

**2012/14 BEGINNING POSTSECONDARY STUDENTS
LONGITUDINAL STUDY: (BPS:12/14) FIELD TEST**

Supporting Statement Parts B and C
Request for OMB Review
OMB # 1850-0631 v.7

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

October 11, 2012

Contents

List of Tables	iii
List of Exhibits	iii
B. Collection of Information Employing Statistical Methods	1
1. Respondent Universe and Sampling.....	1
a. NPSAS:12 Field Test Institution Universe and Sample.....	1
b. NPSAS:12 Field Test Student Universe and Sample.....	3
c. Identification of FTBs in NPSAS:12.....	4
d. BPS:12/14 Field Test Sample.....	5
2. Procedures for the Collection of Information.....	6
a. Statistical methodology for stratification and sample selection.....	6
b. Estimation procedure.....	6
c. Degree of accuracy needed for the purpose described in the justification.....	6
d. Unusual problems requiring specialized sample procedures.....	7
e. Any use of periodic (less frequent than annual) data collection cycles to reduce the burden.....	7
3. Methods for Maximizing Response Rates.....	7
a. Locating.....	7
b. Interviewing Procedures.....	8
c. Quality Control.....	10
4. Tests of Procedures and Methods.....	10
a. NPSAS:12.....	10
b. B&B:08/12.....	11
c. ELS:2002/12.....	13
d. BPS:12/14 Field Test.....	15
5. Reviewing Statisticians and Individuals Responsible for Designing and Conducting the Study.....	22
a. Other Contractors' Staff Responsible for Conducting the Study.....	22
C. Overview of Analysis Topics and Survey Items	22
References	22

List of Tables

Table 1. BPS:12/14 Sample Size by institution characteristics: 2011.....	6
Table 2. Start dates for BPS:12/14 field test interviewing, by institution sector.....	9
Table 3. Summary of Field Test and Full-Scale Data Collection Designs, by Study.....	14
Table 4. Distribution of the BPS:12/14 field test sample, by sector and base year response status and mode.....	16
Table 5. Detectable differences for experimental hypotheses.....	20

List of Exhibits

Exhibit 1. Design of the BPS:12/14 field test data collection experiment.....	21
---	----

B. Collection of Information Employing Statistical Methods

1. Respondent Universe and Sampling

The target population for the 2012/14 Beginning Postsecondary Students Longitudinal Study (BPS:12/14) field test consists of all students who began their postsecondary education for the first time during the 2010–11 academic year at any Title IV-eligible postsecondary institution in the United States. The sample students were the first-time beginners (FTBs) from the 2011-12 National Postsecondary Student Aid Study (NPSAS:12) field test. Because the students in the BPS:12/14 field test sample come from the NPSAS:12 field test sample, this section also describes the NPSAS:12 field test sample design, which was a two-stage sample consisting of a sample of institutions at the first stage, and a sample of students from within sampled institutions at the second stage. The BPS:12/14 field test sample is comprised of students from the NPSAS:12 field test sample who were determined to be FTBs, or were potential FTBs as indicated by the NPSAS institution.

a. NPSAS:12 Field Test Institution Universe and Sample

To be eligible for the NPSAS:12 field test, students must have been enrolled in a NPSAS eligible institution in a term or course of instruction at any time during the 2010–11 academic year. Institutions must have also met the following requirements:

- offer an educational program designed for persons who have completed secondary education;
- offer at least one academic, occupational, or vocational program of study lasting at least 3 months or 300 clock hours;
- offer courses that were open to more than the employees or members of the company or group (e.g. union) that administers the institution;
- be located in the 50 states or the District of Columbia;
- not be a U.S. service academy institution; and
- have signed the Title IV participation agreement with the U.S. Department of Education.¹

Institutions providing only avocational, recreational, or remedial courses or only in-house courses for their own employees or members were excluded. U.S. service academies were excluded because of their unique funding/tuition base.

The institution samples for the NPSAS field test and full-scale studies were selected simultaneously, prior to the field test study, using stratified random sampling with probabilities proportional to a composite measure of size (Folsom, Potter, and Williams 1987). Institution measure of size was determined using annual enrollment data from the most recent IPEDS 12-Month Enrollment Component and first time beginner (FTB) enrollment data from the most recent IPEDS Fall Enrollment Component. Composite measure of size sampling was used to ensure that target sample sizes were achieved within institution and student sampling strata, while also achieving approximately equal student weights across institutions. The institution sampling frame for NPSAS:12 field test was constructed using the 2009 Integrated Postsecondary Education Data System (IPEDS) header, Institution Characteristics

¹ A Title IV eligible institution is an institution that has a written agreement (program participation agreement) with the U.S. Secretary of Education that allows the institution to participate in any of the Title IV federal student financial assistance programs other than the State Student Incentive Grant (SSIG) and the National Early Intervention Scholarship and Partnership (NEISP) programs.

(IC), Fall and 12-Month Enrollment, and Completions files. All eligible students from sampled institutions comprised the student sampling frame

The field test institution sample for NPSAS:12 was selected using statistical procedures rather than purposively as had been done in past NPSAS cycles. This provided more control to ensure that the field test and the full-scale institution samples had similar characteristics. It also allowed inferences to be made to the target population, supporting the analytic needs of the field test experiments and instrument. This ability to make analytic inferences extends to the BPS:12/14 field test sample.

From the stratified frame, a total of 1,970 institutions was selected to participate in either the full-scale or field test study. From the 1,970 institutions selected, a subsample of 300 institutions was selected using simple random sampling within institution strata to comprise the field test sample. The remaining 1,670 institutions comprised the sample for the full-scale study. This sampling process eliminated the possibility that an institution would be burdened with participation in both the field test and full-scale samples, and maintained the representativeness of the full-scale sample. All institutions in the NPSAS:12 sample were eligible for either the full-scale or field test studies.

The institution strata used for the sampling design were based on institution level, control, and highest level of offering and are the following:

1. public less-than-2-year,
2. public 2-year,
3. public 4-year non-doctorate-granting,
4. public 4-year doctorate-granting,
5. private nonprofit less-than-4-year,
6. private nonprofit 4-year non-doctorate-granting,
7. private nonprofit 4-year doctorate-granting,
8. private for-profit less-than-2-year,
9. private for-profit 2-year, and
10. private for-profit 4-year.

Due to the growth of the for-profit sector, private for-profit 4-year and private for-profit 2-year institutions were separated into their own strata unlike in previous administrations of NPSAS.

Within each institution stratum, additional implicit stratification for the full-scale was accomplished by sorting the sampling frame within stratum by the following classifications: (1) historically Black colleges and universities indicator; (2) Hispanic-serving institutions indicator (3) Carnegie classifications of degree-granting postsecondary institutions; (4) 2-digit Classification of Instructional Programs (CIP) code for less-than-2-year institutions; (5) the Office of Business Economics Region from the IPEDS header file (Bureau of Economic Analysis of the U.S. Department of Commerce Region); (6) state and system for states with large systems, e.g., the SUNY and CUNY systems in New York, the state and technical colleges in Georgia, and the California State University and University of California systems in California; and (7) the institution measure of size. The objective of this implicit stratification was to approximate proportional representation of institutions on these measures.

