ONCE RELUCTANT, ALWAYS RELUCTANT? EFFECTS OF DIFFERENTIAL INCENTIVES ON LATER SURVEY PARTICIPATION IN A LONGITUDINAL STUDY

Jennie E. Lengacher, Colleen M. Sullivan, Mick P. Couper, and Robert M. Groves
University of Michigan, Survey Research Center

Keywords: Nonresponse, Incentives, Panel Surveys

Nonresponse is of constant interest to survey researchers because it affects both the costs of data collection and the errors made in inferential statements about their target populations. Survey designs that strive for low nonresponse bias on probability samples often spend more money on converting sample persons (or on failing to do so) than on interviewing compliant respondents to the survey. In addition, given that few general household survey efforts yield 100% participation, the researcher must consider how the remaining nonrespondent cases compare to respondent cases on survey statistics of interest. Nonresponse error is a function both of the rate of nonresponse and the distinctiveness of the nonrespondents relative to the respondents.

A growing practice in U.S. household surveys is the use of incentives to increase response rates. A 1992 OMB-sponsored conference at the Kennedy School (COPAFS, 1993) observed that the practice of offering monetary incentives to sample persons was common among major Federal survey contractors. There was a belief among participants that the use of incentives reduced per-unit interview costs, mainly by reducing callbacks from broken appointments and refusal conversions. There was also the consensus that the research literature supported the use of prepaid incentives (those given to sample persons prior to obtaining an interview) (Berk *et al.,* 1987) relative to incentives promised to the persons on the condition that they provide an interview. Finally, there was widespread support for the finding that the effects of incentives increased as the value of the incentive increased, although probably at a decreasing rate (see James and Bolstein, 1990). Kulka (1994), in a review of the use of incentives for reluctant respondents, notes that their use might raise some ethical issues for the field of survey research. These include the possible coercive influences on behavior that incentives may pose for low income sample persons, for children, and for others who might view the incentive as restriction to their freedom to refuse the survey request. So, too, the James and Bolstein work, with its hint of a backfire effect for very large incentives, raises the possibility that under some circumstances or for some persons, incentives might actually lead to reduced cooperation.

One theoretical perspective on the role of incentives in the decision-making process holds that they act to evoke different cognitive scripts for what other types of interactions with strangers the interview request might resemble (see Groves and Couper, 1995). Small incentives (especially nonmonetary ones) could evoke the script of a gift from a visitor to the home, a token of appreciation, common to interactions among friends who visit one another. Large monetary incentives, in contrast, might evoke economic exchange scripts, the payment for the time and effort of one person by another who purchases the time and effort. Unfortunately, the survey methodological literature is relatively sparse on what cognitive script is the default script in the absence of a survey incentive, so arguments about the scripts active during the survey participation decision-making are speculative.

When the survey design is a one-time survey, which script type dominates, the social exchange or the economic exchange, may be irrelevant to cooperation rates. Both of them should be able to generate high cooperation rates under appropriate circumstances. In longitudinal surveys, however, how incentives shape the viewpoint of the sample person to the survey request may be important to the attrition rate over waves of interviewing. Following this logic, one would expect different attrition rates in survey designs that varied the nature of incentives over waves from those that used a consistent incentive scheme, other things being equal.

Longitudinal surveys are also different from one-time surveys with regard to the amount of information sample persons have about the survey request. In a one-time survey and in the first wave of a longitudinal survey, sample persons probably have quite limited information about what the survey interview might entail -- its cognitive burdens, the time commitment required, the type of information sought, the level of enjoyment of interacting with the interviewer, etc. In the second wave of a longitudinal survey, however, the sample persons are likely to associate the interview request with the experience they had in the first wave. The second wave decision might thus be influenced by how enjoyable or interesting the first wave was.

This paper examines a quasi-experimental design that was part of a large-scale longitudinal household survey. It focuses on how cooperation in the second wave is affected by experiences in the first wave. Specifically, it tests the hypothesis of whether large incentives in the first

wave followed by smaller incentives in the second wave yield different cooperation rates than relatively consistent incentive levels.

The Health and Retirement Study

The Health and Retirement Study (HRS) is a longitudinal survey of persons born between the years of 1931 and 1941 (see Juster and Suzman, 1993, for an overview). The HRS used a national area probability sample with supplemental oversamples of Blacks, Hispanics, and the state of Florida. The primary purpose of the HRS is to provide policy makers with up-to-date information on changes in this aging population's perceptions of and experiences with retirement as it relates to their health, financial situation, and family structure. The baseline wave of the HRS data collection began in April of 1992 and ran through February of 1993. Over 70,000 households were screened to identify age-eligible respondents and conduct the 12,652 face-to-face interviews that were taken in 7,702 households. The final response rate for Wave 1 was 82.0%. The second wave of interviews was collected from April through December 1994, and resulted in 11,596 interviews, with a response rate of 92.1%. Interviews will continue to be conducted every two years.

