2016/17 BACCALAUREATE AND BEYOND (B&B:16/17) FIELD TEST

Supporting Statement Part B

OMB # 1850-0729 v.11

Submitted by
National Center for Education Statistics
U.S. Department of Education

January 2016
Revised April 2016

Contents

Tables

B. Collection of Information Employing Statistical Methods

This submission requests clearance for the 2016/17 Baccalaureate and Beyond Longitudinal Study (B&B:16/17) field test data collection instrument and methods. B&B:16/17 is the first follow-up of sample members from the 2015-16 National Postsecondary Student Aid Study (NPSAS:16) who were baccalaureate recipients during the 2014-15 (field test) and 2015–16 (full scale) academic years. For details on the NPSAS:16 sampling design and the final field test and full-scale study design, see NPSAS:16 Field Test (OMB# 1850-0666 v. 12-14) and NPSAS:16 Full Scale (OMB# 1850-0666 v. 15-17) Supporting Statements Part B. Specific plans are provided below for the B&B:16 cohort.

1. Respondent Universe

The respondent universe for the full-scale B&B study consists of all persons who completed requirements for the bachelor's degree during the 2015–16 academic year, and received their degree by June 30, 2017. These respondents will be surveyed for B&B in 2017. For the B&B field test, the respondent universe is the same except that the survey year is 2016 and respondents will have completed their bachelor's degree in the 2014–15 academic year, and received their degree by June 30, 2016.

2. Statistical Methodology

a. Field Test Design

The B&B:16/17 field test will be implemented to fully test all procedures, methods, and systems planned for the B&B:16/17 full-scale study, in a realistic operational environment, prior to implementing them in the full-scale study. The field test will be designed to test and validate data collection and monitoring procedures that will obtain the most accurate data in the least amount of time. Specific field test evaluation goals include the following:

- determining how to identify actual baccalaureate recipients from the various data sources, including: the NPSAS:16 institution enrollment list; NPSAS:16 student-level institution records; student report of degree receipt in the NPSAS:16 interview; the B&B:16/17 interview; and data obtained from extant data sources, including Central Processing System (CPS), National Student Clearinghouse (NSC), and National Student Loan Data System (NSLDS);
- assessing the quality, completeness, and effectiveness of various types of locating data obtained during the NPSAS:16 base year study;
- evaluating the utility of pre-data collection submission of sample members to Telematch, National Change of Address (NCOA), and CPS as a mechanism for obtaining updated locating information; and
- identifying problematic data elements in the B&B student interview.

Additionally, we will evaluate the time required to complete the interview, and sections of the interview, in order to identify possible instrument modifications that will save time in the full-scale interview. We will also conduct a reliability re-interview to evaluate the temporal consistency of selected interview items.

The interview will determine eligibility, however, as part of the field test data collection, we will assess the value in determining cohort eligibility from each of the other independent data sources listed in the first bullet above. The cases

determined to be eligible will comprise the cohort for continuation into any additional B&B follow-up studies.

The field test sample for B&B:16/17 will consist of all interview respondents from the NPSAS:16 field test who completed or expected to complete the requirements for their bachelor's degree at any time between July 1, 2014, and June 30, 2015, and received their degree by June 30, 2016. For the full-scale study, bachelor's degree requirements must be completed between July 1, 2015 through June 30, 2016, and sample members must have received their degree by June 30, 2017. The NPSAS:16 field test yielded 1,300 interview respondents who indicated in the NPSAS field test interview that they had completed or expected to complete the requirements for their bachelor's degree in the 2014-2015 academic year. Additionally, we plan to include all potentially B&B-eligible NPSAS:16 field test interview nonrespondents in the B&B:16/17 field test sample. These were individuals listed as potential baccalaureate recipients by their NPSAS institution but who did not complete a NPSAS:16 student interview to indicate their B&B-eligible status. The base-year sample included 798 interview nonrespondents who were identified as potential bachelor's recipients according to the initial classification by the NPSAS sample institution at the time of student sampling (prior to base-year data collection). Therefore, the total B&B:16/17 field test sample size will be 2,098.

Because of the number of NPSAS:16 field test interview nonrespondents being included in the B&B field test (see Section 4, Tests of Procedures and Methods for further description of the field test study design plan), the eligibility rate for the field test sample is anticipated to be about 85 percent after interviewing and matching to administrative records, which should yield an eligible sample size of about 1,794. The response rate is expected to be about 75 percent among the eligible sample members, which will yield about 1,351 responding bachelor's recipients. Table 1 provides the expected B&B:16/17 field test eligibility and response rates by base-year response status.

