RESULTS OF THE WAVE 1 INCENTIVE EXPERIMENT IN THE 1996 SURVEY OF INCOME AND PROGRAM PARTICIPATION

Tracy L. James, U.S. Bureau of the Census
U.S. Bureau of the Census, Demographic Statistical Methods Division, Washington, D.C. 20233[1]

KEY WORDS: incentives, longitudinal survey, nonresponse rate

I. INTRODUCTION

Beginning with the 1996 Panel of the Survey of Income and Program Participation (SIPP), the length of the panel, or the time in sample, is being extended from 2 and 2/3 years to 4 years. Without some changes in procedures, nonresponse is expected to rise to 30% or more by the end of the four year panel. A possible effective means of maintaining higher response rates is to offer incentives to SIPP sample households. Research has shown that incentives are effective at reducing nonresponse in mail surveys, but little has been done in personal visit interviews. One of the few studies done on a personal visit survey offered a nice ball point pen, which increased response rates from 76% to 81% (Willimack et al. 1995). SIPP also reported limited success in the 1987 panel with a one-time non-cash incentive (Butler, 1991). More specifically, incentives have been shown to decrease refusal rates (Willimack et al. 1995). Further, research has shown that incentives are most effective with minorities and undereducated persons (Berlin et al. 1992) (Ferber and Sudman 1974). Since this group of persons is more likely to have low incomes, the incentive may have higher value to them. A major objective of the SIPP is to provide measures of economic well-being among the low income population, so it becomes important to keep this population well represented in the SIPP sample. But, nonresponse in the SIPP has been shown to be higher for low income, low educated groups (Census Bureau, 1995). So, methods are needed to keep this nonresponse to a minimum and research suggests that incentives can accomplish this. Another relevant point about incentives is that they can reduce the number of callbacks needed, thereby decreasing the cost per interview. Incentives decrease the initial reluctance to be interviewed, giving way to a higher proportion of interviews done in the first visit. Each interview costs less, since fewer trips are made to the address (Berlin et al. 1992) (Willimack et al. 1995) (Berk et al. 1987) (Ferber and Sudman 1974). Further, prepaid incentives are more effective at reducing nonresponse rates than promised rewards that are conditional on survey participation (Berk et al. 1987). Also, incentives are projected to be most helpful in surveys that place exceptional task burden on respondents (Ferber and Sudman 1974). This experiment is expected to answer the following questions:

1. Do incentives reduce nonresponse at the current interview?
2. Do incentives reduce nonresponse among low income households in the SIPP?
3. Do incentives reduce nonresponse at subsequent interviews? (Does an incentive given at Wave 1 reduce nonresponse at Wave 2, Wave 3, ...)
4. Do incentive reduce nonresponse at subsequent interviews for the low income population?
5. Can the incentive reduce the attrition rate, both overall and for the low income population?
6. Do incentives reduce the number of callbacks needed to obtain an interview?

II. DESCRIPTION OF THE SIPP

The SIPP is a nationally representative longitudinal survey with a multistage sample design. Beginning with the 1996 SIPP panel, low income households were oversampled by means of stratification within each primary sampling unit (PSU). Two strata were created within each PSU, one with a high household poverty rate and one with a low household poverty rate. The high poverty stratum resulted in about 41% households below 150% poverty and the low poverty stratum ended up with about 20% of households below 150% poverty. ("Below 150% poverty" means the household income is less than 1.5 times the household poverty threshold.)

The SIPP uses a longitudinal panel design. A new panel of sample households was introduced in 1996, i.e., the 1996 SIPP Panel. The first round of interviews, or Wave 1, was done April - July , 1996. The panel is divided into 4 subsamples, or rotation groups. For each of the interview months, one rotation group is interviewed. Beginning with the 1996 panel, rotation group 1 is interviewed first, followed by rotation group 2, 3 , and 4. So, rotation group 1 was interviewed in April 1996, rotation group 2 in May 1996, rotation group 3 in June 1996, and rotation group 4 in July 1996. At the initial interview, or Wave 1, of the 1996 panel there were 40,188

[1]This paper reports the general results of research undertaken by the Census Bureau staff. The views expressed are attributable to the author and do not necessarily reflect those of the Census Bureau.

eligible households. At the end of Wave 1, the household response rate was 91.52%.

