



Committee on Economic Statistics

Federal Data Collection Comment Opportunity for AEA Members

National Longitudinal Survey of Youth 1979, Round 28

U.S. Bureau of Labor Statistics

Comments Due February 26, 2018

The Bureau of Labor Statistics (BLS) plans to ask OMB to approve Round 28 of the [National Longitudinal Survey of Youth 1979 cohort](#) (NLSY79). The sample includes 9,964 persons who will be 53 to 60 years old on December 31, 2017. This package also seeks clearance to interview children born to female NLSY79 respondents.

The main NLSY79 is funded primarily by BLS. Funding for the NLSY79 Child and Young Adult surveys is provided primarily by the National Institute of Child Health and Human Development through an interagency agreement with BLS and through a grant awarded to researchers at the Ohio State University Center for Human Resource Research. The interagency agreement funds data collection for children and young adults up to age 22; the grant funds data collection for young adults age 23 or older.

The round 28 questionnaire includes new questions on job characteristics, spouse's health, cognition, pain and use of pain killers, and retirement financing. In addition, a new module that assesses the respondents' health as they turn age 60 will be included for the first time.

AEA members are encouraged to:

- provide opinions on the value of the NLSY79-Round 28 for economic research;
- comment on the data collection instruments and methods; and
- suggest changes that would enhance data quality, value, and accessibility and lower respondent burden and federal cost.

Federal Register notice: [December 26, 2017](#) (includes detailed description of changes and instructions for submitting comments)

- Draft NLSY79-Round 28 Supporting Statement (attached) -- describes needs, uses, plans, methods, sample, revisions, and costs
 - Due date for comments: February 26, 2018
 - More information: Donna S. Rothstein, Research Economist, Office of Employment and Unemployment Statistics, BLS, Rothstein.Donna@bls.gov
-

Information on Information Collection Request (ICR) Process:

- By law, each data collection carried out by a federal agency must be cleared by OMB. Through this *Federal Register* notice, BLS is announcing that it intends to submit a request to OMB for clearance to conduct NLSY79-Round 28 and offers the public a 60-day opportunity to submit comments.
- After the close of the 60-day comment period, BLS will prepare and submit its request to OMB. That request will summarize and respond to each of the public comments it received.

Guidance to AEA Members on Preparing Comments:

- Can comment on any aspect of the proposed data collection. Possible topics, for instance, include needs, uses, methodology, design, cost, schedule, and consultation with data users.
- May frame comments on specific topics in any way, such as:
 - assessments – identifying what you do and do not like and support
 - suggestions – for how BLS might proceed in this or future collections
 - requests – for example, for a change in the design of the survey instrument, to be consulted in the future, to carry out research on an alternative approach
 - observations – for example, implications of the sample size for statistical reliability
- May propose that OMB incorporate a request in its “terms of clearance.” For instance, you could suggest as a term of clearance that BLS research the efficacy of an alternative set of questions and report back to OMB in a year on the results.

Additional AEA Resources:

- [A Primer on How to Respond to Calls for Comment on Federal Data Collections](#)
- After reviewing materials, you may provide your observations, critiques, and requests to AEASTat staff Andrew Reamer at areamer@gwu.edu and he will organize them into a draft letter for your review. Prof. Reamer is experienced in crafting comment letters for impact.

National Longitudinal Survey of Youth 1979
1220-0109
November 2017

**Information Collection Request for
The National Longitudinal Survey of Youth 1979
OMB # 1220-0109
Summary and Part A**

Submitted by the Bureau of Labor Statistics

TABLE OF CONTENTS

SUMMARY	1
SUPPORTING STATEMENT	2
A. <i>Justification</i>	2
1. Necessity for the Data Collection	2
2. Purpose of Survey and Data-Collection Procedures	2
3. Improved Information Technology to Reduce Burden	5
4. Efforts to Identify Duplication.....	6
5. Involvement of Small Organizations	8
6. Consequences of Less Frequent Data Collection.....	8
7. Special Circumstances	9
8. Federal Register Notice and Consultations.....	9
9. Payment to Respondents	132
10. Confidentiality of Data	176
11. Sensitive Questions.....	19
12. Estimation of Information Collection Burden.....	243
13. Cost Burden to Respondents or Record Keepers	265
14. Estimate of Cost to the Federal Government	265
15. Change in Burden	265
16. Plans and Time Schedule for Information Collection, Tabulation, and Publication.....	27
17. Reasons Not to Display OMB Expiration Date	27
18. Exceptions to “Certificate for Paperwork Reduction Act Submissions”	276

SUMMARY

This package requests clearance for Round 28 of the National Longitudinal Survey of Youth 1979 cohort (NLSY79). The sample includes 9,964 persons who will be 53 to 60 years old on December 31, 2017. Approximately 10 percent of the sample members are deceased. The NLSY79 is a representative national sample of adults who were born in the years 1957 to 1964 and lived in the U.S. when the survey began in 1979. The sample contains an overrepresentation of black and Hispanic respondents to include a sufficient sample size to permit racial and ethnic analytical comparisons. Appropriate weights have been developed so that the sample components can be combined in a manner to aggregate to the overall U.S. population of the same ages, excluding those who have immigrated since 1978.

This package also seeks clearance to interview children born to female NLSY79 respondents. The Children of the NLSY79 have been assessed since 1986, when the National Institute of Child Health and Human Development (NICHD) began sponsoring a set of supplemental surveys to gather a large amount of information about the lives of these children. A battery of child cognitive, socio-emotional, and physiological assessments has been administered biennially since 1986 to NLSY79 mothers and their children. Starting in 1994, children who had reached age 15 by December 31 of the survey year (the Young Adults) were interviewed about their work experiences, training, schooling, health, fertility, self-esteem, and other topics. By 2016, the sample includes very few children age 14 and under and so we will no longer conduct a separate child survey; children age 12 and older will join the Young Adults. The Young Adult group will include 1,594 respondents ages 12-24 and 4,732 respondents age 25 and older in Round 28.