The HRS Nonresponse Study

Near the completion of HRS Wave 1 data collection, a second phase probability subsample of reluctant sample persons was drawn. The purpose of this experiment, labeled the "Nonresponse Study" (NRS), was threefold: 1) to raise the overall response rate; 2) to reduce bias due to nonresponse; and 3) to test the effects of differential incentives. The sample for the NRS was selected from three groups of nonresponders in the HRS sample. The first group was comprised of cases that were coded as "final refusals" after many refusal conversion attempts by interviewers as of mid-January 1993. The second group were reluctant cases that had fewer contacts than the first. The third group included all remaining nonresponse cases including the most recent refusals and cases that had no contact to date. These groups were defined with differential response probabilities: low, medium, and high, respectively.

Twelve sample selection strata for the NRS were defined by crossing two cost categories (high/low cost geographic sample areas) by two phone status categories (phone/no phone) by the three response probability categories described above. The cost and phone status categories were derived from the data compiled throughout the field period. The response probabilities were determined by the interviewers who had the most experience with the cases.

Once the sample was selected, several types were excluded. The excluded cases included:

1) cases where the interviewer indicated there was some threat, either to the respondent, the respondent's spouse, or to an interviewer;
2) cases where it was thought to be unethical to make another call;
3) cases where the original respondent had moved out of the sample frame;
4) cases that should have been coded as non-interviews originally; and
5) households in which one eligible person had already completed an interview.

The standard procedure for the HRS was to send a pre-survey letter to the respondent and then provide a post-interview payment of $10 for single respondents or $15 for each respondent in a couple household. In the NRS, letter type and incentive amounts were varied in order to test their effects on response rates.

The NRS sample was then divided into six releases. For each release, a letter with an offer of a much higher incentive was sent via overnight mail to the respondent's home. The letter for Releases 1 and 2 stated the importance of the respondent's participation in the HRS, mentioned that an interviewer would be calling them soon (similar to the standard letter in the HRS), and offered them a $100 incentive for their participation. Releases 3 and 4 were sent the same letter as Releases 1 and 2, but the incentive offered differed for each release: $100 for Release 3 and $50 for Release 4. The incentive offered for Releases 5 and 6 was $100 and $50, respectively; however, the letter asked them to call a 1-800 line if they were interested in participating in the study[1].

For Releases 1-4, interviewers were instructed to contact each household in the NRS two days after the respondent received the overnight mailing. If contact was made with one call and the eligible person(s) refused again, the case was to be coded as a final refusal. However, if contact was not made, or the eligible person(s) wanted to think about their participation before answering, the interviewers were allowed to make up to three more contacts before coding the case as a final refusal.

The NRS achieved a 28% response rate, increasing the overall HRS response rate by four percentage points (not reflecting unequal probabilities of selection).

---

[1] The completed interviews for Releases 5 and 6 (n=11) were substantially smaller than completed interviews for Releases 1-4 (n=597), and were therefore deleted in the analysis that follows.

HRS Wave 2 Sample and Field Procedures

For the second wave of the HRS, interviewers were instructed to organize their sample and contact cases as close as possible to the date of the respondent's interview in 1992. Thus, change could be most accurately measured in two-year intervals. All respondents were to receive a $20 prepaid monetary incentive in Wave 2, regardless of the incentive paid in the previous wave. Consequently, one of the main concerns about approaching the NRS cases was that their propensity to respond would be significantly lower due to the larger incentive offered in Wave 1. No special actions were taken for the NRS cases.

Analytic Plan

The analysis first examines the relationship between the Wave 1 experience and the Wave 2 outcome. Next, it explores some simple three-variable hypotheses, all involving questions of whether the relationship between the Wave 1 experience of the case and the Wave 2 outcome is affected by another attribute. These attributes were chosen from among those variables found to be related to response propensity in earlier research and indicators of the affective and cognitive state of the respondent during the Wave 1 interview. We focus on cooperation rates in Wave 2, excluding no-contact sample cases and other noninterviews from the analyses.

Once those three-variable systems have been examined, the attributes found to be informative to the Wave 2 outcome are included in a multivariate logistic regression model, in order to measure their marginal effects.

