Journal of Official Statistics, Vol. 21, No. 4, 2005, pp. 571–589

# Implications for RDD Design from an Incentive Experiment

*J. Michael Brick[1], Jill Montaquila[1], Mary Collins Hagedorn[1], Shelley Brock Roth[1], and Christopher Chapman[2]*

This article describes a large-scale experiment in the use of monetary incentives in a random digit dial (RDD) telephone survey. The experiment tested providing incentives of different amounts of money and at different stages of the survey. The results show that sending incentives at the refusal conversion stage is about as effective as using them in the initial attempt. Alternative sample designs that incorporate the results of the experiment are evaluated to identify designs that achieve high response rates and are economical. As a result of this evaluation, the design for a future study will use monetary incentives and subsample households that refuse to participate at the refusal conversion stage.

*Key words:* Subsampling refusals; response rates; refusal conversion; prepaid monetary incentives; advance mailing.

## 1. Introduction

Achieving high response rates in sample surveys has become increasingly difficult in recent years, especially in random digit dial (RDD) telephone surveys. Atrostic, Bates, Burt, and Silberstein (2001) show declines in response rates for in-person household surveys conducted by the U.S. Census Bureau. Curtin, Presser, and Singer (2004) show a dramatic decline after 2000 in the response rates of the Survey of Consumer Attitudes (SCA), a repeating national RDD survey conducted by the University of Michigan. These findings are consistent with our experiences in RDD surveys, as described below in more detail for the National Household Education Surveys Program (NHES). A major concern is that if response rates continue to decline, the potential for bias due to nonresponse may imperil valid inferences from these surveys.

Groves and Couper (1998) describe a host of factors that affect response rates in household surveys. For RDD surveys, the most important of these factors are the study characteristics (burden, salience, sponsorship, and content), the general survey environment (call screening technology, level of telemarketing and other calls to households, and the economy), and the survey methodology (use of advance letters, refusal conversion attempts, and monetary incentives). In this article, we describe the results of a large-scale experiment

[1] Westat, 1650 Research Boulevard, Rockville, MD 20850-3195, U.S.A.
[2] National Center for Education Statistics, Institute of Education Sciences, 1990 K Street NW, Washington, DC 20006, U.S.A.
The views expressed in this article are part of ongoing research and analysis and do not necessarily reflect the position of the U.S. Department of Education.

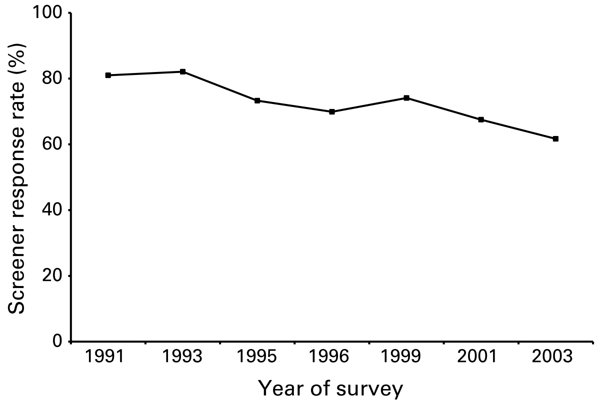
conducted as part of the 2003 survey administration (NHES 2003) to evaluate the effects monetary incentives have on response rates in a federally sponsored RDD survey. We then use these findings to study alternative sample designs that have higher response rates while still being cost-effective.

### *1.1. National Household Education Surveys (NHES)*

NHES is a program of repeating RDD surveys developed by the National Center for Education Statistics (NCES) to collect information on important educational issues through telephone surveys of households in the United States. NCES is part of the U.S. Department of Education's Institute of Education Sciences. Westat conducted these surveys in 1991, 1993, 1995, 1996, 1999, 2001, and 2003. An experiment was conducted in NHES 2003 to determine if monetary incentives should be used in future surveys and to identify a cost-effective level for such incentives. NHES 2003 utilized a list-assisted RDD sample that oversampled areas with high concentrations of minorities and mailable cases (telephone numbers for which addresses could be obtained, as discussed in Brick, Judkins, Montaquila, and Morganstein 2002). The instruments consisted of a screener, in which household members were enumerated and data were collected to sample eligible persons, and two extended interviews on specific educational topics, the Parent and Family Involvement in Education Survey (PFI) and the Adult Education for Work-Related Reasons Survey (AEWR). On average the screener took about 3.5 minutes to complete and an extended interview took about 20 minutes. The monetary incentive experiment focused on the screener, because gaining participation at this initial stage is the most challenging response component of RDD surveys. There was also an incentive experiment for persons sampled for AEWR, but we do not discuss those findings here (for details see Brick, Hagedorn, Montaquila, Roth, and Chapman forthcoming).

NHES 2003 was conducted between January and April 2003. Despite the short field period, NHES uses various procedures in an effort to maximize contact and response rates. Advance mailings and refusal conversion mailings are sent to sampled telephone numbers that can be matched to addresses. The call scheduling algorithm ensures that the multiple call attempts to noncontact telephone numbers are made at different times of the day and on different days of the week. Two rounds of refusal conversion are attempted for nonhostile refusal cases. When answering machines are reached, a message is left but the residential status of the case is determined only through human contact. In NHES 2003, no fewer than 20 call attempts were made before a screener was classified as a "maximum call."

Figure 1 shows that the screener response rates in NHES have been declining over time. These screener response rates are weighted by the inverses of the selection probabilities and are consistent with the response rate definition (RR3) in AAPOR (2004). Screener response rates exceeded 80 percent in the first two administrations, but fell in 1995 and 1996 due to changes in length and content of the screening interview (Brick and Collins 1997). Beginning in 1999, the screener was made more consistent with the earlier surveys and the screener response rate increased somewhat, but the screener response rates for 2001 and 2003 showed sharp declines. This decrease in screener response rates occurred despite the gradual introduction of methods to encourage participation such as mailing letters, leaving answering machine messages, making more call attempts, increasing




NOTE: All response rates were calculated using the business office method. Beginning in 1999 the official response rates are computed using the survival method and differ from those given above. (See Brick, Montaquila, and Scheuren 2002 for details on methods for calculating response rates in RDD surveys.)

*Fig. 1. Weighted screener response rates for NHES: 1991–2003*

the number of refusal conversion attempts, and developing special interviewer training sessions (Brick et al. 1997; Collins et al. 1997; Vaden-Kiernan et al. 1997; Nolin et al. 2000; Nolin et al. 2004; Hagedorn et al. 2004). In 2003, monetary incentives were used for the first time in the NHES program.