Table 1. Expected B&B:16/17 field test eligibility and response rates by base-year response status

NPSAS:16 field test interview response status		Expected eligible		Expected respondents	
	Number	Number	Rate	Number	Rate
Total	**2,098**	**1,794**	**85%**	**1,351**	**75%**
Respondent	1,300	1,235	95%	1,099	88%
Nonrespondent	798	559	70%	252	45%

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2016 National Postsecondary Student Aid Study (NPSAS:16).

b. Full-scale Design

The sample for the first follow-up with the B&B:16 cohort will contain students who completed or expected to complete requirements for the bachelor's degree in the 2015-16 academic year, as indicated by the NPSAS:16 full-scale student interviews (expected to number approximately 25,000). In addition, to have full population coverage of the B&B sample, a subsample of the approximately 10,700 anticipated NPSAS:16 interview nonrespondents who are either listed by the NPSAS sample institution as bachelor's degree candidates or indicated by student records to be

degree candidates, will be selected. The results from the field test examination of data sources will determine the decision rules and data sources to be used to determine eligibility for the full-scale study, as well as the allocation for the subsample of the nonrespondents.

The starting full-scale sample size will be about 25,556, and the expected eligibility rate of the sample members will be about 90 percent after interviewing and matching to administrative records, which will give an eligible sample size of about 23,000. We also expect a response rate of about 87 percent among the eligible sample members, which will yield about 20,053 responding baccalaureate recipients. The sample size and eligibility and response rates are subject to change based upon field test results.

Table 2. Expected B&B:16/17 full-scale study eligibility and response rates by base-year response status

NPSAS:16 full-scale interview response status		Expected eligible		Expected respondents	
	Number	Number	Rate	Number	Rate
Total	**25,556**	**23,000**	**90%**	**20,053**	**87%**
Respondent	25,000	22,611	90%	19,897	88%
Nonrespondent	556	389	70%	156	40%

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2016 National Postsecondary Student Aid Study (NPSAS:16).

3. Methods for Maximizing Response Rates

Response rates in the B&B:16/17 field test and full-scale data collections are a function of success in two basic activities: identifying and locating the sample members involved, then contacting them and gaining their cooperation. Below we present our plans for maximizing response to the student survey.

a. Student Survey: Self-Administered Web and CATI

The following sections outline methods for maximizing response to the B&B:16/17 student survey.

1. Tracing of Sample Members

To achieve the desired response rate, we propose an integrated tracing approach designed to yield the maximum number of locates with the least expense. During the field test, we will evaluate the effectiveness of these procedures for the full-scale study effort. The steps of our tracing plan include the following elements.

- **Advance Tracing.** The advance tracing stage includes tracing steps taken prior to the start of data collection. These include batch database searches and advance intensive tracing (if necessary). Some cases will require more advanced tracing, before mailings can be sent or the cases can be worked in CATI. To handle cases for which mailing address, phone number, or other contact information is invalid or unavailable, B&B staff plan to conduct advance tracing of the cases prior to lead letter mailout and data collection. As lead information is found, additional searches will be conducted through interactive databases to expand on leads found.
- Data collection mailings and e-mails will be used to maintain persistent contact

with sample members as needed prior to and throughout data collection. Initial contact letters will be sent to parents and sample members prior to the start of data collection upon OMB clearance. The letter will alert sample members to their inclusion in the study and request that they, or their parents, provide any contact information updates. Following the initial contact letters, the data collection announcement letter will be sent to all sample members on July 11 (the first day of data collection) to announce the start of data collection. The data collection announcement will include a toll-free number, study website address, and study ID and password, and will request that sample members complete the web survey. Two days after the data collection announcement mailing, an email message mirroring the letter will also be sent to sample members.

- **Telephone Locating and Interviewing.** The telephone locating and interviewing stage includes calling all available telephone numbers and following up on leads provided by parents and other contacts.
- **Pre-Intensive Batch Tracing.** The pre-intensive batch tracing stage consists of the LexisNexis SSN and Premium Phone batch searches that will be conducted between the telephone locating and interviewing stage and the intensive tracing stage.
- **Intensive Tracing.** The intensive tracing stage consists of tracers conducting database searches after all current telephone numbers have been exhausted. In B&B:08/09, about 77 percent of sample members requiring intensive tracing were located, and about 46 percent of those located responded to the interview. In B&B: 08/12, about 89 percent of sample members requiring intensive tracing were located, and about 39 percent of those located responded to the interview. Intensive interactive tracing differs from batch tracing in that a tracer can assess each case on an individual basis to determine which resources are most appropriate and the order in which they should be used. Intensive interactive tracing is also much more detailed due to the personal review of information. During interactive tracing, tracers utilize all previously obtained contact information to make tracing decisions about each case. These intensive interactive searches are completed using a special program that works with RTI's Computer-Assisted Telephone Interviewing Case Management System (CATI-CMS) to provide organization and efficiency in the intensive tracing process. Sources that may be used, as appropriate, include credit database searches, such as Experian, various public websites, and other integrated database services.
- **Other Locating Options.** Other locating activities will take place as needed, including a LexisNexis e-mail search conducted toward the end of data collection for nonrespondents.

2. Training for Data Collection Staff

Telephone data collection will be conducted at the contractor's Research Operations Center (ROC). B&B staff at the ROC will include Performance Team Leaders (PTLs) and Data Collection Interviewers (DCIs). Training programs for these staff members are critical to maximizing response rates and collecting accurate and reliable data.

