# Comparing Incentives at Initial and Refusal Conversion Stages On a Screening Interview For a Random Digit Dial Survey

David Cantor, Patricia Cunningham, Westat; Timothy Triplett, Rebecca Steinbach, The Urban Institute
David Cantor, Westat, 1650 Research Blvd, Rockville, MD 20850

**KEYWORDS:** Pre-notification; Missing Data; Response Bias

## Problem

This paper reports the results of an experiment that tests differences in the timing of the delivery of incentives. More specifically, it compares sending pre-paid incentives at the initial contact call to sending advance incentives to households that initially refuse to cooperate. In a paper presented at last year's AAPOR meetings, we reported results from an experiment comparing these two procedures for a relatively small, geographically restricted, sample. In this paper, we provide the comparisons for a national sample (Cantor, et al., 2002).

Three research questions are addressed in this paper:

1. What is the effect of a pre-paid incentive of $2 at the initial contact compared to a $5 incentive sent at refusal conversion?

2. Do the incentive treatments differentially affect population subgroups?

3. Do the incentive treatments affect data quality?

The amount of the incentive provided in these two alternative procedures ($2 at initial vs. $5 at refusal conversion) is not the same. If one wanted to compare the effectiveness of initial vs. refusal conversion treatments on cooperation, one would want to offer the same amount of money at each stage. The different amounts were chosen for a practical reason — they are equivalent in their expected total cost. Paying money at the initial stage requires sending out money to many more households then at refusal conversion. By equating total cost, this experiment provides a comparison of what a survey designer faces when working with a fixed budget for incentives.

There are other issues related to the use of incentives in an RDD survey. One is understanding the mechanisms around which they work. Use of advance mailing on an RDD survey is not straightforward. Perhaps for this reason, there is only mixed evidence that they increase response rates (Camburn, et al., 1995; Brick, et al., 1997; Singer, et al., 2000; Parsons, et al., 2002). To send an advance letter, each telephone number has to be matched against reverse directories to obtain an address. This is successful for only a portion of the numbers. Even for those telephone numbers where there is a match to the directory, some portion may be out of date. Once sending the letter by regular mail, the letter may be thrown away before it is read (Mathiowetz, et al., 1994). Even if it is read, there is no guarantee that the person who answers the telephone will be the one who read the letter. The result of this process is that many people who are called do not see the letter the interviewer calls.

Given the complicated sequence of events related to mailing an advance letter described above, it would be useful to know how many respondents actually receive and remember getting the advance mailings. It would be helpful to find out whether the respondent who did not receive the advance letter knew about the letter or incentive. Furthermore, it would be interesting to find out if the incentive had an effect on the respondent's recollection of the advanced mailing.

## Experimental Design

The experiment was conducted as part of Cycle 3 of the National Survey of America's Families (NSAF). The NSAF is a RDD survey funded by a consortium of private foundations in the United States. It is conducted by Westat for the Urban Institute. An important goal of the survey is to assess the impact of recent changes in the administration of government assistance programs for children and the poor.

The NSAF consists of both a screening and extended interview. The screening interview consists of a 3-5 minute battery of questions designed to select the person that should be administered the extended interview. It determines if there are any persons under 65 years old in the household and whether or not the family is above or below 200% of poverty. If there is someone in the right age-range and the household is sampled (based on poverty status) a respondent for the extended interview is selected. The extended interview is 25 - 50 minutes in length (depending on the type of interview) and covers a wide range of topics, including health, education, child care, income and receipt of

social services. Approximately 42,000 to 45,000 extended interviews are completed in a typical cycle.

The experiment that is discussed below had three conditions, which are shown in Table 1. These include:

- Control – Respondents were sent an advance letter using 1st class mail. If they refused, they were sent a second "refusal conversion" letter in advance to trying a second time. A third refusal conversion attempt was made to those who continued to refuse.

- $2 pre-notification - Respondents were sent an advance letter using 1st class mail with a $2 bill in it. If they refused, they were sent a second "refusal conversion" letter in advance to trying a second time. This refusal conversion letter was sent using USPS priority mail A third refusal conversion attempt was made to those who continued to refuse.

