EVALUATION OF A MONETARY INCENTIVE PAYMENT EXPERIMENT IN THE NATIONAL LONGITUDINAL SURVEY OF YOUTH, 1997 COHORT

A. Rupa Datta
Michael W. Horrigan
James R. Walker[1]

The evidence on incentives in survey literature shows that the effectiveness of incentives varies considerably with the topic of the survey, the nature of the respondent, the amount of the incentive, the survey sponsorship, and the form and timing of the payment. (See, for instance, Council of Professional Associations on Federal Statistics, September 1993; Ezzait-Rice, White, Mosher and Sanchez, 1995; Kulka 1995; Groves and Couper, 1998 [chapter 10]; Singer, Van Hoewyrk, Gebler, Raghunathan and McGonagle, 1999; and Singer, 2000.) The literature, however, is sparse on the special case of panel surveys, although anecdotal evidence is available—especially on the effect of unchanging incentive payment levels over several years as has been the case in the National Longitudinal Surveys (NLS) Program. We contribute to the literature on the effect of incentive fees in panel surveys by providing preliminary analysis of results from an incentive fee experiment conducted in the fourth round of interviews (conducted between November 2000 and June 2001) of the National Longitudinal Survey of Youth, 1997 Cohort (NLSY97).

Respondents of the NLSY97 were born in the years 1980-84 and were 16 to 20 years old in the fourth round of annual interviews. Although the response rates for the first three rounds of interviewing were above 90 percent, the response rate for Round 3 was slightly lower than that of Round 2. The payment incentive during the first three rounds was constant at $10 and was paid by the field interviewer at the completion of the interview. Concerned over the decline in responses rates and its implication for future rounds, the sponsor of the study, the Bureau of Labor Statistics (BLS), applied for and received approval from the Office of Management and Budget (OMB) to conduct an experiment manipulating the timing of payment and the level of respondent fee paid. The interviews for the NLSY97 are conducted by the National Opinion Research Center (NORC) of the University of Chicago under contract to the BLS. The experiment had two treatments: (1) payment of the fee prior to scheduling the in-person interview; and (2) increases in respondent fees.

Key words: Incentive payments; response rates; advance payment; longitudinal surveys.

1. INTRODUCTION

One principle underlying the payment of incentives is the idea that paying the incentive in advance generates a feeling of obligation on the part of the potential respondent that increases the likelihood of an interview. It may be also operationally preferable to pay respondents in advance of the interview because one of the considerable expenses in the survey is the time required for interviewers to establish contact with respondents. If respondents are paid in advance, they may become more cooperative and make it easier for the interviewers to contact the respondent and schedule an appointment for an interview. The experiment included sending the incentive in a

[1] A. Rupa Datta, National Opinion Research Center, University of Chicago; Michael W. Horrigan, Bureau of Labor Statistics; and James R. Walker, University of Wisconsin-Madison

mail communication before the interviewer attempted contact. This is closer to the principle of advance payment as the payment will arrive before the potential respondent is actually asked to give up his or her time.

The second element of the experiment was to vary the amount of the incentive and compare the impact of response rates for different categories of panel members. Three incentive levels were offered for Round 4: $10, $15 and $20. The rationale underlying an increase in the fee is the belief that the $10 incentive payment has become increasingly ineffectual in gaining respondent cooperation, and if unchanged, maintaining respondent cooperation and participation will become more difficult in future rounds. Older (youth) respondents are more likely to combine school and work activities and consequently have greater demands on their time. They have more independence and greater discretion over the allocation of their time. Indeed, many move out of their childhood residence and establish independent households. As a result, parents, many of whom have a deep interest in the survey, exert less influence on their children's cooperation.

To investigate the effect of the treatments across different levels of respondent willingness to participate, respondents were classified into "easy," "average," and "difficult" types, based on the level of cooperation in previous rounds of the NLSY97. This may allow identification of whether and for which groups incentives have the largest effect. The experiment implements a random split-sample balanced design with two payment options (advance payment and payment at the time of the interview), three payment amounts ($10, $15, and $20), and three classifications of respondent types ("easy", "average", and "difficult"). The factorial design has 18 experimental cells (2 x 3 x 3). The unit of analysis is the respondent, and the balanced design (that is, approximately equal sample sizes in each cell) yields 450-500 respondents per cell.