The main NLSY79 is funded primarily by the Bureau of Labor Statistics (BLS). Funding for the NLSY79 Child and Young Adult surveys is provided primarily by NICHD through an interagency agreement with the BLS and through a grant awarded to researchers at the Ohio State University Center for Human Resource Research (CHRR). The interagency agreement funds data collection for children and young adults up to age 22; the grant funds data collection for young adults age 23 or older. The BLS has overall responsibility for the project. The BLS contracts with CHRR and the National Opinion Research Center (NORC) at the University of Chicago to conduct the surveys. NORC handles the interviewing, initial data preparation, and weighting. Questionnaire design, additional data cleanup and preparation, development of documentation, and preparation of data files are handled by CHRR.

The data collected in this survey are part of a larger effort that involves repeated interviews administered to a number of cohorts in the U.S. Many of the questions are identical or very similar to questions previously approved by OMB that have been asked in other cohorts of the National Longitudinal Surveys. Many of the questions in the NLSY79 also have been designed to reflect the changing nature of institutions and the different problems facing these groups of people. Those data elements of a particularly sensitive nature and those not previously collected are justified in more detail in this document.

SUPPORTING STATEMENT

National Longitudinal Survey of Youth 1979 A Survey of Persons who were Ages 14 to 21 on December 31, 1978 Rationale, Objectives, and Analysis of Content

A. Justification

1. Necessity for the Data Collection

This statement covers main fielding of Round 28 of the National Longitudinal Survey of Youth 1979 (NLSY79). The NLSY79 is a nationally representative sample of persons who were born in the years 1957 to 1964 and lived in the U.S. when the survey began in 1979. The longitudinal focus of the survey requires the collection of identical information for the same individuals over the life cycle and the occasional introduction of new data elements to meet the ongoing analytical needs of policymakers and researchers. Most of the information to be collected this round has been approved by OMB in previous rounds of the NLSY79.

The mission of the Department of Labor (DOL) is, among other things, to promote the development of the U.S. labor force and the efficiency of the U.S. labor market. The BLS contributes to this mission by gathering information about the labor force and labor market and disseminating it to policymakers and to the public so that participants in those markets can make more informed and, thus, more efficient choices. The charge to the BLS to collect data related to the labor force is extremely broad, as reflected in Title 29 USC Section 1:

“The general design and duties of the Bureau of Labor Statistics shall be to acquire and diffuse among the people of the United States useful information on subjects connected with labor, in the most general and comprehensive sense of that word, and especially upon its relation to capital, the hours of labor, the earnings of laboring men and women, and the means of promoting their material, social, intellectual, and moral prosperity.”

The collection of these data contributes to the BLS mission by aiding in the understanding of labor market outcomes faced by individuals in the early stages of career and family development. See attachment 1 for Title 29 USC Sections 1 and 2.

2. Purpose of Survey and Data-Collection Procedures

Through 1984, the NLSY79 consisted of annual interviews with a national sample of 12,686 young men and women who were ages 14 to 21 as of December 31, 1978, with overrepresentation of blacks, Hispanics, and economically disadvantaged nonblacks/non-Hispanics. The sample also included 1,280 persons serving in the military in 1978. Starting in 1985, the military sample was reduced to 201 due to a cessation of funding from the Department of Defense. Starting in 1991, interviews were discontinued with the 742 male and 901 female members of the economically disadvantaged nonblack/non-Hispanic sample. This reduced the eligible pool of sample members to 9,964. The NLSY79 was conducted annually from 1979 to 1994 and has been conducted every two years since 1994.

In addition to the regular interviews, several supplementary data-collection efforts completed during the early survey years greatly enhance the value of the survey to policymakers and researchers. The full Armed Services Vocational Aptitude Battery (ASVAB) was administered to 94 percent of sample members. Also, for a large proportion of the sample, information has been collected about the characteristics of the last high school each respondent attended, as well as the courses taken, grades, and some other personal characteristics about the respondents while attending high school.

These data-collection efforts have enabled researchers to complete studies of the relationship between background environment, employment behavior, vocational aptitudes, and high school quality. The data have helped the Departments of Labor, Defense, Education, and Health and Human Services and many congressional committees to make more knowledgeable decisions when evaluating the efficacy of programs in the areas of civilian and military employment, training, and health.

The NLSY79 is a general-purpose study designed to serve a variety of policy-related research interests. Its longitudinal design and conceptual framework serve the needs of policymakers in a way that cross-sectional surveys cannot. In addition, the NLSY79 allows a broad spectrum of social scientists concerned with the labor market problems of young baby boomers to pursue their research interests. Support for the survey by other government agencies is encouraged, as the increasingly omnibus nature of the survey makes it an efficient, low-cost data set. As noted, the survey has incorporated items needed for program and policy purposes by agencies other than the Department of Labor. In this survey round, we anticipate funding from the Social Security Administration, NICHD, and the National Institute on Drug Abuse.

In this survey round, information once again will be collected about the biological children of the female NLSY79 respondents. For the most part, this collection of data about the children repeats surveys administered to these children biennially from 1986 to 2016. These unique data permit medical and social science researchers to consider a large number of basic research issues relating to the effects of family background, program activities, and infant and maternal health on outcomes from early childhood through adolescence and into early adulthood. While the principal focus of the survey remains the collection of data for labor force analysis, the questionnaires administered to these children and older youths include items needed by other agencies that are not always directly related to employment and training studies. As these children reach adolescence, the focus of the surveys of these young adults returns to the school-to-work transition.

Sample sizes and the expected number of Round 28 respondents for each group are listed in table 1.

Table 1. NLSY79 Sample Size and Expected Response in 2018 (Round 28)

Cohort	Approximate sample size	Expected number of respondents
NLSY79 main	9,055 (9,964 minus 909 deceased sample members)	6,900
Young Adults ages 12-24	1,594	1,276
Young Adults age 25 and older	4,732 ¹	3,784 ¹

¹ Members of the Young Adult grant sample are contacted for interviews every other round once they reach age 31. These numbers do not include Young Adult sample members age 31 and older who will not be contacted for interviews in 2018.