Approximately 300 institutions were sampled for the NPSAS:12 field test. Overall, almost 100 percent of the sampled institutions met the eligibility requirements; of those, approximately 51 percent (or about 150 institutions) provided enrollment lists.

b. NPSAS:12 Field Test Student Universe and Sample

The students eligible for the BPS:12/14 field test are those eligible to participate in the NPSAS:12 field test who were FTBs at NPSAS sample institutions in the 2010-11 academic year. Students eligible for the NPSAS:12 field test were those who attended a NPSAS eligible institution during the 2010–11 academic year and who were

- enrolled in *either*: (a) an academic program; (b) at least one course for credit that could be applied toward fulfilling the requirements for an academic degree; (c) exclusively non-credit remedial coursework but determined by the institution to be eligible for Title IV aid; *or* (d) an occupational or vocational program that required at least 3 months or 300 clock hours of instruction to receive a degree, certificate, or other formal award;
- not currently enrolled in high school; and
- not solely enrolled in a General Educational Development (GED) or other high school completion program.

The NPSAS:12 field test institution sample included all levels (less-than-2-year, 2-year and 4-year) and controls (public, private nonprofit and private for-profit) of Title IV eligible postsecondary institutions in the United States. The field test student sample was randomly selected from lists of students enrolled at sampled institutions between July, 1 2010 and April 30, 2011.

The NPSAS:12 first test study year covers the time period between July 1, 2010 and June 30, 2011, to coincide with the federal financial aid award year. To facilitate timely completion of data collection and data file preparation, institutions were asked to submit enrollment lists for all eligible students enrolled at any time between July 1 and April 30 or, for institutions with continuous enrollment, between July 1 and March 31. The March 31 deadline for continuous enrollment institutions was used for the field test due to the compressed data collection schedule and was not used in the full-scale.

Because previous cycles of NPSAS have shown that the terms beginning in May and June add little to enrollment and aid totals, May-June starters were excluded to allow institutions to provide enrollment lists earlier which, in turn, allowed the student interview process to begin earlier. In the full-scale study, post-stratification of survey estimates based on IPEDS records on enrollment and National Student Loan Data System (NSLDS) records on financial aid distributed was used to adjust for the survey year's inclusion of any terms that begin by April 30 and the consequent exclusion of a small number of students newly enrolled in May or June.

To create the student sampling frame, each participating institution was asked to submit a list of eligible students. The requests for student enrollment lists specifically indicated how institutions should handle special cases, such as students taking only correspondence or distance learning courses, and foreign exchange, continuing education, extension division, and non-matriculated students. The data required for each enrollee were the following:

- student's name;
- student ID;

- Social Security number;
- date of birth;
- date of high school graduation (month and year);
- degree level during the last term of enrollment (undergraduate, masters, doctoral-research/scholarship/other, doctoral-professional practice, or other graduate);
- class level if undergraduate (first, second, third, fourth, or fifth year or higher);
- major;
- CIP code;
- indicator of whether the institution received an Institutional Student Information Record (ISIR) (an electronic record summarizing the results of the student’s Free Application for Federal Student Aid [FAFSA] processing) from the Central Processing System (CPS);
- FTB status; and
- contacting information, such as cell phone number, local telephone number and address, permanent telephone number and address, campus e-mail address, and permanent e-mail address.

Requesting contact information for eligible students prior to sampling allowed for student record abstraction and student interviewing to begin shortly after sample selection which helped to ensure the management of the field test schedule for data collection, data processing, and file development.

Student sample sizes for the field test were formulated to ensure representation of various types of students. Specifically, the sample included a large number of potential first-time beginners to provide a sufficient sample size to obtain a sample yield of at least 1,000 students for BPS field test. The NPSAS:12 field test sample included 4,530 students of which 4,130 were potential FTBs, 200 were other undergraduate students, and 200 were graduate students.

Students were sampled at fixed rates according to student education level and institution sampling strata. Sample yield was monitored and sampling rates were adjusted when necessary, resulting in a statistical sample of the required sample size for the field test. The same approach was to be used for the full-scale study. Student enrollment lists provided by the institutions were reviewed to make sure that required elements were included, and were also compared for consistency with counts from the 2009 IPEDS 12 Month Enrollment Component.

c. Identification of FTBs in NPSAS:12

To be eligible for the BPS field test, students must have begun their postsecondary education for the first time after completing high school on or after July 1, 2010, and before July 1, 2011. Close attention was paid to accurately identify FTBs in the NPSAS field test to avoid unacceptably high rates of misclassification (e.g., false positives)² which can, and have, resulted in (1) excessive cohort loss, (2) excessive cost to “replenish” the sample, and (3) an inefficient sample design (excessive oversampling of “potential” FTBs) to compensate for anticipated misclassification error. To address this concern, participating institutions were asked to provide additional information for all eligible students and matching to administrative databases was utilized to further eliminate false positives prior to sample selection.

Participating institutions were asked to provide the FTB status and high school graduation date for every listed student. High school graduation date was used to remove students from the frame who were

² A student identified by the institution on the enrollment list as an FTB who turns out to not be an FTB is a false positive.

co-enrolled in high school. FTB status, along with class level and student level, were used to exclude misclassified FTB students in their third year or higher and/or those who were not an undergraduate student. FTB status, along with date of birth, were used to identify students older than 18 to send for pre-sampling matching to administrative databases.

If the FTB indicator was not provided for a student on the lists, but the student was 18 years of age or younger and did not appear to be dually enrolled in high school, the student was sampled as an FTB. Otherwise, if the FTB indicator was not provided for a student on the list and the student was over the age of 18, then the student was sampled as an “other undergraduate” but would be included in the BPS cohort if identified during the student interview as an FTB.

Prior to sampling, students over the age of 18 listed as potential FTBs were matched to NSLDS records to determine if any had a federal financial aid history pre-dating the NPSAS year (earlier than July 1, 2010 for the field test). Since NSLDS maintains current records of all Title IV federal grant and loan funding, any student with disbursements from the prior year or earlier could be reliably excluded from the sampling frame of FTBs. Given that about 60 percent of FTBs receive some form of Title IV aid in their first year, this matching process could not exclude all listed FTBs with prior enrollment, but significantly improved the accuracy of the list prior to sampling, yielding fewer false positives. After undergoing NSLDS matching, students over the age of 18 still listed as potential FTBs were matched to the National Student Clearinghouse (NSC) for further narrowing of potential FTBs based on evidence of earlier enrollment.

Matching to NSLDS identified about 19 percent of cases as false positives and matching to NSC identified about 14 percent of cases as false positives. In addition to NSLDS and NSC, a subset of potential FTBs on the student sampling frame was sent to CPS for matching to evaluate the benefit of the CPS match for the full-scale study. Of the 58,690 students sent, CPS identified about 10 percent as false positives. Overall, matching to all sources identified about 32 percent of potential FTB students over the age of 18 as false positives, with many of the false positives identified by CPS also identified by NSLDS or NSC. The matching appeared most effective among public 2-year and private for-profit institutions. While public less-than 2-year institutions have a high percent of false positives, they represent a small percentage of the total sample.

