More Money? The Impact of Larger Incentives on Response Rates in a Two-Phase Mail Survey

Cameron McPhee
American Institutes for Research
cmcphee@air.org

Sarah Hastedt
National Center for Education Statistics
sarah.hastedt@ed.gov

Introduction

A growing literature has consistently demonstrated the effectiveness of prepaid cash incentives in boosting survey response rates across a variety of modes. As a result, survey researchers increasingly rely on incentives to improve response rates in surveys. Understanding how incentives can be used to achieve desired response rates under cost constraints is a critical challenge for survey researchers. In this paper we look at the impact of incentives on response rates in the National Household Education Survey (NHES) 2011 Field Test. The NHES utilizes a two-phase data collection approach in which households are first screened via the mail to determine if there is an eligible child in the household. If the household contains an eligible child, an in-depth topical survey is then sent to the household. To avoid possible biases associated with survey nonresponse and to protect the power of the sample, it is critical to achieve a high response to both surveys. Experiments using different incentive levels were included at both the first phase, referred to as the screener phase, and the second phase, referred to as the topical phase. At the screener level, the effectiveness of $2 and $5 prepaid cash incentives was tested. At the topical level, the effect of including no incentive or a $5, $10, $15, or $20 prepaid cash incentive with the initial topical mailing was tested. An additional experiment at the topical level tested the effectiveness of including a $5 or $15 cash incentive with the final nonresponse mailing.

Background: Incentives in Mail Surveys

The use of incentives has been consistently associated with higher response rates in mail surveys (Petrolia and Bhattacharjee 2009; Lesser, Dillman, Carlson, Lorenz, Mason, and Willits 2001; Church 1993; Armstrong 1975). Past research in this area has focused on examining the relationship between incentive amount, type (monetary or nonmonetary), and timing (prepaid or postpaid). A meta-analysis by Church (1993) found that increased response rates were only associated with incentives that were provided, prepaid, with the initial survey mailing rather than contingent upon completion of the survey. His analysis also found a high correlation between incentive amount and response rate and that prepaid monetary incentives generated the largest increases in response rates relative to controls. Lesser, Dillman, Carlson, Lorenz, Mason, and Willits (2001) examined the results from eight experiments that included financial incentives ranging from $2 to $5. Their analysis also found that financial incentives were associated with higher response. Two of the studies included in the analysis by Lesser et al. tested $2 prepaid incentives against $5 prepaid incentives. In one of the studies, the $5 group had a higher response rate than the $2 group while in the other study there was no significant difference in response rates between the two groups. A key difference between the two studies was that in the former study the incentive was included with the first mailing while in the latter study the incentive was included in the second mailing. This suggests that including incentives with the initial mailing may be more effective than including incentives in nonresponse mailings.

While past research provides strong support for the effectiveness of prepaid cash incentives at increasing response rates in single-phase mail surveys, less is known regarding how incentives operate in a two-phase mail survey. As noted in Brick, Williams, and Montaquila (2011), respondents may view each phase of the survey (screener and topical) as separate survey requests or the topical survey may be viewed as a continuation of the initial screener survey. If respondents view the topical survey as a continuation of one complete survey in combination with the screener survey, incentives sent for the topical phase might not lead to higher topical response rates. In their analysis of data from the NHES: 2009 Pilot Test, Brick et al. found that the $15 dollar prepaid cash incentive sent in the initial mailing at the topical phase of the survey was associated with higher response rates compared to the control group (no incentive) and the group that received a prepaid $5 cash incentive. The $5 incentive was associated with a nominal increase in response relative to the control group but this difference was not statistically significant. In our analysis, we build on this research by examining the impact of incentives on response rates at each phase of a two-phase mail survey and exploring the impact of the timing of incentives on response.

Background: The National Household Education Survey

The National Household Education Survey (NHES) is the primary household-based survey program sponsored by the National Center for Education Statistics (NCES) within the U.S. Department of Education. It covers topics that are difficult to study in institution-based frames, such as early childhood care and education, children's readiness for school, participation in adult and continuing education, parent involvement in education, school choice, and homeschooling. Surveys were conducted approximately every other year from 1991 through 2007, and each of these prior administrations used random digit dial (RDD) sampling and telephone data collection from landline telephones only. Each data collection involved a household screener instrument and two or three longer topical surveys focused on a particular subject. Telephone interviews were conducted using computer-assisted telephone interviewing (CATI) to accommodate the survey's complex skip patterns and automated within-household sampling techniques.

Like most RDD surveys, NHES response rates have been declining over time. The NHES screener response rate fell from over 80 percent in 1991 to 53 percent in 2007 (Van de Kerckhove et al. 2009). In addition to declining response rates, the increase in households converting from landlines to cell phone-only service has raised concerns about population coverage. Recent estimates from the National Health Interview Survey (NHIS) indicate that more than one quarter of American households have only cell phone service (Blumberg and Luke 2010). In addition to concerns about cell phone-only households, Fahimi, Kulp, and Brick (2009) found that the standard list-assisted method used for RDD sampling may fail to include up to 20 percent of landline numbers. The combination of the continuing declines in RDD survey response rates and concerns about population coverage prompted NCES to redesign the NHES program after the 2007 data collection. The NHES redesign involved shifting from an RDD landline sample with CATI data collection, to an address-based sample with a two-phase mail survey as the primary data collection approach. This new methodology was tested in both a small-scale feasibility pilot test in 2009 and a larger-scale field test in 2011.

The aim of the NHES: 2009 Pilot Test, conducted from September to December 2009, was to determine whether a two-phase mail survey design for the NHES was feasible, in terms of both response rates and data quality. The Pilot Test used a nationally representative address-based sample covering the 50 United States and the District of Columbia. The first step of the 2009 Pilot Test was screening a sample of approximately 11,800 households for the presence of eligible children. This was done through a brief self-administered mail questionnaire that asked respondents to enumerate children ages 20 or younger in the household. This questionnaire, referred to as the screener, acted as a screening instrument for second phase eligibility. Households were sent up to two nonresponse follow-up mailings if they did not respond to the initial screener mailing. All households were sent a prepaid cash incentive of $2 with their initial screener mailing (the incentive amount had been tested in previous NHES experiments). In each household with at least one eligible child, one child was selected for one of two topical follow-up surveys. In households with multiple eligible children, only one child was selected for a topical survey to

minimize burden on respondents. Parents or guardians of the sampled child were then mailed a topical questionnaire and asked to respond. The overall screener response rate for the NHES: 2009 Pilot was approximately 59 percent and the overall topical response rate was approximately 75 percent, for an overall response rate of about 44 percent (Brick et al. 2011).