III. DESCRIPTION OF THE EXPERIMENT

An incentive experiment was done to see whether $10 and/or $20 incentives given at the door at the initial interview would help increase response rates at the current interview and subsequent interviews in the 1996 SIPP. At the initial interview of the 1996 SIPP Panel, i.e., Wave 1, the incentives were given as early as possible in the personal visit to the address in the last three of the four rotation groups. The incentives were introduced to the respondent as " a token of appreciation". The respondents were given a paper voucher that resembled cash with the denomination of the incentive printed in the corners. The respondents were instructed to fill in their name and check their address and return it to the Census Bureau in a preaddressed stamped envelope. They were told that they would receive a check for the amount of the incentive in 2 to 3 weeks. Interviewed and noninterviewed households received the incentive, i.e., incentives were given out regardless of the household interview status. Approximately one third of the sample in the last three rotation groups received vouchers for $10, one third received vouchers for $20, and one third did not receive vouchers, which corresponded to a sample size of about 10,000 households per treatment group.

Rotation groups 2, 3, and 4 were used in the study. Treatment groups were assigned at the stratification PSU level. Typically, a stratification PSU is made up of one or more counties. The PSUs were sorted into 11 blocks based on their 1990 measure of size (number of households). Each block was composed of 23 to 39 PSUs. The PSUs were ordered by size within each block then randomly assigned to the $0, $10, and $20 groups. Generally, interviewers were assigned to only one treatment group. The exception came when cases had to be reassigned due to reluctant respondents or interviewers leaving. Interviewers were aware of the experiment and the treatment groups, which probably affected their motivation for getting completed interviews.

Table 1. Number of Eligible Households Assigned to the Incentive Groups

	$0	$10	$20
Overall	10328	9686	10038
High Poverty Stratum	3185	2602	2898
Low Poverty Stratum	7143	7084	7140

IV. METHODOLOGY

Throughout this investigation, two analytic techniques are employed: t tests and logistic regression analysis.

The t tests are used to examine the outcome at Wave 1. In this analysis, the nonresponse rates for rotations 2, 3, and 4, combined, are compared by incentive group. The significance level used is $\alpha = 0.10$. To further support this analysis, rotation 1 is added into the analysis as a control group. Mock incentive assignments were made to the PSUs in rotation 1, consistent with the assignments made in rotations 2-4. Then, the difference between the nonresponse rates in rotation 1 and rotations 2-4 for the $0 incentive group is compared to the difference occuring in the $20 incentive group. If the $20 incentive is indeed effective, the latter difference should be significantly larger. This analysis is repeated for the two mutually exclusive groups, the high poverty stratum and the low poverty stratum. Additionally, the outcomes at Waves 2 and 3 are also examined in a similar manner.

The logistic regression was used to study the outcome at Wave 1, also. Using this method we were able to control for the size of the PSU, which is assumed to be correlated with response rates. In this way, we guard against misinterpreting the results from a "bad" sample, i.e., one in which the response rates are different between the treatment groups before any incentives are given. The logistic regressions were run using the data from rotations 2, 3, and 4.

Throughout this paper the standard errors for the significant tests include a conservative design effect of 1.6, except for the logistic regression, which uses a stratified jackknife replication procedure performed in WESVAR PC to reflect the complex sample design.

V. RESULTS

Nonresponse Rates in Wave 1

Looking at the data in Table 2 from rotation groups 2-4, where the incentive experiment took place, the $20 incentive significantly reduces nonresponse rates (from 9.13% to 7.51% with a t value = 3.3). Now, if we compare the differences that exist between rotation 1, where no incentive was given, and rotations 2-4, in the $0 group and the $20 group, the difference of the differences is not significant at the 10% level. That is, the difference in the nonresponse rate in the $0 group between rotation 1 and rotations 2-4 (-.23) is not significantly different from the difference occurring in the $20 group between rotation 1 and rotations 2-4 (0.61). So, we might conclude that the incentive may not be the sole

reason for the reduction in nonresponse, that a difference might have existed between the treatment groups even before any incentives were given. But, in the logistic regression analysis, after taking into account the size of the PSUs, the $20 appears to be effective. Households given $20 have odds of responding that are about 1.2 times that of households not receiving an incentive. As in the results using the t tests, there was no significant effect with the $10 incentive.