Since this pre-sampling matching was new to NPSAS:12, the FTB sample size was set high to ensure that a sufficient number of true FTBs would be interviewed. In addition, FTB selection rates were set taking into account the error rates observed in NPSAS:04 and BPS:04/06 within each sector (additional information on NPSAS:04 methodology is available in the study methodology report, publication number NCES 2006180, or BPS:04/06 methodology report, publication number NCES 2008184.). These rates were adjusted to reflect the improvement in the accuracy of the frame from the NSLDS and NSC record matching. Sector-level FTB error rates from the field test will be used to help determine the rates necessary for full-scale student sampling.

d. BPS:12/14 Field Test Sample

At the conclusion of the NPSAS:12 field test, 2,003 students had been interviewed and confirmed to be FTBs, and all will be included in the BPS:12:14 field test. In addition, the field test sample will include the 1,493 students who did not respond to the NPSAS:12 field test but were potential FTBs

according to student records or institution lists. The distribution of the BPS:12/14 field test sample is shown in table 1, by institution sector.

Table 1. BPS:12/14 Sample Size by institution characteristics: 2011

Institution characteristics	Confirmed and Potential FTBs from the NPSAS:12 Field Test		
	Total	Confirmed FTBs	Potential FTBs
Total	3,496	2,003	1,493
Institution type			
Public			
Less-than-2-year	17	6	11
2-year	1,526	827	699
4-year non-doctorate-granting	204	142	62
4-year doctorate-granting	436	321	115
Private nonprofit			
Less-than-4-year	37	23	14
4-year non-doctorate-granting	205	155	50
4-year doctorate-granting	139	107	32
Private for-profit			
Less-than-2-year	139	71	68
2-year	211	70	141
4-year	582	281	301

Rounds to zero.

¹Includes master's, doctoral-research/scholarship/other, doctoral-professional practice, and other graduate students.

NOTE: Detail may not sum to totals because of rounding. FTB = first-time beginner.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2011–12 National Postsecondary Student Aid Study (NPSAS:12) Field Test

2. Procedures for the Collection of Information

a. Statistical methodology for stratification and sample selection

The BPS:12/14 field test sample will consist of all interviewed and confirmed FTBs from the NPSAS:12 field test and also all potential FTBs who did not respond to the NPSAS:12 field test; no further sampling or subsampling will be needed. The institution and student sampling strata used in the NPSAS:12 field test will be retained for use in analysis. An experiment to test methods for improving response rates and reducing bias due to nonresponse is being conducted as a part of the field test, and random assignment to treatments will be balanced across the NPSAS:12 strata and response statuses.

b. Estimation procedure

The students in the BPS:12/14 field test were also in the NPSAS:12 field test. Because the NPSAS:12 field test was a probability sample, sampling weights (which are the inverse of the selection probability) are available for each of the students in the NPSAS:12 field test and BPS:12/14 field test. These base weights can be used in analyses of the results of the field test experiment. In addition,

response weight adjustments will be applied to these base weights, and the response-adjusted weights will be used in the analysis of BPS:12/14 field test interview data.

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

One purpose of the field test experiments described in section 3 is to improve response rates and reduce nonresponse bias. The sample will include interviewed FTBs from the NPSAS:12 field test and interview nonrespondents who were potential FTBs. Table 5, at the end of this section, shows that the detectable differences range from 3 percent to 20 percent for the hypotheses being tested using the field test data.

d. Unusual problems requiring specialized sample procedures

No special problems requiring specialized sample procedures have been identified.

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

BPS interviews are conducted no more frequently than every two years.

3. Methods for Maximizing Response Rates

a. Locating

The response rate for the BPS:12/14 data collection is a function of success in two basic activities: locating the sample members and gaining their cooperation. Several locating methods will be used to find and collect up-to-date contact information for the BPS:12/14 sample. During BPS:12/14, batch searches of national databases will be conducted prior to the start of data collection. Follow-up locating methods will be employed for those sample members not found after the start of data collection. The methods chosen to locate sample members are based on the experience gained from the NPSAS:12 field test and full-scale studies, the B&B:08/12 field test and full-scale studies, and other recent postsecondary education studies.

Many factors will affect the ability to successfully locate and survey sample members for BPS:12/14. Among them are the availability, completeness, and accuracy of the locating data collected in the NPSAS:12 field test interview and institution records. The locator database includes critical tracing information for nearly all sample members, including address information for their previous residences, telephone numbers, and e-mail addresses. This database allows telephone interviewers and tracers to have ready access to all the contact information available for BPS:12/14 sample members and to new leads developed through locating efforts.

To achieve the desired locating and response rates, RTI will use a multistage locating approach that capitalizes on available data for the BPS:12/14 sample. RTI's proposed locating approach includes five basic stages:

1. **Advance Tracing** includes batch database searches, contact information updates, and advance intensive tracing conducted as necessary.
2. **Telephone Locating and Interviewing** includes calling all available telephone numbers and following up on leads provided by parents and other contacts.

3. **Pre-Intensive Batch Tracing** consists of the Premium Phone searches that will be conducted between the telephone locating and interviewing stage and the intensive tracing stage.
4. **Intensive Tracing** consists of tracers checking all telephone numbers and conducting credit bureau database searches after all current telephone numbers have been exhausted.
5. **Other Locating Activities** will take place as needed and may include use of social networking sites and additional tracing resources that are not part of the previous stages.

The steps described in the tracing plan are designed to locate the maximum number of sample members with the least expense. The most cost-effective steps will be taken first so as to minimize the number of cases requiring more costly intensive tracing efforts.

BPS:12/14 will also conduct an address update (panel maintenance) procedure to encourage sample members to update their contact information prior to the start of data collection. All cases will be offered an incentive of \$10 to update their contact information via the study website prior to the start of data collection (approximately 2-3 weeks before the 2013 field test data collection begins, upon OMB approval). We are also requesting approval to conduct full scale address updates since little time will be available to conduct address update activities once OMB clearance is received in 2014, in advance of the start of data collection. Receiving approval to conduct full scale panel maintenance in conjunction with this field test OMB package will enable the address update activity to take place approximately two months before data collection begins.

b. Interviewing Procedures

Telephone interviewer training procedures. Training for individuals working in survey data collection will include critical quality control elements. Contractor staff with experience in training interviewers will prepare the BPS:12/14 Telephone Interviewer Manual, which will provide detailed coverage of the background and purpose of BPS, sample design, the questionnaire, and procedures for the telephone interview. The manual will also serve as a reference throughout data collection. Training staff will prepare training exercises, mock interviews (specially constructed to highlight the potential of definitional and response problems), and other training aids.

Interviews. Interviews will be conducted using a single web-based survey instrument for self-administered and telephone data collection. The data collection activities will be accomplished through the Case Management System (CMS), which is equipped with the following capabilities:

- online access to locating information and histories of locating efforts for each case;
- questionnaire administration module with input validation 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 and incorporates the following features:
 - Automatic delivery of appointment and call-back cases at specified times.
 - Sorting of non-appointment cases according to parameters and priorities set by project staff.

- Restriction on allowable interviewers. Complete records of calls and tracking of all previous outcomes. Flagging of problem cases for supervisor action or supervisor review. Complete reporting capabilities.

A system such as the CMS that integrates these capabilities reduces the number of discrete stages required in data collection and data preparation activities. Overall, the scheduler provides a highly efficient case assignment and delivery function, reduces supervisory and clerical time, improves execution on the part of interviewers and supervisors by automatically monitoring appointments and callbacks, and reduces variation in implementing survey priorities and objectives.

