Information Collection Request for
The National Longitudinal Survey of Youth 1997
OMB # 1220-0157
Part B

Submitted by the Bureau of Labor Statistics

TABLE OF CONTENTS

B. Collections of Information Employing Statistical Methods

1. Respondent Universe and Respondent Selection Method

This section summarizes the primary features of the sampling and statistical methods used to collect data and produce estimates for the NLSY97. Additional technical details are provided in the NLSY97 Technical Sampling Report, available online at https://www.nlsinfo.org/content/cohorts/nlsy97/other-documentation/technical-sampling-report. Chapter 2 of the report describes the design of the NLSY97 sample. Chapter 3 describes the sample-selection process. Chapter 4 describes the sample weighting process. Chapters 5 and 6 describe the accuracy and representativeness of the sample.

Additional information about statistical methods and survey procedures is available in the NLSY97 User's Guide at: https://www.nlsinfo.org/content/cohorts/NLSY97/

The initial sample was selected to represent (after appropriate weighting) the total U.S. population (including military personnel) 12 to 16 years of age on December 31, 1996. The sample selection procedure included an overrepresentation of blacks and Hispanics to facilitate statistically reliable analyses of these racial and ethnic groups. Appropriate weights are developed after each round so that the sample components can be combined to aggregate to the overall U.S. population of the same ages. Weights are needed to adjust for differences in selection probabilities, subgroup differences in participation rates, random fluctuations from known population totals, and survey undercoverage. Computation of the weights begins with the base weight and then adjusts for household screener nonresponse, sub-sampling, individual nonresponse, and post-stratification of the nonresponse-adjusted weights. The number of sample cases in 1997, the first round, was 8,984. Retention rate information for subsequent rounds is shown in the table below. BLS anticipates continued declines in retention rate in Round 19. Declines are anticipated for two reasons. First, the Round 17 experience repeated the declines observed in Round 16 and mimicked the previous experience of the NLSY79 when it went to biennial interviewing after a history of annual interviews. Moreover, we saw declines in Round 18 when the NLSY97 transitioned to a primarily phone-administered survey. We suspect that the absence of extensive in-person fielding harmed our ability to retain respondents in Round 18. In Round 16, the retention rate (taking into account the deceased respondents) was close to that of Round 8. We saw an increase in retention rate in Round 10, a modest decline in Round 11, and additional increases in Rounds 12 and 13. Response rates fell slightly in Round 14 and then remained stable in Round 15. Round 16 was the first round after adoption of a biennial (rather than annual) interview schedule. The retention rate dropped markedly in Round 16, although continued to exceed 80 percent when deceased respondents are accounted for. Only sample members who completed an interview in Round 1 are considered in-scope for subsequent rounds (as long as they are not known to be deceased). Even if NORC is unable to complete an interview for an in-scope sample member in one round, they attempt to complete an interview with that sample member in each subsequent round. The interview schedule is designed to pick up crucial information that was not collected in the missed interviews.

The schedule and sample retention rates of past survey rounds are shown in Table 3.

Table 3. NLSY97 Fielding Periods and Sample Retention Rates

Round	Months conducted	Total respondents	Retention rate	Number of deceased sample members	Retention rate excluding the deceased
1	February–October 1997 and March–May 1998	8,984	—	—	—
2	October 1998–April 1999	8,386	93.3	7	93.4
3	October 1999–April 2000	8,209	91.4	16	91.5
4	November 2000–May 2001	8,081	89.9	15	90.1
5	November 2001–May 2002	7,883	87.7	25	88.0
6	November 2002–May 2003	7,898	87.9	30	88.2
7	November 2003–July 2004	7,755	86.3	37	86.7
8	November 2004–July 2005	7,503	83.5	45	83.9
9	October 2005–July 2006	7,338	81.7	60	82.2
10	October 2006–May 2007	7,559	84.1	77	84.9
11	October 2007-June 2008	7,418	82.6	90	83.4
12	October 2008 – June 2009	7,490	83.4	103	84.3
13	September 2009 – April 2010	7,561	84.2	112	85.2
14	October 2010 – May 2011	7,420	82.6	118	83.7
15	September 2011 – June 2012	7,423	82.6	134	83.9
16	October 2013 – June 2014	7,141	79.5	151	80.8
17	September 2015 – May 2016	7,103	79.0	173	80.6
18	October 2017 – October 2018	6,734	75.0	207	76.7

Note 1: The retention rate is defined as the percentage of base year respondents who were interviewed in a given survey year.

