

## CAPI, EVENT HISTORIES, AND INCENTIVES IN THE NSFG CYCLE 5 PRETEST

William D. Mosher, William F. Pratt, National Center for Health Statistics  
Allen P. Duffer, Jr., Research Triangle Institute  
William D. Mosher, NCHS, 6525 Belcrest Rd., Hyattsville, MD 20782

Keywords: Incentives; Event histories; CAPI

### INTRODUCTION

The National Survey of Family Growth (NSFG) is a periodic survey designed to provide estimates of factors affecting the U.S. birth rate and the reproductive health of U.S. women 15-44 years of age. The Pretest for Cycle 5 of the NSFG was more than a pretest. It contained a number of new features and experiments:

(1) Computer-Assisted Personal Interviewing (CAPI) with laptop computers;

(2) event histories and the use of a life history calendar to record all episodes of living arrangements, school attendance, work, and cohabitation, which increase the length and complexity of the interview;

(3) use of Audio Computer-Assisted Self-Interviewing (Audio CASI) to ask questions over headphones and have the respondent enter the answers into the computer herself;

(4) a pre-planned Telephone Reinterview, which allows us to collect data at much lower cost and with a much shorter interview than the initial interview, but makes high response rates critical;

(5) a test of interviewing in the home vs. at non-home sites; and

(6) use of incentives for respondents. In order to meet the data needs of our co-sponsors and users, we made a series of design decisions that made it necessary to use incentives for respondents. This paper describes why we made the design decisions, and why we believe it is necessary to use incentives in the NSFG Main Study in 1995.

The National Survey of Family Growth (NSFG) is conducted periodically by the U.S. National Center for Health Statistics (NCHS). The purpose of the survey is to provide national estimates of factors related to the U.S. birth rate and to the reproductive health of women in the childbearing ages. The age range of the survey is 15-44 years because the number of pregnancies occurring outside that age range in the United States is too small to study with a sample of 8,000-10,000 women.

The NSFG is descended from a line of national surveys that began under private auspices in 1955 (Pratt et al, 1984). These earlier surveys were conducted in 1955, 1960, 1965, and 1970. The NSFG has been conducted by NCHS in 1973, 1976, 1982, and 1988.

The planning described here will result in Cycle 5 of the survey, which will be conducted in 1995. The scope of the NSFG includes the full range of the intermediate fertility variables (Davis and Blake, 1956), or proximate determinants of fertility (Bongaarts, 1978): (a) variables that affect exposure to intercourse, including sexual activity, cohabitation, marriage, and divorce; (b) variables that affect the probability that intercourse will result in pregnancy, including contraception, sterilization, infertility and breastfeeding; and (c) variables that affect the probability that conception will result in live birth, including induced abortion, spontaneous miscarriage, and stillbirth. The survey's subject matter also includes aspects of the social and economic environment, including family background, education, labor force participation, and use of health care, that affect the intermediate variables. These surveys have typically had one-time, cross-sectional interviews lasting about 60 minutes, and achieved response rates of around 80 percent without the use of incentives. In Cycle 5, however, we made a series of design decisions that made it necessary to experiment with incentives.

### DESIGN FEATURES

1. CAPI.--When planning for Cycle 5 began, an initial decision was made to use Computer-Assisted Personal Interviewing (CAPI) on laptop computers, in order to improve the quality and timeliness of the data. CAPI relieves interviewers of their largest burdens in past cycles of the survey, and thus prevents interviewer errors (what Baker (1992) calls "illegal skips"). These burdens include:

(a) deciding which question to ask next,

(b) deciding how to word the question to fit the respondent,

(c) checking to see whether the month and year of one event (like a pregnancy) is consistent with the month and year of another event (such as first intercourse); and

(d) deciding whether and how many times to ask a series of questions that can be repeated several times (such as a series that is asked for each pregnancy, each marriage, or each partner). (Weeks, 1992).