CATI Start Dates. As with NPSAS:12 full-scale study, RTI plans to initiate telephone efforts at varying times in order to maximize participation for sample members from sectors that have historically been challenging. Please see Table 2 below for details.

Table 2. Start dates for BPS:12/14 field test interviewing, by institution sector

Sector	Institution sector	Cases begin CATI...
Base year respondents		
1	Public less-than-2-year	5 days after lead letters are sent
2	Public 2-year	5 days after lead letters are sent
3	Public 4-year non-doctorate-granting	3 weeks + 1 day after lead letter mailing
4	Public 4-year doctorate-granting	3 weeks + 1 day after lead letter mailing
5	Private not-for-profit less-than-4-year	5 days after lead letters are sent
6	Private not-for-profit 4-year non-doctorate-granting	3 weeks + 1 day after lead letter mailing
7	Private not-for-profit 4-year doctorate granting	3 weeks + 1 day after lead letter mailing
8	Private for-profit less-than-2-year	5 days after lead letters are sent
9	Private for-profit 2-year	5 days after lead letters are sent
10	Private for-profit 4-year	5 days after lead letters are sent
All base year nonrespondents		5 days after lead letters are sent

Refusal Conversion. Recognizing and avoiding refusals is important to maximize the response rate. Supervisors will monitor interviewers intensely during the early data collection and provide retraining as necessary. In addition, supervisors will review daily interviewer production reports to identify and retrain any interviewers with unacceptable numbers of refusals or other problems.

After encountering a refusal, comments are entered into the CMS record that include all pertinent data regarding the refusal situation, including any unusual circumstances and any reasons given by the sample member for refusing. Supervisors review these comments to determine what action to take with each refusal; no refusal or partial interview will be coded as final without supervisory review and approval.

If a follow-up is not appropriate (e.g., there are extenuating circumstances, such as illness or the sample member firmly requested no further contact), the case will be coded as final and no additional contact will be made. If the case appears to be a “soft” refusal, follow-up will be assigned to an

interviewer other than the one who received the initial refusal. The case will be assigned to a member of a special refusal conversion team made up of interviewers who have proven especially skilled at converting refusals.

Refusal conversion efforts will be delayed until at least 1 week 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. Project staff sometimes receive refusals via email or calls to the project director's direct line. These refusals are included in the CATI record of events and coded as final when appropriate.

c. Quality Control

Interviewer monitoring will be conducted using RTI's Quality Evaluation System (QUEST) as a quality control measure throughout the field test and full scale data collections. QUEST is a system developed by a team of RTI researchers, methodologists, and operations staff focused on developing standardized monitoring protocols, performance measures, evaluation criteria, reports, and appropriate systems security controls. It is a comprehensive performance quality monitoring system that includes standard systems and procedures for all phases of quality monitoring, including obtaining respondent consent for recording, procedures for interviewing respondents who refuse consent and for monitoring refusals at the interviewer level; sampling of completed interviews by interviewer, evaluating interviewer performance; maintaining an online database of interviewer performance data; and addressing potential problems through supplemental training. These systems and procedures are based on "best practices" identified by RTI in the course of conducting thousands of survey research projects.

As in previous studies, RTI will use QUEST to monitor approximately 7 percent of all completed interviews plus an additional 2.5 percent of recorded refusals. In addition, quality supervisors will conduct silent monitoring for 2.5 - 3 percent of budgeted interviewer hours on the project. This will allow real-time evaluation of a variety of call outcomes and interviewer-respondent interactions. Recorded interviews will be reviewed by call center supervisors for key elements such as professionalism and presentation; case management and refusal conversion; and reading, probing, and keying skills. Any problems observed during the interview will be documented on problem reports generated by QUEST. Feedback will be provided to interviewers and patterns of poor performance (e.g., failure to use conversational interviewing techniques, failure to probe, etc.) will be carefully monitored and noted in the feedback form that will be provided to the interviewers. As needed, interviewers will receive supplemental training in areas where deficiencies are noted. In all cases, sample members will be notified that the interview may be monitored by supervisory staff.

4. Tests of Procedures and Methods

The design of the BPS:12/14 field test experiment expands on data collection experiments designed for three NCES studies which preceded it, including the NPSAS:12 field test which served as the base year interview for the current study. Brief summaries of each experiment are provided below, and in table 3. Although results are not available for the most recent experiments, the design proposed for the BPS:12/14 field test builds on what is known about those experiments and target populations, and how they compare to the BPS:12 cohort.

a. NPSAS:12

The response propensity experiment conducted during the NPSAS:12 field test data collection (March 2011 to June 2011) was designed to reduce nonresponse bias through targeted use of incentives. Using data from NPSAS:04, RTI identified variables available prior to data collection which were predictive of response likelihood, then used the variables to estimate a NPSAS:12 field test sample member's response propensity. Sample members with a low response propensity were sorted at random into either a control group, which was offered the usual \$30 incentive for participation, or an experimental group, which was offered \$45. High response propensity sample members were sorted at random into a control group that was offered \$30 or an experimental group that was offered \$15. Following data collection, RTI evaluated the predictive ability of the response propensity model and determined if bias was reduced in the experimental cases.³

The propensity model successfully distinguished between high and low propensity cases in terms of response rate. The unweighted low propensity response rate was 57.7% and the unweighted high propensity response rate was 67.7%, a statistically significant difference ($\chi^2 = 42.003$, $p < .0001$). However, while the primary goal of the response propensity approach was to reduce bias in key estimates, the weighted estimates in both the low propensity control and treatment groups were virtually identical, suggesting that differential incentives did not have any effect on reducing bias. Although the NPSAS:12 field test experiment was not designed to increase response rates *per se*, response rates by incentive amount within propensity groups were tested. Within the low propensity group, no statistically significant difference between experimental and control groups was noted ($\chi^2 = 2.527$, $p > .05$) while the difference observed between high propensity control and treatment groups was statistically significant, with lower incentives being associated with lower response rates ($\chi^2 = 13.576$, $p < .001$).

Given the equivocal results of the response propensity experiment, RTI adopted a responsive design approach to the NPSAS:12 full-scale data collection (February 2012 to present), dropping pre-data collection modeling of either responses or paradata. All sample members are being offered a \$30 incentive. This approach to data collection has been used during the early response phase varied by institution sector (e.g., public, 2-year) as a substitute for response propensity. For example, students in public, 4-year institutions, with historically higher response rates, were handled with the typical data collection plan: three weeks of online-only interviewing followed by outbound calling to nonrespondents. In contrast, students in institutions with historically lower response rates and a lower likelihood of responding online, were moved almost immediately to outbound calling, shortening the time to initial contact and, when needed, referral to intensive tracing.