In 2011, a larger-scale Field Test was conducted to test strategies for improving response rates, further refine operational procedures, and test different question wordings. For the NHES: 2011 Field Test, three independent samples consisting of 60,000 total addresses were drawn. The nationally representative sample comprised two-thirds of the total number of addresses. In addition to the nationally representative sample, two additional samples were drawn to examine aspects of the redesign associated with English literacy and the bilingual and Spanish mailing materials: (1) a linguistically isolated sample[1] and (2) a Spanish surname sample[2]. The analysis in this paper focuses only on the nationally representative sample of addresses; therefore the remaining discussion will focus on this sample.

In the first phase of the Field Test, households were screened for the presence of eligible children and youths. In order to test strategies for improving screener response rates, the first phase of the NHES: 2011 Field Test contained several experiments designed to investigate the impact of different incentive levels, mailing techniques, and screener forms on screener response rates. These experiments will be discussed in more detail in the "Methodological Experiments Included in the NHES: 2011 Field Test" section. The discussion below will focus on the basic mailing procedures used in the Field Test.

Sampled households were mailed an initial screener package[3] which contained a cover letter, screener questionnaire, postage-paid return envelope, and an incentive. Addresses were randomly assigned a monetary incentive level of $2 or $5. Several screener survey instruments were experimentally tested during the Field Test to investigate the impact of different screener forms on response rates. All screener survey instruments contained a question asking whether any children age 20 or younger lived in the household. If there were no children age 20 or younger living in the household, respondents were asked to check a box indicating this and return the form. If there were children or youth age 20 or younger living in the household, respondents were asked to enumerate each child and provide the child's age, sex, school enrollment status, and grade in school if the child was enrolled in school and return the form. Approximately two weeks after the initial screener mailing package was sent, households were sent a "Thank you"/Reminder postcard. Non-responding households were then sent a first follow-up mailing package which contained a cover letter, screener questionnaire, and postage-paid return envelope. Households that did not respond to the initial screener mailing or first follow-up mailing were sent a second follow-up mailing that contained a cover letter, screener questionnaire, and postage-paid return envelope.[4]

Similar to the screener phase of the NHES: 2011 Field Test, the topical phase of the Field Test included several embedded experiments such as testing the effect of different mailing envelopes, mailing methods, and incentive levels on topical response rates as well as testing different question wordings through the use of a split panel test of two different questionnaires for each topical survey. In households whose returned screener questionnaires indicated the presence of at least one eligible child, one child was selected for a topical follow-up survey. Eligible children fell into two categories: (1) children ages 0 to 6 and not yet enrolled in kindergarten were eligible to receive the Early Childhood Program Participation (ECPP) topical survey and (2) children age 20 or under and enrolled in public or

[1] Linguistically isolated areas were defined as census tracts where at least 13% of households had no household member age 14 or older who "speaks English well."
[2] The Spanish surname sample consisted of addresses identified by the address vendor as having Spanish surnames.
[3] A small group of households was randomly assigned to receive an advance letter to test the effect of an advance letter on response rates. These households then received a screener mailing package which contained a cover letter, screener questionnaire, postage-paid return envelope, and incentive.
[4] The second follow-up mailing was sent either via Federal Express or United States Postal Service (USPS) Priority Mail.

private school for kindergarten through twelfth grade[5] were eligible for the Parent and Family Involvement in Education (PFI) topical survey. In households with multiple children, one child was randomly selected as the focal child for a topical survey in order to minimize burden on respondents.

The initial topical mailing package contained a cover letter, topical questionnaire, and postage-paid return envelope. For some randomly selected respondents, the initial topical mailing contained a monetary incentive of varying amounts. Approximately two weeks after the initial topical mailing, a "Thank you"/Reminder postcard was sent out. Non-responding households were then sent a topical first follow-up mailing which contained a cover letter, replacement questionnaire, and postage-paid return envelope. Households that did not respond to the first topical follow-up mailing were sent a second topical follow-up mailing that contained a cover letter, replacement questionnaire, and postage-paid return envelope. For a subset of households that did not receive a monetary incentive with the initial mailing, the second follow-up mailing also contained an incentive of either $5 or $15.

Data and Methods: Methodological Experiments Included in the NHES: 2011 Field Test

Figure 1 summarizes the experiments at both the screener and topical levels conducted as part of the NHES: 2011 Field Test. As noted, these included experiments focused on incentives and experiments focused on form design, envelope design, and mailing method. Because incentives are the focus of this paper, we will first provide a detailed description of the incentive experiments included in the Field Test. A brief overview of the non-incentive experiments included in the Field Test will then be provided.

Field Test experiments were designed to build on findings from the NHES: 2009 Pilot Test and provide insight into the optimal level of the incentive, considering both response rates and cost. In the 2009 Pilot Test, all households received a $2 prepaid cash incentive with their initial screener mailing. To help isolate effects of changes in incentives and contact approaches relative to the 2009 Pilot Test, a portion of the 2011 Field Test sample was set aside to receive the exact same treatments as the 2009 Pilot Test sample group with the highest overall response rate. In the 2011 Field Test, addresses assigned to the Pilot control group received a $2 incentive. Addresses that were not part of the Pilot control group were randomly assigned to receive either a $2 cash incentive or a $5 cash incentive with their initial screener mailing. This test was designed to determine whether a gain in response rates could be realized by using a higher prepaid incentive at the first screener mailing.

At the topical level, varying the monetary value of the incentive and the timing of the incentive were both tested. Households that received their initial topical mailing via first-class mail were assigned to receive one of the following incentive levels at the initial topical mailing: $0, $5, $10, $15, or $20. All incentives were prepaid cash. This experiment was designed to build on findings from the 2009 Pilot test which tested including no incentive, a $5 prepaid cash incentive, or a $15 prepaid cash incentive with the first topical questionnaire mailing. The results from the Pilot Test demonstrated a substantial increase in response at the topical level by using a $15 cash incentive (topical response rate of approximately 83%) compared to a $5 incentive (topical response rate of approximately 77%) and no incentive (topical response rate of approximately 72%) (Brick et al. 2011). The Pilot results did not provide information on whether an incentive level between $5 and $15 would be as effective as the $15 incentive or whether a further increase in response rates could be attained through an incentive greater than $15. The Field Test topical incentive level experiment allows us to address this issue.

