

2008/12 BACCALAUREATE AND BEYOND
LONGITUDINAL STUDY: (B&B:08/12)

Supporting Statement Part B
Request for OMB Review
OMB # 1850-0729 v.8

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

April 16, 2012

Contents

B. Collection of Information Employing Statistical Methods.....	5
1. Potential Respondent Universe and Sampling.....	5
B&B:08/12 (Second Follow-up)Sample Design.....	5
2. Methods for Maximizing Response Rates.....	9
a. Locating.....	9
b. Interviewing Procedures.....	9
c. Quality Control.....	11
3. Tests of Procedures and Methods.....	11
a. Results of data collection experiment #1: Increasing Survey Participation Using Informational Video	12
b. Results of data collection experiment #2: Response Propensity Approach.....	12
c. Responsive Design Approaches and Metrics.....	14
d. Experimental Design.....	21
4. Reviewing Statisticians and Individuals Responsible for Designing and Conducting the Study.....	25
a. Other Contractors' Staff Responsible for Conducting the Study.....	26
C. Overview of Analysis Topics and Survey Items.....	26
References.....	27

List of Tables

Table 12. Distribution of the B&B:08/12 sample, by interview response status in NPSAS:08 and B&B:08/09.....	6
Table 13. B&B:08/09 nonresponse bias analysis.....	7
Table 14. Summary of Mahalanobis values, by month of data collection – B&B:08/09.....	16
Table 15. Outcome measure estimates, by change in respondent status using a cut point of the third quartile (27.5) – simulation 2.....	19
Table 16. Estimates and confidence intervals ($\alpha = .95$) for postbaccalaureate enrollment and federal financial aid by Mahalanobis distance cut points.....	20
Table 17. Experimental design.....	22
Table 18. Detectable differences for experiment hypotheses.....	25

List of Figures

Figure 1. Response rates by predicted propensity level, B&B:08/12 field test.....	13
Figure 2. Distribution of Mahalanobis distance values among all nonrespondents, the R target group and other nonrespondent groups after 3 months.....	15
Figure 3. Summary of Mahalanobis values by month of data collection – B&B:08/09 data.....	16

Figure 4. Scatterplot of Mahalanobis values for Nonrespondents after 3 months of Data Collection.....17

Figure 5. Summary of Mahalanobis values using a cut point of the third quartile (27.5), by change in likelihood of response - simulation 1 (Average Mahalanobis vs. Percent change.....18

Figure 6. Summary of Mahalanobis values using a cut point of the third quartile (27.5), by change in respondent status - simulation 2.....18

B. Collection of Information Employing Statistical Methods

1. Potential Respondent Universe and Sampling

The B&B:08/12 sample design has four stages. The first two stages occurred during the 2007-08 National Postsecondary Student Aid Study (NPSAS:08), when samples of NPSAS-eligible institutions and students within institutions were selected. The third stage was in the first follow-up, when all confirmed baccalaureate recipients and a subsample of potential baccalaureate recipients from NPSAS:08 were included in the B&B:08/09 sample. The fourth stage occurs during the second follow-up, when all eligible sample members from B&B:08/09 (as determined by the B&B:08/09 interview and the transcripts) are included in the B&B:08/12 sample. The sampling specifications presented here describe the sample design for the second follow-up study¹.

B&B:08/12 (Second Follow-up) Sample Design

The B&B:08/09 sample included all base year respondents and a subsample of 500 base year nonrespondents. For the B&B:08/12 full-scale study, all prior nonrespondents are included in the sample. Thus, the sample will include about 1,500 additional difficult cases. As a result, we anticipate that the overall response rate will be lower than observed in B&B:08/09, at approximately 81 percent². Although the anticipated response rate is less than 85 percent, we expect that the resulting yield of respondents will be larger than would have been obtained under a subsampling scenario (approximately 13,822, rather than 13,140).

There were three types of nonrespondents in B&B:08/09:

- a student who responded to the NPSAS:08 interview but did not respond to the B&B:08/09 interview (referred to henceforth as a first follow-up nonrespondent);
- a student who did not respond to the NPSAS:08 interview but did respond to the B&B:08/09 interview (referred to henceforth as a base year nonrespondent); and
- a student who did not respond to either the NPSAS:08 or B&B:08/09 interviews (referred to henceforth as a double nonrespondent).

Table 12 shows the distribution of the B&B:08/12 sample by prior response status, and the expected response rate and predicted yield. The B&B:08/12 sample will consist of all B&B:08/09 eligible respondents and all B&B:08/09 nonrespondents, resulting in a sample size of 17,058. An eligibility rate of 95 percent among the B&B:08/09 nonrespondents is assumed. Based on an estimated response rate of 81 percent, the expected yield is 13,822.

¹ see Cataldi et al. (2011) for additional information about the sampling design for the prior stages

² In accordance with NCES statistical standards, nonresponse bias analysis, weighting, and imputation will be conducted as appropriate.

Table 12. Distribution of the B&B:08/12 sample, by interview response status in NPSAS:08 and B&B:08/09

NPSAS:08 interview status	B&B:08/09 interview status	Count	Expected eligibility (95%)	Expected response rate	Expected yield
Total		17,164	17,058	0.81	13,822
Respondent	Respondent	14,825	14,825	0.86	12,750
Respondent	Nonrespondent	1,883	1,789	0.50	894
Nonrespondent	Respondent	223	223	0.45	100
Nonrespondent	Nonrespondent	233	221	0.35	77

NOTE: Many of the NPSAS:08 interview nonrespondents were study respondents and therefore have some NPSAS data.

As part of our planning process, some alternative sample designs considered including:

- all B&B:08/09 interview respondents and a subsample of first follow-up and double nonrespondents;
- all B&B:08/09 interview respondents, all first follow-up nonrespondents, and a subsample of double nonrespondents;
- all B&B:08/09 interview respondents and a subsample of first follow-up nonrespondents, and exclude all double nonrespondents; and
- all B&B:08/09 interview respondents and all first follow-up nonrespondents, and exclude all double nonrespondents.

The following section discusses the details of each of these alternative scenarios and why it was not chosen.

NCES longitudinal surveys have taken different approaches to sampling nonrespondents in the follow-up studies. For example, BPS and previous rounds of B&B have typically included either all nonrespondents or a subsample of the various types of nonrespondents. For ECLS-B and ECLS-K, follow-up sample members had to be base year respondents, and for ELS:2002, nonrespondents to both the base year and first follow-up were excluded from the second follow-up but counted as nonrespondents.

Interviewing first follow-up nonrespondents and double nonrespondents will likely be difficult and more costly than interviewing B&B:08/09 respondents, and the response rate among prior nonrespondents is likely to be low. In the field test, the unweighted response rates for first follow-up and double nonrespondents were higher than expected at 49.1 percent and 36.7 percent, respectively.

To determine whether the time, effort, and cost to attempt interviews with these nonrespondents would be well invested, we considered the effects of subsampling nonrespondents and excluding double nonrespondents on nonresponse bias, design effects, and analysis. Findings from these analyses are discussed below.