As each wave of the NPSAS:12 sample has moved from the early response to the production phase of data collection, the approach taken to encourage response continued to depend on institution sector. Other factors about an individual's experience in data collection were also considered. For example, specialized emails are being prepared based on paradata, such as break offs and expressed preference to complete the online interview, and USPS Priority Mail is being used to contact cases sampled as FTBs. As the data collection period is ending, cases are being

³ For the NPSAS:12 field test, the institution sample was selected statistically, rather than purposively, as had been done in past NPSAS cycles, in order to allow inferences to be made to the target population, supporting the analytic needs of the field test experiments.

offered the abbreviated interview depending on their time in data collection and expressed reluctance to commit time to the interview.

b. B&B:08/12

In the B&B:08/12 field-test (July 2011 to October 2011), RTI targeted cases with a low propensity to respond and a high likelihood of contributing to nonresponse bias in order to increase the response rate and yield less biased survey estimates. To begin, frame data, paradata, and indicators of previous response behavior were used to develop a predictive model of a given sample member's propensity to respond. To build the model, RTI estimated logistic regression coefficients using data from the NPSAS:08 base year to predict response in the first follow-up (B&B:08/09). The resulting model produced odds ratios ranging from 0.99 to 2.65 with an r-squared value of 0.19.

Before the start of data collection, response propensities for all sample members were calculated based on the developed model, then used to divide the sample into low and high response propensity groups. Approximately one-third of the cases were in the high propensity group, and two-thirds were in the low propensity group. The low propensity group was comprised of those less likely to complete and most likely to introduce nonresponse bias if they remained nonrespondents. Within each of the propensity levels, cases were randomly assigned into a control group, which received the same incentive offered in the prior field test round (\$35 or \$55), and an experimental group, whose incentive amounts varied by response propensity: \$20 or \$40 was offered to all cases in the high propensity group, and \$50 or \$70 was offered to cases in the low propensity group, depending on the amount they received in the prior field test, with those receiving \$55 in 2009 receiving \$70 in 2011.

Like the NPSAS:12 field test, evaluation of the field test results showed that the propensity model was able to accurately predict relative likelihood to respond. The proportion of nonrespondents in the low propensity group (39 percent) was more than three times the proportion in the high propensity group (11 percent; $\chi^2, (1, N=1,588)=139.0; p < .01$). Analyses of response rates for the treatment and control groups indicated that changes in incentives had the strongest impact on response rates for those individuals in the middle of the propensity score range. Observed response rates were higher in the incentive treatment group for those individuals with the highest propensity scores within the low propensity classification (81.4 and 73.3 percent, $t = 2.04, df = 539, p \leq .05$). Those with the lowest propensity scores within the high propensity classification showed a numerical difference in response rates between treatment and control groups (79 and 89 percent, respectively), but the difference was not statistically significant. However, field test results did not show a reduction in bias as a result of the additional response.

In order to focus on identifying and targeting cases most likely to contribute to nonresponse bias, a revised approach is being tried in the B&B:08/12 full-scale data collection that uses a responsive design and the Mahalanobis distance measure to identify cases for targeted treatments. For the first three months of data collection, which began in August 2012, all sample members are receiving the same treatment in data collection – the web, online interview option with “CATI-light,” during which a small number of calls is being made, mainly to prompt sample members to complete the online interview. Incentive offers during the first three months are determined by a

case's propensity score, calculated prior to the start of data collection. Cases with the highest propensity scores are being offered \$20, midrange \$35, and lowest propensity \$55.

The B&B full-scale sample has been split into a treatment and a control group at random and, in Month 3 (November 2012), Mahalanobis values of treatment group nonrespondents will be calculated. Cases above a threshold value (high-distance) will be offered another \$15 in addition to their original incentive offer of \$20, \$35, or \$55 (once a case becomes eligible for the additional \$15, they will remain eligible for the additional \$15). All other treatment group and control group members will continue at their initial incentive level.

After an additional month of data collection (December 2012), Mahalanobis values will be reevaluated for the remaining treatment group nonrespondents. Those above a new cut point (determined based on the remaining nonrespondents at Month 4) will receive extensive case review in which project staff will review the CMS-CATI events log and paradata available for a particular case (e.g., availability of e-mail address, parent address, etc.) to identify any specific actions that may encourage a sample member's participation. Cases eligible for extensive case review will also be prioritized in the CMS. The high-distance nonrespondents in the control group and all low-distance nonrespondents will receive extensive case review, but on the regular schedule (i.e., six weeks later).

In Month 6 (February 2013), Mahalanobis values will be evaluated again for remaining nonrespondents, and those above the cut point (determined based on the remaining nonrespondents at month 6) will be offered an abbreviated interview. The high-distance nonrespondents in the control group and all low-distance nonrespondents will receive an abbreviated interview, but on the regular schedule (i.e., 6 weeks later).

c. ELS:2002/12

For the ELS third follow-up field test (July 2011 to September 2011), sample members with the lowest response propensities were empirically identified, then targeted with interventions in an attempt to encourage participation. A logistic regression model was fitted with the sample member's ELS:2002 second follow-up field test response status as the dependent variable. As independent variables, a range of information known for *all* respondents and nonrespondents from each prior wave of the longitudinal field test, including information from panel maintenance activities were examined for significance.

Predicted probabilities derived from the logistic regression model were used to get an estimate of a case's response propensity for the field test. Cases were split into two groups of equal size. Field test sample members above the median response propensity were classified as high propensity (528 cases), and those below the median as low propensity (527 cases). For the implementation of the experiment, the low propensity cases were randomly split into experimental and control groups. Low propensity experimental group cases were offered a higher incentive of \$45 at the start of data collection (weeks 1-9), increasing to \$55 starting at week 10. High propensity and low propensity control group cases were offered \$25 until week 10 of data collection, after which the incentive increased to \$35.

The predictive model developed ahead of the field test data collection effectively predicted the eventual response outcome for sample members. The high propensity group's response rate

(67.4%) was significantly higher than that of the low propensity control group (45.4%). This difference was statistically significant ($\chi^2 = 34.9$; $p < .0001$). In examining the effect of the higher incentive treatment for low propensity cases, a numerical difference in participation (51.6% for treatment cases vs. 45.4%) was observed, however, the difference was not statistically significant. The small ELS:2002 field test sample size and the brevity of the data collection period may have contributed to the inability to detect a significant difference in the results. In reviewing the mean relative bias, it appeared that including low propensity cases in the dataset may have helped reduce bias, if only slightly, and the higher incentive for the low propensity experimental cases may have lowered the bias relative to the low propensity control group.