The specific objectives of the NLSY79 fall into several major categories that will be further explained below:

1. to explore the labor market activity and family formation of individuals in this age group
2. to explore in greater depth than previously has been possible the complex economic, social, and psychological factors responsible for variation in the labor market experience (including retirement) of this cohort
3. to explore how labor market experiences explain the evolution of careers, wealth, and the preparation of this cohort for the further education of their children and for their own retirement

-
4. to analyze the impact of a changing socio-economic environment on the educational and labor market experiences of this cohort by comparing data from the present study with those yielded by the surveys of the earlier NLS cohorts of young men (which began in 1966 and ended in 1981) and young women (which began in 1968 and ended in 2003), as well as the more recent NLSY97 cohort of young men and women born in the years 1980-84 and interviewed for the first time in 1997
 5. to consider how intergenerational links between mothers and their children, including the employment-related activities of women affect the subsequent cognitive and emotional development of their children, and how the development of the children affects the activities of the mother
 6. to meet the data-collection and research needs of various government agencies with an interest in the relationships among child and maternal health, drug and alcohol use, and juvenile deviant behavior and child outcomes such as education, employment, and family experiences

The NLSY79 has several characteristics that distinguish it from other data sources and make it uniquely capable of meeting the major purposes described above. The first of these is the breadth and depth of the types of information that are being collected. It has become increasingly evident in recent years that a comprehensive analysis of the dynamics of labor force activity requires an eclectic theoretical framework that draws on several disciplines, particularly economics, sociology, and psychology. For example, the exploration of the determinants and consequences of the labor force behavior and experience of this cohort requires information about (1) the individual's family background and ongoing demographic experiences; (2) the character of all aspects of the environment with which the individual interacts; (3) human capital inputs such as formal schooling and training; (4) a complete record of the individual's work experiences; (5) the behaviors, attitudes, and experiences of closely related family members, including spouses and children; and (6) a variety of social-psychological measures, including attitudes toward specific and general work situations, personal feelings about the future, personality characteristics, and perceptions of how much control one has over one's environment.

A second major advantage of the NLSY79 is its longitudinal design, which permits investigation of labor market dynamics that would not be possible with one-time surveys and allows directions of causation to be established with much greater confidence than cross-sectional analyses. Also, the considerable geographic and environmental information available for each respondent for each survey year permits a more careful examination of the impact that area employment and unemployment considerations have for altering the employment, education, and family experiences of these cohort members and their families.

Third, the oversampling of blacks and Hispanics, together with the other two advantages mentioned above, makes possible more sophisticated examinations of human capital creation programs than previously have been possible. Post-program experiences of "treatment" groups can be compared with those of groups matched not only for preprogram experience and such conventional measures as educational attainment, but also for psychological characteristics that have rarely been available in previous studies.

As has been indicated above, the study has several general research and policy-related objectives.

Attachment 3- Survey Applications

We elaborate on these basic purposes by setting forth a series of specific research themes.

Attachment 4- Analysis of Content of Interview Schedules.

The detailed content of the interview schedule is then related to the themes from Attachment 3.

Attachment 5- New Questions and Lines of Inquiry

Summarizes the new questions and lines of inquiry in the proposed questionnaire.

Attachment 6- Respondent Advance Letters

Provides the advance letters that will be sent to respondents prior to data collection and the questions and answers about uses of the data, confidentiality, and burden that will appear on the back of each advance letter.

Attachment 7-Analysis of Attrition and Its Implication in the NLSY79

Analysis of non-response in NLSY79 and of its effects on estimates of labor market outcomes.

The NLSY79 is part of a broader group of surveys that are known as the BLS National Longitudinal Surveys program. In 1966, the first interviews were administered to persons representing two cohorts, Older Men ages 45-59 in 1966 and Young Men ages 14-24 in 1966. The sample of Mature Women ages 30-44 in 1967 was first interviewed in 1967. The last of the original four cohorts was the Young Women, who were ages 14-24 when first interviewed in 1968. The survey of Young Men was discontinued after the 1981 interview, and the last survey of the Older Men was conducted in 1990. The Young and Mature Women surveys were discontinued after the 2003 interviews. The most recent cohort added to the NLS program is the National Longitudinal Survey of Youth 1997 (NLSY97), which includes persons who were ages 12-16 on December 31, 1984.

The National Longitudinal Surveys are used by BLS and other government agencies to examine a wide range of labor market issues. The most recent BLS news release that examines NLSY79 data was published on August 24, 2017, and is available online at <http://www.bls.gov/news.release/nlsoy.nr0.htm>. In addition to BLS publications, analyses have been conducted in recent years by other agencies of the Executive Branch, the Government Accountability Office, and the Congressional Budget Office. The surveys also are used extensively by researchers in a variety of academic fields. A comprehensive bibliography of journal articles, dissertations, and other research that have examined data from all National Longitudinal Surveys cohorts is available at <http://www.nlsbibliography.org/>.

3. Improved Information Technology to Reduce Burden

The field staff of NORC makes every effort to ensure that the information is collected in as expeditious a manner as possible, with minimal interference in the lives of the respondents. Our success in this regard is suggested by the very high continuing response rate and low item refusal rates that have been attained. More recent efforts also have advanced technologies that lower respondent burden.

During Round 11 (1989) of the NLSY79, about 300 cases were collected using a Computer Assisted Personal Interview (CAPI). During Round 12 (1990) CAPI was again used, this time for about 2,400 cases using a longer and more complex questionnaire. The CAPI efforts in 1989 and 1990 were designed to assess the feasibility of the method and the effect of CAPI methods on data quality. Since 1993, the NLSY79 has been conducted using CAPI for all cases. Note that some cases have been completed over the telephone in each round. In recent rounds, the percentage of cases completed by phone is around 85%. The system has proved to be stable and reliable and is well received by interviewers, respondents, and researchers.

