

Supporting Statement for Paperwork Reduction Act

Jobs Reporting under Section 1512 of the American Recovery and Reinvestment Act of 2009,
Public Law 111-5

B. Collections of Information Employing Statistical Methods

The agency should be prepared to justify its decision not to use statistical methods in any case where such methods might reduce burden or improve accuracy of results. When the question “Does this ICR contain surveys, censuses or employ statistical methods” is checked, “Yes,” the following documentation should be included in the Supporting Statement to the extent that it applies to the methods proposed:

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.

Study Objective: The goal of the study will be to estimate the accuracy of the total jobs reported by ARRA recipients.

Universe: Our universe is comprised of all prime awards required to report under section 1512, totaling 88,568 awards. The sampling frame used to estimate this universe will be all non-zero ARRA job reports submitted for the third quarter of calendar year 2010 by prime recipients of ARRA funds. The exact size of this sampling frame will depend on the number of reports submitted for the quarter ending September 30, 2010. The initial testing with reports from the previous quarter involved 249,295 reports, we expect the quarter tested to approximate this number.

Sample Selection: Recovery Board staff will be pulling a Brewer sample using pairs of observations from 40 strata, for a total sample size of approximately 81 recipients. Due to the wide range of the job numbers reported by recipients, small errors in the larger recipient numbers tend to alter the results significantly. Given this limitation we are expecting a fairly wide range of precision, currently estimated to be +/- 16 percent, assuming a two-sided, 95 percent confidence interval.

Response Rates: We conservatively expect the response rate to exceed 95 percent. This estimate is based on our experience with previous quarterly report response rates from prime recipients of Recovery Awards. In the second quarter of 2010, for instance, there were 312 non reporters out of 88,568 prime awards, constituting a non-response rate of 0.35 percent.

2. Describe the procedures for the collection of information including:

*** Statistical methodology for stratification and sample selection:**

The strata will be assigned based on evenly ascending size increments, with respect to jobs reported, each one having an upper limit approximately 1.3 x the size of the previous one. Exact spacing of the increments will depend on the data filed by recipients at the time.

*** Estimation procedure:**

Strata verification and construction: We will run a SAS program against data from the third quarter to simulate reasonable error scenarios and verify the viability of strata identified in our early testing. Strata will be adjusted as needed to optimize the precision of the results obtainable with third quarter data.

Sample Selection: A SAS program will be used to select a Brewer Sample from the third quarter data. The Brewer method is a probability proportionate to size method and sample observations will be weighted according to the number of jobs listed in that observation.

Estimation of Findings: The previously selected Brewer Sample will be run under a standard process for samples which have been selected with a probability proportionate to size. Means and standard errors will be calculated by SAS and applied to the universe from which the sample was drawn.

*** Degree of accuracy needed for the purpose described in the justification.**

Due to the wide range of the job numbers reported by recipients, small errors in the larger recipient numbers tend to alter the results significantly. Given this limitation we are expecting a fairly wide range of precision. We are currently estimating this to be in the range of +/- 16 percent, assuming a two-sided, 95 percent confidence interval. This estimate is based on test modeling using data from the quarter ending June 30, 2010.

*** Unusual problems requiring specialized sampling procedures.**

The wide numerical range of jobs reported required that we use a highly stratified method in order to group observations according to their likely rates of variance, and to group observations with high rates of variance together.

*** Any use of periodic (less frequent than annual) data collection cycles to reduce burden.**

We will use a single quarterly reporting period ending September 30, 2010.

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.

We are expecting our response rate to exceed 95 percent without any additional actions on our part due to the extensive oversight and reporting requirements established by congress at the time the funds were appropriated.

In addition, we have pulled a randomly selected, alternate sample unit (spare) for 37 of the 41 strata in the sample design, to cover non-responses, though we expect these to be rare. The remaining 4 strata will undergo a 100% audit, because of their size, and we are confident of their response because they are large state agencies.

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.

We used the standard reporting data submitted by recipients for the quarter ending June 30, 2010 to develop and test our statistical sampling and estimation methods. Several sampling methods were compared, including a simple random sample, stratified sampling, heavily stratified sampling, balanced replicated re-sampling (BRR) and a Brewer sample. All of these methods were compared for numbers of samples required and the accuracy yielded by each. The BRR and Brewer sample proved to be the best for this data.

Once the BRR and Brewer were established as the most viable sampling methods, error simulations were run with the test data to compare and validate the accuracy of these methods. The first series of error simulations involved randomly assigned error rates, averaging 10 percent. Both methods proved to be reasonably acceptable in detecting error levels under this simulation, but the BRR had a slightly narrower level of precision, and tended to slightly underestimate the true level of error. A second, more challenging series of error simulations was conducted using rates that varied from group to group, were higher than before in some groups, and left many groups with no errors. Using this second series of tests, the BRR method proved to be insufficiently accurate and its tendency to underestimate became more pronounced. Under these circumstances, variance estimations produced by the BRR were in error more often than the confidence interval would imply. Based on this testing it was determined that the Brewer was the best sampling method for this data.

Additional testing was done with the Brewer sample to determine the best stratification methodology for optimizing the precision available with widely ranging data.

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 the information for the agency.

Regina Van Houten – (202) 254-7990
Deborah Ben-Moshe – (202) 254-7977

Turner Bond - (202) 407-3533

OIRA has produced a number of documents that may serve as useful reference material for completing Supporting Statement B. These can be found at:

http://www.whitehouse.gov/omb/inforeg_statpolicy/