Table 3. Summary of Field Test and Full-Scale Data Collection Designs, by Study

Study	Field Test	Full-Scale*
NPSAS:12	<ul style="list-style-type: none"> • Sample sorted by modeled response propensity into 4 groups: <ul style="list-style-type: none"> ▪ High propensity: \$15 (E) \$30 (C) ▪ Low propensity: \$45 (E) \$30 (C) • Modeling successfully differentiated groups by propensity • Response rate differences observed only in high propensity group • No effect on bias reduction 	<ul style="list-style-type: none"> • \$30 for all sample members • Institution sector used as proxy for response propensity • Different data collection strategies applied depending on institution sector
B&B:08/12	<ul style="list-style-type: none"> • Sample sorted by modeled response propensity into 4 groups: <ul style="list-style-type: none"> ▪ High propensity: \$20/\$40 (E) \$35/\$55 (C) ▪ Low propensity: \$70 (E) \$35/\$55 (C) • Modeling successfully differentiated groups by propensity • Response rate differences observed only in low propensity group among highest propensity scores • No effect on bias reduction 	<ul style="list-style-type: none"> • Calculated response propensities and Mahalanobis distances for all sample members. Sorted sample into experimental and control groups within 3 propensity groups: <ul style="list-style-type: none"> ▪ High propensity: \$20 ▪ Medium propensity: \$35 ▪ Low propensity: \$55 • At Time 1, Mahalanobis recalculated for experimental group – greatest distance cases offered an additional \$15 • At Time 2, Mahalanobis recalculated for experimental group – greatest distance cases to receive intensive case review early • At Time 3, Mahalanobis recalculated for experimental group – greatest distance cases offered abbreviated interview early
ELS:2002/1 2	<ul style="list-style-type: none"> • Sample sorted by modeled response propensity into 3 groups: <ul style="list-style-type: none"> ▪ High propensity: \$25 (\$35 in Week 10) ▪ Low propensity: \$45 (E; \$55 in Week 10) \$25 (C; \$35 in Week 10) • Modeling successfully differentiated groups by propensity • Response rate differences observed only in high propensity group • Slight reduction of bias among low propensity group receiving higher incentive 	<ul style="list-style-type: none"> • \$25 base incentive offered to all sample members • Mahalanobis distance calculated at 3 time points: <ul style="list-style-type: none"> ▪ Time 1: Highest distance cases offered additional \$30 ▪ Time 2: New “highest distance” cases offered additional \$30; all highest distance cases will receive intensive tracing and limited field interviewing ▪ Time 3: New “highest distance” cases offered additional \$30; all highest distance cases will receive \$5 prepaid incentive

(E) = Experimental Group; (C) = Control Group

*Full-scale data collection is underway for all three studies.

An alternative approach, like that being used for the B&B:08/12 full-scale data collection, is being implemented for the ELS:2002/12 full scale study (July 2012 to January 2013). Mahalanobis distances are being calculated to identify those nonrespondent cases which are most unlike existing respondents and, therefore most likely to contribute to response bias. Substantive data (*e.g.*, enrollment status, parent's education, high school completion status) and paradata (*e.g.*, response status, number of contact attempts in the early data collection period) already available from the base year and first and second follow-ups are being used to calculate the Mahalanobis distances.

Distance functions are being measured at three points during data collection: 4 weeks and 9 weeks after the start of data collection (Phases 1 and 2, respectively), and at 8 weeks prior to the end of data collection (Phase 3). Most cases are being offered an initial incentive of \$25. Cases with the largest calculated distance scores at each time point will be offered an increased incentive of \$55 (once at \$55, the incentive offer will not decrease). Additional activities will be conducted to locate and interview targeted cases including, at Time 2, performing pre-data collection intensive tracing and pursuing the cases in person with field locator/interviewers and, at Time 3, including a \$5 prepaid incentive with the mailing.

d. BPS:12/14 Field Test

The two BPS:12/14 field test experiments were developed out of recent experiences with the NPSAS:12 full-scale study and the data collections underway in B&B:08/12 and ELS:2002 third follow-up. The first experiment emerged from anecdotal reports offered by NPSAS:12 interviewers that sample members refused to participate upon hearing the estimated time required to complete the interview (average of 25 minutes); they simply did not have the amount of time available to complete the interview. For the BPS:12/14 field test, participation rates of base year nonrespondents offered the full length interview (35 minutes) will be compared to the participation rates of nonrespondents offered a shorter, modified interview (20 minutes) to determine if the expected level of effort reported to sample members affects willingness to participate.

The BPS:12/14 responsive design experiment builds on the responsive design work currently being conducted in B&B:08/12 and ELS:2002. The basic design will be the same – Mahalanobis distances will be calculated at several time points and specific treatments applied to encourage response among sample members likely to contribute the most to bias if they do not participate. There is an important difference in target populations, however, which warrants examination of the responsive design approach for BPS. B&B and ELS follow cohorts of students who are more homogenous than the BPS cohort will be. That is, while all B&B sample members will have earned a bachelor's degree during the same academic year and all ELS sample members will be the same age, BPS sample members will not have a common postsecondary experience nor will they all be of the same age. The distinctly different populations could result in considerably different Mahalanobis calculations and distributions, and data collection outcomes.

1. Early Response Phase (Weeks 1-3).

All NPSAS:12 base year interview respondents, and all nonrespondents who were potential FTBs, will be included in the BPS:12/14 field test data collection experiment. Data collection will occur in two main phases, an early response phase, during the first 3 weeks of data collection, followed by the main,

production interviewing phase. The distribution of the NPSAS:12 FTB field test sample is shown in table 4 by the mode in which they completed the NPSAS:12 field test interview. At the start of the BPS:12/14 field test interview, the entire field test sample (N=3,496) will be moved to interviewing. Treatment during this early phase will depend on a sample member's base year response status and mode of interview completion as described below.

Base Year Respondents

Among the 2,003 NPSAS:12 field test base year respondents, there were 725 sample members whose NPSAS institution was a public 4-year or a private, nonprofit 4-year institution. Most of these sample members completed the base year interview online, and will be encouraged to complete the full, 35-minute BPS:12/14 interview online during the first 3 weeks of data collection. If a sample member from those institutions contacts the study help desk for any reason, he/she will be offered the option of completing a telephone interview at that time, but may still complete the interview online if preferred.

Table 4. Distribution of the BPS:12/14 field test sample, by sector and base year response status and mode

Type of institution	Sample	Nonrespondents (Potential FTBs)		Respondents (Confirmed FTBs)					
		Number	Percent	Web		Telephone		Partial	
				Number	Percent	Number	Percent	Number	Percent
Total	3,496	1,493	42.7	1,535	79.0	409	21.0	59	2.9
Public less-than-2-year	17	11	64.7	1	16.7	5	83.3	0	0.0
Public 2-year	1,526	699	45.8	603	75.1	200	24.9	24	2.9
Public 4-year non-doctorate-granting	204	62	30.4	121	85.8	20	14.2	1	0.7
Public 4-year doctorate-granting	436	115	26.4	270	87.1	40	12.9	11	3.4
Private nonprofit less-than-4-year	37	14	37.8	19	82.6	4	17.4	0	0.0
Private 4-year non-doctorate-granting	205	50	24.4	140	90.9	14	9.1	1	0.6
Private 4-year doctorate-granting	139	32	23.0	100	93.5	7	6.5	0	0.0
Private for-profit less-than-2-year	139	68	48.9	43	63.2	25	36.8	3	4.2
Private for-profit 2-year	211	141	66.8	38	58.5	27	41.5	5	7.1
Private for-profit 4-year	582	301	51.7	200	74.9	67	25.1	14	5.0

The remaining base year respondents – those whose first institution was a public, less-than 2-year or 2-year institution; a private, not-for-profit less-than-4-year institution; or any of the 3 levels of for-profit institution – had comparatively low online completion rates. Those who did complete the NPSAS:12 field test interview online (N=904) will be given time to complete it online again during the first 3 weeks of data collection, or by telephone if they contact the Help Desk. However, those who completed a telephone interview (N=374) during the base year will be moved to outbound calling within 5 days of the initial data collection mailing, allowing only enough time for the initial contact materials to reach them by regular mail. Calls from interviewers are primarily intended to prompt sample members to complete the online interview, although any willing to complete the interview by telephone will be

encouraged to do so. During the early response phase, base year respondents completing the first follow-up interview, either online or by telephone, will receive a check in the amount of \$30.