In the high poverty stratum, the $20 incentive is effective at reducing nonresponse rates (from 9.32% to 5.94% with a t value of 3.96). Also, the differences that exist between rotation 1 and rotations 2-4 in the $0 group and the $20 group, i.e., the difference of the differences, is significant at the 10% level with a t value of 1.68. That is, the difference in the nonresponse rate in the $0 group between rotation 1 and rotations 2-4 (-.16) is significantly different from the difference occurring in the $20 group between rotation 1 and rotations 2-4 (2.38). This suggests that the $20 significantly reduces nonresponse rates in the high poverty stratum, even after taking into account the nonresponse rates in rotation 1, although the test is rather weak. What is encouraging is that the results are replicated in the logistic regression analysis. In the poverty stratum, households given $20 have odds of responding that are 1.6 times that of households not receiving an incentive. There was no significant effect with the $10 in the poverty stratum.

Either looking at the nonresponse rates for rotations 2-4 alone or taking into account rotation 1 nonresponse, using the t test, the $20 does not appear to be effective in the low poverty stratum (comparing the nonresponse rates in the $0 group to the $20 group). In the logistic regression analysis, taking into account the size of the PSU, the $20 incentive does appear to be effective. In the nonpoverty stratum, households given $20 have odds of responding that are about 1.1 times that of households not receiving an incentive.

Nonresponse Rates at Wave 2

Wave 2 nonresponse rates, as reported here, do not include Wave 1 nonresponse. Wave 1 nonrespondents are not attempted to be interviewed in later waves in the SIPP. So, the nonresponse rates for Wave 2 are for "new" nonrespondents, those occurring in Wave 2 specifically.

The $10 and the $20 incentives are effective at Wave 2 at reducing nonresponse rates, looking at the data in Table 3 for rotations 2-4. But, if we take into account the nonresponse rates in rotation 1, we are unable to detect any significant differences. The difference of the differences in the $0 and $20 between rotation group 1 and rotation groups 2-4 (1.03 - 2.36 = -1.06) is not significantly different from 0 (t value = 1.38).

The same results occur in the poverty stratum; looking at the data for rotations 2-4, the $10 and $20 incentives reduced the nonresponse rate by 1.48% and 1.9%, respectively. But the significance is lost when we take into account the nonresponse rates in rotation 1. The difference of the differences (1.22 - 2.83 = -1.61) has a t value of 0.875, which is in part due to the small sample size in rotation 1 from the high poverty stratum.

In the low poverty stratum, the $20 incentive appears to be only mildly effective. Using only rotations 2-4, the difference in the nonresponse rates in the $0 group and the $20 group is 1.07% with a t value of 1.11.

Nonresponse Rates at Wave 3

At Wave 3 the effectiveness of the $20 incentive is evident even after taking into account the nonresponse rates for rotation group 1. The nonresponse rates shown in Table 4 for rotation 1 appear to be moving in the opposite direction between the $0 group and the $20 group compared to rotations 2-4. Although there is no significant differences in the nonresponse rates in rotation 1 among the incentive groups.

Similarly, in the high poverty stratum, the $20 incentive elicits lower nonresponse even after taking into account rotation 1 nonresponse by incentive group. Again, the nonresponse rates for the incentive groups are moving in the opposite direction comparing rotation 1 to rotations 2-4.

In the low poverty stratum, the $20 incentive is effective if we look at the rotation 2-4 data only. If we take into account rotation 1 nonresponse, the difference of the differences (.75 - 2.81 = -2.06) is not significant.

Effect of Incentives on Attrition

We hope that incentives can help to keep households in the SIPP sample in future waves, particularly low income households. Looking at Graph 1 we see the attrition patterns from Wave 1 to Wave 3 for rotation 1, where no incentive was given, followed by that for rotations 2-4 in the three treatment groups. As shown in the graph, the nonresponse may be similar for all four groups at Wave 1, but at Wave 2 and Wave 3 a distinctive pattern emerges. By Wave 3, the attrition rate for the $0 group is about 20.5%, whereas in the $20 group it's only about 19%. The attrition rate for rotation 1 by Wave 3 is about 22%. If we restrict our analysis to the high poverty stratum, focusing on Graph 2, by Wave 3, the attrition

rate for the $0 group is about 21%, whereas in the $20 group it's only 16%. The attrition rate for rotation 1 at Wave 3 is about 21%.

Effect of Incentives on Callbacks Needed to Obtain Interviews

In order to track the number of visits made to each address, a question was added to the CAPI instrument in Wave 1 that the interviewers were instructed to update each time they visited an address. Both the $10 and the $20 incentives reduce the average number of personal visits required to complete assignments, 3.40 and 3.46, respectively, compared to 3.60 for the no incentive group. (The mean number of visits are not significantly different between the $10 and $20 groups.)