Performance Team Leaders, who are responsible for all supervisory tasks, will attend project-specific training for PTLs, in addition to the interviewer training. They will

receive an overview of the study, background and objectives, and the data collection instrument through a question-by-question review. PTLs will also receive training in the following areas: providing direct supervision during data collection; handling refusals; monitoring interviews and maintaining records of monitoring results; problem resolution; case review; specific project procedures and protocols; reviewing CATI reports; and monitoring data collection progress.

Training for DCIs is designed to help staff become familiar with and practice using CATI-CMS and the survey instrument, as well as to learn project procedures and requirements. Particular attention will be paid to quality control initiatives, including refusal avoidance and methods to ensure that quality data are collected. DCIs will receive project-specific training on telephone interviewing and answering questions from web participants regarding the study or related to specific items within the interview. At the conclusion of training, all B&B ROC staff must meet certification requirements by successfully completing a certification interview. This evaluation consists of a full-length interview with project staff observing and evaluating interviewers, as well as an oral evaluation of interviewers' knowledge of the study's Frequently Asked Questions.

3. Case Management System

Student interviews will be conducted using a single web-based survey instrument for both web and CATI data collection. The data collection activities will be accomplished through the CATI-CMS, which is equipped with the numerous capabilities, including: on-line access to locating information and histories of locating efforts for each case; state-of-the-art questionnaire administration module with full "front-end cleaning" capabilities (i.e., editing as information is obtained from respondents); sample management module for tracking case progress and status; and automated scheduling module which delivers cases to interviewers. The automated scheduling module incorporates the following features:

- *Automatic delivery of appointment and call-back cases at specified times.* This reduces the need for tracking appointments and helps ensure the interviewer is punctual. The scheduler automatically calculates the delivery time of the case in reference to the appropriate time zone.
- *Sorting of non-appointment cases according to parameters and priorities set by project staff*. For instance, priorities may be set to give first preference to cases within certain sub-samples or geographic areas; cases may be sorted to establish priorities between cases of differing status. Furthermore, the historic pattern of calling outcomes may be used to set priorities (e.g., cases with more than a certain number of unsuccessful attempts during a given time of day may be passed over until the next time period). These parameters ensure that cases are delivered to interviewers in a consistent manner according to specified project priorities.
- *Restriction on allowable interviewers*. Groups of cases (or individual cases) may be designated for delivery to specific interviewers or groups of interviewers. This feature is most commonly used in filtering refusal cases, locating problems, or foreign language cases to specific interviewers with specialized skills.
- *Complete records of calls and tracking of all previous outcomes*. The scheduler tracks all outcomes for each case, labeling each with type, date, and time. These

are easily accessed by the interviewer upon entering the individual case, along with interviewer notes.

- *Flagging of problem cases for supervisor action or supervisor review*. For example, refusal cases may be routed to supervisors for decisions about whether and when a refusal letter should be mailed, or whether another interviewer should be assigned.
- *Complete reporting capabilities*. These include default reports on the aggregate status of cases and custom report generation capabilities.

The integration of these capabilities reduces the number of discrete stages required in data collection and data preparation activities and increases capabilities for immediate error reconciliation, which results in better data quality and reduced cost. Overall, the scheduler provides a highly efficient case assignment and delivery function by reducing supervisory and clerical time, improving execution on the part of interviewers and supervisors by automatically monitoring appointments and call-backs, and reducing variation in implementing survey priorities and objectives.

4. Survey Instrument Design

Student interview preparation involved a meeting of the B&B technical review panel in October 2015. The focus of this first meeting (with the second to be held in late fall of 2016 after the field test data collection) was to discuss the key data elements to be included in the field test student interview data collection, where employment outcomes of bachelor's degree recipients and their consideration of a career in pre-kindergarten through 12th grade teaching will feature prominently.

The B&B:16/17 student interview employs a web-based instrument and deployment system, created by RTI, known as Hatteras which has been in use since NPSAS:08. Hatteras is a flexible system that provides multimode functionality, whereby the survey instrument is created one time and can be used for self-administration, including on mobile devices, CATI, CAPI, or data entry. The instrument is provided in appendix D.

In addition to the functional capabilities of the CATI-CMS and web instruments described above, our efforts to achieve the desired response rate will include using established procedures proven effective in other large-scale studies we have completed. These include:

- Providing multiple response modes, including mobile-friendly self-administered and interviewer-administered options.
- Offering incentives to encourage response (see incentive structure described in Section 4, Tests of Procedures and Methods).
- Assigning experienced CATI interviewers who have proven their ability to contact and obtain cooperation from a high proportion of sample members.
- Training the interviewers thoroughly on study objectives, study population characteristics, and approaches that will help gain cooperation from sample members.
- Maintaining a high level of monitoring and direct supervision so that interviewers who are experiencing low cooperation rates are identified quickly and corrective action is taken.

- Making every reasonable effort to obtain an interview at the initial contact, but allowing respondent flexibility in scheduling appointments to be interviewed.
- Thoroughly reviewing all refusal cases and making special conversion efforts whenever feasible (see next section).

5. Refusal Aversion and Conversion

Recognizing and avoiding refusals is important to maximize the response rate. We will emphasize this and other topics related to obtaining cooperation during interviewer training. PTLs will monitor interviewers intensely during the early days of outbound calling and provide retraining as necessary. In addition, the supervisors will review daily interviewer production reports produced by the CATI system to identify and retrain any data collectors who are producing unacceptable numbers of refusals or other problems.