Nonresponse bias can potentially occur when respondents and nonrespondents are different. As part of the B&B:08/09 weighting process, a student nonresponse bias analysis was conducted and nonresponse bias did exist. Nonresponse weighting adjustments were done which reduced the bias. While that bias analysis compared all nonrespondents (both first follow-up nonrespondents and double nonrespondents) with respondents, we have also conducted bias analyses comparing the double nonrespondents with B&B:08/09 respondents and with first follow-up nonrespondents. As shown in table

13, these additional analyses also indicate that bias exists; the double nonrespondents are different from the B&B:08/09 respondents and first follow-up nonrespondents. While weight adjustments in B&B:08/12 could adjust for this bias even if the double nonrespondents are excluded, it is preferable to include some or all of them in the sample so that those who do respond would provide data to strengthen the nonresponse model.

Table 13. B&B:08/09 nonresponse bias analysis

Group	Mean relative bias	Median relative bias	Percent significant bias
Respondents vs. nonrespondents	3.90*	3.14	27.5
Respondents vs. double nonrespondents	3.72*	2.69	30.0
First follow-up nonrespondents vs. double nonrespondents	12.39*	8.90	25.0

* The mean relative bias is significantly different from zero at the 0.05 level.

Any subsampling affects the unequal weighting effect (UWE), which is a component of the design effect (DE). Subsampling would increase the design weights of the subsampled cases and likely cause their weights to be much different from the weights for the other sample members, thereby causing the variance to increase overall. The overall UWE using the interview weight for B&B:08/09 was 2.4, and subsampling would likely cause the UWE to increase above 2.4. While trimming and smoothing of the weights is frequently done to reduce the UWE, it is preferable to not subsample or to subsample at a high rate rather than to introduce a large UWE. For example, subsampling a tenth of the nonrespondents would result in weights for the subsampled cases ten times higher than their initial weight, but a subsample of half of the nonrespondents would result in weights for the subsampled cases only two times higher than their initial weight.

Another important factor to be considered is the analytical use of the data. Including all or a subsample of prior nonrespondents in the sample will likely provide better data given the potential bias of the first follow-up nonrespondents and the double nonrespondents. However, these nonrespondents would not be analyzed independently from the other sample members, so weight adjustments could be sufficient.

Another analytical consideration is how the transcript data will be used for B&B:08/12 analyses and what transcript panel weights may be necessary. Some of the first follow-up nonrespondents and double nonrespondents have transcript data and are included on the B&B:08/09 transcript file but not the interview file. Including all B&B:08/09 nonrespondents will allow for more flexibility for transcript analyses.

Additionally, there will possibly be a third follow-up of this cohort, so future longitudinal analyses also need to be considered. Third follow-up panel weights can be constructed to look at different combinations of respondents, as long as the sample size is sufficient. Including all B&B:08/09 nonrespondents in B&B:08/12 will again allow for more flexibility as a third follow-up study is designed.

Including prior nonrespondents could also have implications for imputation. In B&B:08/09, data were imputed for NPSAS:08 variables that were missing for some B&B cases because they:

- were NPSAS study nonrespondents but B&B interview respondents;

- were determined to be eligible for B&B:08 after the NPSAS study because they were identified in NPSAS as graduate students;
- were not identified as B&B eligible in NPSAS, but were later determined to be eligible via transcript information.

However, for B&B:08/12, data from previous rounds will not be imputed for first follow-up or double nonrespondents who respond in B&B:08/12. Instead, panel weights will be created in addition to a cross-sectional weight, and B&B:08/12 respondents will be analyzed using the appropriate weight, taking into account their response status in previous rounds. Recent changes to PowerStats will facilitate the use of different weights for different analyses.

Including all B&B:08/09 nonrespondents rather than a subsample may improve the imputation donor pool by including a larger number of B&B:08/12 respondents who may have different characteristics from other respondents. That is, when B&B:08/12 items need to be imputed for prior round nonrespondents there should be a sufficient number of similar cases that can be used as donors for imputation.

Given the small number of double nonrespondents, expected response rates for B&B:08/09 nonrespondents, and sufficient resources to pursue nonrespondents, we plan to include all B&B:08/09 nonrespondents in the B&B:08/12 sample.

Because the students in the B&B:08/12 sample are a subset of the NPSAS:08 sample, the B&B:08/12 weights will be derived from the NPSAS:08 weights. Weights will be computed to compensate for the unequal probability of selection of institutions and students in the NPSAS:08 sample as well as for the subsampling of nonrespondents in B&B:08/09. The weights will also adjust for multiplicity at the institutional and student levels and unknown student eligibility for NPSAS:08. The B&B:08/12 base weight is the NPSAS:08 weight prior to nonresponse adjustments and adjusted for the B&B:08/09 subsampling.

Nonresponse and poststratification adjustments for the B&B:08/12 respondents will also be computed. Poststratification will be used to adjust the B&B:08/12 weights so that they match B&B:08/09 weight sums. The poststratification adjustment will also include trimming and smoothing of the weights to reduce unequal weighting.

In addition to a cross-sectional analysis weight, there will be panel weights for analysis of the B&B:08/12 data in conjunction with the NPSAS:08, B&B:08/09, and transcript data. To facilitate computation of standard errors for both linear and nonlinear statistics, a vector of 200 bootstrap sample weights will be computed following NPSAS:08 procedures.

A precision goal for NPSAS:08 was to achieve relative standard errors (RSEs) of 10 percent or less. For key national estimates, an additional goal was to achieve RSEs that were comparable to or less than the NPSAS:2000 RSEs for those estimates. This helped to determine the sample size of the B&B:08 cohort.

2. Methods for Maximizing Response Rates

a. Locating

Several locating methods were used to find and collect up-to-date contact information for the B&B:08/12 sample. During B&B:08/09, batch searches of national databases and address update mailings to sample members were conducted prior to the start of data collection. Follow-up locating methods were employed for those sample members not found after the start of data collection, including CATI locating and intensive tracing.

The response rate for the B&B:08/12 full scale data collection is a function of success in two basic activities: locating the sample members and gaining their cooperation. We will rely on a variety of tracing techniques to locate and survey sample members. The methods used to locate sample members are based on the experience gained from the 2009 round of B&B, the B&B:08/12 field test, and other recent postsecondary education studies.

Many factors will affect our ability to successfully locate and survey sample members for B&B:08/12. Among them are the availability, completeness, and accuracy of the locating data from NPSAS:08 and B&B:08/09. Our 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 B&B sample members and to new leads developed through locating efforts.

To achieve the desired locating and response rates, we will use a multistage locating approach that will capitalize on available data for the B&B:08/12 sample from previous rounds. 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 our 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.

b. Interviewing Procedures

Training procedures. Training will be provided for individuals working in survey data collection and will include critical quality control elements. Contractor staff with extensive experience in training

interviewers will prepare the B&B Telephone Interviewer Manual, which will provide detailed coverage of the background and purpose of B&B, sample design, questionnaire, and procedures for the telephone interview. This manual will be used in training and as a reference during interviewing. Training staff will also prepare training exercises, mock interviews (specially constructed to highlight the potential of definitional and response problems), and other training aids.

Interviews. As with the field test study, 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 and 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.

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.

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 the field test, RTI will use QUEST to monitor approximately 10 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 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.