The questionnaire program and specifications were very complex (O'Reilly, 1993). The median length of the Pretest questionnaire was 104 minutes, compared to

60 minutes in 1982, and 70 minutes in 1988. The questionnaire was programmed in Version 2.38 of BLAISE, a program developed in the Netherlands, and contained 27,000 lines of BLAISE code. It requires at least a 386 processor and a 40 megabyte hard drive. The program was so long and complex that it had to be split into 10 successive sections, with selected data passed from one section to the next. (O'Reilly, 1993)

**2. Event histories.**--An event history is simply a list of all the instances of some event, with the dates they occurred. A work history, for example, is a list of all periods of time when a woman worked for pay. A schooling history is a list of all periods of time when she attended school. Researchers studying fertility and family planning wish to apply event history analytical techniques to fertility survey data to sort out the causal processes that produce fertility trends and differentials. (Allison, 1984). To supply that kind of data to the agencies that co-sponsor the NSFG, we needed to collect event histories of all of our major independent variables, such as living arrangements, education, work, and cohabitation, in addition to the other event histories we have always collected--marriage, pregnancy, and contraception. The need for the data was clear, but converting these histories to CAPI was a challenge.

As one observer said, "It isn't just a long interview; it's a hard interview," because the respondent has to recall and keep consistent the dates (months and years) when marriages, pregnancies, cohabitations, sexual partners, contraceptive use, and other events occurred. In addition, the histories made the interview more variable in length. In sum, the questionnaire was long on average, variable in length, and the event histories were often hard for respondents to answer, and hard for us to program.

**3. Audio CASI.**--To study unintended pregnancy, the NSFG has always tried to collect data on pregnancy outcomes (miscarriage, stillbirth, abortion, live birth). The data on miscarriage, stillbirth, and live birth are of high quality, but data on abortion have always been under-reported--often by 50% or more--in fertility surveys in the U.S and elsewhere (Jones and Forrest, 1992b). In addition, interest in the Human Immunodeficiency virus, the virus that causes AIDS, has prompted requests for data on the numbers and characteristics of sexual partners. We developed a self-administered questionnaire delivered over headphones to collect data on topics such as abortion and the characteristics of sexual partners. We hoped that the self-administered Audio CASI would increase the reporting of sensitive behaviors, and avoid the problems of missing and inconsistent data that are all too common in paper and pencil self-administered questionnaires.

**4. Pre-planned Telephone Reinterviews.**--The last 4 cycles of the NSFG have been done about every 6 years--in 1976, 1982, 1988, and 1995. Our co-sponsors and other data users have expressed a desire for data at more frequent intervals. They have also encouraged us to collect longitudinal data on such topics as the accuracy and stability of expectations for future sexual activity, marriage, and future births. The cost and interview length of a telephone reinterview are both about one-third the cost and length of the original in-person interview. Thus, the telephone reinterview is a very efficient way to collect additional data, once the original interview has been done in person.

However, if we are to avoid potentially damaging response bias, the response rates must be kept as high as possible. The following examples will illustrate this point. The NSFG sample is a list sample derived from a large NCHS sample called the National Health Interview Survey, or NHIS. The response rate to the NHIS is about 95 percent. If 75 percent of the 95 percent respond to the NSFG Main Study and 75 percent of those respond to the Telephone Reinterview, we have a compound response rate of 53 percent, raising concerns about bias in the Telephone Reinterview. If, however, response rates could be increased by 10 percentage points in both the 1995 NSFG and the 1997 Telephone Reinterview, we have 95% times 85% times 85%, or 69%. This result has two benefits: a panel about 1,500 respondents larger in the telephone reinterview, and second, much less concern about response bias in the reinterview sample. Thus, our target response rate for the 1995 NSFG is 80-85 percent.

**5. Interviews outside the home.**--The Cycle 5 Pretest contained a test of interviewing outside the home. This test was done because it appeared that respondents' main concerns about the privacy of the interview were to keep their answers private from the members of their own family or household, not from distant or anonymous governments or organizations. So we wanted to determine whether the reporting of abortion, sexual partners, drug use, and other sensitive topics could be improved by moving the interview outside the home to a designated non-home site, such as a conference room in a library or rented space in an office building. In order to compensate respondents for the added burden of traveling to the non-home site, arranging for child care, or other inconvenience of being away from home, we offered a \$40 reimbursement (plus mileage or taxi fare) for respondents who agreed to go to the non-home site. One aim of the Pretest was to study the effect of non-home interviews on response rates and costs, and

project those response rates and costs to a national survey.