- $5 refusal conversion - Respondents were sent an advance letter using 1st class mail. If they refused, they were sent a second "refusal conversion" letter in advance to trying a second time. This letter had $5 in it and was sent using USPS priority mail A third refusal conversion attempt was made to those who continued to refuse.

To examine the first research question, response rates at the screener were compared across conditions, to address several hypotheses:

Hypothesis 1: Providing an incentive at any stage will increase the response rate.

Hypothesis 2: Sending an incentive at pre-notification will be more effective than sending it at refusal conversion.

Hypothesis 3: Sending an incentive will increase the number of respondents who remember receiving the advance material.

Hypotheses 1 and 2 are examined by comparing the response rates across the three conditions. Analysis for hypothesis 1 separately compares the control condition against the $2 and the $5 conditions. Analysis for hypothesis 2 compares the $2 and $5 condition. Hypothesis 3 is tested using data from questions asked of a sample of respondents about whether they remember receiving the advance materials.

We explore the possibility that the effects of the treatments vary by different population subgroups (Research question 2). This is accomplished using geographic data linked to each of the telephone numbers. Finally, we assess differences in data quality using two measures from the screening interview (Research question 3).

**Methods**

These experiments were conducted at the beginning of Cycle 3 of the NSAF. The field period for the NSAF was from February to October of 2002. Since the experiment was conducted with the initial release groups, most of these cases were finalized between February and April of 2002.

All interviewers administering the NSAF during this period participated. This is approximately 300 individuals, once the survey was fully staffed. Interviewers were aware of the different experimental manipulations. The money was mentioned in the introductions to each of these two experimental treatments.

The data reported below are weighted by the initial probability of selecting the telephone number. The weighted data are used to be able to generalize the results to a national population. The significance tests were calculated using *WESVAR 4*©, in conjunction with the JK2 estimation method.

To calculate screener response rates, the following formula was used:

$$SR = (CS)/(CS + R + .63(AM) + .27(NA) + ONR)$$

Where: CS = completed screener, including eligible and ineligible households; R=refusals; AM = Answering machines; NA = No answer and ONR = Other non-response

**Response Rates and Awareness by Experimental Condition**

Table 2 provides the response rates for the screener by the three different test conditions. The final response rates across conditions were 64.8, 67.9 and 68.2 for the control, $2 pre-note and $5 refusal conversion treatments respectively. With respect to hypothesis 1, there is a significant difference between the control with the two incentive conditions ($p<.05$; one-tailed test). There is no significant difference between the two incentive conditions (Hypothesis 2).

The assumption is that the difference between the incentive and the control is the receipt of the incentive. However, the incentive may also have the effect of drawing attention to the materials that are included in the pre-notification package. If a household member opens the mail and sees cash, they may be more likely to read the material that is included in the letter. It may convince respondents to read the material. Once doing so, there should be a greater likelihood they will cooperate, regardless of any social obligation related to the payment.

To get a sense of how aware respondents were of the different treatments, the questionnaire included items that asked whether the respondent remembered any of the materials that were sent prior to the call. They were also asked if they told anyone else in the household about these materials and whether anyone else had told them. The items were asked after the respondent had completed the interview.[1] Table 3 provides these results by the different experimental conditions.

One striking result is the very low rate of recognition for the control condition. Less than one-third of those cooperating remember seeing the letter at the initial call. This is significantly lower than the 60% reported for another RDD survey (Parsons, et al, 2002).

The variation in this measure across the treatments is consistent with the idea that the incentives serve to bring attention to the pre-notification material (Hypothesis 3). Sixty-one percent of the respondents in the $2 pre-note condition report remembering the material at the initial stage (Table 4) As expected, the rate of recognition for the $5 refusal conversion treatment is equivalent to the control at the initial call, but increases significantly to 68% at refusal conversion.

There also seems to be some communication of the material to other members of the household. When respondents were asked if they had told anyone else in the household about the letter (and incentive), between 40% and 50% said they had.[2] Interestingly, the $2 produced the most communication in this direction, with the percentage of people telling others being statistically different than the other two treatments at the initial stage. Even at first refusal conversion, the $2 treatment is nominally the highest, although none of these differences are statistically significant.