An analysis of the Round 12 (1990) experimental data revealed that the quality of the data, as measured by missing or inconsistent responses, was greatly improved by CAPI. The effects on the pattern of responses were minor, although some answer patterns were affected because the CAPI technology in certain cases changed the way the questions were asked. Production data from Rounds 15-21, based on over 80,000 completed interviews, showed these predictions of high-quality data were correct.

In 1994 we began to administer the Child Survey assessments using CAPI. CAPI simplified the administration of these assessments and reduced interview time and respondent burden. We saw the

average interview time fall throughout the field period, and quality control measures revealed the data to be in excellent condition. The Young Adult survey, begun in 1994, originally had a CAPI interview with an additional paper and pencil self-report booklet. In 2000, this booklet was converted to CAPI. In the 2014 fielding, the entire Young Adult questionnaire was conducted in CAPI format.

CAPI surveying will continue to be used in Round 28. We estimate the average interview would be 10-15 percent longer if paper and pencil were used instead of CAPI.

In Round 28, we will continue to use computer interviewing to reduce the number of validation reinterviews needed to ensure that field interviewing is accurate. The survey includes a number of questions for which interviewers should not be able to fabricate answers, such as the respondent's date of birth, height, and weight. Entries for these questions that are inconsistent with information we already have could signal that an interviewer falsified an interview. We will then conduct reinterviews only in cases where an inconsistency suggests that a check is needed. This approach, first used in Round 20, allowed us to reduce the number of validation reinterviews from about 1,250 in Round 19 to about 200 in Round 20. In rounds 25 and 27, no reinterviews were needed. In round 26, 12 validation reinterviews were conducted.

In Round 28, as in Round 27, we also plan to continue using an additional computerized check to ensure that an interview has taken place. Our laptops are equipped with an internal microphone and recording capability, and the phone system we use during the "Early Bird" phase of data collection also has a recording capability. We will record portions of the interview at points unknown to the interviewer. If we have any suspicions about a given interview, we can listen to the sound files and ensure that an interview is taking place. In addition to identifying any interviews that are completely falsified, the recording of interview segments also can reveal when an interviewer is taking shortcuts or not reading questions as instructed or when respondents are having trouble understanding questions. We will use these recordings only to evaluate interviewer performance and validate suspicious interviews.

We will include in the introduction to the questionnaire a statement that "Parts of this interview may be recorded for quality control purposes. This will not compromise the strict confidentiality of your responses" ("CONSENT-1200" in the main NLSY79 questionnaire and "INTRO" in the young adult questionnaire). If the respondent does not want to be recorded, for interviews conducted during the regular field period the laptop recording capability will be immediately turned off. In the Early Bird phase of data collection (discussed in section A.9), the technology does not enable us to stop recording without breaking off the interview, ending the phone call, and calling the respondent back. If an Early Bird respondent does not consent to having the interview recorded, we will delete the recording without listening to it.

4. Efforts to Identify Duplication

A 1982 study entitled "National Social Data Series: A Compendium of Brief Descriptions" by Richard C. Taeuber and Richard C. Rockwell includes an exhaustive list with descriptions of the national data sets available at that time. A careful examination of all the data sets described in their listing indicates that no other data set would permit the comprehensive analyses of youth and young adult employment that can be conducted using the National Longitudinal Surveys. Indeed, the absence of alternative data sources was the deciding factor in the Department of Labor's determination (in 1977) to sponsor the NLSY79. The longitudinal nature of the survey and the rich background information collected mean that no survey subsequently released can provide data to replace the NLSY79. The expansion in the mid-1980s of the NLSY79 to incorporate information on child outcomes represents a unique intergenerational data-collection effort.

Survey staff have continued to confirm that no comparable data set exists. An investigation into data sets related to wealth by F. Thomas Juster and Kathleen A. Kuester describes the data available in the wealth domain, showing the unique nature of the data available in the NLS.

The 1993 volume *The Future of the Survey of Income and Program Participation* points out a number of contrasts between the Survey of Income and Program Participation (SIPP) and the NLS and other major longitudinal studies (see especially pages 77, 107, and 265–7). This book was written primarily to review the SIPP but helps put the major longitudinal surveys in perspective.

BLS convened a conference in fall 1998 to look at the design of the NLSY79 for 2002 and beyond. In the process, external reviewers assessed the coverage, quality, and duplication of the survey. In its central areas of coverage—event histories on labor supply, major demographic events, and child assessments—this conference concluded that the NLSY79 was well-designed and continued to be a unique resource, unduplicated in the national statistical inventory.

An article by Michael Pergamit, et al. (2001) discusses the strengths of the NLSY79, and prominent research that had been done using NLSY79 data. Many of these studies could not have been done otherwise, because the NLSY79 has such a breadth of information, along with a history of all jobs ever held and a cognitive test score. A later paper by Aughinbaugh, et al. (2015) examines the unique strengths of the NLSY79 and NLS97 surveys, and the resulting important research.

As the NLSY79 respondents are getting closer to their retirement years, the BLS convened a panel of aging, health, and retirement experts to give advice on the content and direction the survey needs to take in the next 10-15 years. The Retirement Working Group, led by Professor Kathleen McGarry, met in June 2016 and August 2017. It proved to be very productive, and provided advice on survey content, linkages to administrative data, and such, which the NLSY79 has already begun implementing. The NLSY79 provides advantages for studying aging and retirement over other longitudinal studies, which tend to begin when respondents are much older. The NLSY79 contains a long history of information collected as events occurred. In addition, over half of the sample members have a sibling in the sample, the NLSY79 Child/Young Adult data contains assessments and interviews with children of the female respondents, and the survey contains data on respondents' cognition and health over the years.