The next two sections provide background and context for understanding the experiment. Section 2 describes the experimental design while section 3 presents the procedures used to implement the experiment in the field.

2. EXPERIMENTAL DESIGN

Two operational issues guided the experimental design. First, we required that respondents in the same household (or more generally those we had reason to believe may be in close communication) be assigned to the same experimental treatment group. Payment of different incentive fee amounts to siblings may create ill will that could have consequences for future rounds of the survey. And the notion of "advanced" notification is not be meaningful unless all respondents in close communication are handled identically. We investigated allocating respondents in the same city or in the same school to the same treatment groups. Since most of the respondents live in larger urban areas the risk of sharing information among respondents in the same city is low. And, upon investigation, only a handful of respondents were (previously) in schools with a large number of other respondents. And, with 40 percent of the respondents beyond high school we decided the risk of contaminant of treatment groups to be small compared to the logistical problems involved.

The second design issue is that we wanted to know whether incentives could be targeted to particular subgroups of respondents and thereby maintain responses rates, data quality and survey costs in the rounds ahead. Indeed, we were most interested in whether incentives can be tailored to reach the "hard-to-interview" cases. Consequently, to address possible interactions between treatments and respondent-types we included "Interview-Effort" (IE) as an experimental factor within our experimental design. We briefly describe next the construction of the respondent-type

based on Interview-Effort.

2.1 Individual Respondent Types

Information on the number of contacts by field interviewers with respondents from Rounds 2 and 3 was used to categorize respondents into three effort groups – low, medium, or high – corresponding to our respondent types of easy, average, and difficult, respectively. The Round 2 information is from NORC's Case Management System (CMS) Record of Calls, which contains a record of contact attempts, the number of times the interviewer spoke to the respondent, appointments made (and broken)--among other measures. The Round 3 information is from field interviewers, who were asked to complete a short questionnaire about the interview experience before transmitting the completed interview record. The data from the Round 2 CMS is more objective than the data provided by the Round 3 field interviewer questionnaire. However, the timing of the Round 4 fielding necessitated using the more readily available Round 3 field interviewer questionnaire data, since the Round 3 CMS files were still being post-processed when the experimental cells were being assigned.

We used two measures of contacts from the CMS for Round 2. The first is the most inclusive definition of contacts counting virtually *any attempt* to speak with a respondent (e.g., personal visit, phone call with respondent, phone with household member or left message on the answering machine). The second measure considers only the number of times the interviewer spoke with the respondent (either in person or on the phone). For Round 3, field interviewers estimated the number of telephone calls (YIR-1765) and the number of personal visits (YIR-1770) needed to complete an interview.

To approximately balance cell sizes and use all available sources of information we employed a hierarchical assignment scheme to define respondent types. That is, we first assign respondents to the "high" effort category and then assign respondents to the "medium" and "low" categories. Effort groups are defined so the three groups are about the same size.

The highest effort group includes any respondent who was not interviewed in either Round 2 or Round 3, or broke an appointment in Round 2[2] or required more than 8 personal interactions in Round 2 or more than 4 personal visits in Round 3.

We define a continuous index of effort, the total number of contacts for Rounds 2 and 3, which was obtained by summing the first contact measure from the Round 2 CMS with the (field interviewer) estimated number of phone calls from Round 3. Midpoints for the interval categories for the Round 3 measure are used without adjustment.[3] The most inclusive measure of contacts from Round 2 and telephone calls in Round 3 had the highest correlation among the contact measures available.[4] There is high correlation among measures within a round and

[2] 1,077 respondents were not interviewed in either round 2 or 3. 7,955 were interviewed in both Round 2 and Round 3. 442 respondents were interviewed in Round 2 but not in Round 3; and 255 were interviewed in Round 3 but not in Round 2, and 370 were not interviewed in both Round 2 and Round 3. About 15 percent of the cases ever broke an appointment in round 2. Of those who broke an appointment, 64 percent broke only one appointment and another 18 percent broke exactly two appointments.

[3] The top coded value, eight or more, is entered as 8.