2. Design and Procedures for the Information Collection

The NLSY97 includes personal interviews with all living Round 1 respondents, regardless of whether they subsequently become institutionalized, join the military, or move out of the United States. We employ a thorough and comprehensive strategy to contact and interview sample members. At each interview, detailed information is gathered about relatives and friends who could assist NORC field staff in locating respondents if they cannot readily be found in a subsequent survey round. Every effort is made to locate respondents. Interviewers are encouraged to attempt to contact respondents until they reach them. There is no arbitrary limit on the number of call-backs.

Preceding the data collection, the NORC interviewers are carefully trained, with particular emphasis placed on resolving sensitive issues that may have appeared in prior rounds. Most of the NORC interviewers have lengthy experience in the field from having participated in earlier NLSY97 rounds as well as from involvement with the NLSY79 and other NORC surveys. All new recruits are given one day of personal training on general interviewing techniques. All interviewers (whether having previous experience on the NLSY97 or NLSY79 or not) must complete a comprehensive and robust online training covering the questionnaire and its administration, data quality, study protocols and procedures and respondent confidentiality. In preparation for Round 19, the project will continue training modules with content targeting the changes to phone administration of the survey to improve respondent experience. Interviewers must attend training calls with their field managers and pass a certification where they must prove a full understanding of the questionnaire and how to correctly administer it. In addition to these trainings, the field staff received weekly memos throughout Round 18 which contained protocol reminders

reinforcing proper procedures, tips for improving field work, and updates from the central office. This will be continued for Round 19.

Field interviewers are supervised by NORC Field Managers and their associates. All Field Managers complete the same online training that their interviewers will complete prior to the start of the study. NORC has divided the U.S. into 10 regions, each supervised by a Field Manager who is responsible for staffing and for the quality of field work in that region. A ratio of 1 supervisor to 15 interviewers is the standard arrangement. Field Managers are, in turn, supervised by one of the two Field Project Managers.

The interview content is prepared by professional staff at BLS, CHRR, and NORC. When new materials are incorporated into the questionnaire, assistance is generally sought from appropriate experts in the specific substantive area.

Because sample selection took place in 1997 in preparation for the baseline interview, sample composition will remain unchanged.

In Round 18, we converted the NLSY97 to a predominantly telephone survey, anticipating that approximately 75 percent of interviews will be completed by telephone, in contrast to a projected 26 percent in Round 17 and 15 percent in Round 16. Instead the telephone rates were higher with approximately 90% of interviews completed by telephone and an additional 21.5 percent of sample members requiring in-person outreach. Within the survey research literature, both unit non-response and item non-response are documented to be higher in telephone administration than in in-person administration. (Safir and Goldenberg (2008) "Mode Effects in a Survey of Consumer Expenditures," Office of Survey Methods Research, Bureau of Labor Statistics retrieved from http://www.bls.gov/osmr/abstract/st/st080200.htm, Groves, Dillman, Eltinge and Little 2002 "Survey Nonresponse" New York: Wiley.)

For the NLSY97, mode differences may come about from lack of coverage (respondents not having telephones), non-contact (telephone technology offers more ways for respondents to avoid interviewers), distraction (telephone respondents may be more likely to engage in other activities during the interview due to social norms governing in-person interactions), privacy (either increased perception of privacy from not being face-to-face with the interviewer or decreased privacy because the interviewer cannot ensure that the respondent is alone and out of the hearing range of others), or other factors.

For several years, respondents have been encouraged to visit the NLSY97 respondent web-site to update their contact information. This practice will continue in Round 19. Between Round 18 and 19, we sought to collect information on contact preferences from a randomly selected one half of our sample and any address updates from the full sample. Through this effort, we sought to better understand mode preferences and uses of technology as they might inform our efforts to contact our sample members and learn about respondents' interests in topics and news sources that might pertain to the NLSY97, as well as their interest in applications of NLSY97 data in different settings to inform development of project communications, as well as, collect updated locator information to facilitate respondent contact and avert sample attrition due to non-contact.

