Note to Reviewer of 1220-0109

OMB approved the data collection of NLSY79 Round 23 on January 11, 2008, with the following terms of clearance.

"TERMS OF CLEARANCE: This collection is approved with the following comments: 1) BLS is working on a report on research being conducted on response bias due to incentives and will provide a copy to OMB when available; 2) BLS is continuing to explore the use of technologies such as T-ACASI to improve the quality of reporting on sensitive items; and 3) BLS is working to offer more equal incentives to all respondents."

1. *BLS is working on a report on research being conducted on response bias due to incentives and will provide a copy to OMB when available.*

That report is attached to this document.

2. *BLS is continuing to explore the use of technologies such as T-ACASI to improve the quality of reporting on sensitive items.*

The number of sensitive questions in the NLSY79 is fairly small. The women in the sample are old enough that there are few abortions to ask about, and the survey has fewer questions about drug use than it had when respondents were younger. As the respondents have aged, the questionnaire has included more questions about health, which can be sensitive, but respondents probably do not perceive questions about health to be as sensitive as questions about abortion, drug use, sexual activity, or criminal behavior.

Telephone-Audio Computer Assisted Self-Interview technology might have greater relevance for the NLSY79 Young Adult survey, which has a high phone completion rate and a greater number of sensitive questions than the main NLSY79. The NLSY79 Young Adult survey includes a section that is self-administered on the laptop for in-person interviews and read by the interviewer for telephone interviews.

Interactive voice response (IVR) technology allows a computer to interpret spoken responses to questions. Because it enables respondents to self-administer sensitive questions, IVR may help improve the quality of data gathered during phone interviews. Our experience with IVR to date indicates that this technology is not yet workable for the NLS program.

Our primary experience with IVR technology was an experiment in the Round 9 NLSY97 pretest (July-August 2005). The NLSY97 pretest sample was split into an IVR group and a control group. The IVR group initially included 101 cases (half of the sample). During the field period, 21 of these cases were determined to require in-person interviewing, so IVR was not tested. Telephone interviews were completed with 59 sample members in the IVR group. Of these, 52 were able to access the IVR portion of the interview. Of the 52 cases, two respondents requested to return to interviewer administration at some point during the IVR section.

The primary issues in the experiment were the quality and reliability of the phone connections needed for the IVR technology to function. To access the IVR system, interviewers needed to use a VoIP (Voice over Internet Protocol) program through their laptop computer. In most cases, calls placed through this program were as clear as those made from a normal phone, but interviewers sometimes encountered poor sound quality or dropped connections. Our subsequent experience with Internet-based phone interviewing on the NLSY79 Early Bird program indicates that sound quality continues to be an occasional problem with VoIP calls, although those problems have declined in each successive round, reflecting the spread of better bandwidth to residences. Problems with sound quality are correctable if the interviewer is on the phone because the interviewer can call back on a land line if voice quality deteriorates. If the respondent is in the IVR section, however, no such option is available. We could circumvent this problem on the interviewer end of the call by only using interviewers with high-quality Internet connections to implement IVR, but that reduces our ability to use some interviewers.

Problems with the IVR technology also may originate on the respondent end of the call. In the NLSY97 experiment, respondents on cell phones experienced more problems than those on land lines because cell phones experience more sound quality problems than VoIP. The number of cell phone-only households is increasing, particularly among younger Americans like the NLSY97 and NLSY79 Young Adult samples. If IVR is less likely to be compatible with cell phones, this would pose a particular problem with implementation in these samples.

We also note that the complexity of the NLS interview means that problems with IVR technology could cause significant respondent irritation if implementation was attempted. The interview contains a significant number of question paths which are driven by previous respondent answers. If the IVR system misinterprets a response, the respondent would be directed down the wrong path, leading to respondent confusion and irritation and the loss of data. One way to avoid this problem is to ask the respondent to verify the recorded response to each question, but that would significantly lengthen the interview and cause further irritation. IVR is a better candidate when the questions used are "linear" with no intervening skips. This would be more robust to sound quality problems because we could record the call and later code answers the IVR system could not interpret. Of course, human coding at least partially negates the confidentiality advantages of IVR.