Base Year Nonrespondents

At the completion of the NPSAS:12 field test there were 1,493 interview nonrespondents with unknown eligibility for BPS. For Experiment 1, these nonrespondents will be randomly assigned into one of two groups. Half of the nonrespondents will be invited to complete the full BPS:12/14 field test interview requiring, on average, approximately 35 minutes of their time (Full Interview group). The other half will be invited to complete a modified BPS:12/14 field test interview, containing just the enrollment, employment, and locating sections, of about 20 minutes' duration (Modified Interview group). All base year nonrespondents will be able to complete the BPS interview online at any time, once the initial notification is received by email. Their cases also will be moved to outbound calling 5 days after the initial data collection mailing, allowing time for initial contact materials to reach them by regular mail. The early response phase will continue for 3 weeks, with all base year nonrespondents receiving a \$30 check for a completed interview. Throughout the BPS:12/14 field test, the assignment of base year nonrespondents to either the full or modified interview will not change.

2. Production Interviewing Phase (Weeks 4 – end of data collection).

Following the 3-week early response phase, all sample members who did not respond will be made available to RTI's Call Center Services (CCS) for outbound calling. While some sample members will have already received calls from CCS during the Early Response Phase, the frequency of calls will increase to the level of routine production interviewing.

At the end of the first two weeks of the production phase, remaining nonrespondents will be divided into a control and experimental group. Each group will have an equal number of base year respondents, base year nonrespondents offered the full interview, and base year nonrespondents offered the modified interview from the early phase.

Time 1 Mahalanobis Calculation. At the end of the first two weeks of production interviewing, a Mahalanobis distance will be calculated across all sample members in the control and experimental groups. Cases with the highest Mahalanobis values will be identified as Time 1 high distance cases, irrespective of group assignment although, with random assignment, we anticipate an approximately equal number of high distance cases in both groups. Those sample members assigned to the control group will be tracked as either Control-High Distance or Control-Normal Distance but, otherwise, data collection will continue as in the production interviewing phase, with sample members able to complete the interview online or by telephone. Those in the control group who complete the interview will receive a \$30 check.

In the Experimental group, Experimental-Normal Distance cases will be treated like the control group in that they will be able to complete the interview online or by telephone, and will receive a \$30 check for a completed interview. The Experimental-High Distance cases will, like the other three groups, be able to complete the interview online or by telephone, and the frequency with which they are contacted will be the same. However, those Experimental-High Distance cases completing the interview will receive a check for \$55, instead of \$30.

Mailings to the four groups will follow what is considered a standard schedule of contacts, in both conventional mail and electronic mail forms. Production interviewing following the Time 1 Mahalanobis calculation will continue for 3 weeks.

Time 2 Mahalanobis Calculation. After the 3 weeks of outbound calling, Mahalanobis values for all remaining BPS interview nonrespondents will be recalculated. From the Control-Normal Distance and Experimental-Normal Distance groups, a subsample of new high-distance nonrespondent cases will be identified. Sample members in the Control-High Distance group will still receive \$30 for a completed interview; sample members in the Experimental-High Distance group will have their incentive offer increased to \$55. Cases already offered the higher \$55 incentive at Time 1 will continue at that incentive level while all others will continue to be offered \$30 for a completed interview whether in the Control-High Distance or Normal Distance group. Production interviewing following the Time 2 Mahalanobis calculation will continue for another 3 weeks.

Mahalanobis Calculations, Time X_{3,4}. Mahalanobis values for all remaining interview nonrespondents will be recalculated after two additional 3-week periods of outbound calling, for a total of 4 calculations covering 12 weeks of data collection. As in prior weeks, nonrespondents with the highest Mahalanobis distances will be assigned to either the Control- High Distance or the Experimental-High Distance group (\$55). The field test design is summarized in exhibit 1.

Responsive Design Research Questions

With the assumption that increasing the rate of response among high-distance cases will reduce nonresponse bias, the BPS:12/14 responsive design experiment will explore the following research questions:

1. Do response rates differ between high-distance cases in the experimental and control groups?
2. Do estimates of key variables differ between high-distance and low-distance cases?
3. Does treatment of high-distance cases reduce nonresponse bias?

Methods and Null Hypotheses

Research question 1: do response rates differ between high-distance cases in the experimental and control groups?

Because of the assumption that yielding a greater number of high-distance cases will reduce nonresponse bias, response rates for the high-distance control and treatment groups will be compared to determine whether the overall response rates for the treatment and control groups differ significantly. Specifically:

- H₀: There will be no difference in response rates between the high-distance experimental and control groups

Research question 2: do estimates on key variables differ between high-distance and low-distance cases?

Using administrative record data available for the entire sample, such as indicators of federal aid data from NSLDS, RTI will compare estimates of key variables between high- and low-distance sample members. They will also compare interview outcome measures between all high-distance respondents and all low-distance respondents. Specifically:

- H_0 : There will be no difference in estimates for key variables known for all cases (e.g., enrollment beyond year 1, federal financial aid applications, federal loan amount) between the high- and low-distance sample members
- H_0 : There will be no difference in survey estimates between the high- and low-distance survey respondents
- H_0 : There will be no difference in survey estimates between all respondents excluding the high-distance experimental group, and all respondents, excluding the high-distance control group

Research question 3: does treatment of high-distance cases reduce nonresponse bias?

In addition to calculating nonresponse bias statistics for the whole sample, RTI will measure the effect of the responsive design approach on nonresponse bias by comparing estimates between the control and treatment groups. Specifically:

- H_0 : There will be no difference in unit nonresponse bias between all eligible cases excluding the high –distance treatment group, and all eligible cases excluding the high-distance control group.

The analysis will rely upon the variables listed below, typically used in BPS to evaluate bias, and will identify significant bias at the $p < .10$ level, if any:

- Type of base year institution
- Region
- Central Processing match during base year
- Federal aid applicant
- Pell Grant recipient
- Total Pell Grant amount received
- Stafford Loan recipient
- Total Stafford Loan amount received
- Base year institution undergraduate enrollment
- Age at base year
- High school graduation year
- Dependency status at base year
- Income level at base year
- Race/ethnicity
- Gender
- Marital status at base year
- Citizenship status at base year

As part of the planning process for developing the experimental design, the differences necessary to detect statistically significant differences have been estimated. That is, how large of a difference between the control and experimental groups is necessary to determine whether the response rates are different in hypothesis 1 or how large of a difference between the groups is necessary to determine whether the estimates are different in hypotheses 2 through 5.

Table 5 shows the expected sample sizes and statistically significant detectable difference for the hypotheses to be tested at both $\alpha = .05$ and $.10$. Several assumptions were made regarding response rates and sample sizes. In general, the closer a rate is to 50 percent (either less than or greater than), the larger the detectable difference. Likewise, smaller sample sizes require larger detectable differences.