[4] Other information from the Interviewer Remarks Section (Round 3) was investigated but found to be less useful. For example, 70 percent of the cases are reported as either 9 or 10 on the 10-point cooperative scale (YIR-100). There is little difference in average level cooperation in round 3 with any of the round 2

(surprisingly small) correlation across rounds. The low correlation across rounds might be due to the different methods of measurement (e.g., self-reported in Round 3 versus record of calls information in round 2). It may also be due that personal circumstances of the respondents vary independently from year to year; a difficult respondent in one year may not be in the next.

We define the medium and low effort groups according to the distribution of the constructed measure of total number of contacts for rounds 2 and 3. Based on the hierarchical ordering of these groups, after the assignments to the 'high' effort group were made, the remaining individuals who were eligible for assignment to either the medium or low effort groups were ones who were interviewed in both rounds 2 and 3. The medium effort group is defined as a total number of contacts for Rounds 2 and 3 of at least 8. The low effort group is defined as a total number of contacts fewer than 8.

Table 1 presents descriptive statistics and the frequency distributions of the total number of contacts for rounds 2 and 3 by respondent type. The medium effort group required about 50 to 100 percent more contacts than did the lowest effort group. The high group is not included in table because the constructed measure of total contacts is only defined for cases interviewed in both rounds 2 and 3, whereas the bulk of cases in the high effort group missed at least one interview. Nevertheless, for the restricted subgroup of the high-effort group that was interviewed in both rounds 2 and 3, the number of contacts is 20 to 40 percent higher than the medium-effort group; the mean is 14.1 and the median is 12 contacts and 17 (versus 11 for the medium group) at the third quartile.

Table 1
Distribution of the Constructed Measure Total number of Contacts Rounds 2 and 3
Number of Contacts by Respondent Type

				Percentile				
Effort	**Count**	**Mean**	**Std dev**	**10**	**25**	**50**	**75**	**90**
Medium	2719	10.5	3.2	8	8	10	11	14
Low	3300	5.8	1.1	4	5	6	7	7

Note:
Number of contacts equals the number of attempts to speak with the respondent in Round 2 plus the number of phone conversations in Round 3. Tabulation is done prior to adjustment for cases in the same household.

2.2 Procedure to assign common treatment to respondents of the same household.

Random assignment of treatments occurred within each of the experimental factor Interviewer Effort groups, subject to the condition that all respondents living in the same (Round 1) household be allocated to the same treatment groups.

To ensure that respondents from the same household are treated identically, we randomly select a respondent in each household and define that respondent to be the "target" respondent. The respondent types of all members of the household are set equal to that of the "target" respondent.

Assignment of treatments within respondent types is made through the use of the household

---

contact measures.

identification number. Respondents selected for receipt of an advance payment have household identification (id) numbers divisible by 2 (i.e., even numbered ids). Assignment of the incentive fee level is done similarly. Table 2 reports the number of respondents by treatment status and experimental factor after adjustment for respondents from the same (Round 1) households.

Table 2
Number of Respondents by Treatment Status
And Experimental Factor Interview-Effort

	Pay at interview			Pay in advance		
Effort	**Fee $10**	**Fee $15**	**Fee $20**	**Fee $10**	**Fee $15**	**Fee $20**
High	528	544	457	477	492	458
Medium	447	448	443	477	464	424
Low	537	565	560	561	522	570
Total	1512	1567	1460	1515	1478	1452

2.3 Implementation of the Incentive Experiment

When developing procedures for the incentive experiment, we focused on making minimal changes to existing field procedures, maximizing field interviewers' ability to implement the experimental design accurately, and building in flexibility to adapt to unforeseen conditions that might evolve.

All NLSY97 respondents receive an advance letter prior to each round's data collection. The advance letter mailing occurs approximately six weeks prior to the start of fielding, so that address updates and potential unlocatables can be identified and communicated to the appropriate field managers and interviewers before interviewing begins. Round 5 advance letters were mailed out in the third week of September 2000.

To simplify the packing process and lessen the chance for error, only one experimental pre-pay group was to be handled each day. On the first three days mail shop clerks and central office staff would stuff letters for one of the three pre-pay categories each day with all the post-pay letters being printed and mailed on the fourth day. This did not work out exactly as planned because the task proved to be far more time consuming than projected. However, the letters were handled in sequence and we finished one experimental group before starting the next. Project staff checked 100 percent of all pre-pay envelopes to ensure that the respondent name on the letter matched the envelope, that the amount mentioned in the letter matched the amount enclosed, and that the enclosed amount was in accordance with the experimental design. Early in the following week, all post-pay envelopes were printed and mailed. Historically, the advance letter has not mentioned compensation. To maintain that tradition, the post-pay letter made no mention of respondent incentive.