Additionally, the effect of including an incentive at the second follow-up mailing (final nonresponse mailing) rather than at the initial topical mailing was tested. Households that were assigned to receive their initial topical mailing by Priority Mail did not receive an incentive with the initial topical mailing, but were eligible to receive an incentive at the final topical mailing if they did not respond to either the initial topical mailing or the first follow-up mailing. These households were randomly assigned to receive either a $5 or $15 cash incentive with the second follow-up

[5] Homeschooled children were not eligible to receive a PFI survey in the NHES: 2011 Field Test.

mailing. This experiment was designed to determine if only sending incentives to nonrespondents to the topical questionnaire could save money without negatively impacting response rates.

In addition to the incentive experiments included in the NHES: 2011 Field Test, several additional experiments were included at both the screener and topical levels. The first screener experiment involved testing the impact of sending an advance letter on screener response rates. To test this, a random subsample of addresses was selected to receive an advance letter prior to the initial screener questionnaire mailing. Another screener experiment involved testing the impact of including a token magnet with the Department of Education logo on it in the initial mailing. A random subsample of addresses received a token magnet in their initial mailing.

There were also several screener experiments that focused on the design of the screener questionnaire. One of these experiments tested two different screener questionnaire types. The first type was the "screenout" screener questionnaire. This was a two-page form which asked households to indicate whether or not there were any children age 20 or under in the household by checking a box and then enumerating any children living in the household. The second was an eight-page "engaging" screener questionnaire that contained a series of education and civic involvement questions. The engaging screener contained the same series of questions about the presence of children age 20 or under in the household and child enumeration as the screenout screener. In addition to testing different screener form types, the effect of requesting children's names on the screener form was tested. Each screener form type had a "name" version and a "no name" version. On the "name" version, the child enumeration section contained a question asking for the child's first name, initials, or nickname. This question preceded the questions asking for the child's demographic information (sex, age, school enrollment status). On the "no-name" version of the screener, the child's name was not asked and each child was denoted by a number on the form (e.g., Child 1, Child 2, etc.). The child enumeration section on these forms started with a request for the child's age. The combination of these experiments and the pilot test form resulted in five different screener mailing forms.

The effect of switching the form type (screenout vs. engaging) at the first follow-up mailing was also tested. This experiment was designed to examine whether changing the form type would elicit response from non-responding households. If the initial questionnaire did not elicit a response, a second form of a different type may lead households to respond. This experiment also tested whether the order in which households received different screener form types impacted response differentially. Finally, two different methods of rush delivery at the second nonresponse mailing were tested. Non-responding households after the first follow-up mailing were randomly assigned to receive either United States Postal Service (USPS) Priority Mail delivery or FedEx delivery for the second and final non-response mailing.

At the topical phase, an experiment was conducted that tested the effect of different first-class carrier mailing envelopes against USPS Priority Mail for the delivery of the first topical survey. Households with eligible children were randomly assigned to receive one of two first-class envelope types or USPS Priority Mail delivery.[6] The topical phase of the Field Test contained a split panel test of each survey instrument (ECPP and PFI) to test the efficacy of different question wordings. Households were randomly assigned to receive either a "mainline" version of the questionnaire (which was very similar to the questionnaire used in the 2009 Pilot Test) or an "alternate" version of the questionnaire which contained revised question wordings and new questions.

Data and Methods: Identifying Response Wave

As noted above, the NHES 2011 Field Test was conducted as a two-phase mail survey. The screener phase involved identifying households with children through a series of up to three screener questionnaire mailings. The topical

[6] Households that received a second follow-up mailing at the screener phase were not eligible for assignment to the USPS Priority Mail treatment group for the initial topical mailing. This was due to the need to distinguish between the last mailing received for the screener and the first mailing received for the topical. Households that received a second follow-up screener mailing received this mailing by special delivery (USPS Priority Mail or FedEx).

phase involved sending households with an eligible child a more detailed survey about the sampled child. The initial topical mailing was sent within three weeks of receiving the completed screener. Non-responders to the initial topical mailing were sent up to two follow-up mailings approximately three and six weeks after the initial topical mailing, respectively.

Since each non-response follow-up mailing adds cost, it was important to determine if any of the experimental treatments, including the incentive level, increased response overall and at the earlier mailing waves. To examine the impact of the experimental treatments on response time, households were classified according to the "wave" in which they returned a completed questionnaire. Wave 1 responders were those who responded to the initial mailing, wave 2 responders were those who responded to the first follow-up (second overall) mailing, and wave 3 responders were those who responded to the second follow-up (third overall) mailing. In order to classify households whose completed screener questionnaire was returned on a date close to one of the follow-up mailings, the date the questionnaire was returned (screener or topical) was compared to the date the most recent mailing was sent. If a questionnaire was received within three days of the date of the second mailing (the first nonresponse follow-up), the household was considered a "wave 1 respondent"; if the questionnaire was received within three days of the date of the third mailing (the second nonresponse follow-up), the household was considered a "wave 2 respondent"; otherwise, if the questionnaire was received more than three days after the third mailing, but within the field period, the household was considered a "wave 3 respondent." This decision was made based on the assumption that a household was unlikely to have received and returned a nonresponse follow-up form within three days of the form being mailed out. Response wave was calculated for both screener and topical questionnaire respondents. Figure 1 illustrates the flow of the mailing procedures for the nationally representative sample.