### *1.2. Survey mailings and incentives*

Before discussing incentives, it is worth noting that the NHES standard practice since 1996 has been to send an advance letter prior to any call attempts to those telephone numbers for which an address can be found using commercial vendors that match telephone and household addresses. This advance mailing was effective in raising screener response rates in NHES 2001 by about 5 percentage points for the mailable numbers. This result confirmed earlier findings from an NHES experiment conducted in NHES 1996 (Brick and Collins 1997) and is consistent with the findings of Goldstein and Jennings (2002) and Dillman (2000). We suspect the use of official U.S. Department of Education envelopes and stationery may contribute to the effectiveness of the mailings and may partially explain the difference from the nonsignificant result of advance mailings reported in Singer, Van Hoewyk, and Maher (2000).

Since 1999, follow-up letters have been sent to households that initially refuse to participate in NHES collections. Letters are sent prior to making a refusal conversion call and are intended to persuade respondents to change their minds and complete the screener. Unfortunately, literature on the effectiveness of mailing strategies for RDD refusal conversion is limited, and this strategy has not been systematically tested in NHES. One analysis by Cantor, Cunningham, Triplett, and Steinbach (2003) regarding experiences with the National Survey of America's Families found that sending refusal conversion letters by Federal Express (courier mail) significantly increases the refusal conversion rate by capturing the attention of potential respondents to a greater extent than a First-Class letter.

Incentives have been used as a way of improving response rates in surveys for decades; yet the theory supporting the use of incentives is still not fully established. One theory proposed to explain the effectiveness of incentives is social exchange (Dillman 2000), which suggests that giving an incentive to a household is considered as an act that the household should reciprocate. Responding to the survey is the reciprocal act. Monetary exchange theory is also sometimes invoked to explain responses when incentives are used. This theory suggests the household may be more likely to respond to the survey because the incentive is payment for this response. Groves, Singer, and Corning (2000) suggested a more complex model in which incentives act to increase response rates by substituting for some study characteristics (e.g., burden, salience, sponsorship, or content) that are otherwise lacking. Singer (2002) summarized these theories and reviewed the intended and unintended consequences of using incentives in surveys. Some of this literature is reviewed briefly below, focusing on issues relevant to incentives in telephone surveys.

A meta-analysis by Church (1993) shows that the payment of incentives in mail surveys is one method that consistently results in higher response rates. The studies of incentives nearly always find that prepaid incentives (those offered without requiring the recipient to respond) are more effective than promised incentives that are contingent on response (e.g., Berk, Mathiowetz, Ward, and White 1987; Church 1993). This literature also shows that monetary incentives are more effective than nonmonetary ones. Trussell and Lavrakas (2004) found that for households that had previously agreed (in a telephone contact) to complete a mail survey, a \$5 prepaid incentive resulted in a significantly higher cooperation rate than a \$1, \$2, \$3, or \$4 prepaid incentive. The evidence on the utility of incentives in RDD surveys is more limited and more recent. Singer, Van Hoewyk, Gebler, Raghunathan, and McGonagle (1999) and Gelman, Stevens, and Chan (2003) agreed that prepaid incentives increase response rates, but came to different conclusions about the effectiveness of promised incentives in RDD surveys. Cantor et al. (1997) also found that prepaid incentives in the initial stages of an RDD survey were effective.

Another way of using incentives is to provide an incentive only to those units that refuse to complete the interview in the initial attempt. Martin, Abreu, and Winters (2001) found that response rates increased when prepaid monetary incentives were used to convert refusers in a later wave of a panel survey. Cantor et al. (2003) found that prepaying incentives at the refusal conversion stage was effective in an RDD survey conducted at Westat. Since these results are from a cross-sectional RDD survey, they are much more relevant for our purposes, but the lack of literature on the subject highlights the need for studying this approach in the NHES experiment.

## 2. NHES 2003 Experiment

The overall objective of the experiment was to determine if there were economical alternatives for improving NHES response rates. Because it had already been determined that advance mailing was effective in increasing response rates, every sampled household for which an address could be identified was sent a First-Class letter on official stationery in a U.S. Department of Education envelope. The experiment did vary the amount of money in the advance letter, the type of mail in the refusal conversion mailing (First-Class Mail® or Priority Mail® in the U.S. Postal Service Priority Mail envelopes), and the amount of money in the refusal mailing. Table 1 shows the ten experimental conditions, with the sample size for each condition. The sample sizes in the table correspond to the numbers of telephone numbers assigned to each group for the initial treatment. A total of 59,365 mailable numbers were included in the experiment. Experimental Group 4 was very similar to the treatments used in the previous NHES surveys and Group 6 was expected to be the most cost-effective; thus, larger sample sizes were assigned to these two groups to increase the power to detect differences. The survey also included a sample of 50,435 cases that did not have mailable addresses and could not be included in the experiment. The results for the cases without mailing addresses are generally omitted from this analysis, but are given in Brick et al. (forthcoming). However, for comparison purposes, the response rate for cases that did not have mailable addresses is given in the discussion of response rates in Section 2.1.

Group 1 is essentially a control group with no incentive treatments, but every case in this group was sent a short color brochure describing NHES. The brochure was not included in any other advance mailings. Among all experimental groups, Group 1 had the lowest initial cooperation rate (shown later), but the difference between Group 1 and the other groups with no incentive in the initial mailing was not statistically significant. Since the inclusion of the brochure did not increase the initial cooperation rate when compared to other groups with no monetary incentive in the initial mailing, we ignore the brochure in the analysis.

The cases were randomly assigned to the groups and the advance letters were mailed. If the household refused to participate, a refusal conversion letter was sent with the appropriate incentives; all refusal mailings included the study brochure. If a refusal was hostile (profane or abusive), no further efforts were made to contact the household. In addition, a small number of cases that directly contacted NCES to refuse received no further contact. Thus, for analyses of the treatments' effects on refusal conversion, the sample sizes in each group are considerably smaller than shown in Table 1 (see Table 2.)

If the household refused the screener a second time, it was subject to another refusal conversion call. To attempt to boost response rates, a random subsample of 75 percent of the households that had not received a Priority Mail letter previously (Groups 1, 2, 3, 6, 9, and 10) were sent a Priority Mail letter prior to the second refusal conversion call.