The HRS sample design includes stratification, clustering, and unequal probabilities of selection. All results presented are based on weighted data. Standard error estimates and statistical tests reflect the complexity of the sample design, calculated using Taylor series approximation (using SUDAAN, Shah *et al.*, 1993).

Results

The primary analysis for this paper relied on the following grouping of cases in Wave 1: 1) interviews completed without the use of any formal refusal conversion techniques (the "compliant" group); 2) interviews completed after an interviewer reassignment or the mailing of a persuasion letter (the "persuasion" group); 3) interviews completed as part of the Nonresponse Study, as described above (the "large incentive" group)[2].

Table 1 shows the bivariate relationship between the Wave 1 grouping and the Wave 2 response. It can be seen that the compliant group cooperated at a higher rate in Wave 2 than either the persuasion group or the large incentive group. However, our interest in this paper is primarily on differences between those who cooperated in Wave 1 only after a large incentive was offered, relative to those who cooperated after additional persuasion but no large incentive. The cooperation rates for these two groups are essentially the same. There appear to be no net effects of the different experiences in the first wave to the behavior of the persuasion and large incentive respondents in the second wave of the survey. Most notably, offering a large incentive to reluctant Wave 1 respondents followed by a much smaller Wave 2 incentive appears not to yield Wave 2 reactions that are dissimilar to those interviewed in Wave 1 after successful persuasion efforts.

Table 1. Wave 2 Cooperation by Wave 1 Grouping

Wave 2 Outcome	Compliant	Persuasion	Large Incentive
Interview completed	95.5%	89.3%	88.3%
Refusal	4.5%	10.7%	11.7%
Total	100.0%	100.0%	100.0%
(n)	(10,428)	(1,172)	(582)

The relationship between the Wave 1 grouping and Wave 2 response was examined for several independent variables. This was done for two reasons: increased power to reflect the marginal effects of the Wave 1 grouping by controlling for various covariates, and tests of various interaction hypotheses. These three-variable systems (not presented here, for reasons of parsimony) are a means to test a set of interaction hypotheses; that the effects of the Wave 1 experience on Wave 2 outcome varies for different subgroups of the population. The motivation for these hypotheses is basically found in the contrast between the implications of a cognitive script of the interview as a social exchange episode or an economic exchange episode. Those who in Wave 1 come to see the interview as a service provided to the interviewer in return for payment may be more sensitive to the reduction of the payment, relative to those who see the interview as a reciprocating act for some kindness bestowed by the interviewer prior to the request. The independent variables chosen as third variables are reported charitable contributions of $500 or more and the

2. A test was run to check for statistical significance between the $50 (Releases 4) and $100 (Releases 1, 2, and 3) cases and found that there was no difference. Given the small number of cases in Release 4 (n=18), these were collapsed with the $100 incentive cases in the "large incentive" group.

Table 2. Logistic Models Predicting Wave 2 Interview Versus Refusal With and Without Interaction Terms

Predictor Variables	Model 1 Coefficient (S.E.)	Model 2 Coefficient (S.E.)
Constant	2.52 (0.15)**	2.44 (0.15)**
Covariates		
Disabled & not working	0.88 (0.24)**	0.88 (0.24)**
R lives alone	-0.15 (0.19)	-0.34 (0.19)
Wave 1 proxy interview	-1.17 (0.19)**	-1.18 (0.19)**
Income (excluded = $1 - $20K)		
$20K - $40K	0.34 (0.16)*	0.33 (0.16)*
$40K - $60K	0.23 (0.13)	0.23 (0.13)
$60K+	0.48 (0.15)**	0.46 (0.15)**
Predictor Sets		
Wave 1 Group (excluded = Compliants)		
Persuasion	-0.71 (0.15)**	-0.54 (0.22)*
Large Incentive	-1.02 (0.18)**	-0.52 (0.32)
≥$500 Given to Charity	-0.008 (0.08)	0.11 (0.10)
R enjoyed Wave 1 Interview	0.57 (0.11)**	0.65 (0.15)**
Interactions for Wave 1 Group and Charity:		
Persuasion, ≥$500 to charity	--	-0.22 (0.26)
Large Incentive, ≥$500 to charity	--	-0.68 (0.31)*
Interactions for Wave 1 Group and Enjoyment:		
Persuasion, enjoyed Wave 1 interview	--	-0.21 (0.32)
Large Incentive, enjoyed Wave 1 interview	--	-0.52 (0.36)

* p<.05, ** p<.01
Unweighted n=12,182

interviewer's rating of whether the interview was enjoyable for the respondent in Wave 1. In addition, disability status, household size, proxy status in Wave 1, and income, major predictors of response propensity, are included in the analysis. In the HRS, those who are not disabled, those who lived alone, those for whom a proxy interview was necessary in Wave 1, and those in lower income groups are less likely to provide a Wave 2 interview.