**6. Incentives for in-home interviews.**--The \$40 incentive for the non-home interview was a compound variable. We became concerned that the non-home interviews with a \$40 incentive might increase response rates and/or data quality, but we would not know whether that improvement was due to the \$40 incentive or to the non-home site. We were also concerned that \$40 incentives for non-home interviews might be too costly to implement in a national survey. Thus, we began to look for an alternative that would cost less, and would shed light on whether it was the non-home site or the incentive that was producing any differences.

## INCENTIVES

The literature cited by Groves et al (1992:480-481) suggests that incentives may work in three ways: first, incentives may create a reciprocation norm--i.e., that people are more willing to comply with a request "if compliance constitutes repayment of a perceived gift." Second, incentives can be understood in terms of social exchange theory: an incentive creates an informal contract between the respondent and interviewer. Thus, incentives may result in both higher response rates and higher data quality if respondents feel obligated to carry out their part of the "contract." Third, incentives may be viewed by some respondents as straightforward compensation. Some respondents may simply want the incentive money, and may cooperate if the task seems reasonable and the request appears legitimate.

A growing number of federal social and health surveys have characteristics that make incentives necessary and cost-effective--because they are long, sensitive, involve repeated interviews, and sometimes ask the respondent to leave their home or keep detailed records. In short, they make demands on respondents that exceed the public's idea of what a "survey" is. (Ezzati-Rice et al, 1994). These include the High School and Beyond Survey, the 1977 National Medical Care Expenditure Survey and the 1980 National Medical Care Utilization and Expenditure Survey. These medical care surveys ask the respondent to go to considerable effort to collect and save medical records and participate in several interviews. The National Health and Nutrition Examination Survey (NHANES) asks respondents to take a long medical exam in a non-home site as well as respond to an interview. Experiments with the NHANES have found that incentives have consistently increased response rates. The National Longitudinal Survey of Youth, conducted every year for 14 years with thousands of young adults, has used incentives to maintain its consistently high

response rates.

A study by Chromy and Horvitz (1978) showed strong effects of incentives on the response rate in the National Assessment of Educational Progress. In 1992, a field test for the National Adult Literacy Survey was conducted with 1,700 respondents in 16 areas across the country (Berlin et al 1993). This experiment showed that a \$20 incentive produced higher response rates and lower costs per completed case than no incentive. The \$20 incentive also got higher response rates from adults with less education and marginal literacy, and thereby produced higher estimates of adult illiteracy than the no-incentive design.

The Office of Management and Budget (OMB) has written regulations enforcing the Paperwork Reduction Act of 1980. These regulations (5 CFR 1320.6) generally prohibit the use of payment for survey respondents in surveys done by federal agencies or their contractors. These rules do allow OMB to permit incentives in "exceptional" circumstances. At a recent conference on this topic, a committee of survey experts suggested to OMB that it seriously consider the use of incentives if a survey is long, sensitive, requires detailed record keeping, is affected by relatives or friends who block access to the respondent (gatekeepers), or is part of a longitudinal panel in which retention of the panel over time is important (COPAFS, 1993, pp. 8-9).

It was apparent that the 1995 NSFG would be longer and more sensitive than the 1988 survey; it would also be part of a panel survey, and would be affected by gatekeepers (such as husbands and parents) who might try to prevent the interviewer from talking to the respondent. Under these conditions, we thought that incentives would be necessary if we were to meet our goals for response rates, costs, and complete reporting of event histories and sensitive behaviors. Accordingly, we chose to test a \$20 incentive for in-home interviews.

## PRETEST DESIGN

Thus, the design of the pretest was:

- 1) in-home interviews with no incentive
  - a) No audio CASI
  - b) Audio CASI
- 2) in-home interviews with \$20 incentive
  - a) No audio CASI
  - b) Audio CASI
- 3) non-home interviews with \$40 incentive

The NSFG contract was awarded about October 1, 1992. Interviewer training for the Pretest occurred October 3-10, 1993. Interviewer training was held in Durham North Carolina. About 30 interviewers were trained. Pretest interviewing occurred from October

11-December 20, 1993. 500 interviews were obtained, out of 787 eligible women, for a response rate of 64%. The primary reason that the response rate was less than 70% is simple: this was a list sample, and 41% of the sample had moved at least once since the NHIS interview. Because the amount of time and money available to trace movers was limited, about one-fourth of the movers (9% of the sample) could not be traced. Since most of this 9% would be found in a main study, and since persons cannot respond to an incentive or advance letter if they cannot be found, that 9% of the sample is excluded from the "completion rates" shown in Table 1. The next section describes the results in each of the 3 main Pretest groups: in-home interviews with no incentive; in-home interviews with \$20 incentive; and non-home interviews with \$40 incentive.