Interviewers report monthly to their regional office on the total number of hours they spend and their total mileage working SIPP cases. Using this data, the $20 incentive reduces the hours and miles needed per case, while the $10 incentive increases hours and mileage needed per case. Please note that this data is incomplete; on average only about 78% of the interviewers reported their information. By regional office, the percent of interviewers reporting their data ranged from 55% to 91%.

Table 5. Hours and Miles per Case by Incentive Group

Incentive Group	Hours per Case	Miles per Case
$0	3.34	42.78
$10	3.37	43.07
$20	3.32	40.61

VI. SUMMARY

Looking at the data from rotation groups 2-4, the $20 incentive significantly reduces overall nonresponse rates and nonresponse rates in the high poverty stratum. These were consistent results from the t tests using the data from rotations 2-4 only and the logistic regression taking into account the size of the PSU. Using rotation 1 as a control group, comparing the differences that existed between rotation 1 and rotations 2-4 in the $0 group and the $20 group, the difference of the differences overall was not significant, indicating that the $20 was not effective over all the population. This result is probably due to the larger variance in rotation 1 caused by the smaller sample size. Ultimately, we conclude that the $20 is effective overall and in both the high poverty stratum and the low poverty stratum. From the logistic regression analysis, in the overall population, households receiving $20 have odds of responding that are 1.2 times that of households not receiving an incentive. In the poverty stratum, households given $20 have odds of responding that are 1.6 times that of househpolds not receiving an incentive. In the nonpoverty stratum, households receiving $20 have odds of responding that are about 1.1 times that of households not receiving an incentive.

It seems the $10 and the $20 incentives are effective at Wave 2 at reducing nonresponse rates both overall and in the high poverty stratum, looking at the data for rotations 2-4. But, if we take into account the nonresponse rates in rotation 1, we are unable to detect any significant differences, again, due to the small sample size in rotation 1 from the high poverty stratum. This analysis could be strengthened by running the logistic regression as was done in Wave 1, to see what results could be replicated.

At Wave 3 the effectiveness of the $20 is evident both overall and in the high poverty stratum, even after taking into account the nonresponse rates for rotation group 1. So, the $20 incentive has the potential for affecting response rates multiple interviews later. In this case, a significant effect is found at the third interview, when the incentive was given only at the initial interview.

There is evidence that incentives reduce the number of callbacks needed to obtain a complete interview. This is consistent with the previous literature.

Evident from the decrease in nonresponse rates in the high poverty stratum, the $20 incentive is very likely to reduce the nonresponse bias associated with low income households in the SIPP.

In terms of the effect on attrition rates, the $20 incentive appears to be have a strong effect in helping to retain households in the high poverty stratum.

VII. Future Research

- Analysis on the differences between the responding population in the $0 and the $20 incentive groups. Although the analysis by high and low poverty stratum in this paper is a good indication of how incentives affect the response propensity of the low income population, a more direct route would be to use the actual income reported in Wave 1, in the Wave 2 analysis.
- Analysis on the differences in item nonresponse rates by incentive group as an indication of the affect of incentives on data quality.
- More indepth analysis of the characteristics of the responding population in future waves, to determine how incentives affect the attrition pattern of different subsets of the population.

REFERENCES

Berlin, M., L. Mohadjer, and J. Waksberg, "An Experiment in Monetary Incentives", 1992, *Proceedings of the Survey Research Section of the American Statistical Association*, 393-398.

Berk, M. , N. Mathiowetz, E. Ward, and A. White, "The Effect of Prepaid and Promised Incentives: Results of a controlled Experiment", 1987, *Journal of Official Statistics*, Vol. 3, No. 4, 449-457.

Butler, D., "SIPP 87: Gift Experiment Results", 1991, Internal Census Bureau Memorandum dated April 2 to R. Singh.

Census Bureau, "Profile of Nonresponse", 1995 Internal Census Bureau Memorandum from the Nonresponse Research Team .

Ferber, R. and S. Sudman, "Effects of Compensation in Consumer Expenditure Studies", 1974, *Annals of Economic and Social Measurement*, 319-331.

Willimack, D., H. Schuman, B. Pennell, and J. Lepkowski, "Nonmonetary Incentives in Face-to-Face Surveys", Spring 1995, *Public Opinion Quarterly*, Vol. 59, 78-92.