*Table 1. Sample sizes for NHES:2003 screener experiment*

| Experimental group | Advance letter | Refusal conversion | Sample size |
|---|---|---|---|
| 1. $0/First-Class $0 | First-Class – no incentive | First-Class – no incentive | 5,765 |
| 2. $0/First-Class $2 | First-Class – no incentive | First-Class – $2 incentive | 5,700 |
| 3. $0/First-Class $5 | First-Class – no incentive | First-Class – $5 incentive | 5,700 |
| 4. $0/Priority $0 | First-Class – no incentive | Priority – no incentive | 6,850 |
| 5. $0/Priority $2 | First-Class – no incentive | Priority – $2 incentive | 5,700 |
| 6. $2/First-Class $0 | First-Class – $2 incentive | First-Class – no incentive | 6,850 |
| 7. $2/Priority $0 | First-Class – $2 incentive | Priority – no incentive | 5,700 |
| 8. $2/Priority $2 | First-Class – $2 incentive | Priority – $2 incentive | 5,700 |
| 9. $5/First-Class $0 | First-Class – $5 incentive | First-Class – no incentive | 5,700 |
| 10. $2/First-Class $2 | First-Class – $2 incentive | First-Class – $2 incentive | 5,700 |

Table 2. *Screener initial cooperation and first refusal conversion rates, by incentive group*

| Incentive group | Initial cooperation | | First refusal conversion | |
|---|---|---|---|---|
| | Sample size* | Rate (s.e.) | Sample size* | Rate (s.e.) |
| 1. \$0/First-Class \$0 | 3,794 | 51.2 (0.8) | 2,166 | 30.7 (1.2) |
| 2. \$0/First-Class \$2 | 3,762 | 52.9 (0.8) | 2,088 | 36.3 (1.1) |
| 3. \$0/First-Class \$5 | 3,761 | 54.6 (0.7) | 2,023 | 38.8 (1.2) |
| 4. \$0/Priority \$0 | 4,530 | 53.4 (0.7) | 2,486 | 31.8 (1.1) |
| 5. \$0/Priority \$2 | 3,761 | 54.8 (0.8) | 2,007 | 35.1 (1.1) |
| 6. \$2/First-Class \$0 | 4,503 | 59.1 (0.7) | 2,188 | 29.5 (1.1) |
| 7. \$2/Priority \$0 | 3,803 | 57.9 (0.7) | 1,850 | 35.1 (1.3) |
| 8. \$2/Priority \$2 | 3,777 | 57.6 (0.8) | 1,839 | 36.7 (1.1) |
| 9. \$5/First-Class \$0 | 3,799 | 61.1 (0.8) | 1,734 | 29.9 (1.3) |
| 10. \$2/First-Class \$2 | 3,808 | 57.9 (0.7) | 1,878 | 34.9 (1.1) |

*The sample size for the initial cooperation rate is the number of telephone numbers with completed screeners plus the number of residential telephone numbers without completed screeners that ever had a refusal result on a screener attempt. The sample size for the first refusal conversion rate is the number of residences that previously refused the screener.

### 2.1. *Effects on response rates*

The key results of the experiments are measures of the effectiveness of the various treatments for improving screener response rates. In addition to the final screener response rates, we present the initial cooperation rates, the first refusal conversion rates, and the second refusal conversion rates to evaluate the effectiveness of the treatments at different stages. The initial cooperation and refusal conversion rates are the percentage of possible interviews completed at that stage of the survey, where the denominator is based on the set of cases that completed or refused to participate at that stage. The denominator for the final screener response rates is all cases, including those that did not complete the interview for other reasons such as language barriers. The denominator includes 19.7 percent of the sampled numbers that were never answered and that were assumed to be residential, based on survival method calculations described by Brick, Montaquila, and Scheuren (2002).

The first step of the analysis was to verify that the percentages of the experimental cases having final dispositions that should not have been affected by the treatments (e.g., the percent nonresidential and the percent never answered) were consistent over the groups. The percentage distributions of these dispositions for all of the experimental groups were within sampling error. Below, we tabulate the screener response rates by the experimental groups and then use logistic regression to analyze the experiments. All of the tables and regression analyses are weighted to account for the differential probabilities of selection of the sampled numbers.

Table 2 summarizes the initial cooperation and first refusal conversion rates for the 10 treatment groups. All estimated sampling errors, including the regression estimates presented later, were computed using a jackknife replication approach (JK1) in WesVar® to account for the complex sample design. The sampling error for the estimate is given in parentheses following the estimate. The initial cooperation rates exhibit the type of pattern expected, with rates in those groups with no incentive lower than those with \$2 or \$5. The initial cooperation rate for all cases with no incentive in the advance mail is 53.4 percent (s.e. = 0.3); for all cases with the \$2 incentive, it is 58.2 percent (s.e. = 0.4); and it is 61.1 percent (s.e. = 0.8) for Group 9 which is the only one with the \$5 advance incentive. Clearly, incentives improve cooperation and the \$5 incentive is more effective in bolstering initial cooperation rates than the \$2 one. But, consistent with the findings of Trussell and Lavrakas (2004), the relative effectiveness–the percentage point increase in the initial cooperation rate per dollar when compared to no incentive–is 2.4 percentage points for the \$2 incentive and 1.5 percentage points for the \$5 incentive.

The two refusal conversion treatments, amount of incentive and type of mail, were used only if a household refused the screener. While refusal conversion is effective in RDD surveys, most households that eventually complete the screener never refuse. In NHES 2003, 70 percent of the completed cases never refused, 23 percent came from first refusal conversion attempts, and only 7 percent came from second conversion attempts.

One treatment considered at this stage is that of sending a letter by First-Class versus Priority Mail prior to the first refusal conversion call attempt. One hypothesis is that households will distinguish Priority Mail from other types of mail, and this might encourage households to respond. To isolate the effect of using Priority Mail for refusal conversion, we paired groups that had the same advance and refusal incentive treatments, but with different first refusal mailing conditions. The four pairs of groups are: Groups 1 and 4, Groups 2 and 5, Groups 6 and 7, and Groups 8 and 10. As shown in Table 2, apart from the Groups 6 and 7 pair, no statistically significant differences in screener response rates were observed, and no observed difference was larger than 2 percentage points. The comparison of Groups 6 and 7 does show a significant difference in the first refusal conversion rate between First-Class and Priority Mail. However, using Priority Mail in the initial refusal conversion letter did not generally result in substantial increases in screener response rates, holding the other conditions constant.

This absence of a statistically significant difference for Priority Mail for the initial refusal conversion letter differs from that reported by Cantor et al. (2003), who found that Federal Express refusal conversion mailings resulted in statistically significant improvements for the National Survey of America's Families (NSAF) in 1997, even though these findings were based on smaller samples. One possible explanation is that NHES 2003 is a Federal Government survey that uses an official government First-Class envelope, whereas the NSAF is not. The official letters may provide more inducement to respond than a mailing from a nonfederal organization. A second consideration is that the effect of Priority Mail might not be the same as the effect of Federal Express for refusal conversion mailings. Another difference is that the percentage of telephone numbers with mailable addresses increased substantially due to improvements in commercial services for matching telephone numbers to addresses between 1997 and 2003. Thus, the types of households that were treated in 1997 may have been different from those that were treated in 2003.

Table 2 also shows that the first refusal conversion rates vary across the treatments from 29 percent to 39 percent. If the use of Priority Mail at this stage is ignored, the ten incentive groups can be classified into three groups by the amount of incentive in the conversion letter, as was done for the initial cooperation rate analysis. The first refusal conversion rate for the \$0 group is 31.4 percent (s.e. = 1.2); for the \$2 group, it is 35.7 percent (s.e. = 1.1); and for the \$5 group, it is 38.8 percent (s.e. = 1.2). The groups with \$2 and \$5 conversion incentives have significantly higher first conversion rates, while the difference between the $2 and $5 groups is not statistically significant. While this analysis does not account for the advance incentive treatments, the regression analysis given later supports the conclusion that these differences are primarily the effect of the refusal conversion incentive amount.