The U.S. Mail returned approximately 381 letters due to incorrect addresses. These letters included $700 of incentive cash, indicating that more post-pay letters came back than we would expect randomly. Where possible, mail-shop staff re-mailed letters to revised addresses. Field managers had access to information about which letters came back to NORC and which were re-mailed.

To communicate each respondent's treatment status to the interviewer, we recorded a variable on the electronic face sheet, a screen in the interviewer laptop that serves as the primary source of respondent-specific fielding information (e.g., contact information, demographic information, last interview dates, etc.). The electronic face sheet identified respondents as 'Pay XX' or 'Pre-pay XX,' with the appropriate dollar amount specified. The interviewers would then know, prior to contacting each respondent, whether or not that respondent had already been paid and how much the respondent was (to be) paid.

We held separate training sessions on the incentive experiment with both the field managers and the field interviewers. Training for Field Managers came first. For the most part, they were not initially particularly receptive to the experiment. They were especially concerned about: respondents becoming hostile if they learned of the experiment; their interviewers losing a conversion tool for those respondents paid in advance (in addition to respondents forgetting that they had received the incentive by the time the interview would take place); and, not knowing what the incentive amount would be in future rounds. By the end of training, they were much more comfortable with the experiment and understood its potential benefit for maintaining response rates in future rounds.

Experienced interviewers were trained with a home study and their concerns (which echoed the concerns voiced by the Field Managers) were addressed by the Field Managers. Interviewers without prior project experience were trained in person. These interviewers—who did not have any experience with the payment procedures employed in previous rounds--did not raise the same levels of concern as did the Field Managers or the experienced home-trained interviewers. The training for all field staff included: a briefing on the purpose of the experiment; what had been done so far (i.e. the advance letter mailing procedures); where to locate each respondent's experimental grouping (Electronic Face Sheet); the importance of not disclosing the existence of the experiment to the respondents; the procedure for paying the post-pay respondents; and handling respondent concerns regarding the change in payment amount (or how to handle respondents that may inadvertently become aware of the experiment).

For the respondents who were to be paid at the time of interview, the field interviewers were allowed to disclose that respondent's amount as an incentive while attempting to schedule an interview. For two-thirds of such respondents, the amount reflected an increase from prior rounds. These respondents, however, did not receive the incentive until completion of the interview. For these post-pay respondents, the interviewers were provided cash advances from their Field Manager and were required to obtain a signed receipt from the respondent upon payment (which was identical to procedures followed in previous rounds).

We initially had two concerns regarding the experiment. The first concern involved the possibility of respondents learning of the experiment from other respondents who received a different incentive amount or timing. This was a valid concern because of the randomness of the experiment assignments. We maintained a household level of selection; therefore everyone in a household was placed in the same experimental group, however persons in the same neighborhood might receive different amounts or timing of payment. In actuality, there were relatively few such cases during the field period. When this situation did arise, the interviewer was authorized to pay the lower-paid respondent the amount of the higher-paid respondent. The interviewers were also instructed to explain to the respondents that they were assigned to the different experimental payment groups randomly, similar to the way they had been chosen randomly to be a respondent in the survey. This proved to be a useful explanation, as the

respondents had been introduced to this concept of random selection (and the fact that they 'represent' persons their age) in previous rounds of the survey.

The second concern related to the possibility that a pre-pay respondent might not receive their incentive amount. This could happen for a number of reasons, including: lost/misdirected mail; the first sibling to get the mail taking all the letters intended for other siblings in multiple respondent households; parents taking the incentive (especially in low-income families); etc. This proved to be a far more costly and common occurrence than was expected. In the end, we paid (or re-paid) one-third of the pre-pay respondents at the time of the interview. Explanations for these discrepancies include the reasons cited above, as well as (one of the field personnel's initial concerns) the respondent forgetting that they had received the incentive and most likely some respondents dishonestly reporting that they had not received the advance incentive.

An additional problem, which was not highly anticipated, was that many parents disliked the fact that we had sent cash in the mail. Also, many parents disliked the advance payment in principle and saw it as an attempt to coerce the respondent into cooperating. Some went so far as to return the money.