As these studies show, there is no other longitudinal data set available that can address effectively the many research topics highlighted in attachment 3. This data set focuses specifically and in great detail on the employment, educational, demographic, and social-psychological characteristics of a nationally represented sample of younger baby boomers and their families and measures changes in these characteristics over long time periods. The survey gathers this information for both men and women, and for relatively large samples of black and Hispanic adults. The repeated availability of this information permits consideration of employment, education, and family issues in ways not possible with any other available data set. The combination of (1) longitudinal data covering the time from adolescence; (2) national representation; (3) large black and Hispanic samples; and (4) detailed availability of education, employment and training, demographic, health, child outcome, and social-psychological variables make this data set, and its utility for social science policy-related research, unique.

In addition to the content of the interviews, the survey is also distinctive because of its coverage of the respondents' lives for more than 35 years and the linkage between concurrently collected data on mothers and their children. These aspects attract the thousands of users who rely on this survey for their studies of American workers, their careers, and their families.

References:

Aughinbaugh, Alison, Charles R. Pierret, and Donna S. Rothstein. "The National Longitudinal Surveys of Youth: Research Highlights." *Monthly Labor Review* (September 2015).
<https://doi.org/10.21916/mlr.2015.34>.

Citro, Constance C. and Kalton, Graham, eds. *The Future of the Survey of Income and Program Participation*. Washington, DC: National Academy Press, 1993.

Juster, F. Thomas and Kuester, Kathleen A. "Differences in the Measurement of Wealth, Wealth Inequality and Wealth Composition Obtained from Alternative U.S. Wealth Surveys." *Review of Income and Wealth* Series 37, Number 1 (March 1991): 33-62.

Pergamit, Michael R., Charles R. Pierret, Donna S. Rothstein, and Jonathan R. Veum. "Data Watch: The National Longitudinal Surveys." *Journal of Economic Perspectives* 15, 2 (Spring 2001): 239-253.

Taeuber, Richard C. and Rockwell, Richard C. "National Social Data Series: A Compendium of Brief Descriptions." *Review of Public Data Use* 10, 1-2 (May 1982): 23-111.

5. Impact on Small Businesses

The NLSY79 is a survey of individuals in household and family units and therefore does not involve small organizations.

6. Consequences of Less Frequent Data Collection

The core of the National Longitudinal Surveys is the focus on labor force behavior. It is very difficult to reconstruct labor force behavior retrospectively and still maintain sufficient precision and data quality. This is the single most important reason we strive to maintain regular interviews with these respondents, who on average have frequent transitions in employment, income and earnings, and family and household structure. Historic dates relating to these transitions are difficult to reconstruct when one focuses on events earlier than the recent past. For those who are employed, retrospective information on wages, detailed occupations, job satisfaction, or other employment-related characteristics cannot be easily recalled.

As with employment-related information, data about a respondent's education and training history are also difficult to recall retrospectively. Completion dates of training and education programs are subject to severe memory biases. Thus, causal analyses that require a sequencing of education, training, and work experiences cannot be easily or accurately accomplished with historical data. Not only are completion dates of educational and training experiences frequently difficult to recall, but there is also evidence that misreporting of program completion is not unusual.

The precise timing and dating of demographic, socio-economic, and employment events, so crucial to most labor force analysis, is in most instances impossible to reconstruct accurately through retrospective data collection that extends very far into the past. For example, we have evidence from the NLS that dates of events of fundamental importance, such as marriage and birth histories, are subject to considerable error at the disaggregated level when collected retrospectively. Respondents have difficulty recalling when their marriages began or ended. Also, accurate information about household structure, how it changes over time, and how this relates to changes in family income and labor force dynamics is difficult to reconstruct retrospectively, as is the information on the health and related behaviors of the respondents, their spouses, and their children.

Finally, it is important to emphasize that information of a subjective nature can only be accurately reported and collected on a contemporaneous basis. Recollection of attitudes may be colored by subsequent experiences or reflect a rationalization of subsequent successes or failures. Attitudes as widely diverse as one's ideas about women's roles or how one perceives one's health as of an earlier

period can be recollected inaccurately, even when respondents are trying to be as honest as they can. In addition, the older the events that one tries to recall, either objective or subjective in nature, the greater the likelihood of faulty recall. The recall of events or attitudes is often biased either by a tendency to associate the event with major life-cycle changes (that may or may not be in temporal proximity to what one is trying to recall) or to move the event into the more recent past. The cognitive and socio-emotional information collected for the children of the NLSY79 respondents is, of course, sensitive to the age and life-cycle stage through which the particular children are progressing and, in many instances, changes over time. This is the reason why we need to repeat some of the assessments, in order to measure the extent to which the results are related to the age of the child as well as intervening family and non-family activities.

While more frequent interviewing is desirable, financial limitations prompted the NLSY79 to move to a biennial interview cycle beginning in 1994. The data loss due to reduced frequency is somewhat ameliorated by the fact that the cohort is more established, having negotiated the school-to-work transition with varying degrees of success. The NLSY79 uses bounded interviewing techniques and is designed so that when respondents miss an interview, information not collected in the missed interview is gathered in the next completed interview. In this way, the event history on work experience is very complete.

A study was conducted to assess the impact of the longer recall period by using an experimental design in the 1994 interview. About 10 percent of the respondents who were interviewed in 1993 were given a modified instrument that was worded as if the respondents were last interviewed in 1992. Using this experimental structure, we examined the respondents' reports on experiences between the 1992 and 1993 interviews using information from their 1993 and 1994 reports on that same reference period. As expected, recall was degraded by a lower interview frequency. Events were misdated and some short duration jobs were not reported when the reference period was moved back in time. Based on this evidence, it is clear that less frequent data collection adversely affects longitudinal surveys.

A second potential problem caused by the move to a biennial interview is a decline in our ability to locate respondents who move. We have been able to compensate for this so far, but a change to less frequent interviewing would likely have a more negative impact.

7. Special Circumstances

None of the listed special circumstances apply.

8. Federal Register Notice and Consultations

XXX comments were received as a result of the Federal Register notice published in 82 FR 61030, on December 26, 2017.