Based on our experience with the NLSY97 experiment and the state of IVR technology, the VoIP system used on the NLSY79, and the increasing use of cell phones, we do not propose implementation of IVR at this time in the NLS program. We recognize the desirability of improving data quality and providing respondents with a greater sense of anonymity with sensitive questions, and we continue to monitor the state of VoIP technology. However, our experience suggests cell phone calls may be the larger problem as voice quality is connected to the power level on the phone, distance from a tower, and whether the cell phone is in the open or behind dense walls. Voice recognition technology has improved, but sound quality is still the limiting factor.

3. *BLS is working to offer more equal incentives to all respondents*.

BLS recognizes OMB's desire to offer equal incentives to all respondents participating in the NLSY79. We must balance this concern with the legacy of past incentive experiments, our experience of the most effective ways to motivate respondents, and the constraints of the project budget. In Round 23 (2008), we took several steps in the direction of equalizing incentives: (1) offering the Early Bird option to all respondents rather than a selected group of historically cooperative respondents; (2) reducing from $20 to $10 the difference between the two historical Early Bird groups; and (3) reducing from $40 to $30 the difference between the regular respondent incentive and the special $80 Round 19 experimental incentive. Over the next several rounds we intend to eliminate the Early Bird differential and continue to reduce the differential caused by the Round 19 experiment.

We do not propose raising incentive amounts for any group in Round 24, so these small differences would remain constant from Round 23. We have found that it can harm long-term respondent cooperation to increase incentive amounts in consecutive rounds because respondents then come to expect an increase in every single round. That would quickly drive incentives higher than the project can afford. Therefore, BLS requests that the current incentive structure be allowed to continue for at least one more round so that we can avoid raising incentives in consecutive rounds. We make this proposal with the understanding that we will continue to work toward a more level incentive structure over the next two or three rounds, as appropriate within the constraints of our relationship with the respondents and the overall project budget.

An analysis of the impact of respondent incentives on respondent cooperation in the NLSY79

Introduction

As a result of an earlier OMB review, the National Longitudinal Surveys program agreed to examine the impact of respondent incentives on response bias. A rigorous examination of respondent answers versus actual facts is very difficult, however. Executing such a cross-check usually entails accessing and matching to administrative data, which can create confidentiality and security issues. In addition, administrative data are not available for most of the topics covered in the NLSY79. Bias is frequently the outcome of factors such as satisficing behavior, failure to perform the cognitive "work," or failure to provide the level of cooperation needed to generate a quality response. In this report, we analyze two measures of respondent cognitive work and cooperation that are reasonable proxies for bias. These measures are item nonresponse and interviewer reports of respondent cooperativeness during the interview. Briefly, we find evidence that increasing respondent incentives reduces item nonresponse and increases respondent cooperativeness. In other reports, respondent incentives can be a powerful tool to curtail the data loss to attrition. This analysis suggests those benefits are complemented by an improved quality of response to the NLSY79.

To counter the increasing difficulty of obtaining cooperation from sample members, many survey organizations offer monetary incentives to complete an interview (Singer et al. 1998). A number of papers have studied the effectiveness of monetary incentive payments by investigating how these payments affect response rates (Church 1993, Shettle and Mooney 1999, Singer et al. 1998, Singer et al. 2000, Olsen 2005, Yu and Cooper, 1983). However, these studies have only focused on how incentive payments affect unit nonresponse rates. Little is known about the relationship between monetary incentives and item nonresponse.

Basically, there are two types of survey nonresponse. Unit nonresponse occurs when a sampled individual does not participate in the survey as a whole. Item nonresponse occurs when a respondent does not provide valid responses to questions that the respondent is eligible to answer. Item nonresponse may be caused by the respondent's unwillingness to answer some survey items that he or she perceives as sensitive.

Another measure that could be used to evaluate the effectiveness of incentive payments is the attitude of the respondents when being interviewed. Since respondents' attitude may influence data quality in terms of truthfulness or accuracy of responses, understanding the relationship between monetary incentives and respondents' attitudes is important to survey researchers.

The objective of this report is to assess the effectiveness of respondent incentive payments. Specifically, the report is aimed at investigating whether monetary incentives have any impact on respondent cooperation as measured by item nonresponses and respondents' attitude.