3. Maximizing Response Rate

A number of the procedures that are used to maximize response rate already have been discussed in items 1 and 2 above. The other component of missing data is item nonresponse. Nonresponse includes respondents refusing to answer or not knowing the answer to a question. Almost all items in the NLSY97 have low levels of nonresponse. For example, in prior rounds there was virtually no item nonresponse for basic questions like the type of residence respondents lived in (YHHI-4400) or the highest grade of school respondents had ever attended (YSCH-2857).

Cognitively more difficult questions, such as "How many hours did you work per week?" (YEMP-23901) have low levels of nonresponse. In the hours per week example, 8 individuals out of 1,584 (0.5%) did not answer the question in Round 13.

Sensitive questions have the highest nonresponse. Table 4a and Table 4b present examples of Round 17 and Round 18 questionnaire items that are most sensitive or cognitively difficult. In Round 17, almost all respondents (0.6% nonresponse rate) were willing to reveal whether they had earned money from a job in the past year, but many did not know or refused to disclose exactly how much they had earned (13.6% nonresponse rate). Because high nonresponse rates were expected for the income amount question, individuals who did not provide an exact answer were asked to estimate their income from a set of predetermined ranges. This considerably reduces nonresponse on the income question. Only 8.6% of those who were asked to provide a range of income did not respond. These individuals represent about 1% (67/5770) of all individuals requested to provide income data in that round. The patterns for non-response to these items in Round 18 are similar with about 1% of individuals not providing income information for the previous year, though we see lower rates of nonresponse to the question on the exact amount of income they had earned from their jobs in the previous year (YINC-1700).

Table 4a. Examples of Nonresponse Rates for Some Round 17 Sensitive Questions

Q Name	Question	Number Asked	Number Refused	Number Don't Know	% Nonresponse
YINC-1400	Receive Work Income in 2014?	7103	37	9	0.6%
YINC-1700	How Much Income from All Jobs in 2014?	5770	78	713	13.6%
YINC-1800	Estimated Income from All Jobs in 2014?[1]	783	46	21	8.6%

[1]Asked of respondents who were unable or unwilling to answer the previous question (YINC-1700).

Table 4b. Examples of Nonresponse Rates for Some Round 18 Sensitive Questions

Q Name	Question	Number Asked	Number Refused	Number Don't Know	% Nonresponse
YINC-1400	Receive Work Income in 2016?	6734	42	9	0.8%
YINC-1700	How Much Income from All Jobs in 2016?	5523	70	362	7.8%
YINC-1800	Estimated Income from All Jobs in 2016?[1]	432	48	14	14.4%

[1]Asked of respondents who were unable or unwilling to answer the previous question (YINC-1700).

To reduce the proportion of "don't know" or "refused" responses to questions on income or assets (such as YINC-1700, shown in Table 4a and Table 4b), respondents who do not provide exact dollar answers are asked follow-up questions designed to elicit approximate information. For many income categories, the respondents are asked to select the applicable category from a predefined list of ranges. The approach for asset questions is slightly different: The initial question asks the respondent to provide an exact value, but if he or she is unable or unwilling to do so, interviewers are instructed to ask the respondent to define a range for the value using whatever values he or she feels are appropriate. If the respondent does not know or refuses to provide either an exact value or a range, a follow-up question asks him or her to select the appropriate range from a predefined list. This method provides researchers with some information on income, asset, and debt amounts when the respondent is reluctant or unable to furnish an exact figure.

Face-to-face interviewing has been found to lead to under-reporting of sensitive items relative to telephone interviewing. Thus, sensitive items that had previously been interviewer-administered in-person but will now be

interviewer-administered by telephone may experience decreases in item non-response. Income and other financial items would be the chief examples of such items. (de Leeuw E.D., van der Zouwen J. (1988). "Data quality in telephone and face to face surveys: a comparative metaanalysis." In: Groves RM, Biemer PP, Lyberg LE, Massey JT, Nicholls WL II, Waksberg J, eds. Telephone Survey Methodology. New York: Wiley: 273:99).