Data and Methods: Regression Analysis

In order to examine the impact of the incentives on response while controlling for the other experimental treatments included as part of the 2011 Field Test, logistic regression was used to model response propensity overall and by questionnaire mailing wave. These analyses were conducted separately for screener and topical response. At the screener phase, separate models were run predicting overall response to the screener questionnaire (at any time point), and response at each mailing wave. The independent variables used in the analysis included the following indicator variables (coded 0/1):

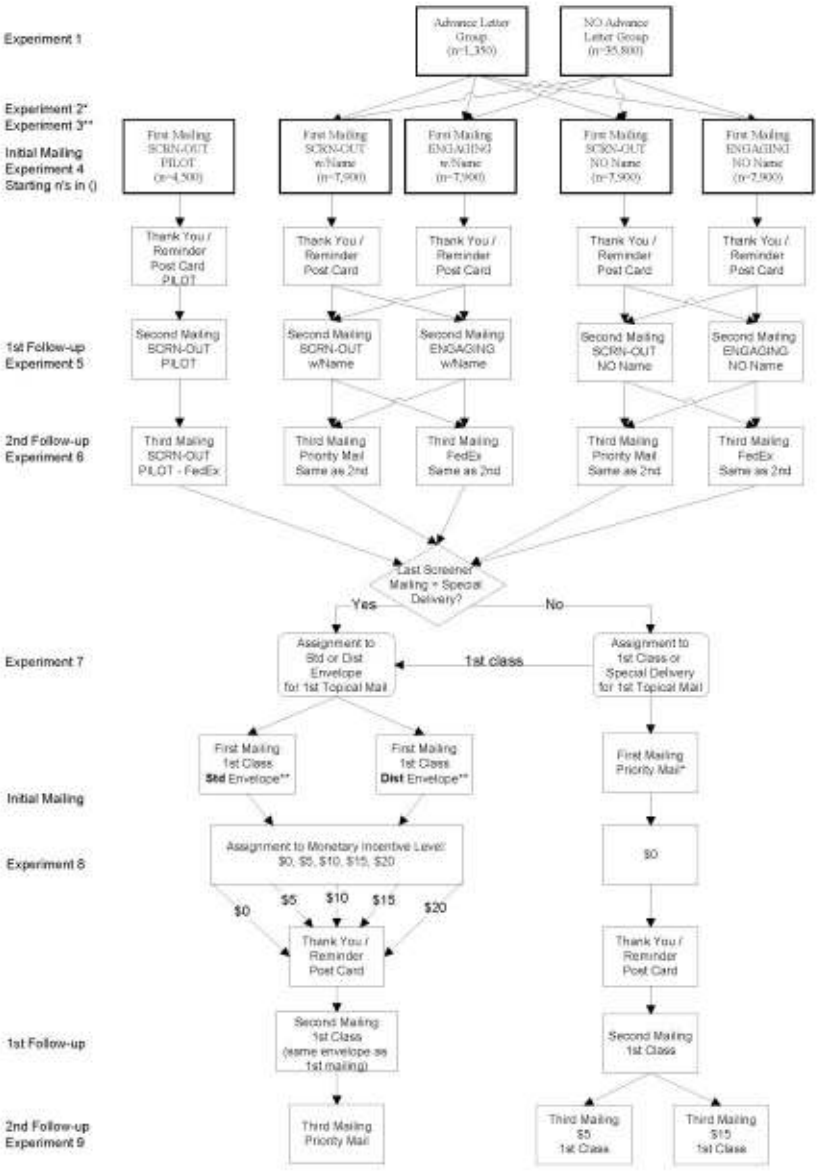
- Prepaid incentive ($2 or $5)
- Whether the household received a prenotice letter prior to the initial questionnaire mailing
- The screener type mailed at each mailing wave[7]
- Whether the household received a Department of Education magnet in the envelope with the initial screener mailing
- Whether the screener form asked the respondent to report the names or nicknames of the children in the household
- Whether the screener form asked the household to report the home phone number[8]
- Whether the third mailing (second nonresponse follow-up) was sent via FedEx or USPS Priority Mail[9]

[7] Some households received the same form at all three mailing waves (either a short "screenout" form or a longer, more detailed, "engaging" form), while others received a different form for the second and third mailing (the 1st and 2nd nonresponse follow-ups) than they did in the initial mailing. Mailing type for the 2nd and 3rd mailings was only included models predicting response at the 2nd and 3rd mailing waves.

[8] A question requesting the household's phone number was included at the end of the "screenout" and "engaging" forms but not the Pilot form.

[9] This variable was only included in models predicting response at the 3rd screener mailing.

Figure 1. Experiments in the NHES: 2011 Field Test



*Experiment 2: Addresses were randomly assigned to monetary incentive levels for the initial screener mailing; $2 (16,300) or $5 (16,300). The pilot path only received $2 (4,500).

**Experiment 3: Addresses were randomly assigned to receive a token magnet (1,700) or not receive a magnet (35,450) with the initial screener mailing. The pilot path did not receive a token magnet.

***Experiment 10: For each topical survey two versions of each were fielded for the NHES: 2011 Field Test. Each version was randomly distributed among the experimental conditions.

Note: Numbers included reflect 10% address ineligibility rate and are rounded to the nearest 50.

For the topical response rate analyses, eligible households were separated into two analytic groups based on the study design. Group 1 included a random selection of cases who responded to the screener after the first or second mailing wave as well as all cases who responded to the screener after the third mailing wave (i.e., late screener responders). These cases were randomly assigned to one of five prepaid incentive levels ($0, $5, $10, $15, or $20) mailed with the initial topical questionnaire. Group 2 included only cases who responded to the screener after the first or second mailing. These cases received the first topical mailing via USPS Priority Mail and no incentive with this initial mailing. Cases in this group who had not responded to the topical survey by the third mailing received either $5 or $15 with the third mailing (the second nonresponse follow-up).

The first set of analyses of the topical incentives examined the impact of the different initial mailing incentive amounts on response rates for households in Group 1. Logistic regression models were used to predict the likelihood of response to the topical survey overall and at each topical mailing wave. Independent variables included those described above[10] plus the following:

- Wave at which the household responded to the screener (first, second, third, or CATI)[11]
- Incentive level ($0, $5, $10, $15, or $20)
- Topical survey questionnaire type (ECPP or PFI)
- Topical survey form type (original or revised question wording)
- Envelope type ("plain" or "distinctive")

The second set of analyses included only households who responded to the screener after either the first or second mailing wave and examined the effect of receiving an incentive with the initial topical mailing or the second nonresponse follow-up.