As mentioned above, the field staff in general did not like the experiment. Their main concern was that by pre-paying the respondents they were losing a conversion tool. For those pre-paid respondents that completed the interview the interviewers reported feeling awkward without having something to give to the respondents in appreciation for their time and cooperation. In the end, there was no statistical difference between the completion rates of the experimental groups based on the timing of the payment. It could be interpreted that the lack in statistical difference between the pre-pay and post-pay respondents was affected by two factors. First, the reactions by field interviewers toward the experiment suggest that they may have treated their pre-pay respondents in a manner that was different than their post-pay respondents. Second, the unexpected necessity of having to pay a third of the pre-pay respondents a second time (effectively becoming a post-payment). Both of these factors may serve, in combination or isolation, to minimize the natural differences between the two methods.

3. RESULTS

3.1 Evaluation factors

We compare response rates and measures of interviewing effort. Given the high levels of response rates attained in Rounds 1 through 3 (and more generally within the NLS program) we expected there may be little direct effect on the response rate, and instead, anticipated more effect on the effort needed to obtain a completed interview and/or on the level of effort put forward by the respondent in the interview process. Besides the completed response rate, we measure the number of interviewer contacts with the respondent and members of the respondent's household, the number of phone calls required to complete a case, and the length of the interview.

3.2 Findings

Table 1 provides the response rates associated with each of our experimental conditions. The overall response rate was 89.7 percent. Based on our pre-classification of cases, the response

rates were 77.2 percent, 94.4 percent, and 97.1 percent for cases pre-classified as difficult, average and easy, respectively. This pattern of response rates showing substantial differences between cases pre-classified as difficult as compared to easy or average are also evident for comparisons by the timing or the amount of incentive payments. Within each pre-classification group there appears to be little difference in response rates based on either the timing or amount of incentive payments.

A more formal examination of the contribution of our experimental conditions to differences in the likelihood of completing a case was conducted using a logistic regression specification. The results of our analysis are presented in Table 2. The first column of Table 2 presents the results based on the contribution to the log likelihood of completing (or not completing) a case of the three payment levels, the timing of the incentive payment, the three original pre-classification types, and controls for sex, race and ethnicity, and the eligible respondents' age on December 31, 1996. The excluded categories for explanatory indicator variables are the $10 incentive payment level, cases pre-classified as easy, payment at the time of the interview, females, and non-black/non Hispanics.

Based on these controls, respondents that were pre-classified as either difficult or average had significantly lower likelihoods of completing a case as compared to respondents pre-classified as easy to convert. Blacks were more likely than non-black/non-Hispanics to complete a case, and the likelihood of completing a case decreases significantly with age. In this specification neither the timing nor the amount of the incentive payment had a significant impact on the likelihood of completing a case.

The second column of the table includes a redefined incentive payment variable that contrasts an incentive payment of $20 against payments of either $10 or $15. As well, an interaction term is included which examines the impact of the $20 incentive payment on the cases that were pre-classified as the hardest to convert.

Once again, the timing of the payment does not have an impact. Those pre-classified as either difficult or average have significantly lower likelihoods of completing a case as compared to those pre-classified as easy. In this case, those receiving the $20 payment were more likely to complete an interview as compared to those receiving either $10 or $15. However, the interaction term measuring the impact of the $20 incentive payment on the hardest to convert cases did not have a significant impact. This implies that a strategy of targeting higher fees to the hardest cases will not be an effective policy for maintaining response rates in future rounds. As in the first specification, blacks were more likely to complete a case, and the likelihood of completing a case again falls significantly with age.[5]

3.3 Timings, calls, contacts, don't knows and item refusals

We turn now to an examination of the impact of the timing and the amount of incentive payments on 5 separate dependent variables: case timings; number of calls required to complete a case;

[5] We also examined a third specification, which included a new respondent type variable in which respondents pre-classified as difficult are compared to respondents pre-classified as either easy or average. This specification included the contrast between a $20 incentive payment against payments of either $10 or $15. An interaction term was included which examined the impact of the $20 incentive payment on the cases that were pre-classified as the hardest to convert. The results were the same as the ones just reported for the second specification.