A much lower percentage of the respondents who said they had not gotten the material say that someone else in their household had told them about them the study. The lack of symmetry between these two measures may be indicative of some measurement error. One would have expected that about as many people who said they told someone else would have said that someone else had told them. This may mean that those who are initially responding that they had seen the material may be those who were told about it, rather than seeing it directly.

**Differences in Data Quality**

Two different measures were used to assess data quality (Table 5). The first was the amount of missing data to the income item on the screener. This item asked respondents whether their total household income was above or below 200% of the poverty line for that household (depends on size and presence of children). As can be seen from these data, there are no significant differences in this measure across the different treatments.

The second indicator was whether the measure of poverty for that household switched between the screener and extended interview. This measure involved comparing the measure of 200% of poverty status taken at the screener to the measure collected on the extended interview. The extended interview contains a detailed battery of questions that ask about specific income sources. The income for the entire household is then computed and a variable created for whether the household was above or below 200% of the poverty line. This second measure of poverty status was then compared to the single item on the screener. Higher data quality is indicated by a lower percentage of households switching between these two points in the interview.

As seen from these data, there are no statistically significant differences across the three treatment groups on this measure.

**Differences Across Population Subgroups**

As noted in the discussion above, there is

[1] If the household was deemed not eligible or if the respondent was not selected for an extended interview, these questions were asked right after completing the screener. If the respondent was selected for the extended interview, the questions were asked after they completed the extended interview.

[2] Only households with more than one person were asked about intra-household communication.

some reason to believe that the effect of incentives may be different across sub-populations. One possibility is that incentives will be more effective for low income households. This has been found in other research related to incentives (Singer, 2002), although the evidence is mixed. Another possibility is that incentives would be more effective for reluctant respondents. The rationale for the latter is that normal refusal conversion efforts (without incentive) may suffice in convincing those who are marginally reluctant to cooperate. It takes something like an incentive, or the provision of a more tangible benefit, to convince those who are particularly reluctant.

Data from the sampling frame was used to explore whether the incentives were differentially effective over population sub-groups. Response rates were tabulated by population groups, which were formed based on the phone numbers' geographic area information. Geographic area is defined by the county or state associated with the number.

Generally, the effects of the treatments do not differ greatly by type of geographic area. Nonetheless, there were a few significant effects that are consistent with the above discussion (Table 6). For example, there is a significant difference between the incentive and no-incentive conditions for the low response rate states ($p<.10$; two tailed test), while no difference exists for the high response rate states. Several of the measures related to socio-economic status show weak statistical significance. This includes high migration, low owner occupancy, High foreign born, Low employment and High black population. The family income measure does not show as strong a pattern as one might expect, given the expectation that income is an important correlate of incentives. The incentives were significantly less effective in geographic areas where travel time to work was longer. This relationship is consistent with the time-use research (Robinson & Geoffrey Godbey, 1999) findings that Americans with less free time are more likely to choose to have a day off over an extra day's pay.

The weak relationships found here may be due to the measures. Counties are large geographic units and there could be quite a bit of variation within these areas. Refining these measures to smaller geographic units may reveal stronger correlations.

**Discussion**

This paper discussed the results of an experiment that explored the use of incentives at different stages of an RDD survey. A primary question addressed was the relative effectiveness of using an incentive at the initial call or during refusal conversion. The results show that providing $2 at the initial attempt to complete the screener works about as well with respect to response rates as a $5 treatment at refusal conversion.

The actual processes incentive affect is not entirely clear. It does seem to be the case that both types of incentives increase the number of persons that report seeing the letter. Of those in the no-incentive group, only about 30% of the persons that completed a screener report remembering the letter. This number is essentially doubled when sending $2 at pre-notification or $5 at refusal conversion. A benefit of the incentive is that it draws attention to the advance material. It is unclear how much of the effect of the incentive adds to the perceived benefits of participating on the survey. This could not be disentangled in this experiment.

There doesn't seem to be a big difference by incentive groups with respect to whether household members tell other members of the household about the study. While the data presented above are ambiguous, there are no clear patterns across the three experimental groups for whether respondents report telling someone else.