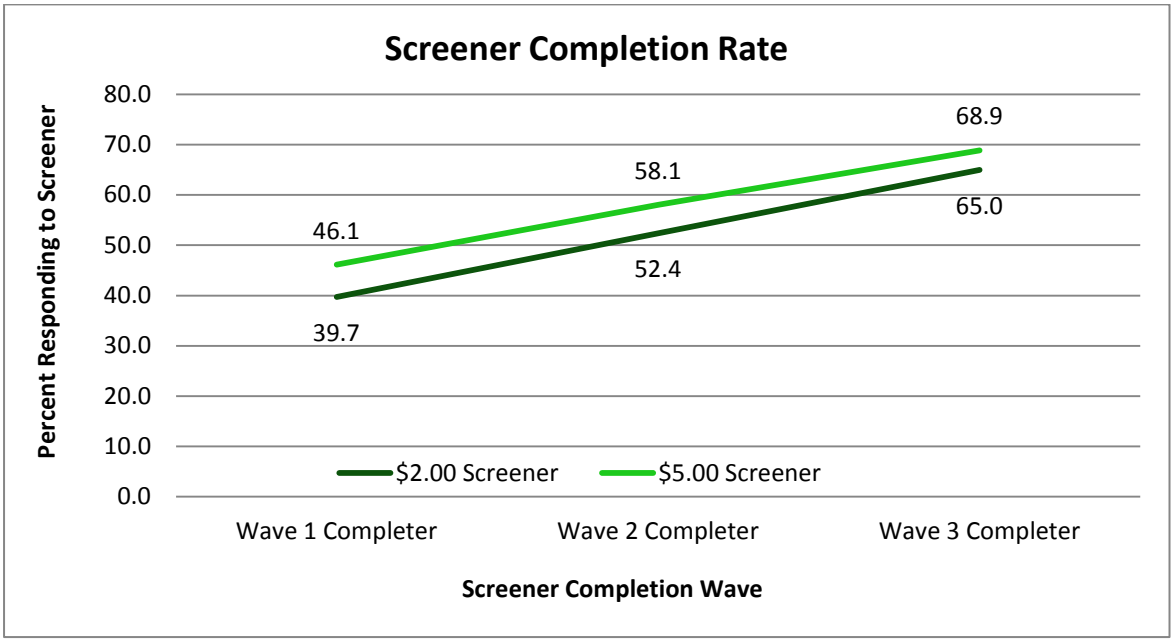
Results: Screener Response Rates

Controlling for all other treatments at the screener level, logistic regression results in table A1 show that receiving the $5 incentive was a strong predictor of overall response (odds of response were about 1.2 times higher for this group compared to those receiving the $2 incentive) (table A1). Looking at the odds of response to each screener mailing independently, the odds of response were higher among those who received a $5 incentive compared to those who received a $2 incentive at both the first and second screener mailings. Specifically, the odds of responding to the first screener mailing were approximately 1.3 times higher among those who received $5 compared to $2 and the odds of responding to the second screener mailing were about 1.1 times higher among those who received $5 compared to those who received $2. The odds of responding to the third screener mailing were not measurably different for households who received $5 compared to those who received $2 (table A1). As illustrated in Figure 2, the higher odds of responding translated into an overall final screener response rate that was approximately 4 percentage points higher for the households that received $5 than for the households that received $2.

[10] Experimental treatments specific only to the 2nd or 3rd screener mailings were not included in the topical response propensity analysis.

[11] Limited telephone nonresponse follow-up was used at the screener phases of the 2011 Field Test. A subsample of households with a telephone number available (through a reverse match to the address provided by the address vendor) was included in telephone nonresponse follow-up effort. Households that responded to the screener through the telephone operation and had an eligible child were mailed a topical questionnaire.

Figure 2. Screener Completion Response Rate, by Response Wave and Incentive Level



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES): 2011 Field Test.

Furthermore, while the overall difference in response rate between the households receiving $5 and those receiving $2 was approximately 4 percentage points, this difference was more pronounced for non-White racial and ethnic groups. The difference in response rates between households receiving a $2 incentive compared to a $5 incentive for households identified by the sampling frame as being Black/African American was 8 percentage points, 4.3 percentage points for Hispanics, and 7.3 percentage points for Asian/Pacific Islanders (table A2). White households also experienced an increase in screener response rates when shifting from a $2 to a $5 incentive, but the difference was 3.3 percentage points. This indicates that in addition to increasing the overall response rate on the screener, the $5 incentive brought in more racial/ethnic minority households than the $2 incentive (table A2).

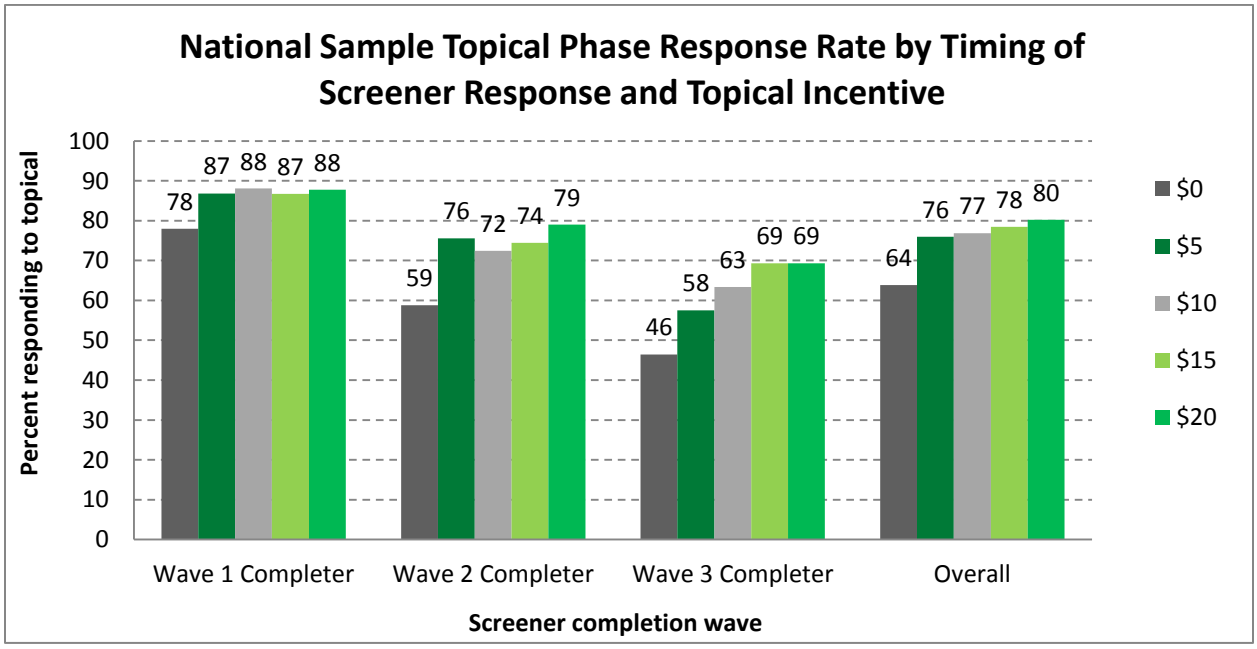
Results: Topical Questionnaire Response Rates