3. Tests of Procedures and Methods

Two experiments were conducted during the B&B:08/12 field test. The first tested whether viewing a short informational video to describe the study had any impact on response rates. The second experiment evaluated the use of an approach designed to model response propensity and target cases with low likelihood of response, with the goal of improving weighted response rates, minimizing nonresponse bias, and improving data quality. Each of these experiments and the plans derived from their outcomes are described in detail below.

a. Results of data collection experiment #1: Increasing Survey Participation Using Informational Video

In a prior clearance package, we received permission (approved 8/18/2010) to test whether a short informational Lego video increased a sample member's likelihood of visiting a website to confirm or update locating information. Results of this experiment showed no significant difference in the rate of address update completions between the group that saw the video and the group that did not. The Lego video was also included in contact materials distributed at the start of field test data collection. Field test results indicated that those who received the video were not more likely to complete the survey instrument.

We propose to extend the previous experimental design to include an additional treatment. This additional treatment will allow us to evaluate the effectiveness of multiple exposures to informational videos on interview participation rates. Sample members will be randomly assigned to control and treatment groups within the control and treatments groups used for the panel maintenance video experiment. The interview treatment group will receive a link to the video with the data collection announcement, and with subsequent reminders. The control group will receive the study materials without the video link. This design will allow examination of the effectiveness of the video for improving interview participation while taking into account effects of the panel maintenance video experiment and will allow the impact of the interview invitation video to be tested conditionally within the address update video groups; that is, the four cells created by the interaction of the two experiments can be evaluated (e.g., control 1 vs. control 2, control 1 vs. experiment 2, experiment 1 vs. control 2, and experiment 1 vs. experiment 2).

b. Results of data collection experiment #2: Response Propensity Approach

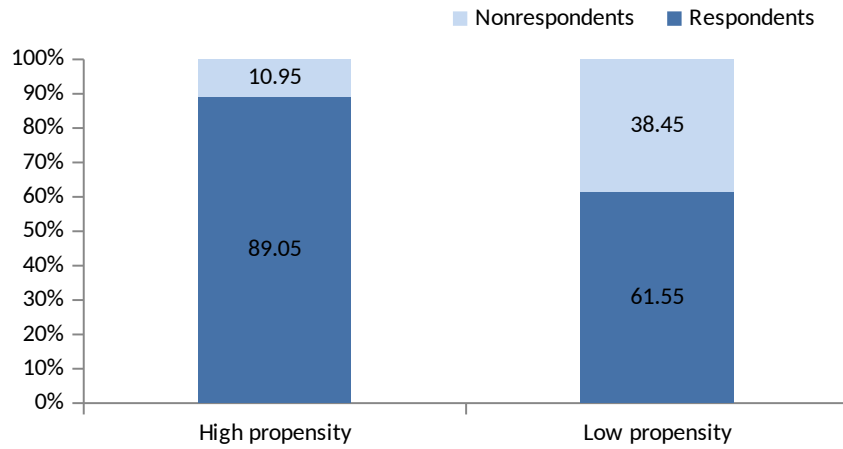
Nonresponse bias in sample surveys can lead to inaccurate estimates and compromise data quality. In the B&B:08/12 field test, we tested a new methodology, developed by RTI, with the goal to minimize nonresponse bias by targeting cases that have a low likelihood of responding and a high likelihood of contributing to nonresponse bias. We describe the results of this experiment in this section.

Survey organizations commonly address nonresponse bias by attempting to increase the survey response rate. This step is usually accomplished by pursuing nonrespondents believed to be most likely to complete an interview. However, this approach may not be successful in reducing nonresponse bias even if higher response rates are achieved (Merkle and Edelman, 2009). To the extent that response propensity and key survey estimates are related, nonresponse bias could even be increased by yielding more heterogeneous, higher response propensity cases. In contrast, if lower response propensity cases are brought into the response pool, we anticipated that this would increase the weighted response rate and result in less biased survey estimates. This is the hypothesis we tested with this experiment.

As outlined in previous submissions (1850-0729 v.7, approved 7/1/2011), this experiment had several key steps: identify those cases that are *least* likely to respond to the interview; develop an incentive program that could be used to bring in those cases; and evaluate the accuracy of the predicted response rates and the potential for reduction of bias in the full scale study. These steps have been outlined in detail in previous submissions.

Figure 1 shows that our model was able to accurately predict relative propensity 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).

Figure 1. Response rates by predicted propensity level, B&B:08/12 field test



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, $t = 2.04$, $df = 539$). Additionally, those with the lowest propensity scores within the high propensity classification showed a large difference in response rates between treatment and control groups (79 and 89 percent, respectively), but the difference was not significant, due to small group sizes ($t = 1.29$, $df = 78.6$).

In summary, results showed that a higher monetary incentive did increase response for the majority of the sample members, but not those individuals near the highest and lowest propensity scores. Based on these field test findings, the full scale study will have three initial incentive amounts. The 30% of cases that have the highest response propensity scores will receive an initial incentive offer of \$20. The 30% of cases with the lowest response propensity scores will receive an initial incentive offer of \$55. All others will receive an initial offer of \$35.

Despite these efforts, field test results did not show a reduction in bias as a result of the additional response. Analyses based on full scale B&B:08/09 data indicated that those respondents identified as least likely to respond would have significantly increased nonresponse bias had they not responded. In contrast, those cases that were estimated to be most likely to respond would not have had a significant impact on nonresponse bias, had they not responded.

Due to the equivocal evidence for the benefits of the *a priori* estimation of propensity scores, we propose a revised approach that focuses on an iterative process of identifying and targeting cases most likely to contribute to nonresponse bias. RTI is currently undertaking an initiative, modeled on Responsive Design methodologies developed by Groves (Groves and Heeringa, 2006), to develop new approaches to improve survey outcomes that incorporate different responsive and adaptive features. Although still in the development phase, RTI has implemented several of these procedures on recent

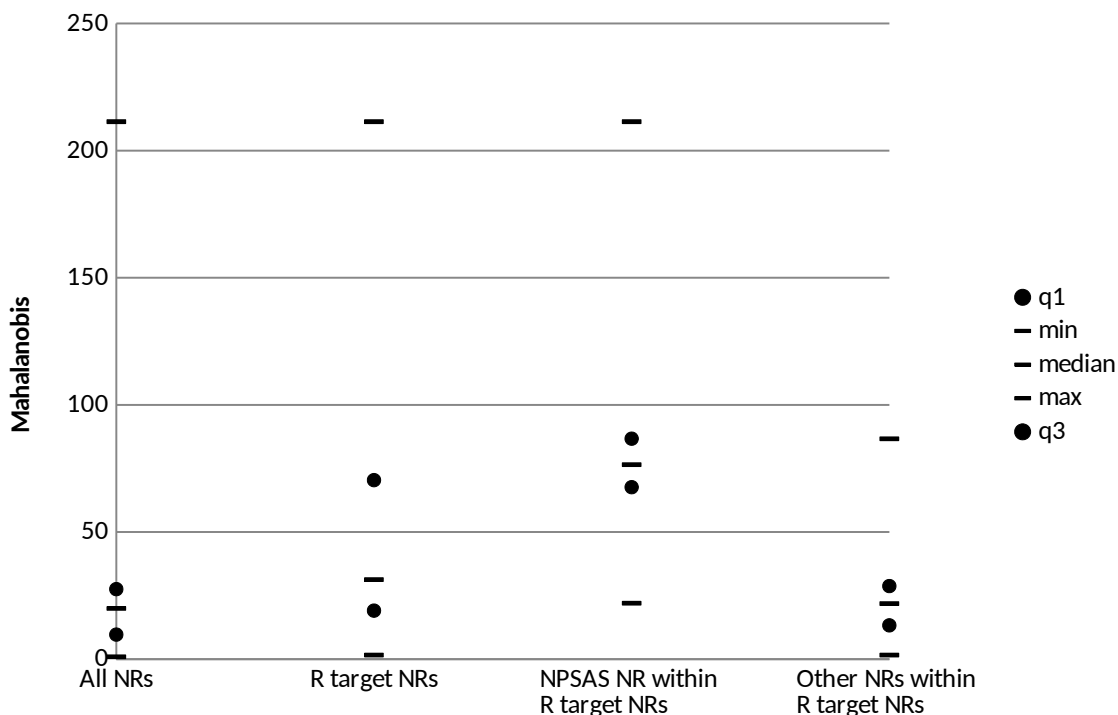
studies and have published preliminary results (Rosen, et al., 2011; Peytchev, et al., 2010). An approach modeled on the Responsive Design methodologies is described in detail in below.