Background

The NLSY79 was designed to represent the population born between January 1, 1957, and December 31, 1964, and living in the U.S. when the survey began in 1979. The initial sample included 12,686 young men and women. In 1985, most of the supplemental sample of military service members was dropped. The supplemental sample of disadvantaged nonblack, non-Hispanic youths was discontinued in 1991. The remaining 9,964 respondents were eligible for interview annually through 1994 and biennially thereafter. The NLSY79 is an omnibus study, focusing on labor market activity, while also addressing questions about education, marital and fertility histories, income and assets, health, insurance coverage, and alcohol and substance use.

Method

The focus of this report is to analyze how respondent incentives affect respondent cooperation as measured by item nonresponse and respondent attitude. The literature shows that item nonresponse occurs mainly with highly sensitive questions such as questions on income and wealth (Riphahn and Serfling 2005, Shoemaker et al 2002). Therefore, a block of 18 questions in the income and asset section of the NLSY79 is examined here, and the number of Refusal and Don't Know answers are counted to represent the number of item nonresponses. Another variable of interest is the interviewer's assessment of the respondent's attitude during the interview. Individual fixed-effects models are employed for the analysis. The econometric models are written as:

$$I_{it} = \alpha_I + \beta'X_{it} + \delta_i + \varepsilon_{it}$$

$$C_{it} = \alpha_C + \zeta'X_{it} + \delta_i + v_{it}$$

I_{it} is the number of Refusal and Don't Know answers in the income and asset section of individual i at round t. C_{it} is the attitude of respondent i during the round t interview. X_{it} is a vector of explanatory time-varying variables including marital status, employment status, and number of children. δ_i is a vector of individual fixed effects. Respondent attitude is a categorical variable coded as follows: (1) Friendly and interested; (2) Cooperative but not particularly interested; (3) Impatient and restless; (4) Hostile. Marital status is a dichotomous variable indicating whether the individual was married at the time of the interview. Similarly, employment status is a dichotomous variable indicating whether the individual was employed at the time of interview.

The two survey rounds used for this analysis are Round 17 (1996) and Round 23 (2008). A total of 8,636 respondents completed Round 17 interviews and 7,755 respondents completed Round 23 interviews. The final dataset consists of 7,424 respondents who were interviewed in these two rounds.

Table 1: Frequency of respondent incentives for Round 17

Amount	Frequency	Percent
$20	5,311	77.97
25	210	3.08
26	1	0.01
28	1	0.01
29	3	0.04
30	264	3.88
35	76	1.12
38	1	0.01
40	170	2.5
41	1	0.01
45	15	0.22
50	669	9.82
55	7	0.1
60	25	0.37
65	2	0.03
70	6	0.09
75	21	0.31
80	2	0.03
85	3	0.04
90	1	0.01
95	1	0.01
100	22	0.32

Table 2: Frequency of respondent incentives for Round 23

Amount	Frequency	Percent
$50	1,977	29.02
59	1	0.01
60	121	1.78
70	3,097	45.46
75	2	0.03
80	1,514	22.23
90	65	0.95
100	26	0.38
110	8	0.12
120	1	0.01

Table 1 and Table 2 present frequencies of respondent incentives for Round 17 and Round 23, respectively. Respondents received from $20 to $100 for completing the interview in Round 17, in which 75 percent of them received the base incentive of $20. In Round 23, the base incentive increased to $50, but a majority of the respondents received $70 or $80 due to an Early Bird bonus incentive opportunity. Only 29 percent of Round 23 respondents received the $50 base

incentive, although they were offered the Early Bird bonus incentive for prompt completion of the interview.