Households eligible to receive the topical questionnaire and selected to be included in Group 1 received a prepaid incentive of $0, $5, $10, $15, or $20. The initial set of logistic regression results indicated that, controlling for other experimental treatments (including those fielded as part of the screener mailing), households who received any incentive with the first mailing were significantly more likely to respond to the topical questionnaire than households that did not receive money in the first mailing. Additionally, the regression analysis indicated that one of the strongest predictors of response to the topical questionnaire was the wave at which the household responded to the screener; the earlier households responded to the screener, the more likely they were to respond to the topical questionnaire (table A3 in appendix A presents complete logistic regression results). Initial analyses also suggested that there was an interaction between the wave in which a household responded to the screener and the effect of higher incentives on response rates. Therefore, logistic models were run separately by screener response wave to predict overall topical questionnaire response.[12] These models suggested that prepaid incentives of more than $5 related to increased response rates primarily for late screener responders (i.e., those households responding to the

[12] Unlike at the screener level, consistent differences were not seen between the predictors of overall topical response and response at each topical mailing wave separately. As a result this paper presents only the results for the overall models at the topical level.

screener after the second nonresponse follow-up). Specifically, among wave 3 screener completers, the odds of topical response were higher for those who received $15 or $20 in the initial topical mailing compared to those who received $5 in the initial topical mailing. There was no measureable difference in the odds of topical response between those who received $5 in the initial topical mailing and those who received a $15 incentive or a $20 incentive among wave 1 screener respondents or wave 2 screener respondents. Additionally, for no screener completion group did $20 relate to a measurably higher response rate than $15. Figure 3 presents the bivariate distribution of response rates by screener completion wave and topical incentive level.

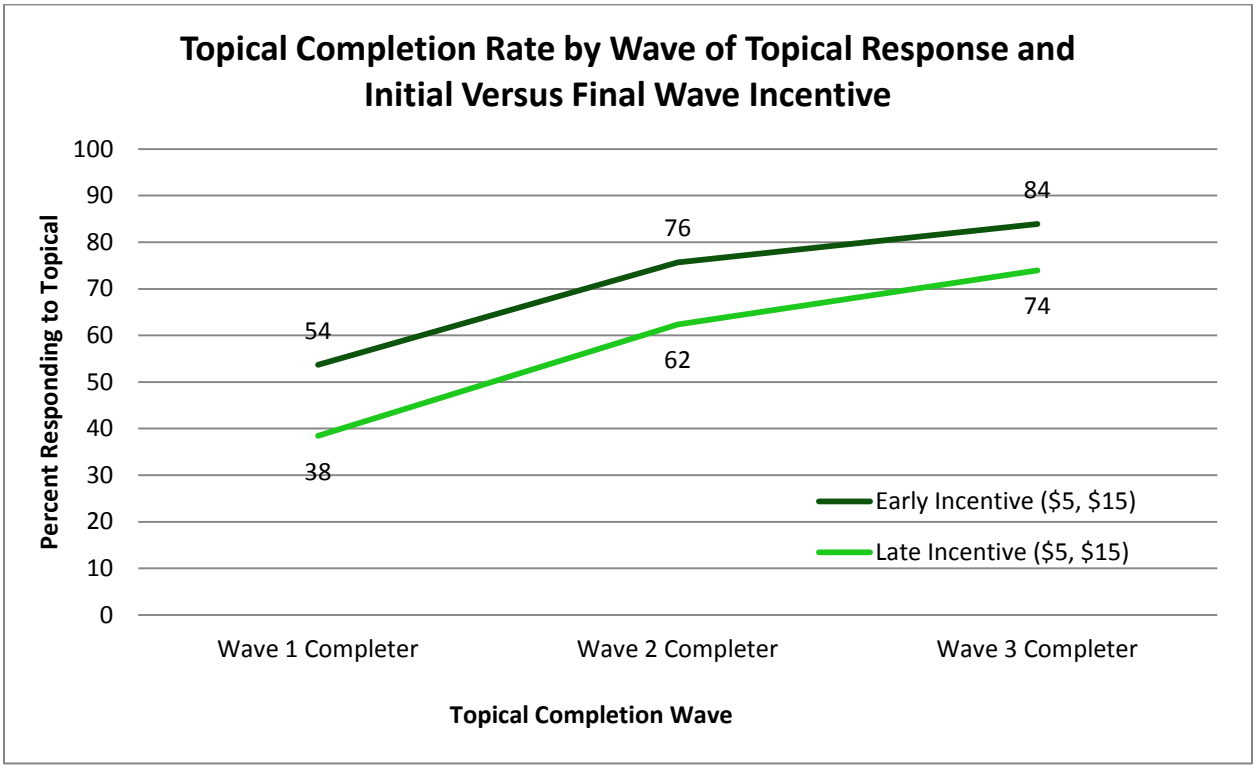
Figure 3



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES): 2011 Field Test.