Table 2. Nonresponse Rates at Wave 1

Incentive Group	Rot 1	Rot 2-4	Differences in Rot 1 vs. Rot 2-4
Overall			
$0	8.90%	9.13%	-.23
$10	9.01%	9.11%	-.10
$20	8.12%	7.51%*+	0.61
High Poverty Stratum			
$0	9.16%	9.32%	-0.16
$10	8.49%	8.11%	0.38
$20	8.77%	5.94%*+	2.38*
Low Poverty Stratum			
$0	8.78%	9.04%	-0.26
$10	9.19%	9.47%	-.28
$20	7.84%	8.15%+	-.35

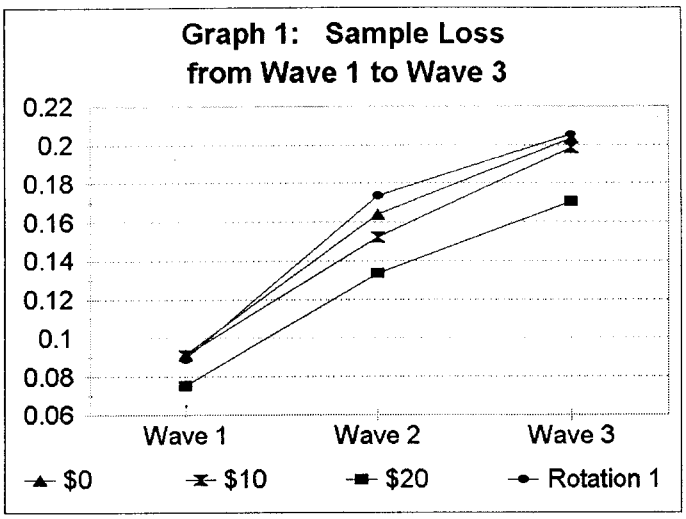
*Significantly different from the $0 group at the .10 level
+Significantly different from the $10 group

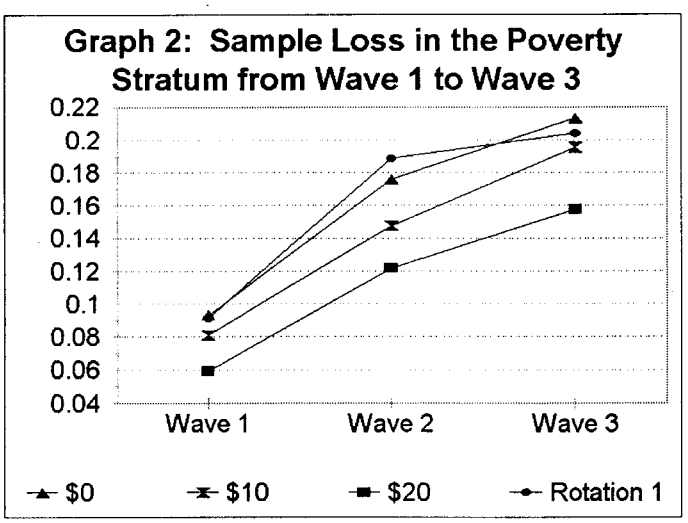
Table 3. Nonresponse Rates at Wave 2

Incentive Group	Rot 1	Rot 2-4	Differences in Rot 1 vs. Rot 2-4
Overall			
$0	8.04%	7.01%	1.03
$10	7.82%	6.00%*	1.82
$20	8.04%	5.68%*	2.36
High Poverty Stratum			
$0	9.16%	7.94%	1.22
$10	8.98%	6.46%*	2.52
$20	8.87%	6.04%*	2.83
Low Poverty Stratum			
$0	7.52%	6.60%	0.92
$10	7.41%	5.83%	1.58
$20	7.70%	5.53%*	2.17

Table 4. Nonresponse Rates at Wave 3

Incentive Group	Rot 1	Rot 2-4	Differences in Rot 1 vs. Rot 2-4
Overall			
$0	10.76%	10.43%	0.33
$10	10.47%	10.06%	0.41
$20	12.16%	8.98%*+	3.18*
High Poverty Stratum			
$0	10.47%	11.08%	-.61
$10	11.57%	10.59%	0.98
$20	13.24%	9.19%*	4.05*
Low Poverty Stratum			
$0	10.89%	10.14%	0.75
$10	10.08%	9.86%	0.22
$20	11.70%	8.89%*	2.81


Graph 1: Sample Loss from Wave 1 to Wave 3
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
Wave 1
Wave 2
Wave 3
$0
$10
$20
Rotation 1


Graph 2: Sample Loss in the Poverty Stratum from Wave 1 to Wave 3
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
Wave 1
Wave 2
Wave 3
$0
$10
$20
Rotation 1