Refusal conversion efforts will be delayed for at least one week to give the respondent time after the initial refusal. Attempts at refusal conversion will not be made with individuals who become verbally aggressive or who threaten to take legal or other action. Refusal conversion efforts will not be conducted to a degree that would constitute harassment. We will respect a sample member's right to decide not to participate and will not impinge this right by carrying conversion efforts beyond the bounds of propriety.

4. Tests of Procedures and Methods

The B&B:16/17 field test will include two data collection experiments: the first set of experiments focuses on survey participation and aims at reducing nonresponse error, and the second experiment focuses on minimizing measurement error to further improve data accuracy.

a. Experiment #1: Finding the optimum strategy to minimize sampling design weight variation and nonresponse bias and boost response rates

Introduction. The B&B cohort, a historically highly-responsive, relatively-homogenous group, presents favorable, yet unique, challenges for data collection:

- The last several B&B cohorts have participated in the B&B student interview at high response rates, ranging from 85%-92%;
- B&B sample members who responded in the NPSAS (base-year) studies have responded at higher rates in B&B follow-up studies than the base-year interview nonrespondents;
- In past B&B administrations, great sampling design weight variation (e.g., in B&B:08/12, the mean weight in some sectors exceeded 100) could partially be attributed to the low subsampling rate (about 10%) for NPSAS (base-year) nonrespondents who were potentially eligible for B&B;
- B&B:08 base-year nonrespondents showed significant nonresponse bias in some survey estimates based on data collected during the first follow-up (B&B:08/09), while base-year respondents did not.

Given the historical attributes of the B&B cohorts, our approach for the B&B:16/17 field test involves increasing the subsampling rate of base-year nonrespondents and

separating base-year nonrespondents and base-year respondents into four groups targeted for different intensities of data collection protocols. At the same time, we also plan to conduct an observational Mahalanobis modeling procedure that will help us better understand how individual Mahalanobis values change over times, as data collection proceeds. The field test design proposed may allow us to better minimize sampling design weight variation, increase response rates for base-year nonrespondents, and minimize the risk of nonresponse bias in B&B:16/17 full scale estimates and those of follow-up waves. The details of the design are described below.

B&B:16/17 field test design. Our experimental design brings into the sampling pool all of our NPSAS:16 field test interview nonrespondents who are potentially B&B-eligible (n=798), and assigns them either to an "aggressive" data collection protocol, or a "default", more standard, protocol used for base-year interview nonrespondents.

Our approach treats both the NPSAS and B&B data collection efforts as one continuous data collection. That is, we have already learned something about the sample members by the way they have responded to the initial interview request during NPSAS:16, and we will intervene at our second request (B&B:16/17). Considered in this way, our design is analogous to double-sampling for nonresponse (Deming, 1953; Hansen & Hurwitz, 1946) where a random subsample of nonrespondents (in this case, NPSAS:16 field test interview nonrespondents) is selected for a more aggressive, more costly, data collection protocol. In addition to reducing the risk of nonresponse bias, this approach is expected to contribute to a higher response rate for base-year interview nonrespondents and can inform the nonresponse adjustments for the full-scale study.

Implementation of the experimental design includes placing the B&B-eligible sample into four groups based on their NPSAS:16 field test interview response status:

- Group 1: Randomly-selected half of the NPSAS:16 field test interview nonrespondents
- Group 2: Randomly-selected half of the NPSAS:16 field test interview nonrespondents
- Group 3: NPSAS:16 field test (late) interview respondents
- Group 4: NPSAS:16 field test early interview respondents (responded in the first three weeks of the NPSAS:16 field test interview data collection effort)

Each of these groups will receive one of three data collection protocols: relaxed, default, or aggressive. Table 3 below shows which groups receive which protocol:

Table 3. B&B:16/17 field test data collection protocols by base-year respondent type

NPSAS:16 field test interview response status	Data collection protocol		
	Aggressive	Default	Relaxed
Nonrespondent	Group 1 (n=399)	Group 2 (n=399)	n/a
Respondent	n/a	Group 3 (n=604)	Group 4 (n=696)

The B&B field test will include three distinct phases of data collection: **the early completion phase** (first four weeks of data collection), **the production phase** (second 10 weeks of data collection), and the **nonresponse conversion phase** (final four weeks of data collection).