Assumptions:

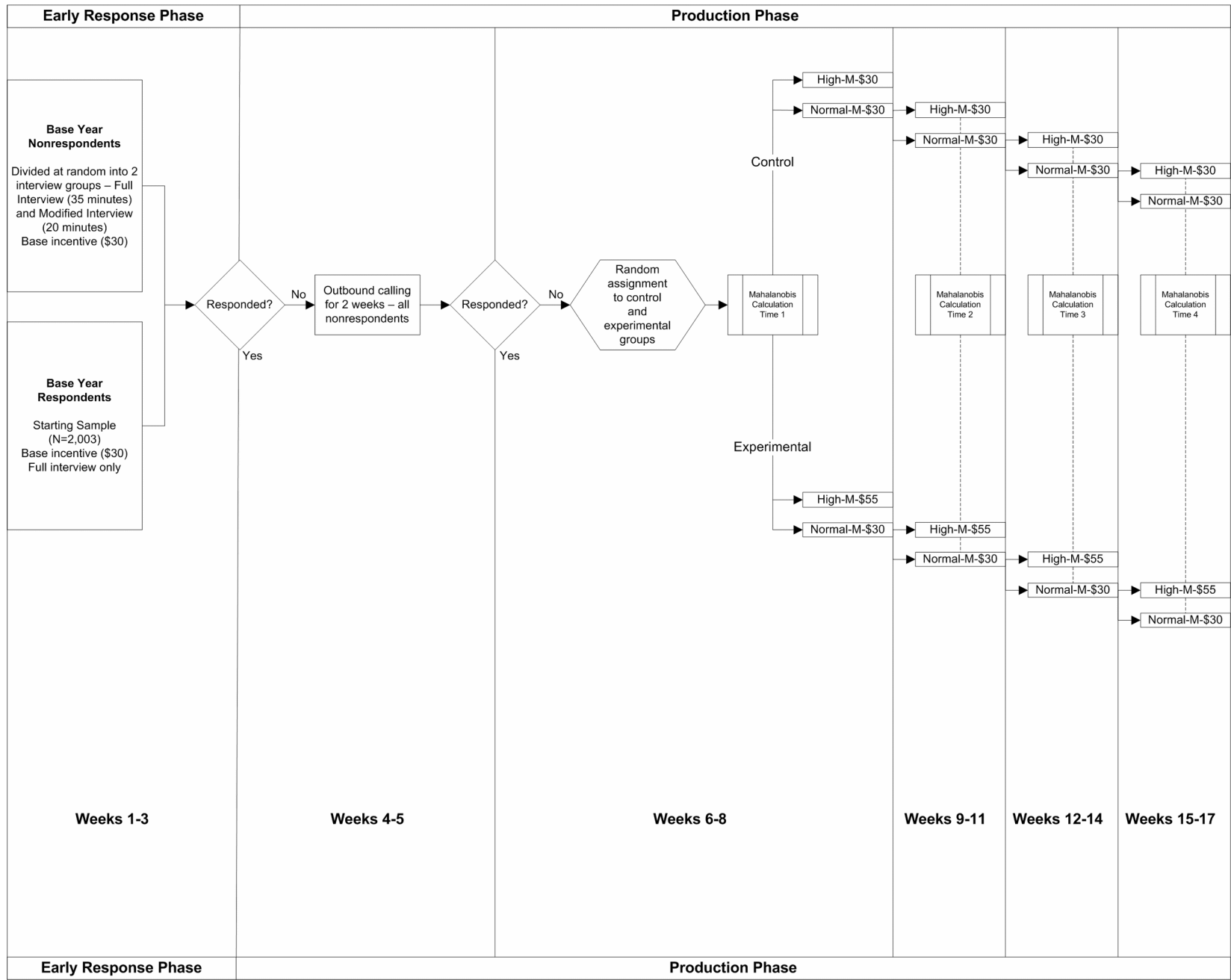
1. Detectable differences with 90 percent confidence were calculated with a two-tailed test for all hypotheses.
2. Initially, the sample will be equally distributed across experimental cells.
3. All eligible sample members, that is, those who are nonrespondents at the time of the first Mahalanobis calculation, will be included in the analyses of hypotheses 1 and 2a.
4. All eligible sample members will be included in the analysis of hypothesis 3.
5. Only respondents will be included in the analyses of hypotheses 2b and 2c because outcome measure data will only be known for respondents.
6. The top 30 percent will be used for determining the cut point for high and normal distance cases.
7. The response rate for the control group for hypothesis 1 will be 30 percent.
8. Unit nonresponse bias for the control group for hypothesis 3 will be ten percent.⁴
9. The statistical tests will have 80 percent power with an alpha of 0.10.
10. The statistical tests will use weighted data.
11. A design effect of 2.0 is assumed.

Table 5. Detectable differences for experimental hypotheses

Hypothesis	Control group		Experimental/Comparison group		Detectable difference at alpha	
	Definition	Sample size	Definition	Sample size	$\alpha = .05$	$\alpha = .10$
1	High-distance cases that did not receive treatment	341	High distance cases that ever received treatment (\$55)	341	9.3	7.8
2a	Normal distance cases	1,154	High distance cases	682	8.1	6.9
2b	Normal distance respondents	593	High distance respondents	350	11.7	1.0
2c	All respondents, excluding high-distance experimental cases,	2,405	All respondents, excluding high-distance control cases	2,446	4.5	3.8
3	Eligible cases, excluding high-distance treatment cases	2,616	Eligible cases, excluding high-distance control cases	2,616	2.5	2.2

⁴ Ten percent is generally considered the maximum acceptable value for unit nonresponse bias analysis.

Exhibit 1. Design of the BPS:12/14 field test data collection experiment



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

Names of individuals consulted on statistical aspects of study design, along with their affiliation and telephone numbers are:

Name	Affiliation	Telephone
Dr. John Riccobono	RTI	(919) 541-7006
Dr. Jennifer Wine	RTI	(919) 541-6870
Dr. James Chromy	RTI	(919) 541-7019
Dr. Natasha Janson	RTI	(919) 316-3394
Mr. Peter Siegel	RTI	(919) 541-6348
Dr. Sara Wheeless	RTI	(919) 541-5891
Dr. Alexandria Radford	MPR	(202) 478-1027

In addition to these statisticians and survey design experts, the following statisticians at NCES have also reviewed and approved the statistical aspects of the study: Dr. Tracy Hunt-White, Ted Socha, Dr. Matt Soldner, Dr. Sean Simone, and Dr. Sarah Crissey.

a. Other Contractors' Staff Responsible for Conducting the Study

The study is being conducted by the Postsecondary, Adult, and Career Education (PACE) division of the National Center for Education Statistics (NCES), U.S. Department of Education. NCES's prime contractor is RTI. RTI is being assisted through subcontracted activities by MPR Associates. Principal professional staff of the contractors, not listed above, who are assigned to the study are provided below:

Name	Affiliation	Telephone
Dr. Bryan Shepherd	RTI	(919) 316-3482
Ms. Donna Anderson	RTI	(919) 990-8399
Mr. Jeff Franklin	RTI	(919) 485-2614
Ms. Chris Rasmussen	RTI	(919) 541-6775
Mr. Michael Bryan	RTI	(919) 541-7498
Dr. Jennie Woo	MPR	(510) 849-4942
Ms. Shirley He Chan	MPR	(510) 849-4942

C. Overview of Analysis Topics and Survey Items

The BPS:12/14 field test data collection instrument is presented in Appendix H. Many of the data elements to be used in BPS:12/14 appeared in the previously approved NPSAS:12 and BPS:04/09 interviews. New items, as well as items that are to be included in the re-interview and in the modified interview used in the experiment, are identified in Appendix H.

References

Folsom, R.E., Potter, F.J., & Williams, S.R. (1987). Notes on a Composite Size Measure for Self-Weighting Samples in Multiple Domains. *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, 792-796.