **Unusual problems requiring specialized sampling procedures, and**
 - **Any use of periodic (less frequent than annual) data collection cycles to reduce burden.**

2a. Sample Allocation and Selection Procedures

The ARS sampling frame used by BLS will be the most current QCEW file. A schedule was developed to balance the workload as evenly as possible so that around one-third of the businesses are surveyed each year. Certain NAICS industries that traditionally have low volatility will be put on a six-year cycle as described previously.

The selection criterion is based on the Federal Employer Identification Number (EIN) of the respondent, which provides a more random sample across all industries. The seventh and eighth positions of the nine-digit EIN are examined with ranges set for respondents to be surveyed. If the EIN is blank and unavailable, a range of digits within the UI account number will be used.

2b. Estimation Procedure

These data are being collected to ensure that the longitudinal QCEW data and the QCEW Report have the appropriate industrial and geographical codes. After data are edited and reconciled for apparent inconsistencies and completeness, these data will be the micro data on the longitudinal QCEW data. Estimates of totals at higher tabulation levels for the QCEW Report will be the sum of the appropriate micro level data. Since this is a census, the estimates are not subject to any sampling error. Therefore, no standard errors will be calculated.

2c. Accuracy

Because a census is being conducted, the estimates are not subjected to any sampling error.

2d. Problems

There are no unusual problems requiring specialized sampling procedures.

2e. Frequency

This survey will consist of a three-year cycle, with approximately one-third of all business establishments sampled each year. Establishments in certain NAICS industries that traditionally have low volatility will be put on a six-year cycle. Establishments assigned a NAICS code 999999 (Unclassified Establishments) are surveyed annually.

- 3. Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.**

3a. Response

The States are required to attain a 70% response rate in establishments or 80% in employment.

If necessary, two non-response follow-up email blasts and printed mailings of the ARS letters as well as telephone contact of key non-respondents are used to increase the response rates.

3b. Non-response Adjustment

Establishments that fail to respond after extensive follow-up efforts will be assigned NAICS codes by staff that have carefully reviewed establishment information gathered from company websites, business directories, and similar reputable sources.

3c. Reliability

Because this survey is a census, no sampling errors will be calculated.

To control non-sampling errors, quality control procedures are incorporated into the survey's design. These procedures include follow-up of all non-respondents and validation of all edit failures. The States and the BLS regional offices also receive industrial coding training and procedural assistance in conducting this survey. Regional offices conduct yearly quality assurance reviews in each State.

Automated ARS processing systems are used for several tasks, such as selecting the employers to be surveyed each year, preparing mailings, and editing the data. These systems further reduce the incidence and impact of non-sampling errors.

- 4. Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of test may be submitted for approval separately or in combination with the main collection of information.**

Work is proceeding on a text analysis tool called the NAICS autocoding tool. The tool will utilize inputs such as Legal and Trade Names, employer-provided information submitted via ARSweb written descriptions, and other pertinent items to determine the likelihood of a particular NAICS code appropriate for an establishment. Additional supplementary data like employment, wages, and firm age could also be used to refine this NAICS autocoding tool.

Potential uses for the tool would be for our State partners to assign NAICS codes to unclassified units, aid in decreasing nonresponse, and to identify units that could be misclassified. Development efforts are in the early stages but BLS has hopes to pilot the autocoding tool in the near future. This effort is consistent with BLS continually pursuing options to reduce employer burden and costs and to take advantage of technological innovations. BLS will test the NAICS autocoding tool and plans to put it into production upon successful completion of the development and testing phases.

5. **Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze person(s) who will actually collect and/or analyze the information for the agency.**

Mr. Ed Robison, Division Chief, Statistical Methods Staff, Office of Employment and Unemployment Statistics is responsible for the statistical aspects of this survey.