c. Responsive Design Approaches and Metrics

We tested responsive data collection designs modeled on two types of statistical distancing measures as alternatives to the *a priori* nonresponse bias reduction approach evaluated in the field test. Two potential metrics for identifying cases most likely to contribute to nonresponse bias, the R-indicator (Schouten et al., 2011) and the Mahalanobis distance measure, were estimated and used in simulations with B&B:08/09 data. We also looked at outcome measures from other data sources to determine if the two nonresponse bias identification metrics were able to capture bias as discussed in the previous section. The sections below discuss how each metric was created and evaluated.

A key requirement of a responsive data collection design is the ability to identify nonrespondents who are most likely to contribute to nonresponse bias. For this reason, it was important to determine which metric was better able to identify potentially biasing cases and the extent of overlap between the sample cases identified by the two measures. Results indicated a negative relationship between Mahalanobis distance and the R-indicator; those identified as high-distance cases via the Mahalanobis calculation also tended to be from groups with low representativeness (low R-values). In other words, R and M generally identified the same cases as potential contributors to nonresponse bias. Figure 2 presents the distributions of Mahalanobis distances for all nonrespondents, the group of nonrespondents identified as target cases by the R-indicator, NPSAS nonrespondents within the R target group and other nonrespondents within the R target group. The cases with the highest distance according to Mahalanobis are mostly the same cases the R-indicator suggests targeting.

Figure 2. Distribution of Mahalanobis distance values among all nonrespondents, the R target group and other nonrespondent groups after 3 months



Note: Targeted estimates are made up of NPSAS nonrespondents and others as indicated by the calculation of partial R. NR = nonrespondents.

The discussion below focuses primarily on analyses using the Mahalanobis distance. The same simulations were conducted using the R-indicator and were presented to NCES and OMB via email and phone conference on 2/7/12.³ Results of simulations using both measures produced similar findings. Given similar abilities to identify cases likely to contribute to nonresponse bias, the Mahalanobis distance, which can be calculated for an individual case, provides more flexibility for implementation than does the R-indicator, which is generally a group (or subgroup) level measure. Because findings from the R-indicator analyses have been presented elsewhere, and because our recommendation is to proceed using the Mahalanobis distance, we do not discuss the R-indicator results in depth below.

Using the B&B:08/09 full-scale data, the Mahalanobis distance was computed for each sample member as the difference between a multivariate vector containing the covariates⁴ and the mean (or expected value) of the vector for the full sample. As Table 14 and Figure 3 indicate, the average Mahalanobis distance for respondents approaches the overall sample average over the course of data collection. However, the average value for nonrespondents *increases* during the same time period, and moves further away from the overall sample average. This is expected if the high-distance cases are not converted to respondents by the end of data collection. The differences between the average respondent Mahalanobis values, the average nonrespondent Mahalanobis values, and the average full sample Mahalanobis values are not significantly different.

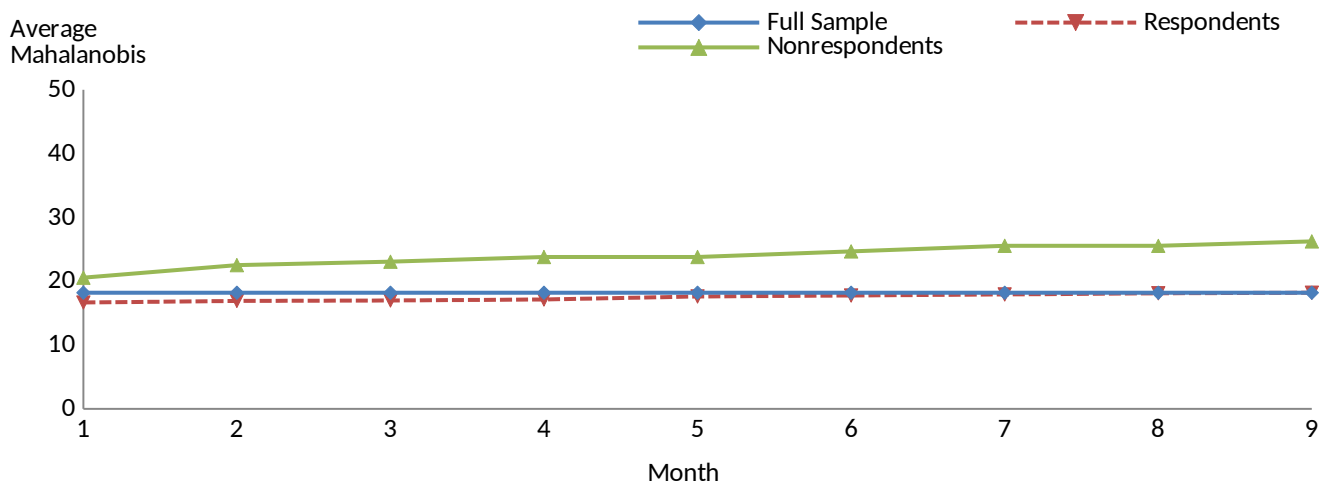
Table 14. Summary of Mahalanobis values, by month of data collection – B&B:08/09

Month	Response rate	Average Mahalanobis overall	Average Mahalanobis for respondents	Average Mahalanobis for nonrespondents
1	0.353	18.2	16.7	20.6
2	0.599	18.2	16.9	22.5
3	0.642	18.2	17.0	23.1
4	0.697	18.2	17.1	23.8
5	0.750	18.2	17.6	23.8
6	0.798	18.2	17.8	24.7
7	0.837	18.2	17.9	25.6
8	0.855	18.2	18.1	25.5
9	0.877	18.2	18.2	26.2

³ Supporting documents for the 2/7/2012 meeting are included as Appendix H. Recommendations in Appendix H are superseded by those in Parts A and B of this document.

⁴ The covariates included in the model and a discussion of how they were chosen were included in a previous package (1850-0729, approved 7/5/2011).