There were no strong differential effects found of the treatments across different subpopulations. There was some suggestion that incentives work better in states with low cooperation rates and in areas that have high migration, low owner occupancy, High Foreign Born, Low Employment and High Black populations. These effects were small, at best, generally reaching statistical significance at the 10% level. The weakness of these patterns, however, may be function of the relatively broad geographical areas represented by the measures (counties). Future research should consider refining these measures by narrowing the geographic areas.

Data quality, as measured by the amount of missing data on income and income switching, did not differ across the different treatments.

Overall, the effects of the incentive treatments relative to the no-incentive group were quite small. Providing either type of incentive increased response rates by about 3% - 4%. This is considerably lower than that reported by Singer et al., (2000) and Cantor, et al., (1998) of 10% and 6%, respectively. This may be an indication that the public is increasingly becoming resistant to doing surveys, regardless of the

use of incentives, at least at the levels that were tested in this experiment.

It should also be noted that the effects discussed above only apply to that portion of the sample for which an address was found. In the case of this particular survey, that turns out to be approximately 80% of the residential households in the sample.

**References**


Brick, J.M., Flores-Cervantes, I., and D. Cantor (1999), *1997 NSAF Response Rates and Methods Evaluation*, Report No. 8, March 1999, Washington, D.C.: Urban Institute.

Camburn, D.P., Lavrakas, P.J., Battaglia, M.P., Massey, J.T. and R.A. Wright (1995), "Using Advance Respondent Letters in Random Digit Dialing Telephone Surveys," *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 969-974.

Cantor, D., Cunningham, P., Wang, K., Singer, E. and F. Scheuren (2002) "An Experiment on the Timing of Incentives and Different Staging Procedures on a Random Digit Dial Survey" Paper presented at the 2002 Annual Meeting of the American Association for Public Opinion Research, St. Petersburg, FL., May 16-19.

Cantor, D., Cunningham, P., and Giambo, P. (1998), "Testing the Effects of a Pre-Paid Incentive and Express Delivery to Increase Response Rates on a Random Digit Dial Telephone Survey," Paper presented at the 1998 Annual Meeting of the American Association for Public Opinion Research, St. Louis, MO., May 14-17.

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

Mathiowetz, N.A., Couper, M.P., and E. Singer (1994), "Where Does All the Mail Go? Mail Receipt and Handling in U.S. Households." U.S. Bureau of the Census: unpublished paper.

Parsons, J., Owens, L. and W. Skogan (2002) "Using Advance Letters in RDD Surveys: Results of Two Experiments." Survey Research, Vol 33, #1, pp. 1-2, Newsletter from the Survey Research Laboratory, College of Urban Planning and Public Affairs, University of Illinois, Chicago.

Robinson, John P., and Geoffrey Godbey. 1999. *Time For Life: The Surprising Way Americans Use Their Time*. Pennsylvania: Penn State University.

Singer, E. (2002) "The Use of Incentives to Reduce Nonresponse in Household Surveys" pp. 163 – 178 in Groves, R., Dillman, D.A., Eltinge, J.L. and R.J.A. Little Survey Nonresponse. John Wiley: New York.

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

Table 1. Experimental Design

| | Condition | | |
|---|---|---|---|
| | Control | $2 Pre-Note | $5 Refusal Conversion |
| Initial Attempt | Letter, 1st Class | Letter, 1st class; Pre-pay $2 | Letter, 1st class |
| Refusal Conversion | Letter, 1st Class | Letter, Priority | Letter, Priority; Advance $5 |

Table 2. Response Rates by Experimental Treatment and Stage

| | Control | $2 Pre-Note | $5 Refusal Conversion |
|---|---|---|---|
| Initial Attempt | 45.9 | 51.2+# | 43.0# |
| After Refusal Conversion | 64.8+* | 67.9+ | 68.2* |
| (Unweighted N) | (1,814) | (16,499) | (3,665) |

+Control vs. $2 significant at p <.05 (one-tailed)
*Control vs. $5 significant at p <.05 (one-tailed)
#$2 vs. $5 significant at p <.05 (one-tailed test)