The sample was drawn from households that had responded to the 1991 National Health Interview Survey (NHIS), in New York City, Nassau-Suffolk (Long Island--NY), Dallas and Austin, Texas, and 4 areas in North Carolina: 2 rural counties, and the cities of Greensboro and Winston-Salem. These areas provide some variation by region, socioeconomic status, and size of place-- from very large cities like New York and Dallas to rural areas.

## PRETEST RESULTS

Group 2: \$20 incentive at home:

The response rate was 67% in both incentive groups compared with 59% in the no-incentive group (table 1). The completion rate is a response rate as a percent of those located; it is a better estimate of what the Main Study response rate will be. The completion rate was 81% for the \$20 incentive group, compared with 73% for the no-incentive group. Overall, most of the difference in response rates between incentive and non-incentive cases was a result of lower refusal rates for incentive cases (table 1).

The percent of women who broke 1 or more appointments with an interviewer was substantially lower for the \$20 in-home cases (24%) than for non-incentive cases (37%). Partly as a result of this large difference in broken appointments, interviewers worked an average of 2.1 hours less to get a completed interview with a \$20 incentive (group 2, 8.8 hours) than without an incentive (group 1, 10.9 hours).

The time of an interviewer costs over \$10 an hour, including wages and fringe benefits. So the incentive saved the interviewers 2 hours of labor per case (8.8 in group 2 vs. 10.9 in group 1), which saves the \$20 cost of the incentive, and makes the \$20 incentive interviews slightly cheaper than the non-incentive interviews. In short, the \$20 incentives paid for

themselves in the Pretest.

Group 3: \$40 Incentive at Non-home sites:

The Non-home \$40 interviews had higher response rates than the non-incentive cases (67 vs 59%), but no higher than the \$20 incentive at-home interviews. Unfortunately, it was very expensive to arrange for the non-home sites; the cost in the pretest was \$101 per case for direct costs. Most of the expense was not for taxi fare or child care for respondents; it was for the labor of the contractor's staff to arrange for office space for the non-home sites. We estimate that the cost of implementing non-home site interviews for all or most respondents in the Main Study would be \$1.0 to \$1.5 million. Since we do not have the funds to pay an extra 1.0-1.5 million dollars, we prefer to use a \$20 incentive for in-home interviews.

## ABORTION REPORTING

The NSFG is a survey about pregnancy, and abortions account for 1 out of 4 pregnancies, and half of all unintended pregnancies. One crucial indicator of data quality is the completeness of reporting of abortions. Abortion reporting has been highly deficient in all US fertility surveys for decades (Jones and Forrest, 1992b). Abortion reporting affects the quality of data on pregnancy rates, unintended pregnancy, and the failure rates for contraceptive methods (Jones and Forrest, 1992a).

Comparing NSFG data with complete counts of abortions from abortion providers (clinics and hospitals who do abortions) and national surveys of abortion providers, women in the NSFG have been reporting about 40% of the abortions they actually had (Jones and Forrest, 1992b). About 13% of the 8,450 women in the 1988 NSFG sample reported that they had ever had at least one abortion (table 2). If 13% was about 40% of the actual number, then complete reporting would be about (13% divided by .40, or) 33%. In the Pretest, in the no-incentive, no-SAQ group (comparable to the 1988 survey), 14% reported an abortion (almost identical to the 1988 results). But abortion reporting was nearly complete in both (a) the group that received a \$20 incentive and an audio Self-administered questionnaire (30%), and in (b) the \$40 non-home interviews (29%), which are much more costly (table 2).

## CONCLUSION

Cycle 5 of the NSFG has a number of new design features that have made the data much more useful for the data user, but these same features have made the interview longer and more burdensome for the respondent. The 1993 NSFG Pretest demonstrates

clearly that incentives are helpful in raising response rates, cutting costs, and increasing the completeness of the data on abortion. For these reasons, we asked OMB for permission to pay each respondent a \$20 cash incentive in the 1995 NSFG, and that permission was granted. So a \$20 incentive will be used in the 1995 NSFG.