The last set of treatments tested the effectiveness of Priority Mail for those cases that refused to participate in the screener twice. A random sample of 75 percent of the households that refused twice and that had no previous priority mailing (Groups 1, 2, 3, 6, 9, and 10) was sent a third mailing by Priority Mail. The remaining 25 percent (like the households in the other groups) were not sent a letter at this stage. Table 3 shows the second refusal conversion rates overall, and by whether the household was sent a Priority Mail letter for each of the six groups treated. The second refusal conversion rates for households sent the Priority Mail letter were 4 to 15 percentage points higher than those not sent a letter. The differences between the Priority Mail subsample and the subsample that was not sent a letter are statistically significant for Groups 1 and 6, but even though the differences for the other groups are not statistically significant, they are large enough to be substantively important (up to 7 percentage points). Because of the variety of other treatments, the overall effect of the use of Priority Mail on the second refusal conversion rates is discussed in the next section.

Thus, contrary to the general finding about the use of Priority Mail at the first refusal conversion, sending a priority letter after a second screener refusal did improve refusal conversion rates. It is also worth noting that the difference in rates at this stage is between those sent a letter by Priority Mail and those sent no mail. The experiment at first refusal conversion compared Priority Mail to First-Class Mail. The data do not indicate whether sending Priority Mail at this second refusal conversion stage would be more effective than sending a First-Class letter. Since Priority Mail is relatively expensive and the second refusal conversion does not account for a large percentage of the completed screeners, it is

*Table 3. Second refusal conversion rates, by incentive group*

| Incentive group | Sample size* | Second refusal conversion rate (s.e.) | | |
|---|---|---|---|---|
| | | Overall | Priority mail | No priority mail |
| 1. $0/First-Class $0 | 1,245 | 21.2 (1.1) | 25.2 (1.3) | 10.5 (1.8) |
| 2. $0/First-Class $2 | 1,062 | 21.5 (1.4) | 23.1 (1.8) | 18.9 (2.9) |
| 3. $0/First-Class $5 | 982 | 21.4 (1.3) | 23.7 (1.5) | 16.8 (2.7) |
| 4. $0/Priority $0 | 1,349 | 14.0 (1.1) | – | – |
| 5. $0/Priority $2 | 1,019 | 18.2 (1.1) | – | – |
| 6. $2/First-Class $0 | 1,228 | 18.7 (1.2) | 20.8 (1.5) | 13.3 (2.0) |
| 7. $2/Priority $0 | 977 | 17.1 (1.3) | – | – |
| 8. $2/Priority $2 | 909 | 16.3 (1.4) | – | – |
| 9. $5/First-Class $0 | 959 | 19.3 (1.4) | 21.1 (1.5) | 15.6 (2.9) |
| 10. $2/First-Class $2 | 970 | 19.8 (1.4) | 21.4 (1.7) | 16.9 (2.8) |

*Sample size is the number of households that refused twice. The sample sizes by use of Priority Mail are Group 1: 915 were sent Priority Mail and 286 were not sent Priority Mail; Group 2: 772 were sent Priority Mail and 247 were not sent Priority Mail; Group 3: 694 were sent Priority Mail and 241 were not sent Priority Mail; Group 6: 877 were sent Priority Mail and 303 were not sent Priority Mail; Group 9: 710 were sent Priority Mail and 214 were not sent Priority Mail; and Group 10: 714 were sent Priority Mail and 215 were not sent Priority Mail.

not clear that Priority Mail at this stage is cost-effective. We discuss this issue in more detail later.

Table 4 summarizes the final screener response rates by incentive group. Since many RDD surveys attempt to convert nonrespondents at most one time, the screener response rates that would have been obtained without a second conversion attempt are also presented. The last column gives the actual screener response rates observed in NHES 2003 for the groups. By comparison, the telephone numbers without addresses, which were not part of the incentive experiment, have a significantly lower screener response rate (49.3%).

### 2.2. *Logistic regression analysis*

To refine the analyses, the initial cooperation and refusal conversion rates were examined using logistic regression, where the outcome for a sampled number was considered a "success" if the household completed the screener at that stage. After examining many variables unrelated to the experiments that were available for all telephone numbers, two good predictors of initial cooperation were identified: region of the country, and metropolitan status. The following is a complete list of the characteristics examined: census region (combined Northeast and South/combined Midwest and West), interview language (English/Spanish), presence of person under 21 in household, home ownership (own/rent/other arrangement), median home value in telephone exchange, median income in telephone exchange, MSA status, minority stratum, percent of households with incomes of $75,000 to $100,000, percent of households with income above $100,000, percent Asian in exchange, percent Black in exchange, percent Hispanic in exchange, percent college graduates in exchange, and percent homeowners in exchange. Table 5 shows the final model for the initial cooperation rates. The model is parameterized so that the last level of each predictor is the reference cell and the parameter for that level is set equal to zero.

*Table 4. Screener response rates after first and second conversion attempts, by incentive group*

| Incentive group | Screener response rate after first refusal conversion[1] (s.e.) | Final screener response rate[2] (s.e.) |
|---|---|---|
| 1. $0/First-Class $0 | 57.9 (0.8) | 64.1 (0.7) |
| 2. $0/First-Class $2 | 61.9 (0.7) | 67.3 (0.7) |
| 3. $0/First-Class $5 | 64.6 (0.8) | 69.5 (0.6) |
| 4. $0/Priority $0 | 59.9 (0.7) | 63.7 (0.7) |
| 5. $0/Priority $2 | 62.4 (0.7) | 66.7 (0.7) |
| 6. $2/First-Class $0 | 63.4 (0.7) | 67.9 (0.7) |
| 7. $2/Priority $0 | 65.0 (0.8) | 68.9 (0.7) |
| 8. $2/Priority $2 | 65.6 (0.7) | 69.1 (0.8) |
| 9. $5/First-Class $0 | 65.3 (0.7) | 69.7 (0.7) |
| 10. $2/First-Class $2 | 65.4 (0.6) | 69.9 (0.6) |

[1] Screeners converted on second refusal conversion attempts are counted as nonrespondents in this computation.
[2] The final screener response rates include the results of second refusal conversion.