The data collection strategies to be employed in the varying protocols have been previously approved by OMB in prior B&B studies and include: an initial email and letter in the early completion phase and reminder emails and postcards to complete the survey throughout data collection, prepaid incentives and promised incentives of no more than $30, CATI locating and tracing efforts, and the offer of an abbreviated interview (although an abbreviated interview is not always offered in a field test design, it is important to include in this field test as a measure of utility for encouraging response, for the full-scale design). The B&B:16/17 field test is unique in its proposal to combine data collection tools into these overall protocols, with varying levels of intensity for the different groups as indicated below:

- The “aggressive” data collection protocol, offered only to Group 1 (a randomly assigned half of the NPSAS:16 field test interview nonrespondents) includes a $10 prepaid incentive. A PayPal prepaid incentive notice will be sent to Group 1 via email, but would also state that, if preferred, the $10 can be requested by the sample member in the form of a check, which would take approximately 4 weeks to process (see appendix C for the language used in the email). Offering the prepaid incentive via PayPal allows us to test the email addresses for Group 1, who are base-year nonrespondents. In addition, any member of Group 1 who neither accepts the PayPal prepaid incentive nor asks for a check instead, but who still completes the survey, would receive the full $30 upon completion. Those members of Group 1 who do accept the prepaid incentive, will receive $20 upon completion of the survey. Other features of the aggressive data collection protocol include CATI interviewing beginning in the early completion phase (two weeks after data collection begins), and a 10-minute abbreviated interview beginning in the production phase of data collection.
- Data collection plans for Groups 2 and 3 are based on standard protocols for base year interview nonrespondents and base year interview respondents, respectively. The “default” data collection protocol for base-year nonrespondents is reflected in Group 2: a CATI collection effort beginning at the start of the production phase, a 10-minute abbreviated interview in the nonresponse conversion phase, and a $30 incentive offer to complete the survey. The “default” data collection protocol for base year respondents is reflected in Group 3 and differs from Group 2’s treatment only in that the CATI collection effort would begin 2 weeks after the start of the production phase and be conducted with less intensity (“Light” CATI) during the production phase of data collection than for Group 2. The default protocol for Group 3 also includes the 10-minute abbreviated interview in the nonresponse conversion phase and the $30 incentive offer to complete the survey.
- The “relaxed” data collection protocol, offered only to Group 4 (NPSAS:16 field test interview early respondents) includes a $20 incentive offer to complete the survey; this protocol does not include any CATI contact and also does not include an abbreviated interview offer.

Table 4 below provides comparisons for how data collection protocols would vary among the B&B:16/17 field test sample member groups.

Table 4. B&B:16/17 field test data collection protocols by data collection phase

	Group 1: NPSAS:16 FT interview nonrespondents-- Aggressive protocol (N=399)	Group 2: NPSAS:16 FT interview nonrespondents-- Default protocol (N=399)	Group 3: NPSAS:16 FT interview respondents-- Default protocol (N=604)	Group 4: NPSAS:16 FT interview early respondents-- Relaxed protocol (N=696)
Early Completion Phase – First 4 weeks of Data Collection July 11, 2016 – August 8, 2016	Initial email and letter offering $10 prepaid incentive via PayPal or check on first day of data collection Begin CATI interviewing/ locating 2 weeks after data collection begins	Initial email and letter Email reminders	Initial email and letter thanking for previous participation Email reminders	Initial email and letter thanking for previous participation Email reminders
Production Phase August 8, 2016 – October 14, 2016	Frequent email reminders Postcard reminders Offer abbreviated interview	Begin CATI interviewing/ locating at the **start** of the production phase Frequent email reminders	Begin "Light" CATI interviewing/ locating to begin ***2 weeks after*** the start of production phase Frequent email reminders	No CATI contact Frequent email reminders
Nonresponse Conversion Phase October 14, 2016 – November 11, 2016	Continued CATI interviewing/ locating efforts Frequent email reminders Postcard reminders Continue offering abbreviated interview	Continued CATI interviewing/ locating efforts Frequent email reminders Postcard reminders Offer abbreviated interview	Continued CATI interviewing/ locating efforts Frequent email reminders Postcard reminders Offer abbreviated interview	E-mail reminders Postcard reminders No CATI contact No abbreviated interview
Incentive amount	$10 prepay + $20 upon interview completion ($30 total)	$30 upon interview completion	$30 upon interview completion	$20 upon interview completion

Observational Mahalanobis in the B&B:16/17 field test design. In addition, during the field test, we propose to conduct a daily Mahalanobis calculation to observe how the group data protocols affect individual cases, although no resulting interventions would be conducted during the field test due to the small sample size.

A B&B:08/12 full-scale study responsive design experiment (discussed in detail in the B&B: 08/12 Supporting Statement Part B, OMB# 1850-0729 v.8) used Mahalanobis calculations to identify sample members who were most likely to contribute to nonresponse bias if they did not respond and to reduce nonresponse bias by increasing response among the identified cases. B&B:08/12 tested a responsive data collection design based on a multivariate distancing measure known as the Mahalanobis distance. Prior to the start of data collection, all sample cases were randomly assigned to control and treatment groups. A Mahalanobis calculation based on the multivariate distance between the baseline respondent average and an individual nonrespondent was calculated at three points during the B&B:08/12 data collection. Nonrespondents with high Mahalanobis distances were considered the most likely to contribute to nonresponse bias, and those in the

treatment group were selected for targeted interventions. Cases targeted during the first intervention point received a $15 increase in their promised incentive. Cases targeted as part of the second intervention received a $5 prepay via FedEx. Cases targeted at the third, and final, intervention were offered an earlier abbreviated interview.