4. Testing of Questionnaire Items

BLS is cautious about adding items to the NLSY97 questionnaire. Because the survey is longitudinal, poorly designed questions can result in flawed data and lost opportunities to capture contemporaneous information about important events in respondents' lives. Poorly designed questions also can cause respondents to react negatively, making their future cooperation less likely. Thus, the NLSY97 design process employs a multi-tiered approach to the testing and review of questionnaire items.

When new items are proposed for the NLSY97 questionnaire, we often adopt questions that have been used previously in probability sample surveys with respondents resembling the NLSY97 sample. We have favored questions from the other surveys in the BLS National Longitudinal Surveys program to facilitate intergenerational comparisons. We also have used items from the Current Population Survey, the Federal Reserve Board's Survey of Consumer Finances, the National Science Foundation-funded General Social Survey, and other federally funded surveys.

Existing questions are also reviewed each year. Respondents' age and their life circumstances change, as does the societal environment in which the survey is conducted. Reviews of the data help us to identify questions that may cause respondent confusion, require revised response categories, or generate questionable data. Sources of information for these reviews include the questionnaire response data themselves, comments made by interviewers or respondents during the course of the interview, interviewer remarks after the interview, interviewer inquiries or comments throughout the course of data collection, other-specify coding, recordings of items during the interviews, and comparison of NLSY97 response data to other sources for external validation. We also watch carefully the "leading edge" respondents, who answer some questions before the bulk of the sample – for example, the first respondents to attend graduate school or to get a divorce. These respondents are often atypical, but their interviews can reveal problems in question functionality or comprehensibility.

In this round, for the second time, the majority of NLSY97 will be interviewed by phone. Because the previous round ran longer than expected, we have focused on shortening and streamlining the questionnaire for Round 19.

Although further edits to questionnaire wording are extremely rare, we monitor the first several hundred interviews each round with particular care. Based on this monitoring, field interviewers receive supplemental training on how best to administer questions that seem to be causing difficulty in the field or generating unexpected discrepancies in the data. This review continues at a lower level throughout the field period, with interviewers receiving ongoing training until the end of the field period.

All changes made between the Round 18 and Round 19 questions are listed in Attachment 5. Round 19 questions that have not appeared in previous rounds of the NLSY97 include:

Questions on income tax filing in the Income section. We propose to ask whether the respondent filed taxes and their filing status. These items come from the Survey of Consumer Finances. Collecting these same items in the NLSY97 is valuable because the longitudinal data permit researchers to examine how tax filing relates to respondent reports on program participation over time.

Questions on pain and use of painkillers in the Health section. We propose to add questions on whether and how often the respondent experiences pain. The question on how often the respondent experiences pain comes from the 2018 NHIS, though our timeframe is different. We propose to ask about pain in the last 30 days compared to last 3 months. These questions have been asked in Round 28 of the NLSY79. The questions on painkillers come from the American Time Use Survey. It is important to know how pain relates to employment over time. Neither the NHIS nor ATUS collects longitudinal data on work over the life-cycle.

Questions on Mental health scales in Health section and Health at age 38/39 section. We propose to ask Center for Epidemiological Depression Scale (CESD) and Generalized Anxiety Disorder Screener (GAD-7) questions in Round 19. The CESD would be asked of all respondents. The GAD-7 would be asked of respondents born in years 1980 and 1981 in Round 19. Both the CESD and GAD-7 are well-known psychological scales. They have been asked in the NLSY79 as well as numerous other surveys. Understanding mental health is important to understanding work outcomes.

Questions on Device ownership in End of Interview section. We propose to ask four questions about device ownership: whether respondents own a cell phone, whether their cell phone is a smartphone, whether they own a tablet computer, and whether they own a laptop or desktop computer. These items have been asked repeatedly since 2011 by Pew. These items are important to help understand how the digital divide relates to success in the labor market.

A list of all changes to the NLSY97 questionnaire from rounds 18 to 19 is contained in attachment 5.

5. Statistical Consultant
> Kirk M. Wolter
> Executive Vice President for Survey Research
> NORC
> 55 East Monroe Street, Suite 3000
> Chicago, IL 60603

The sample design was conducted by NORC, which continues the interviewing fieldwork.