Table 5. *Logistic regression estimates for initial cooperation**

| Parameter | Estimate | s.e. | P | Odds ratio |
|---|---|---|---|---|
| Intercept | 0.56 | 0.049 | 0.000 | |
| Region: Northeast or South | −0.20 | 0.020 | 0.000 | 0.82 |
| Metro status: In county in central city | −0.13 | 0.035 | 0.001 | 0.88 |
| Metro status: In county not in central city | −0.13 | 0.031 | 0.000 | 0.87 |
| Metro status: Subcounty of MSA | −0.12 | 0.038 | 0.003 | 0.89 |
| Metro status: MSA in its own county | −0.29 | 0.050 | 0.000 | 0.75 |
| Initial incentive: $0 | −0.37 | 0.034 | 0.000 | 0.69 |
| Initial incentive: $2 | −0.14 | 0.034 | 0.000 | 0.87 |

NOTE: The reference categories are as follows: Region: Midwest or West; Metro status: not MSA; Initial incentive: $5.
*The sample size is the number of telephone numbers with completed screeners plus the number of residential telephone numbers without completed screeners that ever had a refusal result on a screener attempt. The modified likelihood ratio statistic proposed by Estrella (1998) as a goodness of fit measure for logistic regression models is .008.

In Table 5, the estimated regression coefficients for the $0 and $2 incentive amounts indicate that households receiving these treatments have a lower probability of responding initially to the screener than those receiving the $5 incentive. The odds ratios show that compared to a household in the $5 incentive group, a household in the $2 incentive group is 0.87 times as likely to respond initially, and a household in the $0 incentive group is only 0.69 times as likely to respond initially. A contrast test shows that the effect of the $0 incentive is statistically different from the effect of the $2 incentive. These results are consistent with the results on initial cooperation rates given in Table 2.

Next, logistic regression was used to examine the first refusal conversion rates. In this analysis, predictor variables were included for the two first refusal conversion experimental treatments: whether Priority Mail was sent (no/yes) and the refusal incentive amount ($0/$2/$5). A model containing a term for the interaction between the advance incentive and the refusal incentive was considered, but this interaction term was dropped because it was not statistically significant ($p$-value $= 0.53$). Other models were also considered (e.g., including an interaction between Priority Mail and a level of the monetary incentive), but these are not discussed because they added little to the predictive power of the model. Table 6 gives the estimated regression coefficients and odds ratios. There are three important results from this analysis. First, the estimated coefficient for the Priority Mail treatment is not significantly different from zero at first refusal conversion. Second, the monetary incentives at the refusal conversion stage do increase the conversion rates. The $2 incentive results in a higher conversion rate than no incentive, and $5 is more effective than $2. Third, the monetary incentives in the initial mailing do not have a statistically significant effect on the refusal conversion rates. In other words, the only treatment at the first refusal conversion stage that results in higher rates of completing the screener is the amount of money sent prior to the refusal conversion attempt. This analysis supports statements made earlier in the tabular analysis.

The final logistic regression analysis examines the effectiveness of sending Priority Mail at the second refusal conversion stage. Since only second refusal households that had not been sent a Priority Mail in the first refusal conversion stage were treated, the sample

*Table 6. Logistic regression estimates for first refusal conversion**

| Parameter | Estimate | s.e. | *P* | Odds ratio |
|---|---|---|---|---|
| Intercept | −0.21 | 0.102 | 0.041 | |
| Region (1) | −0.08 | 0.036 | 0.032 | 0.92 |
| Metro status (1) | −0.11 | 0.044 | 0.015 | 0.90 |
| Metro status (2) | −0.01 | 0.050 | 0.850 | 0.99 |
| Metro status (3) | −0.03 | 0.038 | 0.416 | 0.97 |
| Metro status (4) | −0.11 | 0.078 | 0.149 | 0.89 |
| Initial incentive ($0) | 0.04 | 0.069 | 0.604 | 1.04 |
| Initial incentive ($2) | 0.06 | 0.070 | 0.411 | 1.06 |
| First refusal Priority Mail (no) | −0.07 | 0.035 | 0.066 | 0.94 |
| Refusal incentive ($0) | −0.37 | 0.063 | 0.000 | 0.69 |
| Refusal incentive ($2) | −0.18 | 0.063 | 0.005 | 0.83 |

*The sample size is 18,797 households that initially refused the screener and either completed or refused the conversion. The Estrella (1998) goodness of fit statistic is .005.

for this analysis only includes Groups 1, 2, 3, 6, 9, and 10. The predictor variable for using Priority Mail at the first conversion stage is dropped. Table 7 gives the estimated regression coefficients and odds ratios that show the use of a Priority Mail is effective at this stage over no letter at all. A household that was not sent a Priority Mail letter at this stage was only 0.6 times as likely to respond as a household that was sent a Priority Mail second refusal conversion letter. Interestingly, the incentives at the initial and first refusal conversion stages are not statistically significant, again suggesting that there is little or no carry-over effect from incentives at earlier stages.

### 2.3. *Effects on cost and quality*

In addition to affecting cooperation and response rates, it is possible that the monetary incentives and mailing conditions might affect the cost or quality of the data collection. Data collection costs include the direct cost of mailing to the sampled units and the monetary incentives, plus the standard costs associated with RDD surveys such as

*Table 7. Logistic regression estimates for second refusal conversion**

| Parameter | Estimate | s.e. | *P* | Odds ratio |
|---|---|---|---|---|
| Intercept | −0.85 | 0.172 | 0.000 | |
| Region(1) | −0.04 | 0.067 | 0.599 | 0.96 |
| Metro status (1) | −0.53 | 0.089 | 0.000 | 0.59 |
| Metro status (2) | −0.30 | 0.115 | 0.010 | 0.74 |
| Metro status (3) | −0.33 | 0.098 | 0.001 | 0.72 |
| Metro status (4) | −0.37 | 0.172 | 0.035 | 0.69 |
| Initial incentive ($0) | 0.10 | 0.120 | 0.421 | 1.10 |
| Initial incentive ($2) | −0.06 | 0.126 | 0.640 | 0.94 |
| Refusal incentive ($0) | −0.03 | 0.102 | 0.774 | 0.97 |
| Refusal incentive ($2) | 0.04 | 0.107 | 0.740 | 1.04 |
| Second refusal Priority Mail (no) | −0.51 | 0.084 | 0.000 | 0.60 |

*The sample size is 5,951 households that refused the screener twice and had not been sent Priority Mail previously and completed or refused on the last attempt. The Estrella (1998) goodness of fit statistic is .014.

interviewer time, supervisor time, telephone time, and computing. In the next section, these costs are examined for different design options. Here, we limit the investigation to the number of call attempts by the different experimental conditions to determine effects on data collection costs.

The number of call attempts required to obtain a completed screener interview in the experiment varied widely depending on the final status of the telephone number, as is typical in RDD surveys. For example, telephone numbers that ended as completed screeners had an average of 4.7 (s.e. = 0.03) call attempts, those that were finalized as refusals had an average of 12.9 (s.e. = 0.09) call attempts, and those that were never answered had an average of 22.3 (s.e. = 0.10) call attempts. There was substantial variability in the number of call attempts among the households that completed the screener by whether they ever refused. Households that completed without ever refusing had an average of 3.7 (s.e. = 0.03) call attempts, while those that refused and then later completed averaged 10.1 (s.e. = 0.06) call attempts. These factors have important implications for the cost of data collection under different designs, as described in the next section. However, there was little variation from the overall average of 7.0 calls per telephone number across the ten incentive groups. The range was from 6.7 (s.e. = 0.10) call attempts in Group 9 to 7.3 (s.e. = 0.10) call attempts in Groups 4 and 5. This level of variation in average number of call attempts has a minor effect on the overall cost of data collection.