There have been numerous consultations regarding the NLSY79. Preceding the first round of the NLSY79, the Social Science Research Council sponsored a conference at which academic researchers from a broad spectrum of fields were invited to present their views regarding the value of initiating a longitudinal youth survey and what the content of the survey should include. The initial survey development drew heavily on the suggestions made at this conference, which were published in a proceeding under the auspices of the SSRC.

In 1988, the National Science Foundation sponsored a conference to consider the future of the NLS. This conference consisted of representatives from a variety of academic, government, and nonprofit research and policy organizations. There was enthusiastic support for the proposition that the NLS should be

continued in the current format, and that the needs for longitudinal data would continue over the long run. The success of the NLS, which was the first general-purpose, longitudinal labor survey, has helped reorient survey work in the United States toward longitudinal data collection and away from simple cross sections.

Also, on a continuing basis, BLS and its contractors encourage NLS data users to suggest ways in which the quality of the data can be improved and to suggest additional data elements that should be considered for inclusion in subsequent rounds. We encourage this feedback through the public information office of each organization and through the ‘Suggested Questions for Future NLSY Surveys’ (available online at <https://www.nlsinfo.org/nlsy-user-initiated-questions>).

Individuals from other Federal agencies who were consulted regarding the content of the 2018 survey include:

Regina Bures
Demographic and Behavioral Sciences Branch
Center for Population Research
National Institute of Child Health and Human Development

Marsha Lopez
Chief, Epidemiology Research Branch
Division of Epidemiology, Services, and Prevention Research
National Institute on Drug Abuse

John Phillips
Associate Commissioner, Office of Research, Evaluation, and Statistics
Social Security Administration

The NLS program has a Technical Review Committee that advises BLS and its contractors on questionnaire content and long-term objectives. The committee typically meets twice a year. Table 2 below shows the current members of the committee.

Table 2. National Longitudinal Surveys Technical Review Committee (December 2017)

Shawn Bushway School of Criminal Justice University at Albany, State Univ. of NY	Lingxin Hao Department of Sociology Johns Hopkins University	David Johnson Research Professor, Survey Research Center University of Michigan
Lance Lochner Department of Economics Western University	KATHLEEN MCGARRY Department of Economics University of California, Los Angeles	Jesse Rothstein Goldman School of Public Policy University of California, Berkeley
John Cawley	Judith Hellerstein (Chair)	Teresa Seeman Department of Medicine

Department of Policy Analysis & Management and Department of Economics Cornell University	Department of Economics University of Maryland College Park, MD	David Geffen School of Medicine University of California, Los Angeles
Kristen Olson Department of Sociology University of Nebraska-Lincoln	Rebecca Ryan Department of Psychology Georgetown University	John Phillips Associate Commissioner Social Security Administration

The NLS Technical Review Committee convened a conference in 1998 to review the current and future design of the NLSY79. This conference indicated that the central design of the NLSY79 remained strong, although changes in the nation's welfare program required changes in the program reciprocity section of the survey. Many of these changes were implemented in the 2000 and 2002 interviews. Some health section modifications were introduced in 2006 (cognitive functioning model), and the 2008 survey included a new health module for respondents who had reached age 50 (mirroring the age 40 module).

In addition to the Technical Review Committee, the decisions concerning which child outcome measures to include in the child assessment sections of the NLSY79 were carefully considered from a number of perspectives. NICHD has collaborated with BLS on the NLSY79 for many years, with NICHD providing funds for topical modules that are added to the core interview. This collaboration reduces the total cost of data collection for the government. NICHD staff consults with experts outside NICHD to determine priorities for the modules the institute funds. NICHD staff and nationally recognized panels of experts jointly made the decisions about NICHD-sponsored survey topics from medicine and the social sciences.

The NICHD also has convened groups of outside experts to review the progress and content of the data collected for NICHD within the NLSY79 program. A brief description of outside experts consulted with respect to the child instruments and their affiliations may be found in table 3 below.

Table 3. Advisors and Experts Contributing to NLSY79 Child and Young Adult Surveys

Kenneth Wolpin Department of Economics Rice University Kristi Williams Department of Sociology Ohio State University James R. Walker Department of Economics University of Wisconsin Christopher Browning Department of Sociology Ohio State University Wendy Manning Department of Sociology Bowling Green State University	Joe Rodgers Department of Psychology University of Oklahoma Sandra L. Hofferth Department of Family Studies University of Maryland Freya L. Sonenstein Professor & Director Center for Adolescent Health Johns Hopkins Bloomberg School of Public Health Scott Stanley Department of Psychology Center for Marital and Family Studies University of Denver
---	---

9. Payment to Respondents

Because this is a long-term study requiring the subjects to be reinterviewed regularly, respondents are offered financial incentives to secure their cooperation. Respondent payments are appropriate given the continuing nature of the survey; 2018 will be the 28th round and the 39th year since the survey began. We conducted an experiment toward the end of the Round 19 (2000) field period to determine the effect of higher respondent incentives on reluctant sample members. During Round 20 (2002) we conducted an “Early Bird” experiment that focused on cooperative sample members to determine whether respondent incentives can motivate these sample members to cooperate to a degree that significantly reduces our field costs. We documented the results of the first experiment in our clearance statement for the Round 20 field effort and the second in our clearance statement for Round 21. We summarize those findings below:

1. *Respondent incentives (Round 19)* - Offering a higher incentive to reluctant sample members toward the end of the field period is cost-effective when the sample member refused to participate in the previous round. When the sample member had participated in the previous round and the choice was between \$40 and \$80, we were able to convert a sufficiently high fraction of reluctant sample members by normal means and by a \$40 incentive that the cost in respondent incentives per incremental case was about \$270. When the sample member had refused in the previous round, in looking at the yield from \$40 versus \$80, an additional case cost about \$133 in additional incentives. These costs per incremental case are calculated by dividing the change in total incentives paid by the change in cases interviewed for the \$40 versus \$80 treatments.
2. *Early Bird Experiment (Round 20)* - Offering our historically cooperative sample members a higher incentive if they called the interviewer rather than waiting for the interviewer to call the sample member was very effective in reducing field costs. This experiment allowed us to attain several months of production with fewer hours per case than we had observed in the best week in the past fifteen years. We tried both \$60 and \$80 treatments in this experiment. While the \$60 incentive was more cost-effective, field costs were so low for both treatments that either incentive in return for cooperative behavior is highly favorable. This experiment was conducted in twelve replicates fielded sequentially. We were able to evaluate the experiment for each of the randomly selected replicates within a few weeks of mailing out the materials.