Table 3: Means and standard deviations for respondents interviewed in Rounds 17 and 23 (n=7,424)

Variable	Mean	Std Dev
Respondent incentive Round 17 ($)	24.9567	11.265
Respondent incentive Round 23 ($)	49.284	8.6954
Item nonresponses Round 17	0.3782	1.0514
Item nonresponses Round 23	0.2663	0.6853
Attitude Round 17	1.2073	0.4676
Attitude Round 23	1.1669	0.4381
Number of children Round 17	1.7284	1.3887
Number of children Round 23	2.0091	1.4566
Married Round 17	0.5611	0.4963
Married Round 23	0.5508	0.4974
Employed Round 17	0.7873	0.4093
Employed Round 23	0.7871	0.4094

Table 3 reports the means and standard deviations for Rounds 17 and 23. The respondent incentive for Round 23 was adjusted by the Consumer Price Index. Adjusted for inflation, the average respondent incentive in Round 17 was $24 lower in 1996 dollars than the average in Round 23.

Results

Results for the individual fixed-effects model for number of Refusal and Don't Know answers are presented in Table 4. The coefficient for the respondent incentive has a negative sign and is statistically significant at the 1% level. This means that an increase in respondent incentive results in a decrease in number of Refusal and Don't Know answers. Among other explanatory variables, only the coefficient for the marriage variable is statistically significant. Married respondents have a higher number of item nonresponses. This may be explained by the time constraint that married respondents experience. Employment status and the number of children do not have any significant impacts on the number of item nonresponses.

Table 4: Individual fixed-effects model for number of Refusal and Don't Know answers (n=7,424)

Variable	Coefficient	Std Error	t-statistics
Intercept	0.2962	0.0422	7.01
Incentive	-0.0026	0.0005	-5.37
Married	0.2692	0.0265	10.18
Employed	0.0165	0.0262	0.63
No of children	-0.0128	0.0187	-0.69
F statistic	29.93		

Fixed-effects results for respondent attitude are reported in Table 5. Note that the lower the value of the attitude variable, the more cooperative the respondent is. As can be seen from Table 5, the amount of the respondent incentive is negatively correlated with respondent attitude. This indicates that higher respondent incentives lead to more cooperative attitudes from the respondents. Perhaps surprisingly, respondents with more children are more cooperative during the interview.

Table 5: Individual fixed-effects model for respondent attitude (n=7,424)

Variable	Coefficient	Std Error	t-statistics
Intercept	1.3148	0.0229	57.43
Incentive	-0.0009	0.0003	-3.61
Married	-0.0182	0.0143	-1.27
Employed	-0.0128	0.0142	-0.9
No of children	-0.0307	0.0102	-3.02
F statistics	7.44		

Individual fixed-effects models are used to control for time-invariant unobserved individual heterogeneity. However, fixed-effects models do not address the potential problem of reverse causality. That is, the number of item nonresponses and the respondent's attitude may cause changes in respondent incentives over time. For example, a respondent with a high number of item nonresponses in previous rounds might be offered a higher incentive in the following rounds.

In order to control for reverse causality, we included the lagged values of the number of item nonresponses and respondent attitude into the models. Specifically, numbers of item nonresponses and respondent attitude for Round 12 (1990) and Round 19 (2000) were used to represent the lagged values for Round 17 and Round 23, respectively. After eliminating missing values for noninterviews in Round 12 and Round 19, a new dataset was formed consisting of 6,812 respondents. Means and standard deviations of all variables are provided in Table 6.

As can be seen from Table 7, lagged item nonresponses have a significantly negative impact on current item nonresponses. In other words, those respondents with high numbers of item nonresponses in previous rounds are likely to have low numbers of item nonresponses in the next round. The addition of lagged item nonresponses to our models alters only slightly the absolute values of coefficients on incentives and marriage. Item nonresponse is still negatively associated with the amount of the respondent incentive and positively associated with marriage.

Table 6: Mean and standard deviation for respondents interviewed in Rounds 12, 17, 19 and 23 (n=6,812)

Variable	Mean	Std Dev
Respondent incentive Round 17 ($)	24.9567	11.265
Respondent incentive Round 23 ($)	49.284	8.6954
Item nonresponses Round 17	0.3782	1.0514
Item nonresponses Round 23	0.2663	0.6853
Item nonresponses Round 12	0.2026	0.7254
Item nonresponses Round 19	0.5098	1.2127
Attitude Round 17	1.2073	0.4676
Attitude Round 23	1.1669	0.4381
Attitude Round 12	1.2371	0.5998
Attitude Round 19	1.2245	0.4924
Number of children Round 17	1.7284	1.3887
Number of children Round 23	2.0091	1.4566
Married Round 17	0.5611	0.4963
Married Round 23	0.5508	0.4974
Employed Round 17	0.7873	0.4093
Employed Round 23	0.7871	0.4094