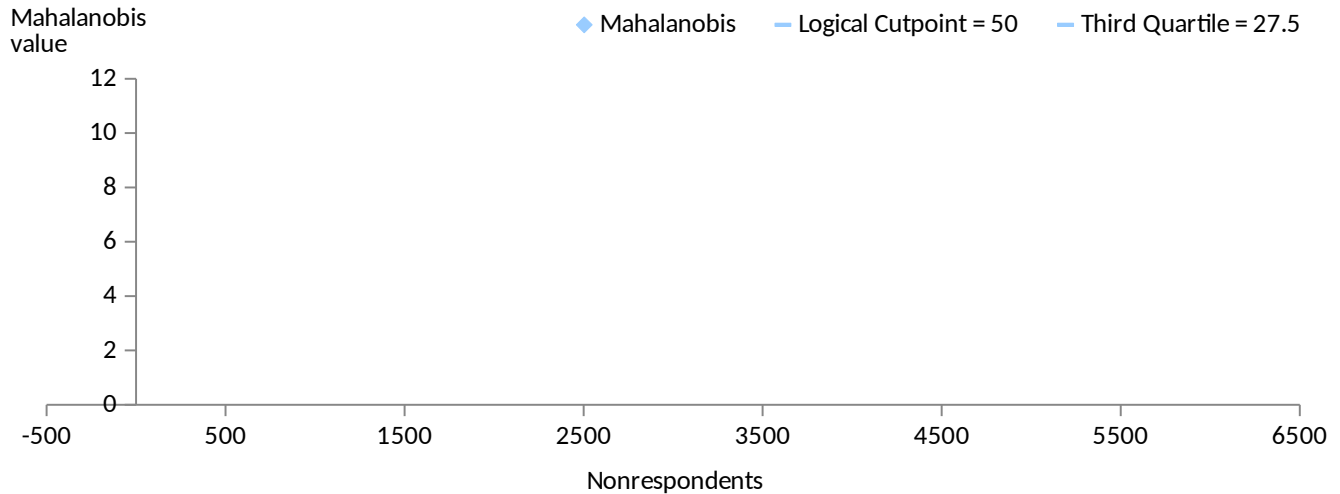
Figure 3. Summary of Mahalanobis values by month of data collection – B&B:08/09 data



Because responsive design approaches seek to target specific cases for intervention during the data collection period, to use the Mahalanobis distance within a framework requires the identification of one or more treatment cut points. Once identified, those above the cut point, e.g. high-distance cases or those presumed to be most likely to contribute to nonresponse bias if they remain nonrespondents, are eligible to receive special data collection procedures to increase the likelihood of response. We identified two potential cut points by examining a scatterplot of the Mahalanobis values three months into data collection using B&B:08/09 full-scale data (Figure 4). Cut points were set at 50, where there appears to be a logical separation of the data, and 27.5, the third quartile. The first cut point, 50, yields 513 high-distance nonrespondents after three months of data collection. The third quartile cut point, 27.5, yields 1,556 high-distance nonrespondents after three months of data collection. Based on these two cut points, we ran two simulations using B&B:08/09 data⁵.

⁵ Although these simulations demonstrate that we can estimate and apply the Mahalanobis calculations as planned, the results for the B&B:08/12 full scale study, which will include experimental treatments based on distance classification, may differ from those observed in the simulations since no experimental treatments were used in B&B:08/09.

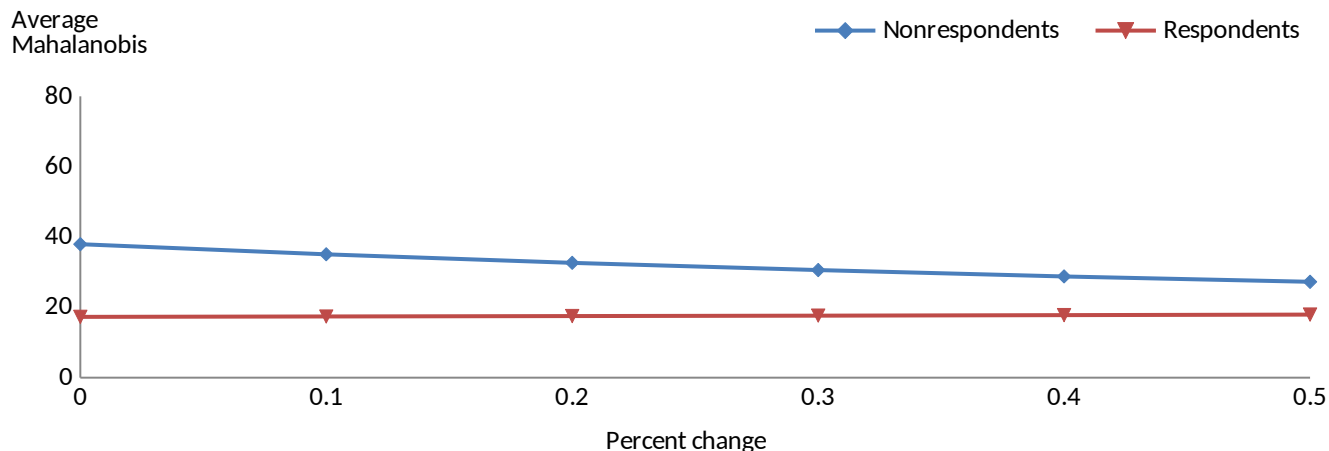
Figure 4. Scatterplot of Mahalanobis values for Nonrespondents after 3 months of Data Collection



The first simulation evaluated the effect on average Mahalanobis distances of increasing the number of high Mahalanobis distance respondents while keeping the overall response rate similar. To keep the overall response rate similar to the observed B&B:08/09 response rate, as the likelihood of response for the targeted group was increased by 10, 20, 30, 40, and 50 percent the likelihood of response of the non-high-distance cases was decreased by the same percentage. The simulation was run 1,000 times and the average Mahalanobis distance after three months was calculated for each iteration.

Average Mahalanobis distance values for first follow-up respondents and nonrespondents were not significantly different, likely due to the high variability in Mahalanobis distance represented in Figures 2 and 4. However, the average distance for respondents remained relatively constant across the simulations, whereas the value for nonrespondents decreased as more high-distance cases were converted to respondents, ultimately converging towards the average for respondents. Figure 5 presents the results of the simulations using the third quartile (27.5) as the cut point; results were similar when the cut point of 50 was used.

Figure 5. Summary of Mahalanobis values using a cut point of the third quartile (27.5), by change in likelihood of response - simulation 1 (Average Mahalanobis vs. Percent change



The second simulation assessed whether decreasing the response rate among high-distance respondents would have affected key outcome measures in B&B:08/09. High-distance cases that were nonrespondents after three months of data collection but ultimately became respondents by the end were randomly assigned to nonrespondent status for the purpose of simulating outcome distributions. Using the cut point of 50, the number of high-distance nonrespondents after three months of data collection that became respondents by the end was 265. Using the third quartile, that number was 920. The percentage of cases that were switched from respondent to nonrespondent was varied in the simulations with 0, 10, 20, 30, 40, 50, and 100 percent of the of the high-distance respondents being treated as nonrespondents.

The simulation was run 500 times and the survey outcomes were averaged over the 500 simulations using the new set of final respondents to calculate them; results were similar when the cut point of 50 was used. Figure 6 shows that the average Mahalanobis value for nonrespondents increased as the number of high-distance nonrespondents increased. No significant differences were observed in the outcome measures for any of the scenarios (table 15).