To provide perspective, toward the end of Round 20 in 2002, the weekly field costs per case were over \$500. Respondent incentives represent only a fraction of the total field costs, and higher incentives can be a cost-effective means of increasing response while constraining the overall budget. We face a growing pool of sample members who are reluctant to cooperate. In this section we propose a set of measures to encourage cooperation that involve more than just higher respondent incentives.

Marketing - We have put a great deal of effort into respondent materials that describe the importance of the study. These materials are colorful, easy to read, and produced by experienced marketers. For example, in Round 20 one brochure used jellybeans as a visual theme illustrating the importance of each respondent. We prepare an array of materials aimed at different concerns and types of respondents.

Pulling these three strands together (respondent incentives, Early Bird promotion, and the marketing campaign), we request clearance for the following, integrated conversion strategy:

Main NLSY79 incentive categories

Incentive Type	Completed Round 27	Missed Round 27
Base	\$70	\$70
Early Bird	\$100	\$100
Final Push - Standard	\$20	\$20
Final Push – Enhanced	\$40	\$40
Missed Rounds		\$20 to 40
In-kind	Up to \$10	Up to \$10
Min	\$70	\$90
Max	\$110	\$150
Typical	\$70	\$90

NLSY79 Young Adult incentive categories

Incentive Type	Age 12-13 ¹	Age 14-17	Age 18+, completed R27	Age 18+, missed R27
Base	\$50	\$70	\$70	\$70
Early Bird			\$100	\$100
Fertility Section		\$10 per child	\$10 per child	\$10 per child
Final Push		\$20 only if older sibling also gets offer	\$20	\$20
Missed Rounds				\$20 to 40 ²
In-kind		Up to \$10	Up to \$10	Up to \$10
Min	\$50	\$70	\$70	\$90
Max	\$50	\$100	\$110+\$70 ³	\$150+\$70 ³
Typical	\$50	\$70	\$70 ³	\$100 ³

¹Note that incentive payments for the age 12-13 YAs will be given to the mother and passed to the child at her discretion. This is in keeping with our past practice of giving the mother the payment for child interviews.

²Young Adult grant respondents are eligible for interview only every 4 years once they reach age 30. The missed round payment will be made to these respondents only for rounds for which they were eligible but did not participate; it will not take into account rounds in which they were not included in the eligible sample.

³\$100 or \$150 is the maximum regular incentive and \$70 is the maximum amount anticipated for the additional per-child fertility section fee.

Main NLSY79:

- We request permission to offer a base respondent incentive of \$70. This is the same amount offered in round 27.
- As OMB requested in past rounds, we will make an Early Bird offer to all main NLSY79 respondents. In Round 27, the Early Bird offer was \$100. In Round 28, we request permission to maintain that amount. Because our approach is to tie higher incentives to cooperative behavior, bringing more respondents under the Early Bird rubric will reduce our overall cost structure, hence extending higher incentives is beneficial for both the project and respondents.

- We request permission to repeat an additional incentive payment for respondents who were not interviewed in the previous survey round(s). This is appropriate because these respondents tend to have longer interviews as they catch up on reporting information from missed rounds. Specifically, we would offer respondents \$20 for the first missed round and \$10 per consecutive additional rounds missed, up to \$40 (3 rounds). This premium is still smaller than the \$70 a respondent would have received in base fee for completing the round 27 interview. In prior rounds we have offered increasing incentives of \$10 per missed rounds up to \$30. Because, in the NLSY79 main youth sample, we have the greatest chance to bring sample members back to the survey after a single missed interview, we propose to increase the incentive to \$20 for that first missed round. We propose to continue a graduated premium of \$10 for up to 2 additional missed rounds. A larger premium after the first missed round concentrates our efforts to get respondents to return on the primary re-entry point. We would be careful to inform respondents that this is a one-time additional incentive in appreciation of the additional time needed to catch up in this interview round and that they will not receive this additional amount next round. Based on our experience in the NLSY79 and NLSY97, we anticipate that respondents will appreciate this noninterview premium and will understand the distinction between the base amount and the additional incentive.
- Final push incentive (standard). In rounds 26 and 27 we requested and received permission to offer a 'final push' incentive to encourage cooperation of the most difficult respondents; we would like to continue that incentive this round. Starting after the first 12 weeks of the round, cases that have had at least 6 contact attempts or at least one refusal will be eligible for a final push incentive of up to \$20 (bringing the total incentive to \$90). In order to facilitate a timely close to the fielding, starting 6 months after the start of fielding, all cases will be eligible for this incentive.
- Final push incentive (enhanced). In order to support sample representativeness, we introduce the concept of an enhanced final push value. In round 28, we propose to evaluate sub-group-specific response rates when the full sample is at 60 percent completion. Subgroups may be defined by the following measures in the most recent completed interview: educational attainment, weeks worked (defined in 4 categories), poverty status, rural/urban residence, marital history, presence or absence of a health condition that limits work activity, and each of these with gender or race/ethnicity. Any subgroup having a response rate 8 percent or more below the sample average would be offered an enhanced \$40 final push amount instead of the \$20 final push amount. The enhanced final push incentive would be offered only after the overall sample had reached the 60 percent completion mark or 6 months after the start of fielding. Respondents will only ever be offered a single final push value, to discourage appearance of bargaining.
- Small in-kind incentive. Sometimes, a small gift such as a cup of coffee or a toy for a child can facilitate an in-person interview. We request the ability to give up to 200 respondents small in-kind gifts such as these. We would spend no more than \$10 on any respondent and average only \$5.