number of contacts required to complete a case; percentage of questions answered with a response of 'don't know'; and the percentage of questions answered with an item refusal recorded. Analysis of variance decomposition techniques were used to make our assessments of statistically significant differences between treatment levels of each explanatory variable in the model. For each of the dependent variables, the following explanatory variables used were:

Fee payment levels	$10 or $15, $20
Payment timings	Advance payment, payment at the time of interview
Respondent types	Pre-classified as easy, average, or difficult
Race ethnicity	Black, Hispanic, mixed race, non-Black/non-Hispanic
Sex	Male, female

3.3.1 Timings

Table 3 provides evidence on the impact of our experimental conditions on the average case timings. An average time of 67.3 minutes was recorded across all interviews. Analysis of variance tests (not shown) indicate that, controlling for the impact of fee payment levels, payment timings, race/ethnicity, and gender, the average time for cases pre-classified as difficult (68.8 minutes) was significantly higher than cases pre-classified as either easy (66.1 minutes). As well, cases pre-classified as average (67.7 minutes) were significantly longer than easy cases. No statistically significant differences in average timings were observed either from the amount or the timing of the incentive payment. Blacks had significantly longer interview times than Hispanics who in turn had significantly longer interview times than non-black/non-Hispanics. Finally, females had significantly longer interview times than males.

3.3.2 Calls

Table 4 examines the impact of the incentive experiment on the average number of calls recorded among complete interviews.[6] Overall, there was an average of 4.6 calls associated with a complete interview. The comparable numbers for cases pre-classified as easy, average, and difficult were 3.9, 4.8 and 5.5 calls respectively. Analysis of variance tests indicate there are statistically significant differences between each of these pre-classification groups based on the number of calls associated with a completed interview. The number of calls associated with a $20 incentive payment was only slightly lower than the number associated with a $10 or $15 incentive. The timing of the payment did not make a difference. Blacks required more calls than non-blacks/non-Hispanics who in turn required more calls than Hispanics. And males averaged more calls than females.

3.3.3 Contacts

The number of actual contacts required to complete a case was 1.9 contacts overall (see Table 5). An analysis of the number of contacts associated with easy (1.6), average (1.8), and difficult cases (2.5) indicate statistically significant differences based on pairwise comparisons of each pre-classification group. No statistically significant findings emerge across any pairwise comparisons either based on the timing or the amount of the incentive payment. Blacks and Hispanics averaged more contacts than non-blacks/non-Hispanics, as did males as compared to females.

[6] The next version of the paper will examine separately the number of calls attempts associated with cases that did not result in an interview.

3.3.4 Don't knows

Tables 6 and 7 show results for the percentages of questions recording a "don't know" response (table 6) or an item refusal (table 7). On average, 1.40 percent of all questions were accompanied by the answer "don't know". Cases pre-classified as difficult had a higher percentage of don't know responses (1.58 percent) than cases pre-classified as either easy or average (1.31 and 1.36 percent, respectively). Again, neither the timing nor amount of incentive payment had a statistically significant impact. Blacks were more likely to provide a "don't know" response than Hispanics who in turn were more likely to do so than non-blacks/non-Hispanics.

3.3.5 Item refusals

Overall, survey respondents refused to answer 0.28 percent of all questions. Item refusals as a percentage of questions answered were significantly higher for cases pre-classified as difficult (0.46 percent) as compared to either easy or average cases (0.19 and 0.25 percent, respectively). Once again, neither incentive payment level nor timing exhibited any statistically significant differences. Hispanics refused to answer questions at a higher rate than non-blacks/non-Hispanics, as did males as compared to females.

4. DISCUSSION

Overall, we did not find any discernable evidence of a strong and consistent impact from either the timing or the amount of the incentive payment on response rates. Throughout we encountered many difficulties in carrying out the experiment. Costs increased from having to re-pay one-third of the pre-paid respondents, and from paying respondents in the pre-pay group that would never complete the interview. As well, neither the field interviewer staff nor the field management were initially receptive to the experiment. Although most became more comfortable with the experiment after training, they continued to voice concerns, especially over the loss of the incentive payment at the time of the interview as a conversion tool for potential respondents who were paid in advance. In fact, the morale of some interviewers was so affected that a few even threatened to quit as a result of its implementation. Finally, for horizontal equity across respondents we are now committed to paying all respondents $20 to bring everyone to the same incentive amount without lowering the incentive amount received in round four for any of the respondents. In sum, the challenges to mounting this experiment were many, and in some cases, unexpected, and in the end, the results did not point to a clear strategy to adopt in terms of using incentive payments to maintain response rates in the future.