Table 3. Percent Remembering Letter by Experimental Group

| | Control | $2 Pre-Note | $5 Refusal Conversion |
|---|---|---|---|
| With Addresses | | | |
| Initial | 27.4+ | 61.0+# | 33.6# |
| (Unweighted N) | (227) | (1966) | (425) |
| At First Refusal Conversion | 35.2+ | 43.2# | 67.6*# |
| (Unweighted N) | (65) | (454) | (197) |
| Total After all Refusal Conversion | 29.2 | 56.2 | 44.9 |
| (Unweighted N) | (309) | (2595) | (666) |
| Without Addresses | | | |
| Total After all Refusal Conversion | 9.5 | 2.1 | 10.8 |
| (Unweighted N) | (19) | (157) | (38) |

+Control vs. $2 significant at p <.001 (one-tailed)
*Control vs. $5 significant at p <.01 (one-tailed)
#$2 vs. $5 significant at p <.001 (two-tailed)

Table 4. Percent Who Communicated Receipt of Materials by Experimental Group

| | Control | $2 Pre-Note | $5 Refusal Conversion |
|---|---|---|---|
| Tell others in the household? | 45.5 | 52.9 | 34.8# |
| (Unweighted N) | (78) | (1,317) | (241) |
| Others tell you? | 2.1 | 4.0 | 4.4 |
| (Unweighted N) | (192) | (1,016) | (357) |

#$2 vs. $5 significant at p <.01 (one-tailed test)

Table 5. Indicators of Data Quality

| | Control | $2 Pre-Note | $5 Refusal Conversion |
|---|---|---|---|
| Missing on Income | | | |
| % Don't know | 3.6 | 2.9 | 3.4 |
| % Refused | 2.5 | 2.7 | 2.9 |
| (Unweighted N) | (746) | (6952) | (1553) |
| % Switched Poverty Status (Total) | 12.2 | 11.3 | 10.5 |
| (Unweighted N) | (494) | (4673) | (1045) |

Table 6. Final Response rate by Experimental Condition and selected area characteristics

| | Control (1) | $2 Pre-Note (2) | $5 refusal (3) |
|---|---|---|---|
| Hi Response rate state | 64.26 | 66.45 | 65.84 |
| Lo Response rate state | 55.30[+*] | 58.56[+] | 59.85[*] |
| | | | |
| Inside central city of MSA | 58.02 | 62.08 | 60.64 |
| Outside a central city of MSA but in central city's county | 59.92 | 61.27 | 64.89 |
| Inside a suburban county of an MSA | 60.61 | 62.29 | 62.64 |
| In an MSA that has no central city | 53.63 | 57.52 | 59.40 |
| Non-MSA | 68.41 | 71.02 | 69.95 |
| | | | |
| Hi Travel to Work | 61.14 | 60.51 | 60.34 |
| Lo Travel to Work | 60.48[#x] | 68.87[#] | 69.41[+] |
| | | | |
| Hi Migration | 59.09[#*] | 64.12[#] | 63.24[*] |
| Lo Migration | 64.68 | 62.27 | 64.20 |
| | | | |
| Hi Owner Occupancy | 65.33 | 66.10 | 68.85 |
| Lo Owner Occupancy | 59.02[+] | 62.47[+] | 61.57 |
| | | | |
| Hi Foreign Born | 57.93[+] | 61.16[+] | 61.38 |
| Lo Foreign Born | 67.43 | 68.92 | 69.24 |
| | | | |
| Hi Employment | 62.71 | 63.72 | 64.94 |
| Lo Employment | 57.94[+] | 63.15[+] | 61.55 |
| | | | |
| Hi Family Income | 60.39 | 62.25 | 62.71 |
| Lo Family Income | 63.23 | 68.05 | 66.83 |
| | | | |
| Hi Black Population | 56.20[+] | 60.84[+] | 59.93 |
| Lo Black Population | 65.49 | 66.12 | 67.06 |
| | | | |
| Hi HH size | 59.24 | 61.80 | 60.47 |
| Lo HH size | 62.61 | 65.25 | 66.69 |
| | | | |
| Hi Presence of Children | 60.31 | 64.06 | 63.55 |
| Lo Presence of Children | 61.50 | 62.94 | 63.56 |

[+] Control vs. $2 significant at p < .10 (two-tailed test)
[*] Control vs. $5 significant at p < .10 (two-tailed test)
[#] Control vs. $2 significant at p < .01 (two-tailed test)
[x] Control vs. $5 significant at p < .01 (two-tailed test)