Results of the experiment showed that the Mahalanobis distance had some success identifying cases with different outcome measures and that the $5 prepaid incentive and abbreviated interview offer increased response rates. However, nonresponse bias analyses did not show significant reductions in mean nonresponse bias across variables between the treatment and control groups across sampling frame variables or those variables included in the Mahalanobis model. Bias may have been reduced for some individual variables, though. For more details about the B&B:08/12 Mahalanobis experiment and results, see the B&B:08/12 Data File Documentation.

While the Mahalanobis experiment had mixed results for B&B:08/12, it may be worth re-considering using Mahalanobis for B&B:16/17 full-scale. If the proposed experiment described above is successful in obtaining responses from base year NPSAS nonrespondents, then an increased number of these nonrespondents could potentially be subsampled. Mahalanobis could be used to target the B&B:16/17 nonrespondents most likely to contribute to nonresponse bias. In B&B:08/12, the cases with the highest distance measure were, generally, the nonrespondents to the previous two rounds.

For the B&B:16/17 field test, we would identify variables from the NPSAS:16 field test sampling frame and NPSAS:16 field test data that are likely to be related to nonresponse bias, and use those variables to compute the Mahalanobis distance each day. This measure would be tracked to determine if the distribution of Mahalanobis changes over time. A reduction in the average value among B&B:16/17 nonrespondents may indicate that the data collection protocol is helping to obtain responses from cases with higher Mahalanobis values. Due to the small field test sample size, we cannot do a formal Mahalanobis experiment or test for significant differences. However, an observed reduction along with a successful data collection experiment may suggest that is worthwhile to try Mahalanobis in the full-scale.

If any effects from the Mahalanobis modeling are shown during the B&B:16/17 field test, we could conduct Mahalanobis calculations during the full-scale study with actual interventions for the cases with high Mahalanobis distances.

Proposed evaluation of the design. The B&B:16/17 field test design will allow us to evaluate whether there is a reduction in bias, increased response rates, and cost savings in data collection gained by employing the different protocols by respondent type. Results of particular interest are listed below.

1. We will evaluate the response rates for the NPSAS: 16 field test interview nonrespondents who receive the aggressive data collection protocol (Group 1) and those who receive the default data collection protocol (Group 2). Data collection efforts that target nonrespondents typically use any type of effort,

within budget, to convert such cases. Because of that, we are mostly interested in the overall aggressive protocol effectiveness, rather than the effectiveness of its components separately. Nevertheless, because of the timing of the different components (prepaid incentives, CATI) in the three phases of data collection, we will be able to isolate, for example, the effect of the prepaid versus no prepaid incentive. If the aggressive protocol is found to be successful in the field test, and we want to understand which component is mostly driving this success, we can plan a full experimental design in the full-scale data collection, where sample size will be less of an issue (in the field test, we would have a starting sample of about 60 cases in some cells and our tests of the effectiveness of each treatment would be severely underpowered).

2. Difference in substantive responses in the B&B:16/17 field test interview (e.g., post baccalaureate employment, employer/job changes, earnings, dependents, marriage) between NPSAS: 16 field test interview nonrespondents receiving the aggressive data collection protocol (Group 1) and those receiving the default data collection protocol (Group 2), and between NPSAS:16 field test interview nonrespondents (Groups 1 and 2) and respondents (Groups 3 and 4), will help us to understand more about whether bringing more nonrespondents into the sample reduces bias.
3. We are interested, as a cost containment measure, in whether the lesser incentive amount for the NPSAS:16 field test early respondents in the relaxed data collection protocol (Group 4) affects response rates compared with the higher incentive amount for the NPSAS:16 field test respondents in the default data collection protocol (Group 3).
4. Observing the Mahalanobis distance measure daily will help us understand how Mahalanobis works with the B&B:16/17 cohort. As described above, we will track the distribution of Mahalanobis to observe how it changes during data collection. We will also see what proportion of nonrespondents have high distance measures, and if the proportion is low, then this may show that Mahalanobis is not necessary for the full-scale study.

If the proposed interventions are successful in encouraging base-year nonrespondents to respond to the field test survey, we can begin with these interventions for base-year nonrespondents during the full-scale study, and use the Mahalanobis responsive-design calculation to target with additional interventions, the cases which most contribute to bias.

It is important to note that if the field test does not indicate there is value in using a responsive design technique involving Mahalanobis to reduce nonresponse bias in the B&B:16/17 full-scale study, an alternative responsive design being employed in the NPSAS:16 base-year full-scale study may be considered for the B&B:16/17 full-scale study depending on results from NPSAS. The NPSAS:16 study aims at reducing variance rather than bias by using multiple imputation to identify study members who are interview nonrespondents and who are likely to have variability in their imputed data for key items. These study members are then targeted with an abbreviated interview containing these key items. For more detail on the NPSAS:16 full-scale data collection see the NPSAS:16 Supporting Statement Part B, OMB# 1850-0666 v.16. A similar approach could improve the quality of the B&B data. Key

interview items would be determined, and respondents with missing data who are likely to have variability in their imputed data would be identified. Then, B&B nonrespondents with similar characteristics to the respondents with high variability would be targeted in data collection to improve the imputation donor pool and reduce variance due to imputation.

b. Experiment # 2: Questionnaire Design

Three sections of the B&B:16/17 field test questionnaire contain a set of looping items that are repeated for survey respondents based on their responses to gate items. A set of looping items can be structured in two ways: *interleafed*, in which the loop items are asked immediately after a gate item and, *grouped*, in which the loop items are asked after multiple gate items. Research has found the grouped format provides more accurate responses than the interleafed format (Eckman et al., 2014). We propose dividing the survey respondents into two groups to determine if the grouped loop structure will increase accuracy and respondent efficiency. We will evaluate accuracy by comparing aggregate responses to external benchmarks for the survey topics covering the looping items, including: undergraduate education institutions, postbaccalaureate institutions, and employers and jobs.