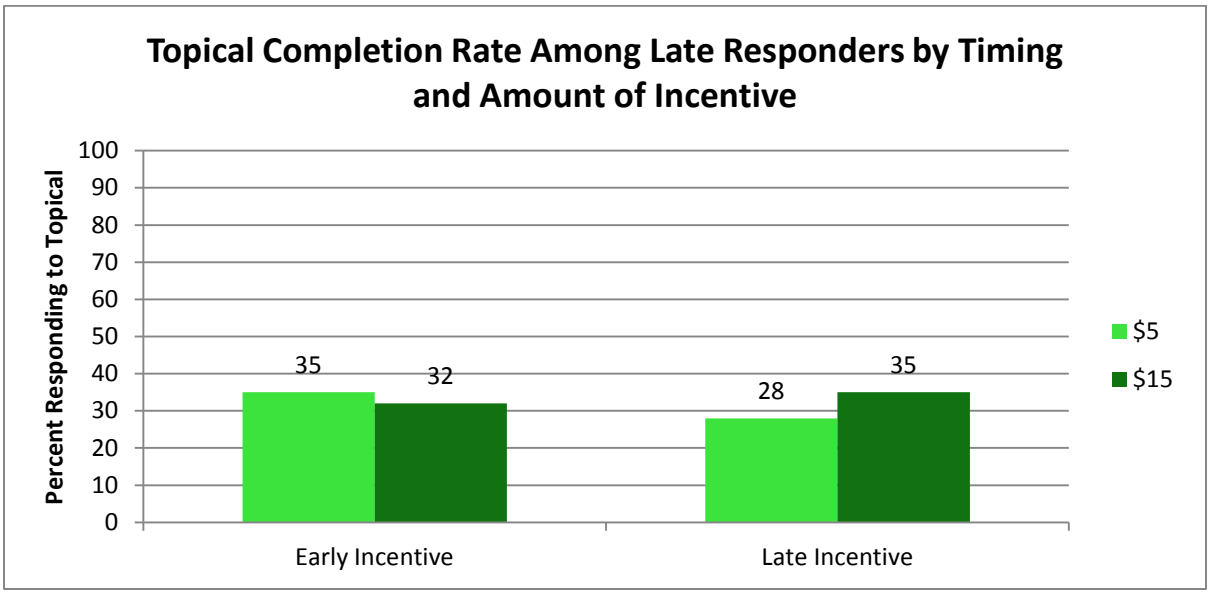
In order to evaluate the impact of the "late" incentive (mailed with the second nonresponse follow-up mailing) compared to the "early" incentive (sent with the initial topical mailing), analyses included eligible households that had responded to the screener in the first or second mailing wave. The analysis of the relationship between incentive timing and response rates excluded cases who responded to the screener after the third mailing wave. Since these cases received the final screener mailing via special delivery mail (either USPS Priority Mail of FedEx), they were ineligible to receive the first topical mailing via Priority Mail. Due to the experimental design of the Field Test (see Figure 1), these cases were also ineligible to receive the "late" topical incentive and, therefore, could not be included in these analyses. The results of these analyses demonstrate that waiting to send the incentive only to the later responding households is associated with a lower final response rate compared to sending an incentive of the same monetary value with the first topical mailing. Figure 4 compares the response rate at each topical wave for the group of households that received either $5 or $15 at the initial mailing with those that received $5 or $15 as part of the second nonresponse follow-up. The group that received the early incentive had a final overall response rate ten percentage points higher than those who received no incentive in the first two mailings and then either $5 or $15 with the third mailing. Figure 5 illustrates the differences by incentive level for those cases who had not yet responded by the third topical mailing. This graph shows that sending $5 early is just as effective at getting households who have not responded by the third mailing to respond as sending $15 just to this group in the final mailing.

Figure 4



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES): 2011 Field Test.

Figure 5



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES): 2011 Field Test.

Summary

Consistent with previous research, sending prepaid incentives with a mail questionnaire increases the likelihood of response and, in general, higher incentive levels are associated with higher response rates. However, our analysis highlights that this is a nuanced relationship in which the dual-phase study design complicates the relationship between incentive level and response rate. While the impact of more money at the screener level was directly related to increased screener response and earlier screener response, the effectiveness of the incentive at the topical level was moderated by the timing of the screener response. This suggests that for two-phase designs, basing the second phase incentive amount on the rapidity of response at the first phase could help minimize cost without sacrificing response rate. More research is needed to perfect the balance between first and second phase incentive amounts, but for the NHES, these data suggest that between $5 and $10 is an appropriate amount to send with the second phase questionnaire for early screener respondents, while those who responded to the final screener mailing would need to receive at least $10 to $15 in order to maximize response at the second phase (though there were response rate increases associated with $5 and $10 second phase incentives compared to no incentive for this group as well). These data also suggest that increasing the incentive to $20 has little additional effect on response rate.

An important extension of this work is to examine whether the patterns observed in the national sample are mirrored in different demographic subgroups. As noted above, the impact of the $5 screener incentive was larger for households identified by the address vendor as Black/African American than for the sample as a whole. Given that other demographic groups such as Hispanics, households with lower incomes, or households with lower education levels tend to have consistently lower response rates, it is possible that increased incentives may have a larger impact on response rates among these groups than among easier to reach populations. It is also not known whether the relationship between the timing of the screener response and the optimal incentive level for the topical questionnaire is the same among different demographic groups. These analyses would be enhanced by a subgroup analysis based on demographic data appended to the ABS sampling frame.

References

Armstrong, J.S. (1975). Monetary Incentives in Mail Surveys. *Public Opinion Quarterly*, *39(1)*: 111-116.

Blumberg, S.J. and Luke, J.V. (2010). Wireless substitution: Early release of estimates from the National Health Interview Survey, January-June 2010. National Center for Health Statistics. http://www.cdc.gov/nchs/nhis.htm.

Brick, J.M., Williams, D., and Montaquila, J.M. (2011). Address-Based Sampling for Subpopulation Surveys. *Public Opinion Quarterly*, *75(3):* 409-428.

Fahimi, M., Kulp, D., and Brick, J.M. (2009). A Reassessment of List-Assisted RDD Methodology. *Public Opinion Quarterly*, *73(4)*: 751-760.

Church, A.H. (1993). Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta-Analysis. *Public Opinion Quarterly, 57(1)*: 62-79.

Lesser, V.M., Dillman, D.A., Carlson, J., Lorenz, F., Mason, R., and Willits, F. (2001) Quantifying the Influence of Incentives on Mail Survey Response Rates and Nonresponse Bias. Paper presented at the annual meeting of the American Statistical Association, Atlanta, GA.

Petrolia, D.R. and Bhattacharjee, S. (2009). Revisiting Incentive Effects: Evidence from a Random-Sample Mail Survey on Consumer Preferences for Fuel Ethanol. *Public Opinion Quarterly, 73(3)*: 537-550.