Marginal Effects of Persuasion and Incentives on Wave 2 Response. Table 2 (Model 1) shows the results of a multivariate logistic regression, including the controls of the covariate variables (disability, living alone, a Wave 1 proxy interview, and income). This shows that the reluctant respondents in Wave 1 (p<.01) and those accepting the large incentive to provide a Wave 1 interview (p<.01) are each less likely to provide a Wave 2 interview than those freely cooperating with the Wave 1 request (the "compliants"). There is no statistically reliable difference between the persuasion group and the large incentive group.

Charitable Contributions. Kennickell and McManus (1993) found some support for the hypothesis that those who give larger amounts to charity, controlling on income and assets, have higher response propensities in a one-time survey. We expected that those who give

larger amounts to charity would be more likely to view the survey request from the perspective of a civic duty, an act of altruism, or an example of a reciprocating favor as part of the loose social exchange between themselves and the polity. In a bivariate analysis (not shown here), the amount of charitable contributions shows only minor effects on the Wave 2 response behavior, in the hypothesized direction. In the multivariate model (Model 1 of Table 2) we find no marginal effect of charitable contributions on Wave 2 cooperation.

Further, the use of charitable giving provides no insight into how the "persuadeds" differ from the "large incentive" cases in their Wave 2 outcomes (see Model 2 of Table 2). That is, just as in the bivariate analysis, the two groups are indistinguishable in their Wave 2 cooperation rates among persons giving similar amounts to charity.

Charitable giving does, however, illuminate another aspect of Wave 2 behavior. Among compliants, those giving large amounts to charity tend to have higher Wave 2 cooperation rates than those who don't ($p<.05$). This was the originally hypothesized outcome -- altruism in charitable giving would be consistent with continued cooperation with survey requests. This is not true for those expressing reluctance prior to granting the Wave 1 interview (either the "persuasion" or the "large incentive" groups). Among those initially resisting the first interview, the persons with large charitable donations tend to drop out of the panel ($p<.04$ for persuadeds, $p<.02$ for large incentive group).

In summary, charitable giving might be viewed as a proxy for altruistic activities more generally or for acceptance of civic duty. We've seen a small positive tendency for those who give larger amounts to charity to respond at higher rates in Wave 2. But this tendency is only found among those who freely and easily cooperate with the Wave 1 request. If those giving larger amounts to charity do not code the Wave 1 request as attractive (i.e., they express reluctance eliminated only by persuasion or a large incentive), then they are less likely to respond in Wave 2 relative to those who give less to charity. We believe that the intermediate experience acts to define the Wave 2 request as outside the domain of altruistic activity. From that perspective, those giving to charity are even less likely to find the Wave 2 request attractive.

Enjoyment of the Wave 1 Interview. Does enjoyment of the Wave 1 interview disclose any unanticipated effects of the incentive? Can enjoyment of that interview in some way neutralize the potentially harmful effects of persuasion or make less salient to high incentive respondents that their Wave 2 incentive is much lower?

Wave 1 respondents who enjoyed the interview are more likely to cooperate in Wave 2 ($p<.01$). This in itself is not surprising. For the large incentive group, however, enjoyment of the interview appears to play a different role in influencing Wave 2 behavior. Enjoyment of the interview for those receiving a large incentive appears to have no effect on cooperating with the next interview request. In this regard, they are different from the "persuadeds," who exhibited the expected relationship (enjoyment of the first interview leading to higher cooperation rates in the second wave). This relationship reaches statistical significance ($p<.03$) in a model with covariates and the enjoyment interaction terms only (not shown in Table 2); and is marginally insignificant ($p<.08$) in Model 2, containing both charity and enjoyment interactions.

One interpretation of this result is that the memory of the large incentive diminished the salience of the enjoyment of the Wave 1 interview when the person was faced with the Wave 2 request. This is consistent with a cognitive dissonance perspective on the experience. The larger the incentive, the greater the likelihood that the respondent will infer that he/she performed the interview because of the incentive. In contrast, those persuaded into the Wave 1 interview, who then enjoyed the experience might have greater tendencies to recall the enjoyment as justification for their prior actions.

If the incentive dominates the recalled memory of Wave 1, then the contrast with the lower incentive in Wave 2 versus Wave 1 would tend to depress Wave 2 cooperation rates. The usual positive effects of pleasant memories of the first interview are nullified.