Figure 6. Summary of Mahalanobis values using a cut point of the third quartile (27.5), by change in respondent status - simulation 2

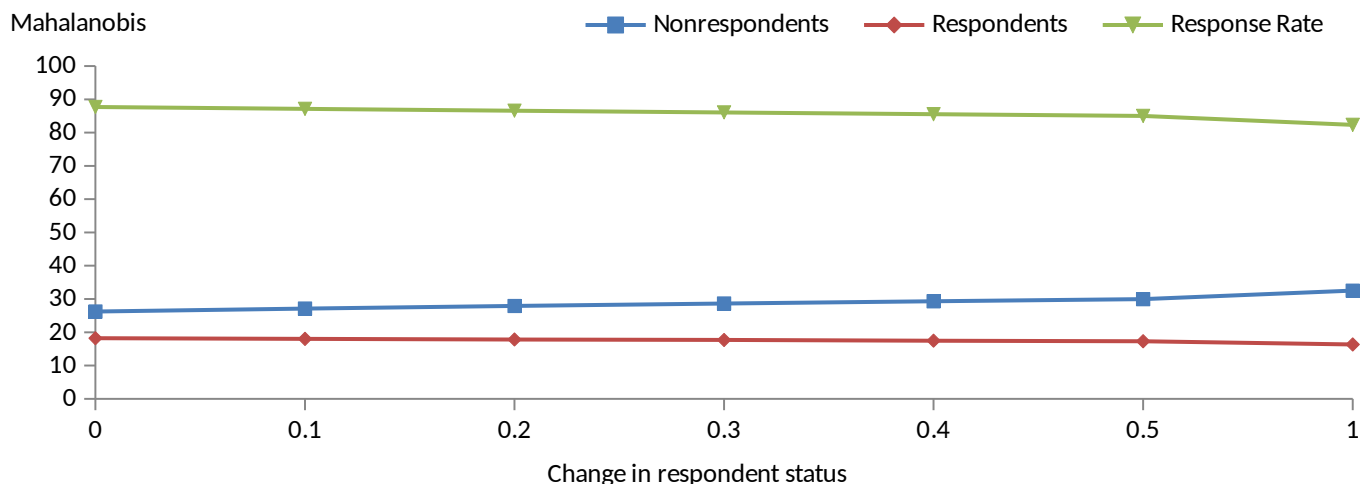


Table 15. Outcome measure estimates, by change in respondent status using a cut point of the third quartile (27.5) – simulation 2

Outcome measure	Percent of high-distance respondents treated as nonrespondents and confidence intervals ($\alpha = .95$)			
	None	10 percent	50 percent	All
Bachelor's degree major – STEM major	0.164 (0.158, 0.170)	0.164 (0.157, 0.17)	0.162 (0.155, 0.169)	0.159 (0.152, 0.167)
Cumulative undergraduate grade point average (multiplied by 100, mean)	326.3 (325, 327.5)	326.5 (325.2, 327.7)	327.4 (326.1, 328.7)	328.8 (327.5, 330.1)
First institution sector – 2-year or less	0.298 (0.287, 0.31)	0.299 (0.288, 0.311)	0.303 (0.291, 0.315)	0.309 (0.296, 0.321)
Number of institutions attended before bachelor's completion	0.551 (0.538, 0.564)	0.553 (0.539, 0.566)	0.558 (0.545, 0.572)	0.568 (0.555, 0.581)
Time to 2007-08 bachelor's degree (mean time in months)	78.7 (76.8, 80.6)	78.8 (76.9, 80.7)	79 (77, 81)	79.3 (77.3, 81.4)
Cumulative total amount borrowed (mean)	16299 (15843, 16755)	16371 (15916, 16826)	16678 (16216, 17140)	17158 (16687, 17629)
Cumulative amount owed as of 2008-09 (mean)	15841 (15365, 16317)	15915 (15440, 16390)	16232 (15749, 16715)	16727 (16234, 17220)
Cumulative federal amount borrowed (mean)	11304 (10992, 11616)	11338 (11025, 11651)	11475 (11152, 11799)	11694 (11355, 12033)
Debt burden in 2008-09 (mean)	3.408 (3.098, 3.718)	3.428 (3.118, 3.738)	3.524 (3.206, 3.842)	3.666 (3.342, 3.991)
Ever received Pell grant	0.372 (0.358, 0.385)	0.373 (0.359, 0.386)	0.375 (0.361, 0.389)	0.379 (0.363, 0.395)
Loan status in 2008-09 – not repaying	0.178 (0.168, 0.187)	0.178 (0.169, 0.188)	0.182 (0.172, 0.192)	0.187 (0.177, 0.198)
Enrollment status in degree program in 2009 – master's	0.011 (0.0085, 0.0136)	0.011 (0.0085, 0.0135)	0.011 (0.0082, 0.0133)	0.01 (0.0078, 0.0129)
Highest degree program enrollment after bachelor's degree, as of 2009 – master's	0.194 (0.184, 0.204)	0.194 (0.184, 0.204)	0.197 (0.187, 0.207)	0.201 (0.191, 0.211)
Number of jobs held since bachelor's degree – one	0.501 (0.489, 0.514)	0.501 (0.488, 0.514)	0.501 (0.488, 0.514)	0.501 (0.488, 0.514)
Employment status in 2009 – one job	0.703 (0.692, 0.714)	0.702 (0.692, 0.713)	0.701 (0.69, 0.712)	0.698 (0.687, 0.709)
Satisfied with employment in 2009 – compensation	0.558 (0.549, 0.572)	0.557 (0.544, 0.571)	0.553 (0.54, 0.567)	0.547 (0.533, 0.561)
Employer benefits in 2009 offered medical or health insurance	0.763 (0.752, 0.774)	0.762 (0.751, 0.773)	0.758 (0.747, 0.769)	0.752 (0.74, 0.763)
Earned income in 2009 (mean)	29140 (28526, 29753)	29086 (28474, 29698)	28853 (28236, 29469)	28480 (27864, 29096)
Job not part of career in industry	0.165 (0.153, 0.177)	0.165 (0.153, 0.177)	0.166 (0.154, 0.178)	0.168 (0.156, 0.181)
Job unrelated to major	0.272 (0.259, 0.284)	0.272 (0.26, 0.284)	0.273 (0.261, 0.285)	0.275 (0.263, 0.287)
Highest education attained by either parent – bachelor's degree	0.26 (0.25, 0.271)	0.26 (0.249, 0.271)	0.259 (0.248, 0.27)	0.257 (0.246, 0.268)

See notes at end of table.

Table 15. Outcome measure estimates, by change in respondent status using a cut point of the third quartile (27.5) – simulation 2—Continued

Outcome measure	Percent of high-distance respondents treated as nonrespondents and confidence intervals ($\alpha = .95$)			
	None	10 percent	50 percent	All
Age at bachelor's degree receipt (mean)	25.27 (25.08, 25.46)	25.28 (25.09, 25.47)	25.3 (25.1, 25.5)	25.34 (25.13, 25.54)
Has disability in 2007-08	0.082 (0.075, 0.089)	0.082 (0.075, 0.089)	0.08 (0.073, 0.087)	0.077 (0.07, 0.084)
Marital status and dependents – unmarried with no dependents	0.653 (0.64, 0.666)	0.652 (0.64, 0.665)	0.651 (0.637, 0.664)	0.647 (0.633, 0.662)
Volunteered in last 12 months as of 2009	0.409 (0.397, 0.421)	0.409 (0.398, 0.421)	0.411 (0.399, 0.423)	0.414 (0.4, 0.427)
Ever voted as of 2009	0.875 (0.866, 0.883)	0.875 (0.866, 0.884)	0.878 (0.869, 0.886)	0.882 (0.874, 0.89)

Note: Outcome measures were estimated for 0, 10, 20, 30, 40, 50, and 100 percent of respondents treated as nonrespondents. Results are presented only for 0, 10, 50, and 100 percent.