REFERENCES

Council of Professional Associations on Federal Statistics (1993) *Providing Incentives to Survey Respondents. Final Report.* Submitted to the Regulatory Information Service Center, General Service Administration. (Contract Number GS0092AEM0914).

Ezzait-Rice, T., A. White, W. Mosher and M. Sanchez (1995) "Time, Dollars, and Data: Succeeding with Remuneration in Health Surveys." Seminar on New Directions in Statistical Methodology, Federal Committee on Statistical Methodology. Statistical Policy Working Paper 23. pp. 225-255.

Kulka, R. (1995) "The Use of Incentives to Survey 'Hard-to-Reach' Respondents: A Brief Review of Empirical Research and Current Research Practice." Seminar on New Directions in Statistical Methodology, Federal Committee on Statistical Methodology. Statistical Policy Working Paper 23. pp. 256-287.

Groves, R. and M. Couper 1998 *Nonresponse in Household Interview Surveys*, John Wiley & Sons.

Singer, E., J. Van Hoewyrk, N. Gebler, T. Raghunathan and K. McGonagle (1999) "The Effect of Incentives on Response Rates in Interviewer-Mediated Surveys," *Journal of Official Statistics* 85: 217-230.

Singer, E., (2000) "The Use of Incentives to Reduce Nonresponse in Household Surveys." Unpublished manuscript, University of Michigan.

Table 1. Response rates		Respondent type		
Experimental Cell	**Total**	**Easy**	**Average**	**Difficult**
All	89.7%	97.1%	94.4%	77.2%
Timing of payment				
At time of interview	89.7%	97.2%	94.8%	77.2%
Advance pay	89.7%	97.0%	94.1%	77.2%
Amount				
$10	89.5%	97.4%	94.6%	76.1%
$15	88.8%	96.3%	93.0%	77.2%
$20	90.9%	97.5%	95.7%	78.3%
Amount				
$10 or $15	89.1%	96.8%	93.8%	76.7%
$20	90.9%	97.5%	95.7%	78.3%
Race Ethnicity				
Black	90.7%	98.0%	93.0%	82.0%
Hispanic	89.0%	96.2%	95.0%	77.4%
Mixed Race	94.0%	100.0%	94.4%	80.0%
Non-black/Non-Hispanic	89.4%	96.9%	94.8%	74.1%
Sex				
Male	89.2%	97.0%	93.3%	77.5%
Female	90.2%	97.2%	95.5%	76.8%

Table 2. Logistic regression results
Dependent variable: Completing or not completing an interview

Explanatory variable	Model 1*	Model 2*
Intercept	6.3781	6.3
Timing of payment		
In advance of interview	-0.0413	-0.0389
At time of interview	…..	….
Incentive payment (A)		
$10	…..	x
$15	-0.0577	x
$20	0.1361	x
Incentive payment (B)		
$10 or $15	x	….
$20	x	0.3375*
Respondent type		
Easy	…..	x
Average	-0.6351**	-0.6325**
Difficult	-2.2451**	-2.1730**
Interaction of respondent type and incentive payment		
$20 and Difficult	x	-0.249
Other	x	….
Race/Ethnicity		
Black	0.3210**	0.3200**
Hispanic	0.0900	0.0881
Mixed Race	0.4235	0.4187
Non-black/Non-Hispanic	….	….
Sex		
Male	-0.0494	-0.0497
Female	….	….
Age as of December 31, 1996	-0.2104**	-0.2104**

* indicates the coefficient is statistically different from zero at the 95% level of confidence
** indicates the coefficient is statistically different from zero at the 99% level of confidence

Table 3. Average timings in minutes		Respondent type		
Experimental Cell	**Total**	**Easy**	**Average**	**Difficult**
All	67.3	66.1	67.6	68.8
Timing of payment				
At time of interview	67.3	66.1	67.7	68.4
Advance pay	67.4	66.1	67.5	69.1
Amount				
$10	67.6	66.6	68.6	67.9
$15	66.9	65.7	67.2	68.3
$20	67.4	66.0	67.0	70.2
Amount				
$10 or $15	67.3	66.1	67.9	68.1
$20	67.4	66.0	67.0	70.2
Race Ethnicity				
Black	70.6	68.9	71.1	72.1
Hispanic	68.0	66.2	68.2	70.0
Mixed Race	69.5	67.1	76.2	55.8
Non-black/Non-Hispanic	65.3	64.7	65.6	66.0
Sex				
Male	66.5	65.4	66.8	67.6
Female	68.2	66.8	68.5	70.1

