SUPPORTING STATEMENT PART B
FOR PAPERWORK REDUCTION ACT SUBMISSION
Gainful Employment Recent Graduates Employment and Earnings
Survey Pilot Test

B. Collection 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. The following documentation should be provided with the Supporting Statement Part A to the extent that it applies to the methods proposed. For further information, please consult OMB's Standards and Guidelines for Statistical Surveys.

1. Describe the potential respondent universe (including a numerical estimate) and any sampling or other respondent selection method 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 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.

    The sample frame for the RGEES pilot test is NSLDS, a transactional data base containing information about those who have been awarded federal student aid. The respondent universe consists of graduates from gainful employment programs who received federal student aid and who completed their program between July 1, 2009 and June 30, 2011. The target population is approximately 2,500,000.[1]

    Sampling and Expected Yield

    We will conduct a pilot study with a sample of 3,400 federal aid recipients selected from NSLDS. The pilot will be used to compare median earnings collected through the appeals survey to median earnings for graduates from comparable programs based on a match to the Social Security Administration as part of the 2012 gainful employment informational rates. The results of the pilot will also be compared to earnings estimates in the CPS and the ACS. The anticipated distribution of the sample by program area is provided in Table 1. The expected yield is based upon an assumed 60% response rate. The specific programs, listed by the code assigned in the NCES Classification of Instructional Programs (CIP), were selected based on the assumed likelihood of their graduates being surveyed under the gainful employment regulations.

---

[1] "Program Integrity: Gainful Employment; Final Rule." *Federal Register* 79 (31 October 2014): 65061.

Table 1: Pilot test starting sample and yield, by program

Program	CIP Code*	Sample size	Expected Yield
Total		**3,400**	**2,040**
Cosmetology and related personal grooming services	**12.04**	**700**	**420**
Somatic bodywork and related therapeutic services	**51.35**	**700**	**420**
Practical nursing, vocational nursing, and nursing assistants	**51.39**	**700**	**420**
All others	**All others**	**1,300**	**780**

*Includes all 6-digit CIP codes within each general category.

Nonresponse Bias Analysis

To the extent that those who respond to surveys and those who do not differ in important ways, there is a potential for nonresponse bias in estimates from survey data. Overall unit response rates for the RGEES are expected to be at least 50 percent of an identified cohort and, in keeping with the draft standards for its administration, a nonresponse bias analysis (NRBA) will be required when unit response rates are less than 80 percent. At the completion of the pilot study, the NRBA will compare respondents and nonrespondents within the program areas targeted (see Table 1). Demographic and other characteristics of both the respondents and nonrespondents needed for the NRBA will be obtained from NSLDS from which the sample will have been selected.

To be a respondent, a graduate must answer at least one of the earnings questions. Nonresponse may be due to refusal, noncontact, noninterview due to incapacity, language difficulties, etc. Nonresponse bias will be computed for the pilot study similarly to how it will be computed by institutions that conduct the survey. To measure nonresponse bias, we will compare respondents and nonrespondents. within program areas, on information available from graduates' student records. Analysis of recent program level data identified four variables as those most correlated with earnings—students with Pell grants, students with mothers with a bachelor's degree or above, students who are independent, and gender. The NRBA will examine these attributes of the program graduates to determine whether response rates vary on the attributes and whether the characteristics of respondents and nonrespondents differ on these characteristics. The measure of nonresponse bias will be calculated as the ratio of the difference between respondents and nonrespondents to the reported survey mean for a particular attribute. For a mean, an estimate of the bias of the respondent mean is given by:

$$B(\bar{y_r}) = \bar{y_r} - \bar{y_t} = (n_{nr}/n)(\bar{y_r} - \bar{y_{nr}})$$

Where:

y_t = the mean based on all cases;
y_r = the mean based only on respondent cases;
y_{nr} = the mean based only on the nonrespondent cases;
n = the total number of cases; and
n_{nr} = the number of nonrespondent cases.

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.

This section describes the data collection procedures to be used in the RGEES pilot test. Once selected for participation in the pilot survey, sample members will receive several mailings, depending on the speediness of their reply and the stage of data collection. In early September, sample members will receive an initial mailing containing introductory letters from RTI and NCES and a paper copy of the survey with a business reply envelope. They will also be provided with log in credentials to complete the survey on line.

Nonrespondents to the initial mailing will receive a series of follow-up mailings, including a postcard and a second letter, sent from RTI by Priority Mail, with another copy of the survey and a business reply envelope. In addition to hard copy mailings, sample members with email addresses on the NSLDS data base will receive an email reminder with log in credentials for direct access to the web survey, approximately every 7 to 10 days throughout the course of data collection. The survey instrument is provided in Appendix 2 and respondent contact materials are provided in Appendix 3.

Survey Monitoring

Survey returns will be processed upon receipt, and reports from the survey management system will be prepared at least weekly. The reports will be used to continually assess the progress of data collection.

3. Describe methods to maximize response 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.

The RGEES incorporates a number of features to maximize response rates. This section discusses those features.

Total Design Method/Respondent-Friendly Design. Surveys that take advantage of respondent-friendly design have demonstrated increases in survey response (Dillman, Smyth, and Christian 2008; Dillman, Sinclair, and Clark 1993). As noted previously, we will include mention of the respondent incentive in the initial mailing. The data collection contractor will maintain an email address and a toll-free questionnaire assistance (TQA) line to answer respondent questions or concerns. If a respondent chooses to provide their information to the TQA staff, staff will be able to collect the respondent's information over the phone.

Engaging Respondent Interest and Cooperation. The content of respondent letters, postcards, and email is focused on communicating the legitimacy and importance of the study, how pilot data will be used, and that there is minimal burden required in completing the survey (5 minutes). To further encourage participation within the timeline of the pilot test, all participants will be offered a $25 incentive for completing the survey. Respondents will be able to choose between receiving the incentive payment either by check or by PayPal, sent to an email address provided in the survey, once receipt of the survey is confirmed. Choice of payment method will be included at the end of the survey where participants will be asked to provide an email address if they choose PayPal.

Nonresponse Follow-up. As listed above, the data collection protocol includes several nonresponse follow-up activities throughout the 2-month data collection. In addition to the number of contacts, changes in method of the follow-up contact (e.g., Priority Mail) are designed to capture the attention of sample members. Emails will provide credentials to allow sample members to jump directly to the survey by clicking a link provided.

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 tests may be submitted for approval separately or in combination with the main collection of information.

   Not applicable.

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 persons who will actually collect and/or analyze the information for the agency.

   The persons listed below participated in the study design and are responsible for the collection and analysis of the data:

   Sharon, Boivin, NCES

Chris Chapman, NCES
Marilyn Seastrom, NCES
Sean Simone, NCES
T. Austin Lacy, RTI International
Peter Siegel, RTI International
Jennifer Wine, RTI International
Erin Dunlop Velez, RTI International