RTI also investigated the potential bias in the high-distance cases by analyzing information obtained from external resources known for both respondents and nonrespondents: the National Student Clearinghouse (NSC) and the National Student Loan Data System (NSLDS). Estimates of postbaccalaureate enrollment and attainment rates, as well as federal loan application status and amounts borrowed were calculated for all sample members, and compared by distance groupings. Results are presented in Table 16. High-distance cases (based on the third-quartile cut point) were significantly less likely than low-distance cases to have enrolled in postsecondary education since receiving their bachelor's degree. High-distance cases were also significantly less likely to have applied for federal financial aid and also borrowed less than low-distance cases (at both cut points).

Table 16. Estimates and confidence intervals ($\alpha = .95$) for postbaccalaureate enrollment and federal financial aid by Mahalanobis distance cut points

	Cut point 1 (50)		Cut point 2 (27.5)	
	Low distance (≤ 50)	High distance (> 50)	Low distance (≤ 27.5)	High distance (> 27.5)
NSC				
Enrolled ^b	0.199 (0.1900 - 0.2091)	0.186 (0.1449 - 0.2269)	0.218 (0.2076 - 0.2289)	0.181 (0.1590 - 0.2032)
Attained	0.013 (0.0108 - 0.0159)	0.013 (0.0041 - 0.0218)	0.015 (0.0126 - 0.0179)	0.012 (0.0069 - 0.0169)
NSLDS				
Total amount guaranteed ^b	\$14,590.31 (14156 - 15014)	\$13,488.49 (11870 - 15101)	\$16,018.77 (15521 - 16394)	\$13,182.08 (12295 - 14081)
Applied for aid ^{a,b}	0.639 (0.6275 - 0.6505)	0.568 (0.5124 - 0.6119)	0.69 (0.6760 - 0.6984)	0.573 (0.5453 - 0.5978)
Log of total amount guaranteed ^{a,b}	6.246 (6.1311 - 6.3605)	5.567 (5.0221 - 6.0046)	6.761 (6.6211 - 6.8469)	5.599 (5.3247 - 5.8439)

^a Significant difference ($p < .05$) between high and low distance cases for cut point 1.

^b Significant difference ($p < .05$) between high and low distance cases for cut point 2.

In conclusion, we propose a responsive design using the Mahalanobis distance measure to identify cases for targeted treatments with the goal of maximizing response among cases presumed to be most likely to contribute to nonresponse bias. Based on analyses using NSC and NSLDS data, it appears that groups defined by Mahalanobis distance do exhibit statistically significant differences on key metrics. While our simulation results do not yield unequivocal support for using Mahalanobis-based treatment decisions during data collection, we propose continued exploration of this approach during the B&B:08/12 full scale study. Due to already high response rates and generally low rates of nonresponse bias, this presents a low-risk—and potentially high-reward—opportunity for NCES and the larger federal statistical community.

Several steps remain to be taken between this submission and full-scale data collection. First, we will revisit the list of variables to be included in the calculation of Mahalanobis to be sure that we have included all important covariates. Additional variables that will be considered for inclusion in the model are the respondent's race, Hispanicity, NSC enrollment and attainment information, and NSLDS loan status (e.g., whether the respondent has applied for loans and the amount they have received). This work is currently scheduled to be completed by the end of May. Then, after the Mahalanobis values have been calculated with the refined model, we will then review a scatterplot of the Mahalanobis values for the B&B:08/12 sample before data collection to determine the appropriate starting cut point. During data collection we will monitor distance values for respondents and nonrespondents to determine the best cut point for each evaluation point. Cut points will be set such that there are a sufficient number of cases for the experiment described below. We will provide OMB with the final model results and cut points as part of a non-substantive change memo when they are available. This work is currently scheduled to be completed in early June. To evaluate the effectiveness of our efforts, we propose the experiment described in the next section.

d. Experimental Design

Our proposed experiment is predicated on the assumption that, to reduce nonresponse bias, response rates among high-distance cases must be increased. To that end, prior to the start of data collection all sample cases will be randomly assigned to control and treatment groups (see table 17 for a description of the experimental groups). Then, treatment group cases with a Mahalanobis value above a to-be-determined cut point will be targeted during data collection at three points in time. At each:

- Mahalanobis values will be evaluated for all remaining nonrespondents;⁶
- cases will be assigned to low- and high-distance groups on the basis of the cut point; and
- treatment cases within the high-distance group will be eligible for interventions as defined below.

⁶ While the Mahalanobis values will not change during data collection, the cases who are nonrespondents will change. Therefore, the cut point could potentially change and an individual respondent's classification as high- or low-distance may change as well.

Table 17. Experimental design

Data collection step	Description	Time frame	Mahalanobis distance	High-distance (based on Mahalanobis distance)		Low Distance (based on Mahalanobis distance)	
				Control	Treatment	Control	Treatment
1	Initial Invitation and CATI-Light	July 2012 –March 2013	Calculate Mahalanobis for all cases. Assign cases to high/low distance based on a cut point	\$20, \$35, or \$55 (based on predicted response propensity)			
2	Full CATI	Begins in October 2012	Evaluate Mahalanobis for nonrespondents. Assign cases to high/low distance based on a cut point		Additional \$15 ¹		
3	Extensive Case Review	Begins in November 2012	Evaluate Mahalanobis for nonrespondents. Assign cases to high/low distance based on a cut point	Late review	Additional \$15 +Early review	Late review	Late review
4	Abbreviated interviews	Begins in January 2013	Evaluate Mahalanobis for nonrespondents. Assign cases to high/low distance based on a cut point	Late abbreviated	Additional \$15 +early abbreviated	Late abbreviated	Late abbreviated

¹ Once a case becomes eligible for the additional \$15, they remain eligible for the additional \$15 even if they move into the low-distance group.

Treatments

Treatment 1 (Month 3) – additional incentive. The first three months of data collection will include web data collection and “CATI-light”, which involves a minimal number of phone calls, mainly to prompt web response. After the first three months, Mahalanobis values will be evaluated for the remaining nonrespondents, and cases above the cut point will be offered a \$15 incentive in addition to their original offer (\$20, \$35, or \$55 based on their response propensity score). Once a case becomes eligible for the additional \$15, they remain eligible for the additional \$15 even if they move into the low-distance group later.

Treatment 2 (Month 4) – extensive case review. After an additional month of data collection, Mahalanobis values will be evaluated again for remaining nonrespondents, and those above the new cut point (determined based on the remaining nonrespondents at month 4) will receive early extensive case review. Project staff will review the CMS-CATI events log, along with any paradata available for a particular case (e.g., availability of e-mail address, parent address, etc.), to identify any specific actions that may be considered to encourage the 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., 6 weeks later).

Treatment 3 (Month 6) – abbreviated interview. After an additional two months of data collection, 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).