Table 4. Average number of calls		Respondent type		
Experimental Cell	**Total**	**Easy**	**Average**	**Difficult**
All	4.6	3.9	4.8	5.5
Timing of payment				
At time of interview	4.7	4.0	4.9	5.6
Advance pay	4.6	3.8	4.7	5.5
Amount				
$10	4.7	3.9	5.0	5.6
$15	4.7	3.9	4.8	5.6
$20	4.5	3.9	4.5	5.4
Amount				
$10 or $15	4.7	5.6	4.9	3.9
$20	4.5	5.4	4.5	3.9
Race Ethnicity				
Black	5.0	4.2	5.1	5.8
Hispanic	4.3	3.5	4.4	5.3
Mixed Race	5.0	2.9	5.9	7.8
Non-black/Non-Hispanic	4.6	3.9	4.8	5.5
Sex				
Male	4.8	4.0	4.9	5.7
Female	4.5	3.8	4.6	5.3

Table 5. Average number of contacts		Respondent type		
Experimental Cell	**Total**	**Easy**	**Average**	**Difficult**
All	1.9	1.6	1.8	2.5
Timing of payment				
At time of interview	1.9	1.7	1.8	2.5
Advance pay	1.9	1.6	1.8	2.4
Amount				
$10	1.9	1.7	1.8	2.4
$15	2.0	1.7	1.8	2.5
$20	1.9	1.6	1.8	2.5
Amount				
$10 or $15	1.9	1.7	1.8	2.4
$20	1.9	1.6	1.8	2.5
Race Ethnicity				
Black	2.2	1.9	2.2	2.7
Hispanic	2.2	1.8	2.0	2.7
Mixed Race	2.2	1.4	1.9	5.1
Non-black/Non-Hispanic	1.7	1.5	1.6	2.1
Sex				
Male	2.0	1.7	1.9	2.5
Female	1.9	1.6	1.8	2.4

Table 6: Percentage of questions answered with a response of "don't know"				
		Respondent type		
Experimental Cell	**Total**	**Easy**	**Average**	**Difficult**
All	1.40	1.31	1.36	1.57
Timing of payment				
At time of interview	1.42	1.30	1.44	1.56
Advance pay	1.38	1.32	1.29	1.59
Amount				
$10	1.44	1.38	1.34	1.63
$15	1.38	1.27	1.37	1.55
$20	1.39	1.29	1.38	1.54
Amount				
$10 or $15	1.41	1.32	1.35	1.59
$20	1.39	1.29	1.38	1.54
Race Ethnicity				
Black	1.64	1.50	1.70	1.76
Hispanic	1.45	1.33	1.29	1.74
Mixed Race	1.23	1.10	1.14	1.86
Non-black/Non-Hispanic	1.26	1.22	1.24	1.36
Sex				
Male	1.40	1.29	1.39	1.57
Female	1.40	1.34	1.33	1.59

Table 7: Percentage of questions respondents refused to answer

Experimental Cell	Total	Respondent type		
		Easy	Average	Difficult
All	0.28	0.19	0.25	0.46
Timing of payment				
At time of interview	0.28	0.16	0.24	0.49
Advance pay	0.29	0.22	0.26	0.43
Amount				
$10	0.30	0.21	0.19	0.55
$15	0.27	0.18	0.27	0.39
$20	0.28	0.18	0.29	0.43
Amount				
$10 or $15	0.29	0.19	0.23	0.47
$20	0.28	0.18	0.29	0.43
Race Ethnicity				
Black	0.32	0.24	0.27	0.32
Hispanic	0.38	0.25	0.35	0.39
Mixed Race	0.33	0.49	0.00	0.33
Non-black/Non-Hispanic	0.23	0.14	0.20	0.23
Sex				
Male	0.33	0.22	0.32	0.48
Female	0.24	0.16	0.18	0.44