Table 7: Individual fixed effects model for number of refusal and don't know answers with lagged value (n=6,812)

Variable	Coefficient	Std Err	t-statistics
Intercept	0.2818	0.0435	6.48
Lagged item nonresponses	-0.0735	0.0105	-7.03
Incentive	-0.0019	0.0005	-3.76
Married	0.2573	0.0270	9.52
Employed	0.0062	0.0270	0.23
No of children	-0.0062	0.0194	-0.32
F statistics	33.99		

As shown in Table 8, the coefficient for lagged attitude is statistically insignificant. This suggests that respondent attitude in a previous round does not affect respondent attitude in the next round. The level of significance and the signs of coefficients for the amount of the respondent incentive and the number of children remain the same when adding the lagged value of respondent attitude. The strong effect of marriage on item nonresponse may well reflect one spouse or the other handling the finances for the household, and if the respondent happens not to be that person, higher rates of nonresponse would likely result. We see little effect of marriage on respondent attitude toward the interview, suggesting specialized knowledge about finances is at work as marriage does not have much of an effect on cooperation *per se*.

Table 8: Individual fixed effects model for respondent attitude with lagged value (n=6,812)

Variable	Coefficient	Std Error	t-statistics
Intercept	1.3018	0.0261	49.97
Lagged attitude	0.0006	0.0096	0.06
Incentive	-0.0008	0.0003	-2.88
Married	-0.0190	0.0146	-1.30
Employed	-0.0279	0.0146	-1.91
No of children	-0.0291	0.0105	-2.78
F statistics	5.95		

Conclusion

Many studies have demonstrated that monetary payments increase survey response rates. However, it remains unclear how incentive payments affect respondent cooperation given that the respondents agreed to participate in the survey. This question becomes more important in longitudinal studies where respondent cooperation in the present round may influence outcomes in subsequent rounds.

The findings from the individual fixed-effects model using the NLSY79 data reveal that respondent incentive payments have a positive impact on respondent cooperation. An increase in the amount of the respondent incentive results in a decrease in the number of Refusal and Don't know answers in the income and asset section. An increase in the amount of the respondent incentive also leads to an improvement in respondent attitude during the interview. The findings also indicate the existence of reverse causality between the amount of respondent incentives and item nonresponse.

In conclusion, respondent incentive payments have been effective in increasing respondent cooperation by reducing the number of Refusal and Don't Know answers and improving respondent attitude during the survey. In a longitudinal survey such as the NLSY79, increasing respondent cooperation not only contributes to improved data quality in terms of completeness and accuracy but also may be important in maintaining productive contact with respondents for subsequent interviews.

References

Church, Allan H. 1993. "Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta-Analysis." *Public Opinion Quarterly* 57:62-79.

Olsen, Randall J. "The Problem of Survey Attrition: Survey Methodology is Key ," *Monthly Labor Review* 128(2005), 63-70.

Riphahn, Regina and Oliver Serfling. 2005. "Item non-response on income and wealth questions." *Empirical Economics* 30:521-538.

Shettle, Carolyn and Geraldine Mooney. 1999. "Monetary Incentives in Government Surveys." *Journal of Official Statistics* 15:231-50.

Shoemaker, Pamela, Martin Eichholz, and Elizabeth Skewes. 2002. "Item non-response: Distinguishing between Don't know and Refuse." *International Journal of Public Opinion Research* 14:193-201.

Singer, Eleanor, John Van Hoewyk, and Mary P. Maher. 1998. "Does the Payment of Incentives Create Expectation Effects?" *Public Opinion Quarterly* 62:152-64.

Singer, Eleanor, John Van Hoewyk, and Mary P. Maher. 2000. "Experiments with Incentives in Telephone Surveys" *Public Opinion Quarterly* 64:171-188.

Yu, Julie and Harris Cooper. 1983. "A Quantitative Review of Research Design Effects on Response Rates to Questionnaires." *Journal of Marketing Research* 2036-44.