Research Questions

Because our assumption is that increasing the rate of response among high-distance cases will reduce nonresponse bias, we will explore the following research questions:

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

Methods and Null Hypotheses

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

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

- H_0 : There will be no difference in response rates between the high-distance treatment and control groups

Research question two: do key outcome measures differ between high-distance and low-distance cases?

Using administrative record data available for the entire sample, such as indicators of post-baccalaureate enrollment and federal aid data from NSC and NSLDS, we will compare outcome measures between high- and low-distance sample members. We 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 outcome estimates known for all cases (postbaccalaureate enrollment, federal financial aid applications, federal loan amount) between the high- and low-distance sample members
- H_0 : There will be no difference in survey outcome estimates between the high- and low-distance survey respondents
- H_0 : There will be no difference in estimates between all respondents, excluding the high-distance treatment group, and all respondents, excluding the high-distance control group

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

In addition to calculating nonresponse bias statistics for the whole sample, we 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 respondents, excluding the high-distance treatment group, and all respondents, excluding the high-distance control group.

The analysis will rely upon the variables listed below and will identify significant bias at the $p < .05$ level, if any:

- institution type;
- region;
- institution enrollment from IPEDS file (categorical);
- Pell grant receipt (yes/no);
- Pell Grant amount (categorical);
- Stafford Loan receipt (yes/no);
- Stafford Loan amount (categorical);
- Parent Loan for Undergraduate Students (PLUS);
- federal aid receipt (yes/no);
- institutional aid receipt (yes/no);
- state aid receipt (yes/no);
- any aid receipt (yes/no);
- postbaccalaureate enrollment; and
- postbaccalaureate degree attainment.

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

Table 18 shows the expected sample sizes and statistically significant detectable difference for the hypotheses to be tested. 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 95 percent confidence were calculated with a two-tailed test for all hypotheses.
2. The sample will be equally distributed across experimental cells.
3. All eligible sample members will be included in the analyses of hypotheses 1, 2 and 5.
4. Only respondents will be included in the analyses of hypotheses 3 and 4 because outcome measure data will only be known for respondents.
5. The third quartile will be used for determining the cut point for determining high and low distance cases.
6. The response rate for the control group for hypothesis 1 will be 30 percent.

7. Unit nonresponse bias for the control group for hypothesis 5 will be ten percent.⁷
8. The statistical tests will have 80 percent power with an alpha of 0.05.
9. The statistical tests will use weighted data.

Table 18. Detectable differences for experiment hypotheses

Hypothesis	Group 1		Group 2		Detectable difference with 95 percent confidence
	Definition	Sample size	Definition	Sample size	
1	High-distance cases with no additional or earlier treatment	780	High-distance cases with additional or earlier treatment	780	9.5
2	Low distance cases	11,330	High-distance cases	2,400	4.6
3	Eligible cases, excluding high-distance cases with additional or earlier treatment for the high-distance cases	16,360	Eligible cases, excluding high-distance cases with no additional or earlier treatment for the high-distance cases	16,360	1.5
4	All respondents, excluding high-distance cases with additional or earlier treatment for the high-distance cases	13,430	All respondents, excluding high-distance cases with no additional or earlier treatment for the high-distance cases	13,570	2.4
5	Eligible cases, excluding high-distance cases with additional or earlier treatment for the high-distance cases	16,360	Eligible cases, excluding high-distance cases with no additional or earlier treatment for the high-distance cases	16,360	0.9

4. 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 provided below.

Name	Affiliation	Telephone
Dr. John Riccobono	RTI	(919) 541-7006
Dr. Jennifer Wine	RTI	(919) 541-6870
Dr. James Chromy	RTI	(919) 541-7019
Ms. Melissa Cominole	RTI	(919) 990-8456
Mr. Peter Siegel	RTI	(919) 541-6348
Dr. Susan Choy	MPR	(510) 849-4942
Dr. Robin Henke	MPR	(510) 849-4942
Dr. Jennie Woo	MPR	(510) 849-4942

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, Dr. Sarah Crissey, and Dr. Tom Weko.

⁷ Ten percent is generally considered the maximum acceptable value for unit nonresponse bias analysis.

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
Mr. Joe Simpson	RTI	(919) 541-5941
Ms. Emily Forrest-Cataldi	MPR	(510) 849-4942
Ms. Stephanie Nevill	MPR	(510) 849-4942
Ms. Vicky Dingler	MPR	(510) 849-4942

C. Overview of Analysis Topics and Survey Items

The B&B:08/12 data collection instrument is presented in appendix G. Many of the data elements to be used in B&B:08/12 appeared in the previously approved B&B:08/09. Additional items will also be included in B&B:08/12. These items have been tested in cognitive interviews, and the report describing cognitive testing results is included with this submission.

References

- Bradburn, E.M., Berger, R., Li, X., Peter, K., and Rooney, K. (2003). *A Descriptive Summary of 1999–2000 Bachelor’s Degree Recipients 1 Year Later, With an Analysis of Time to Degree* (NCES 2003–165). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.
- Bradburn, E.M., Nevill, S., and Cataldi, E.F. (2006). *Where Are They Now? A Description of 1992–93 Bachelor’s Degree Recipients 10 Years Later* (NCES 2007–159). U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- Cataldi, E.F., Green, C., Henke, R., Lew, T., Woo, J., Shepherd, B., & Siegel, P. (2011). *2008–09 Baccalaureate and Beyond Longitudinal Study (B&B:08/09). First Look* (NCES 2011-236). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.
- Choy, S.P., and Li, X. (2006). *Dealing With Debt: 1992–93 Bachelor’s Degree Recipients 10 Years Later* (NCES 2006-156). U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- Groves, R. M., & Heeringa, S. (2006). Responsive design for household surveys: tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 169(Part 3), 439-457.
- Henke, R.R., Chen, X., and Geis, S. (2000). *Progress Through the Teacher Pipeline: 1992–93 College Graduates and Elementary/Secondary School Teaching as of 1997* (NCES 2000–152). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.
- McCormick, A.C., Nuñez, A.-M., Shah, V., and Choy, S.P. (1999). *Life After College: A Descriptive Summary of 1992–93 Bachelor’s Degree Recipients in 1997, With an Essay on Participation in Graduate and First-Professional Education* (NCES 1999–155). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.
- McCormick, A., and Horn, L.J. (1996). *A Descriptive Summary of 1992–93 Bachelor’s Degree Recipients: 1 Year Later, With Essay on Time to Degree* (NCES 96–158). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.
- Merkle, D. M., & Edelman, M. (2009). An Experiment on Improving Response Rates and Its Unintended Impact on Survey Error. *Survey Practice* (March)
- Nevill, S.C., and Chen, X. (2007). *The Path Through Graduate School: A Longitudinal Examination 10 Years After Bachelor’s Degree* (NCES 2007-162). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

- Peytchev, A., S. Riley, J.A. Rosen, J.J. Murphy, and M. Lindblad. (2010). Reduction of Nonresponse Bias in Surveys through Case Prioritization. *Survey Research Methods*, 4(1), 21-29
- Rosen, J., Murphy, J. J, A. Peytchev, Riley, S., & Lindblad, M;. (2011). The Effects of Differential Interviewer Incentives on a Field Data Collection Effort. *Field Methods*.
- Schouten, B., Shlomo, N., and Skinner, C. (2011). Indicators for Monitoring and Improving Representativeness of Response. *Journal of Official Statistics*, Vol.27, No.2, 2011. pp. 231–253