A conjecture of Singer, Groves, and Corning (1999) related to cost and quality is that a potential effect of incentives is that respondents who receive incentives may be more willing to perform other survey-related tasks. In NHES 2003, we assessed this hypothesis by examining the completion rates for the PFI and AEWR extended interview surveys, which were conducted after the screening survey. (The completion rate for an extended interview survey is the response rate for that survey, conditional on having completed the screener.) The completion rates for PFI vary somewhat by experimental group, from a low of 80 percent for Group 5 to a high of 89 percent for Group 8. The AEWR completion rates are all within 5 percentage points of each other, but the rates for AEWR are confounded by another incentive experiment conducted within this survey. While the groups with higher extended completion rates are predominately those groups in which the household received larger incentives, the differences in the rates are relatively small. Thus, if there is any effect on cooperation rates at the extended interview level that is carried over from giving incentives at the screening interview level, then this effect is relatively small.

The last set of findings from the screener experiment considers whether the advance incentives might have encouraged different groups of people to respond to the survey, and whether this might translate into differential nonresponse bias. To examine this hypothesis, comparisons were made between the characteristics of cases sent an advance incentive and those cases that were not sent an advance incentive. The measure of potential nonresponse bias we used is the difference between estimates of characteristics for completed cases that ever refused and cases that never refused. If these differences are constant regardless of whether they were sent an advance incentive, then it is an indication that the advance incentive had no substantial effect on nonresponse bias.

The variables in the logistic regression analysis of the initial cooperation rates were examined, along with a few other variables collected in the screener interview. We found that differences between the "ever refused" and the "never refused" households were statistically and substantively significant for some characteristics. There were statistically significant differences between the ever refused and never refused households by region, interview language, presence of a person under 21 in the household, and home ownership. However, these differences were essentially the same for those sent an advance incentive and those not sent an advance incentive, except for home ownership and presence of a person under 21 in the household. When these characteristics were further examined by incentive group, no statistically significant differences were detected. It is worth noting that the differences noted above after initial call attempts are mitigated by the refusal conversion efforts in NHES. Thus, while NHES:2003 did not exhibit obvious differential nonresponse bias from the use of incentives, RDD surveys that do not use intense refusal conversion efforts might find differences in characteristics of respondents depending on the use of incentives.

## 3. Sample Design

When the results from the incentive experiment are combined with information on the components of data collection costs for RDD surveys, sample design options that might not otherwise be cost-effective become more attractive. Below, we describe relevant components of data collection costs for an RDD survey and examine the contribution of refusal cases to the total costs. We then evaluate sample designs that subsample refusals to give a cost-effective approach and still achieve relatively high screener response rates. Subsampling refusals have a rich history beginning with Hansen and Hurwitz (1946) and continuing through Elliott, Little, and Lewitzky (2000). This research uses the standard paradigm of minimizing data collection costs while achieving a specified precision level. In our application, we also consider nonresponse error, albeit indirectly, by restricting attention to designs that have higher screener response rates.

Many of the most costly data collection activities in NHES and other RDD surveys involve special procedures that are used to increase contact and screener response rates. For example, procedures used for refusals may include special mailings, multiple refusal conversion attempts, special training for interviewers chosen specifically for refusal conversion, and paying refusal conversion interviewers higher wage rates for working these more difficult cases. As noted earlier, the average number of call attempts for refusals is much larger than for those that never refuse. Similarly, special procedures for the hard to contact cases that increase data collection costs include increasing the number of attempts to reach numbers that were not answered, and making additional callbacks to complete interviews after making initial contact with the household.

For the purposes of simplifying the presentation of alternative designs that permit subsampling, we restrict attention to subsampling for refusal cases. Subsampling other expensive types of cases is discussed later. We also restrict attention to surveys that have exactly one refusal conversion attempt since this is probably the standard practice in RDD surveys, and because we have experienced relatively low yields from second refusal conversion attempts in NHES. The extension to surveys with more than one refusal conversion attempt is straightforward.

A critical parameter for evaluating alternative designs is the data collection cost for those households that complete the screener without ever refusing as compared to the cost for those that refuse the screener at least once. Using NHES as a model, the cost to complete a screener for a household that never refuses is roughly one-half of the cost for a household that refuses at least once. In absolute terms, the interviewing operations cost for the never refused households is about $12 (USD) and for the ever refused households is $25 (USD). These costs exclude the costs of any special mailing or monetary incentives. First-Class mailings cost about $0.50 (USD) and special mailings cost about $5.50 (USD), including postage, materials, and the cost of inserting a letter or incentive. The cost of the incentive is either $2 or $5 (USD). The expected costs for each of the conditions in the experiment are computed under various refusal subsampling alternatives, where the costs include interviewing, mailing, and incentive costs. Again, the analysis is restricted to mailable cases only.

In general, subsampling reduces the precision of the estimates because the sample size is reduced and because the retained refusal cases are weighted to account for those subsampled out, inducing a design effect due to the differential weights. The design effect is largely a function of the subsampling rate and, in the application to NHES considered here, the design effect ranges from 1.00 when 90 percent are subsampled, up to 1.09 when 50 percent are subsampled. To account for the design effect and loss of sample due to subsampling, all comparisons presented below are based on designs with the same expected effective number of completed screeners.

Table 8 gives projected screener response rates and relative costs of treatment for each incentive group in the NHES 2003 experiment under different refusal subsampling scenarios. The screener response rates are constant for a given design because the screener response rates are assumed to be independent of subsampling once they are properly weighted to account for subsampling. Table 8 shows that refusal subsampling is cost-effective under the cost model only when incentives are included in the refusal mailing. For example, refusal subsampling reduces costs by only about 3 percent for Group 1 that

*Table 8. Projected cost ratios* and screener response rates, by incentive group*

| Incentive group | Expected screener response rate (s.e.) | Cost ratio for the given refusal subsampling rate | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1.0 | 0.9 | 0.8 | 0.7 | 0.6 | 0.5 |
| 1. $0/First-Class $0 | 57.9 (0.8) | 1.00 | 0.98 | 0.97 | 0.97 | 0.99 | 1.03 |
| 2. $0/First-Class $2 | 61.9 (0.7) | 1.10 | 1.07 | 1.05 | 1.05 | 1.06 | 1.10 |
| 3. $0/First-Class $5 | 64.6 (0.8) | 1.23 | 1.19 | 1.15 | 1.14 | 1.14 | 1.17 |
| 4. $0/Priority $0 | 59.9 (0.7) | 1.15 | 1.12 | 1.10 | 1.08 | 1.09 | 1.12 |
| 5. $0/Priority $2 | 62.4 (0.7) | 1.24 | 1.20 | 1.17 | 1.15 | 1.15 | 1.17 |
| 6. $2/First-Class $0 | 63.4 (0.7) | 1.21 | 1.20 | 1.19 | 1.20 | 1.22 | 1.27 |
| 7. $2/Priority $0 | 65.0 (0.8) | 1.36 | 1.33 | 1.31 | 1.30 | 1.32 | 1.36 |
| 8. $2/Priority $2 | 65.6 (0.7) | 1.44 | 1.40 | 1.37 | 1.36 | 1.37 | 1.40 |
| 9. $5/First-Class $0 | 65.3 (0.7) | 1.57 | 1.56 | 1.56 | 1.57 | 1.60 | 1.65 |
| 10. $2/First-Class $2 | 65.4 (0.6) | 1.30 | 1.28 | 1.26 | 1.26 | 1.28 | 1.32 |