## REFERENCES

Allison, Paul D. 1984. *Event History Analysis. Quantitative Applications in the Social Sciences*, No. 46. Beverly Hills and London: Sage Publications.

Baker, Reginald P. 1992. *New Technology in Survey Research: Computer-Assisted Personal Interviewing (CAPI)*. *Social Science Computer Review* 10 (2): 145-157, Summer, 1992.

Berlin, M. et al. 1993. *An Experiment in Monetary Incentives*. In: *American Statistical Association (editor), 1992 Proceedings of the Section on Survey Research Methods*. Alexandria, VA. Pages 393-398.

Bongaarts, John. 1978. *A Framework for Analyzing the Proximate Determinants of Fertility*. *Population and Development Review* 4: 105.

Chromy, J. M. and D. Horvitz. 1978. *The Use of Monetary Incentives in National Assessment Household Surveys*. *Journal of the American Statistical Association* 73 (363): 473-478.

Council of Professional Associations on Federal Statistics (COPAFS). 1993. *Providing Incentives to Survey Respondents: Final Report*. Unpublished. COPAFS, 1429 Duke Street, Alexandria, VA.

Davis, Kingsley and J. Blake. 1956. *Social Structure and Fertility: An Analytic Framework*. *Economic Development and Cultural Change*. 4: 211-235, April, 1956.

Groves, R.M., R. Cialdini, and M. Couper. 1992. *Understanding the Decision to Participate in a Survey*. *Public Opinion Quarterly* 56: 475-495.

Jones, Elise and J. Forrest. 1992a. *Contraceptive Failure Rates Based on the 1988 NSFG*. *Family Planning Perspectives* 24 (1): 12-19, Jan/Feb 1992.

Jones, Elise and J. Forrest. 1992b. *Underreporting of Abortion in Surveys of U.S. Women: 1976-1988*. *Demography* 29 (1) 113-126, February, 1992.

O'Reilly, James. 1993. *Lessons Learned Programming a Large, Complex CAPI instrument*. Paper presented to the International BLAISE Users Conference, October 1993, London.

Pratt, William et al. 1984. *Understanding U.S. Fertility: Findings from the National Survey of Family Growth, Cycle III*. *Population Bulletin* 39 (5), December 1984. Washington, D.C.: Population Reference Bureau.

Ezzati-Rice, Trena, A. White, W. Mosher and M.

Sanchez. 1994. *Time, Dollars, and Data: Succeeding with Remuneration in Health Surveys*. *Proceedings of the Conference on New Directions in Statistical Methodology*. Alexandria, VA: Council of Professional Associations on Federal Statistics.

Weeks, Michael. 1992. *Computer-Assisted Survey Information Collection: A review of CASIC Methods and Their Implications for Survey Operations*. *Journal of Official Statistics* 8 (4): 445-466.

TABLE 1: SELECTED RESULTS OF THE NATIONAL SURVEY OF FAMILY GROWTH PRETEST, OCTOBER-DECEMBER 1993, NY, TEXAS, AND NC:

	Group 1 (No \$)	Group 2 (Home \$20)	Group 3 (non-home \$40)
Number eligible	333	227	227
Completed interviews	196	153	151
Response rate	59%	67%	67%
Completion rate	73%	81%	80%
Refusal rate	16%	10%	11%
% who broke an appointment	37%	24%	31%
Interviewer hours per complete	10.9 hrs	8.8 hrs.	6.4 hrs
Interviewer cost per complete	\$161	\$147	\$136
Cost to setup nonhome sites	0	0	\$101
Direct cost/complete	\$161	\$147	\$237

TABLE 2: PERCENT OF WOMEN 15-44 YEARS OF AGE WHO REPORTED EVER HAVING HAD AN ABORTION: 1988 NSFG AND 1993 NSFG PRETEST

	%	N
1988 NSFG	13%	8450
1993 Pretest		
No \$, No ACASI	14%	(96)
\$20, No ACASI	22%	(72)
No \$, ACASI	20%	(98)
\$20, ACASI	30%	(80)
\$40, Non-home	29%	(147)