Sample members will be assigned at random into one of two conditions described below.

1. *Treatment Group 1*: Sample members will receive the *grouped* format for all three looping sets in the B&B:16/17 field test survey instrument. The gate items will determine the total number of times the respondent is routed through each looping set. For each looping set, respondents will be asked to provide:
    a. *Undergraduate postsecondary institutions*: whether they attended a postsecondary institution, other than the NPSAS institution, prior to completing bachelor's degree; then, total number and name(s) of postsecondary institutions attended; finally, detail on degrees at each institution attended.
    b. *Postbaccalaureate postsecondary institutions*: whether they attended a postsecondary institution after completing bachelor's degree; then, total number and name(s) of postsecondary institutions attended after completing bachelor's degree; finally, detail on degrees at each institution attended.
    c. *Postbaccalaureate employers and jobs*: whether they worked for pay after completing bachelor's degree; then total number and name(s) of employers after completing bachelor's degree and the number of jobs held at each employer; finally, detail on each employer and job held after completing bachelor's degree.
2. *Control Group*: Similar to previous B&B administrations, sample members will received the *interleafed* format for all three looping sets in the B&B:16/17 field test survey instrument. For each looping set, respondents will be asked to provide:
    a. *Undergraduate postsecondary institutions*: whether they attended a postsecondary institution, other than the NPSAS institution, prior to completing bachelor's degree; then, detail on the additional postsecondary

institution attended; and then again, whether they attended any *other* postsecondary institution prior to completing bachelor's degree. Respondents will be looped through the set of items for as many undergraduate institutions and degrees as they have.

b. *Postbaccalaureate postsecondary institutions*: whether they attended a postsecondary institution after completing bachelor's degree; then, detail on the postsecondary institution and degrees attended; and then again, whether they attended any *other* postsecondary institution after completing bachelor's degree. Respondents will be looped through the set of items for as many postbaccalaureate institutions and degrees as they have.

c. *Postbaccalaureate employers and jobs*: whether they worked for pay after completing bachelor's degree; then, detail on the employer and jobs held; and then again, whether they had any *other* employer after completing bachelor's degree. Respondents will be looped through the set of items for as many employers and jobs as they have.

The amount of detail collected for each institution/degree and employer/job will be the same for both groups. The two groups will be evaluated on accuracy and response rates, breakoffs, and timing.

c. Experimental Design

The two experiments described above will test the numerous hypotheses outlined below. The experimental design includes estimation of the minimum differences between the control and treatment groups necessary to detect statistically significant differences.

The control and treatment groups with the null hypotheses to be tested are defined as follows:

Experiment # 1: Finding the optimum strategy to minimize sampling design weight variation and nonresponse bias and boost response rates

- *Treatment group 1* – NPSAS:16 field test interview nonrespondents who receive the aggressive data collection protocol
- *Control group 1* – NPSAS:16 field test interview nonrespondents who receive the default data collection protocol

1. The response rates for the NPSAS: 16 field test interview nonrespondents who receive the aggressive data collection protocol and those who receive the default data collection protocol will not differ.
2. There will be no difference in response rates among the NPSAS: 16 field test interview nonrespondents who are assigned to a prepaid incentive (treatment) and those who are assigned to a promised incentive (control).
3. There will be no difference in substantive responses (e.g., post baccalaureate employment, employer/job changes, earnings, dependents, marriage) between NPSAS: 16 field test interview nonrespondents receiving the aggressive data collection protocol and those receiving the default data collection protocol.

- *Treatment group 2* – NPSAS:16 field test interview respondents who receive the relaxed data collection protocol
- *Control group 2* – NPSAS:16 field test interview respondents who receive the default data collection protocol

4. There will be no difference in response rates for the NPSAS:16 field test interview respondents, assigned to $30 (control) vs. $20 incentive (treatment) amounts.

- *Treatment group 3* – NPSAS:16 field test interview nonrespondents
- *Control group 3* – NPSAS:16 field test interview respondents

5. There will be no difference in substantive responses (e.g., postbaccalaureate employment, employer/job changes, earnings, dependents, marriage) between NPSAS: 16 field test interview nonrespondents (treatment) and NPSAS: 16 field test interview respondents (control).

Experiment # 2: Questionnaire Design

- *Treatment group 1* – Sample members will receive the grouped format for all three loops in the B&B:16/17 survey instrument.
- *Control group* – Sample members will receive the interleafed format for all three loops in the B&B:16/17 survey instrument.