Van de Kerckhove, W., Montaquila, J.M., Carver, P.R., and Brick, J.M. (2009). *An Evaluation of Bias in the 2007 National Households Education Surveys Program: Results from a Special Data Collection Effort* (2009-029). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Appendix A: Tables

Table A1: Logistic regression results predicting NHES: 2011 Field Test screener response, by experimental treatment

	Odds Ratios							
	Overall		Response at Wave 1		Response at Wave 2		Response at Wave 3	
	N = 36,600	p value	N = 36,600	p value	N = 21,000	p value	N = 16,500	p value
Screener Experiments								
Incentive								
$2.00	--	--	--	--	--	--	--	--
$5.00	1.233	0.0001 *	1.309	0.0001 *	1.107	0.0049 *	1.001	0.9839
Prenotice letter								
Yes	1.062	0.3178	1.243	0.0001 *	0.882	0.1964	0.874	0.1872
No	--	--	--	--	--	--	--	--
Magnet								
Yes	1.063	0.2545	1.023	0.6564	1.022	0.7901	1.118	0.1845
No	--	--	--	--	--	--	--	--
Name								
Yes	0.955	0.0540	0.967	0.1405	0.918	0.0197 *	1.009	0.8091
No	--	--	--	--	--	--	--	--
Phone								
Yes	0.933	0.1091	0.961	0.2958	0.849	0.0083 *	1.027	0.6934
No	--	--	--	--	--	--	--	--
Screener Type (1st mailing)								
Screenout	--	--	--	--	--	--	--	--
Engaging	1.022	0.3602	0.955	0.0440 *	1.025	0.4934	1.097	0.0179
English/Spanish	0.803	0.0949	0.981	0.8272	0.925	0.6852	0.818	0.3201 *
Screener Type (1st & 2nd/3rd mailing)								
Screenout	--	--	†		--	--	--	--
Engaging	0.889	0.0001 *	†		0.937	0.0767	0.767	0.0001
English/Spanish	1.196	0.0677	†		1.174	0.2546	1.073	0.6345 *
Delivery type (3rd mailing)								
Priority Mail	--	--	†		†		--	--
FedEx	1.146	0.0001 *	†		†		1.363	0.0001 *
Non-FedEx deliverable	0.687	0.0001 *	†		†		0.823	0.0048 *

-- Reference Category

* Indicates odds ratio at $p < .05$

NOTE: Sample sizes were rounded to the nearest 50.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES): 2011 Field Test.

Table A2: NHES: 2011 Field Test screener response rates, by household race/ethnicity

Race/Ethnicity	Response Rate			Racial/Ethnic Distribution		
	$2 Completers	$5 Completers	Marginal Change	$2 Completers	$5 Completers	Marginal Change
Total	65	68.9	3.9	100	100	--
White	72.6	75.9	3.3	48.5	47.6	-0.9
Black, African American	57.1	65.1	8	5.2	5.8	0.6
Asian, Hawaiian, Pacific Islander	70.3	77.6	7.3	2.4	2.4	0
Hispanic	54	58.3	4.3	5.2	5.4	0.2
American Indian	66.7	74.1	7.4	0.2	0.2	0
Other, Unknown, Not provided	59.5	63.3	3.8	38.5	38.6	0.1

-- Not applicable.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES): 2011 Field Test.

Table A3: Logistic regression results predicting NHES: 2011 Field Test topical response, by experimental treatment and screener completion wave

	Odds Ratios							
			Screener Completion Wave					
	Overall		Wave 1 completer		Wave 2 completer		Wave 3 completer	
	N = 4850	p value	N =2250	p value	N =950	p value	N =1500	p value
Variable								
Screener completion wave								
Wave 1	--	--						
Wave 2	0.420	0.0001 *	†	†	†	†	†	†
Wave 3	0.251	0.0001 *	†	†	†	†	†	†
CATI	0.076	0.0001 *	†	†	†	†	†	†
Screener Experiments								
Incentive								
$2.00	--	--	--	--	--	--	--	--
$5.00	0.795	0.0018 *	0.755	0.0290 *	0.950	0.7481	0.858	0.1927
Prenotice letter								
Yes	1.011	0.9555	1.165	0.6200	0.836	0.6782	1.019	0.9519
No	--	--	--	--	--	--	--	--
Magnet								
Yes	1.279	0.1563	1.208	0.5383	2.718	0.0281 *	1.010	0.9695
No	--	--	--	--	--	--	--	--
Name								
Yes	1.209	0.0106 *	1.145	0.2911	1.579	0.0047 *	1.145	0.2570
No	--	--	--	--	--	--	--	--
Phone								
Yes	0.928	0.5623	0.682	0.1522	1.549	0.1233	0.835	0.3975
No	--	--	--	--	--	--	--	--
Screener Type (1st mailing)								
Screenout	--	--	--	--	--	--	--	--
Engaging	1.004	0.9537	0.980	0.8764	1.001	0.9953	1.035	0.7757
English/Spanish	1.102	0.7269	1.230	0.6771	0.413	0.4752	1.151	0.8099
Screener Type (1st & 2nd/3rd mailing)								
Screenout	†	†	†	†	--	--	--	--
Engaging	†	†	†	†	0.870	0.3826	0.894	0.3500
English/Spanish	†	†	†	†	2.003	0.5274	0.880	0.7606
Delivery type (3rd mailing)								
Priority Mail	†	†	†	†	†	†	--	--
FedEx	†	†	†	†	†	†	1.004	0.9737
Non-FedEx deliverable	†	†	†	†	†	†	0.759	0.2387

Table A3 continued: Logistic regression results predicting NHES: 2011 Field Test topical response, by experimental treatment and screener completion wave

	Odds Ratios							
			Screener Completion Wave					
	Overall		Wave 1 completer		Wave 2 completer		Wave 3 completer	
	N = 4850	p value	N =2250	p value	N =950	p value	N =1500	p value
Topical Experiments								
Incentive level								
$0	0.569	0.0001 *	0.520	0.0003 *	0.443	0.0003 *	0.635	0.0078 *
$5	--	--	--	--	--	--	--	--
$10	1.140	0.2369	1.092	0.6646	0.806	0.3872	1.279	0.1533
$15	1.264	0.0389 *	0.963	0.8516	0.908	0.6984	1.708	0.0027 *
$20	1.352	0.0069 *	1.063	0.7584	1.195	0.4737	1.662	0.0041 *
Questionnaire type								
ECPP	1.023	0.7698	1.043	0.7609	1.116	0.5006	0.989	0.9302
PFI	--	--	--	--	--	--	--	--
Survey form								
Mainline	--	--	--	--	--	--	--	--
Alternate	1.059	0.4543	1.039	0.7760	1.025	0.8777	1.142	0.2696
Form Language								
English	--	--	--	--	--	--	--	--
Spanish	1.086	0.7945	‡	‡	‡	‡	0.948	0.9349
Envelope type								
Regular	--	--	--	--	--	--	--	--
Distinct	1.132	0.0757	1.089	0.4837	0.961	0.7901	1.307	0.0151 *

-- Reference Category
* Indicates odds ratio at $p < .05$
NOTE: Sample sizes were rounded to the nearest 50.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES): 2011 Field Test.