*The cost ratio is the cost for the alternative design divided by the cost for incentive Group 1 without subsampling.

has no incentives. In addition, for groups with incentives at the refusal stage, the relative costs decline as the refusal subsampling rates decline to a point and then begin to increase. This is due to the fact that at some point, the cost benefits of subsampling refusals are outweighed by the increased costs due to having to field more cases to offset the design effect. The inflection point is around a subsampling rate of 0.6 to 0.7, depending on the amount of the refusal incentive. Figure 2 illustrates this for a few groups.

We have not found any evidence of variation in nonresponse bias by incentive group, but the data available to study bias are very limited. Since the response rate is an important quality measure in surveys, the findings of this study are relevant if the goal is to increase the response rate above and beyond what standard procedures will allow. If an increase of screener response rates by at least 3 percentage points is an important objective, then the estimates in Table 8 may be useful. For example, Groups 1, 2, 4, and 5 (with expected screener response rates ranging 58 percent to 62 percent) have rates that may be unacceptable for a survey like NHES. Group 3 yields a higher expected screener response rate at a lower cost than Group 6 for each refusal subsampling alternative other than no subsampling, so Group 6 may be eliminated. Similarly Group 10 has about the same expected screener response rate and a lower cost than Groups 7, 8, and 9. This logic reduces the choice to be between Groups 3 and 10. The difference in expected screener response rates between these two groups is small and not statistically significant, while Group 3 has a lower cost. Figure 2 shows the relative costs for these two groups, along with the control Group 1. A refusal subsampling rate of about 60 percent is optimal for Group 3 and a rate of about 70 percent is optimal for Group 10. In the next section we discuss how this procedure was implemented in the planning of the NHES 2005 data collection.

## 4. Discussion

The incentive experiment confirmed previous research showing that sending modest prepaid monetary incentives of $2 or $5 (USD) prior to calling a household is effective in increasing cooperation in the difficult screening stage of an RDD survey. The experiment also showed that screener response rates could be increased by giving households these prepaid monetary incentives at the first refusal conversion stage rather than at the initial

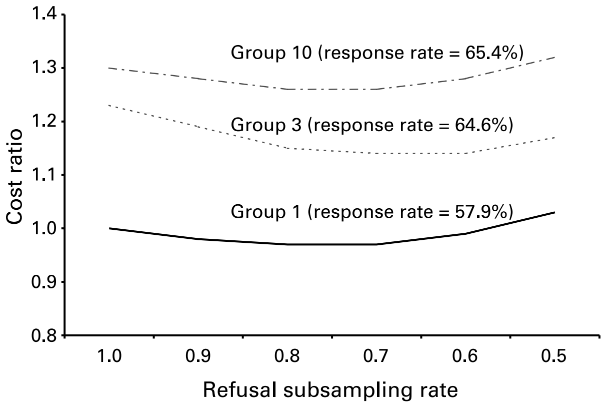



*Fig. 2. Relative costs of selected mailing and incentive strategies*

stage. Priority Mail at the first refusal conversion stage was not effective in increasing refusal conversion and screener response rates. On the other hand, the rates in NHES were improved by sending a Priority Mail letter prior to a second refusal conversion attempt as compared to having no other treatment. The Priority Mail was only sent to households that had not been sent Priority Mail at an earlier stage. Another important finding of the experiment was that treatments at refusal conversion stages were virtually independent of treatments at earlier stages.

The potential effects of monetary incentives on other aspects of cost and quality of the data collection were examined. Little variation was found in the average number of calls per telephone number required to complete a screener across the 10 incentive conditions in the experiment. Similarly, giving incentives at the screener stage did not have substantial carry-over effects on cooperation rates at the extended interview level. We also examined the potential for nonresponse bias associated with the use of incentives. No evidence of nonresponse bias could be linked to the incentives. Some characteristics of households that initially cooperated differed depending on the use of incentives, but refusal conversion reduced these differences and the consequent possibility for biases. However, RDD surveys that do not convert refusals might find some differences due to the use of incentives that may bias estimates.

These findings led to considering refusal subsampling in conjunction with the use of incentives at the refusal conversion stage. This is a cost-effective procedure that gives a higher screener response rate than could be obtained without incentives. In addition to the potential cost benefits of refusal subsampling reported here, the refusal subsampling procedure may have operational benefits such as allowing for more effective management of staffing, boosting interviewer morale, and allowing for earlier and more efficient survey close-out than would be the case without refusal subsampling.

A subsampling design is being implemented for the next NHES in 2005. In that survey, 60 percent of both refusals and hard to contact screeners will be subsampled. The hard to contact cases are those that are never answered after the standard number of attempts or those that are never completed after many attempts because the household is not available. In previous NHES administrations, these types of cases were called many times beyond the standard number to try to increase response rates. The 2005 survey will send a Federal Express letter to subsampled refusals before the second screener refusal conversion attempt to try to further increase screener response rates.

The subsampling of refusals is not appropriate for all RDD surveys. In NHES 2005, the cost of screening is expected to be relatively high compared to the cost of the extended interviews because there is a relatively low eligibility rate (only two-thirds of screened households will have a sampled person), and the extended interviews are relatively short (15 to 18 minutes each). As noted previously, to compensate for the design effect due to subsampling, additional screening and extended interviews must be completed to attain the same effective sample size as a design without subsampling. If the extended interview is much more expensive than the screener, then subsampling to achieve the same effective sample size will not be cost-effective. For example, in longitudinal surveys or surveys that require expensive data collections such as in-person testing or physical examinations, subsampling will not be effective because the increased data collection costs for a larger sample of extended interviews or lengthier interviews will outweigh the benefits of the costs saved in screening. In NHES, like many cross-sectional RDD surveys, the extended interview costs are much smaller in relative terms and refusal subsampling is attractive.

## 5. References

American Association for Public Opinion Research (AAPOR) (2004). Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. Lenexa, KS: AAPOR.