6. Mean time to complete the survey will not be higher for the treatment group than for the control group.
7. Item missingness rates will not be different for the treatment group than for the control group.
8. The estimates for the number of undergraduate education institutions, postbaccalaureate institutions, and employers and jobs will not be different for the control and treatment group.
9. Breakoff rates will not be different for the control group than for the treatment group.

d. Detectable Differences

The differences between the control and treatment group(s) necessary to detect statistically significant differences are shown in table 5. The following assumptions were made in computing detectable differences:

- Detectable differences with 95 percent confidence were calculated as follows:
    a. Hypothesis 6 assumes a one-tailed test.
    b. Hypotheses 1, 2, 3, 4, 5, 7, 8 and 9 assume a two-tailed test.
- The sample will be equally distributed across the experimental groups for hypotheses 1, 2, 3, 6, 7, 8, and 9.
- The sample will be determined by the number of cases from the NPSAS:16 field test for hypotheses 4 and 5.
- Analysis of hypotheses 1, 2, and 4 will include all eligible sample members.
- Hypothesis 2 will be tested after the first 2 weeks of data collection, just before the CATI intervention, for the NPSAS:16 field test interview nonrespondents who receive the aggressive protocol.

- Hypothesis 4 will be tested after the first 6 weeks of data collection, just before the CATI intervention, for the NPSAS:16 field test interview respondents who receive the default protocol.
- Analysis of hypotheses 3, 5, 6, 7, 8, and 9 will include respondents.
- The eligibility rate will be 85 percent.
- The response rate for hypotheses 3, 5, 6, 7, 8, and 9 will be 75 percent.
- The response rate for the control group in hypotheses 1, 2, and 4 will be 45, 20, and 50 percent, respectively.
- The proportion estimates for substantive responses in hypotheses 3 and 5 will be 30 percent.
- The mean time to complete the interview for the control group for hypothesis 6 will be 35 minutes.
- The item missingness rate for the control group for hypothesis 7 will be 2 percent.
- The mean estimates in hypothesis 8 will be 2.0.
- The breakoff rate for the control group in hypothesis 9 will be 1 percent.
- The statistical tests will have 80 percent power with an alpha of 0.05.
- The study design effect will be 1 because the sample is purposive with no weights.
- The intraclass correlation will be about 0.2 when comparing the control group with the treatment group.

Table 5. Detectable Differences

	Group 1		Group 2		
Hypothesis	Definition	Sample size	Definition	Sample size	Detectable difference with 95 percent confidence
1	Aggressive Data Collection Protocol	279	Default data collection protocol for NPSAS interview nonrespondents	279	10.6%
2	Aggressive Data Collection Protocol	279	Default data collection protocol for NPSAS interview nonrespondents	279	8.5%
3	Aggressive Data Collection Protocol	126	Default data collection protocol for NPSAS interview nonrespondents	126	14.5%
4	Relaxed Data Collection Protocol	661	Default data collection protocol for NPSAS interview nonrespondents	574	7.2%
5	NPSAS:16 field test interview nonrespondents	252	NPSAS:16 field test respondents	1,099	8.2%
6	Grouped format	675	Interleafed format	675	3.6 minutes
7	Grouped format	675	Interleafed format	675	1.9%
8	Grouped format	675	Interleafed format	675	0.35%
9	Grouped format	675	Interleafed format	675	1.4%

5. Reviewing Statisticians and Individuals Responsible for Designing and Conducting the Study

The study is being conducted by the National Center for Education Statistics (NCES), U.S. Department of Education. The following statisticians at NCES are responsible for the statistical aspects of the study: Mr. Ted Socha, Dr. Tracy Hunt-White, Dr. David

Richards, Dr. Sean Simone, and Dr. Elise Christopher. NCES's prime contractor for B&B:16/17 is the RTI International (RTI). The following staff members at RTI are working on the statistical aspects of the study design: Dr. Jennifer Wine, Dr. Natasha Janson, Ms. Lesa Caves, Ms. Jennifer Cooney, Dr. T. Austin Lacy, Dr. Emilia Peytcheva, Mr. Peter Siegel, Dr. Sandra Staklis, and Ms. Nicole Tate.

Subcontractors include Coffey Consulting; Hermes; HR Directions; Research Support Services; Shugoll Research; and Strategic Communications, Inc. Consultants are Dr. Sandy Baum, Ms. Alisa Cunningham, and Dr. Stephen Porter. Principal professional RTI staff, not listed above, who are assigned to the study include Ms. Donna Anderson and Ms. Gayathri Bhat.

C. References

Deming, W. E. (1953). On a probability mechanism to attain an economic balance between the resultant error of response and the bias of nonresponse. *Journal of the American Statistical Association, 48*(264), 743-772. doi:10.2307/2281069.

Eckman, S., Kreuter, F., Kirchner, A., Jackle, A., Tourangeau, R., and Presser, S. (2014). Assessing the mechanisms of misreporting to filter questions in surveys. *Public Opinion Quarterly,* 78: 721-33.

Hansen, M. H., & Hurwitz, W. N. (1946). The problem of non-response in sample surveys. *Journal of the American Statistical Association, 41*(236), 517-529. doi:10.2307/2280572.