Summary and Conclusions

Wave 2 cooperation rates are higher for those persons who provide the Wave 1 interview without need for strong persuasive action by the interviewer (the "compliants"). Those who provide the interview after normal persuasion efforts (the "persuaded" group) and those who accept the large incentive (the "large incentive" group) exhibit lower Wave 2 cooperation rates. These effects were measured net of the effects of several other correlates of attrition in the longitudinal survey.

How should these measured effects be interpreted from the practical field perspective? Were there net positive or net negative effects of providing the large incentive on nonresponse rates?

The better contrast for the large incentive group is the persuasion group rather than the compliants. Indeed, the majority of those eligible for the large incentive experiment were persons who did not provide an interview based on persuasive efforts alone. At that crude level of analysis, there appears to be no higher panel attrition due to using a large incentive to bring in the most reluctant cases to Wave 1 requests. That is, the large

incentive group is indistinguishable from the "persuadeds" on their Wave 2 cooperation rates.

After checking on differential incentive effects across income, disability status, and household size groups, the variable that appears to exhibit an interaction with the incentive is the level of enjoyment of the interview. When the charity interaction is dropped from the full model (Model 2), the enjoyment interaction approaches traditional levels of significance (p=.073). For this variable the incentive group behaves differently than the persuaded group. It does not show the common effects of enjoyment of one interview leading to higher levels of cooperating with the next interview request. Instead, it appears that the fact that they responded in Wave 1 only because of a large incentive neutralizes the usual enjoyment effects. Their memories of the incentive might dominate their approach to the Wave 2 request.

Another attribute of sample persons appears to offer other insight into the use of large incentives on the most reluctant -- their charitable giving. Here, the expected tendency is that those who give lower amounts to charity will tend to drop out of the panel versus those who give higher amounts. This appears true only among those freely giving the Wave 1 interview. The intermediate experience of reluctance to Wave 1 (whether it was removed by persuasion or a large incentive) is sufficient to reverse the relationship. That is, those initially objecting to the Wave 1 interview exhibit no positive effects of their charitable giving on panel retention.

Finally, some speculation about the theoretical underpinnings of these results is in order. Overall, there is support for the conclusion that commitment to a longitudinal survey is not marginally harmed by large incentives in the first wave as a method to induce entry into a panel. This implies that the large incentives do not set in place expectations that large incentives will also be offered in later waves of the panel. The reaction of these Wave 1 respondents to the Wave 2 smaller incentive is not distinguishable from those of other respondents who were reluctant to respond in Wave 1 (but entered the panel after succumbing to persuasion). One vehicle through which this could take place is that the cognitive reconstruction of the Wave 1 interview profits from the positive affect of the $100 incentive.

This argument speculates that few respondents engage in the computation of relative benefit between Wave 1 and Wave 2 in making their Wave 2 decision. Instead, they either believe that they face a surplus of reward (from the $100) relative to the Wave 1 burden, and thus owe some reciprocation to the survey organization, or feel a general positive affect toward the first experience and approach the decision on the Wave 2 interview with that perspective.

References

Berk, M., Mathiowetz, N., Ward, E., and White, A. (1987), "The effect of prepaid and promised incentives: Results of a controlled experiment," *Journal of Official Statistics*, 3 (4): 449-457.

Council of Professional Associations on Federal Statistics (COPAFS) (1993), "Providing incentives to survey respondents." Washington, DC: COPAFS (Final report submitted to the Regulatory Information Service Center).

Groves, R.M. and Couper, M.P. (1995), "Contact-level influences on cooperation in face-to-face surveys." *Journal of Official Statistics*, forthcoming.

James, J.M. and Bolstein, R. (1990), "The effect of monetary incentives and follow-up mailings on the response rate and response quality in mail surveys." *Public Opinion Quarterly*, 54: 346-361.

Juster, F.T. and Suzman, R. (1993), "The Health and Retirement Study: an overview." HRS Working Paper Series, Core Paper No. 94-1001.

Kennickell, A.B. and McManus, D.A. (1993), "Sampling for household financial characteristics using frame information on past income." *Proceedings of the American Statistical Association, Survey Research Methods Section*, 88-97.

Kulka, R.A. (1994), "The use of incentives to survey 'hard-to-reach' respondents: A brief review of empirical research and current practice." Paper presented at the COPAFS Seminar on New Directions in Statistical Methodology, Bethesda, MD. May.

Shah, B.V., Barnwell, B.G., Hunt, P.N. and LaVange, L.M. (1993), *SUDAAN User's Manual, Release 6.34*. Research Triangle Park, NC: Research Triangle Institute.