Atrostic, B.K., Bates, N., Burt, G., and Silberstein, A. (2001). Nonresponse in U.S. Government Household Surveys: Consistent Measures, Recent Trends, and New Insights. Journal of Official Statistics, 17, 209–226.

Berk, M.L., Mathiowetz, N.A., Ward, E.P., and White, A.A. (1987). The Effect of Prepaid and Promised Incentives: Results of a Controlled Experiment. Journal of Official Statistics, 3, 449–457.

Brick, J.M. and Collins, M. (1997). An Experiment in Random-Digit-Dial Screening. U.S. Department of Education, National Center for Education Statistics, NCES 1998-255, Washington, DC.

Brick, J.M., Collins, M.A., Nolin, M.J., Davies, E., and Feibus, M.L. (1997). Design, Data Collection, Monitoring, Interview Administration Time, and Data Editing in the 1993 National Household Education Survey (NHES:93). Department of Education, National Center for Education Statistics, NCES 1997-04, Washington, DC.

Brick, J.M., Judkins, D., Montaquila, J., and Morganstein, D. (2002). Two-Phase List-Assisted RDD Sampling. Journal of Official Statistics, 18, 203–216.

Brick, J.M., Montaquila, J., and Scheuren, F. (2002). Estimating Residency Rates for Undetermined Telephone Numbers. Public Opinion Quarterly, 66, 18–39.

Brick, J.M., Hagedorn, M.C., Montaquila, J., Roth, S.B., and Chapman, C. (forthcoming). Monetary Incentives and Mailing Procedures in a Federally Sponsored Telephone Survey. U.S. Department of Education, National Center for Education Statistics. Washington, DC.

Cantor, D., Allen, B., Cunningham, P., Brick, J.M., Slobasky, R., Giambo, P., and Kenny, G. (1997). Promised Incentives on a Random Digit Dial Survey, In Nonresponse in Survey Research, Proceedings of the Eighth International Workshop on Household Survey Nonresponse, A. Koch and R. Porst (eds), ZUMA: Mannheim, Germany, 219–228.

Cantor, D., Cunningham, P., Triplett, T., and Steinbach, R. (2003). Comparing Incentives at Initial and Refusal Conversion Stages on a Screening Interview for a Random Digit Dial Survey. Proceedings of the American Statistical Association, Section of Survey Research Methods.

Church, A.H. (1993). Estimating the Effect of Incentives on Mail Surveys: A Meta-Analysis. Public Opinion Quarterly, 57, 62–79.

Collins, M.A., Brick, J.M., Loomis, L.S., Nicchitta, P.G., Fleischman, S., and Chandler, K. (1997). Design, Data Collection, Interview Timing, and Data Editing in the 1995 National Household Education Survey (NHES:95). U.S. Department of Education, National Center for Education Statistics, NCES 1997-08, Washington, DC.

Curtin, R., Presser, S., and Singer, E. (2004). Recent Response Rate Changes on the Michigan Survey of Consumer Attitudes. Paper presented at the Annual Meeting of the American Association for Public Opinion Research, Phoenix, AZ.

Dillman, D.A. (2000). Mail and Internet Surveys: The Tailored Design Method. New York: John Wiley and Sons.

Elliott, M.R., Little, R.J.A., and Lewitzky, S. (2000). Subsampling Callbacks to Improve Survey Efficiency. Journal of the American Statistical Association, 95, 730–738.

Estrella, A. (1998). A New Measure of Fit for Equations with Dichotomous Dependent Variables. Journal of Business and Economic Statistics, 16, 198–205.

Gelman, A., Stevens, M., and Chan, V. (2003). Regression Modeling and Meta-Analysis for Decision Making: Cost-Benefit Analysis of Incentives in Telephone Surveys. Journal of Business and Economic Statistics, 21, 213–225.

Goldstein, K.M. and Jennings, M.K. (2002). The Effect of Advance Letters on Cooperation in a List Sample Telephone Survey. Public Opinion Quarterly, 66, 608–617.

Groves, R.M. and Couper, M.P. (1998). Nonresponse in Household Interview Surveys. New York: John Wiley and Sons.

Groves, R.M., Singer, E., and Corning, A. (2000). Leverage-Saliency Theory of Survey Participation. Public Opinion Quarterly, 64, 299–308.

Hagedorn, M.C., Montaquila, J., Vaden-Kiernan, N., Kim, K., Li, L., and Chapman, C. (2004). National Household Education Surveys Program of 2003: Data File User's Manual, Volume 1. U.S. Department of Education, National Center for Education Statistics. NCES 2004-101, Washington, DC.

Hansen, M.H. and Hurwitz, W.N. (1946). The Problem of Nonresponse in Sample Surveys. Journal of the American Statistical Association, 41, 517–529.

Martin, E., Abreu, D., and Winters, F. (2001). Money and Motive: Effects of Incentives on Panel Attrition in the Survey of Income and Program Participation. Journal of Official Statistics, 17, 267–284.

Nolin, M.J., Montaquila, J., Nicchitta, P., Kim, K., Kleiner, B., Lennon, J., Chapman, C., Creighton, S., and Bielick, S. (2000). National Household Education Survey of 1999: Methodology Report. U.S. Department of Education, National Center for Education Statistics, NCES 2000-078, Washington, DC.

Nolin, M.J., Montaquila, J., Nicchitta, P., Hagedorn, M.C., and Chapman, C. (2004). National Household Education Surveys Program: 2001: Methodology Report. U.S. Department of Education, National Center for Education Statistics. NCES 2004-071, Washington, DC.

Singer, E. (2002). The Use of Incentives to Reduce Nonresponse in Household Surveys. In Nonresponse in Household Interview Surveys, R.M. Groves, D.A. Dillman, J.L. Eltinge, and R.J.A. Little (eds). New York: John Wiley and Sons.

Singer, E., Groves, R.M., and Corning, A.D. (1999). Differential Incentives: Beliefs About Practices, Perceptions of Equity, and Effects on Survey Participation. Public Opinion Quarterly, 63, 251–260.

Singer, E., van Hoewyk, J., Gebler, N., Raghunathan, T., and McGonagle, K. (1999). The Effect of Incentives on Response Rates in Interviewer-Mediated Surveys. Journal of Official Statistics, 15, 217–230.

Singer, E., Van Hoewyk, J., and Maher, M.P. (2000). Experiments with Incentives on Telephone Surveys. Public Opinion Quarterly, 64, 171–188.

Trussell, N. and Lavrakas, P.J. (2004). The Influence of Incremental Increases in Token Cash Incentives on Mail Survey Response: Is There an Optimal Amount? Public Opinion Quarterly, 68, 349–367.

Vaden-Kiernan, N., Nicchitta, P.G., Montaquila, J., and Collins, M.A. (1997). Design, Data Collection, Interview Administration Time, and Data Editing in the 1996 National Household Education Survey (NHES:96). U.S. Department of Education, National Center for Education Statistics, NCES 1997-35, Washington, DC.

Received June 2004
Revised November 2004