Young Adult Survey:

- We request permission to offer \$70 as the base incentive for young adult respondents age 14 and older. This is the same as round 27 and mirrors the main youth incentive.
- We request permission to offer \$50 as the base incentive for young adult respondents ages 12-13. We feel that this lower amount is appropriate for the very youngest respondents.
- We request permission to offer the same \$20-\$40 missed round premium for young adult respondents not interviewed in round 27 as we have proposed for the main youth. The rationale for and administration of this incentive are described above. To avoid offering large incentive

amounts to younger respondents, this premium will be offered only to respondents ages 18 and older.

- We request permission to offer an additional \$10 per child to respondents who go through the lengthy fertility section, similar to the additional incentive offered in past rounds to main youth mothers for the mother supplement. This additional fertility section, and hence the incentive, only apply to young adults age 14 and older (in Round 26, only four respondents under age 18 had a co-resident child). The section is asked about each co-resident child age 14 and younger.
- We request permission to offer the same standard final push premium of up to \$20 as proposed for the main youth. This offer would be made following the same rules described in the main youth section above. This offer will be available only to respondents ages 18 and older, and to 14-17 year olds with a sibling getting the offer. We will not offer the enhanced final push incentive for the young adult sample.
- Small in-kind incentive. Sometimes, a small gift such as a cup of coffee or a toy for a child can facilitate an in-person interview. We request the ability to give up to 200 respondents small in-kind gifts such as these. We would spend no more than \$10 on any respondent and average only \$5.

Gatekeepers:

- Some “gatekeepers” are particularly helpful in tracking down reluctant sample members. For example, a parent or sibling not in the survey may provide locating information or encourage a sample member to participate. (Note that we never reveal the name of the survey to these gatekeepers; it is described simply as “a national survey.”) We often return to the same gatekeepers round after round. To maintain their goodwill, we would like to be able to offer gatekeepers who are particularly helpful a small gift worth about \$5. This gift would most likely be a box of chocolates or a small plant. NLS has offered a similar incentive in Rounds 26 and 27 of the NLSY79 and the Round 15 through 17 of the NLSY97. It is used very infrequently, but has proved very effective in finding respondents and getting them to complete interviews. Over Round 26 of the NLSY79 and Rounds 15 and 16 of the NLSY97, only 20 gatekeepers were paid the incentive, but the incentive has resulted in 10 completes.

Electronic Payment of Incentives

In round 27, we offered respondents the opportunity to receive their incentive payment via electronic methods (specifically, PayPal). During the NLSY97 Round 28 main field period, we will continue offering this option to telephone interview respondents. The attraction of electronic payment is large for both respondents and the NLS program. For respondents, electronic payment will generally involve less effort to receive funds than would a paper check or money order. For the NLS program, we can save administration costs and mailing costs by making electronic payments. The cost is also much lower if a payment needs to be tracked or a respondent does not recall having received it. Electronic transaction options are expanding markedly, for example, PayPal, GoogleWallet, ApplePay, ChasePay, etc., and we can expect the options during NLSY79 Round 28 main fielding to be even greater than those available today.

Summary

In our experience, there is no single strategy for securing respondent cooperation. The primary need is for flexibility in how we approach sample members and how we follow up with reluctant sample members. Overall, there are about 2,500 sample members who are hard to interview, either because of busy schedules or a mindset that ranges from indifference to hostility. Our Early Bird efforts attempt to reduce our total costs (incentive plus fielding costs) for cooperative sample members so that we can devote the resources necessary for difficult cases.

We reiterate that incentive payments are only part of our approach to encouraging response. An equally important part of the effort is designing an effective marketing campaign, including conversion materials that interviewers can use to respond to sample members who provide a variety of reasons for not completing the interview. This portfolio of respondent materials backs up the interviewer, providing a variety of approaches to converting refusals. We also encourage national teamwork among the interviewers and periodic calls among interviewers to share their successful “tricks of the trade” with each other that turn reluctant sample members into completed cases. Conversion materials and the ability to employ flexible respondent incentives also have important effects on interviewer morale. With a combination of a marketing campaign to “sell” the survey and the ability to personalize their approach to each respondent, interviewers will not feel they are alone and forced to deal with a difficult task of persuasion without the tools to do the job.

Locating is also a key aspect of obtaining respondent cooperation. Over the past several rounds, we have developed continuous locating strategies to keep track of sample members between interviews. For example, we monitor area code changes, which have become more frequent and play havoc with the accuracy of the phone numbers for our sample members. We review the record of calls to identify subsets of sample members for whom a particular style of advance conversion letter will respond to their particular concerns. We also will continue to include questions in the survey that solicit the views and opinions of respondents. Respondents often find such questions interesting and engaging, and the questions help to build credibility with respondents that the survey is a serious scientific and policy-related endeavor.

Our primary goal must be to continue in the good graces of the respondents. When we feel respondents are under heavy stress and suspect that additional contacts will be unproductive, we will set the case aside and try again in two years. Angering sample members is not an option in the face of their ability to screen and reject our calls. Our incentive efforts and contacting approach will continue our efforts to motivate sample members, assuage their concerns, and convey our interest in them as individuals, not numbers.

10. Confidentiality of Data

a. BLS Confidentiality Policy

The information that NLSY79 respondents provide is protected by the Privacy Act of 1974 and the Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA).

CIPSEA safeguards the confidentiality of individually identifiable information acquired under a pledge of confidentiality for exclusively statistical purposes by controlling access to, and uses made of, such information. CIPSEA includes fines and penalties for any knowing and willful disclosure of individually identifiable information by an officer, employee, or agent of the BLS.

Based on this law, the BLS provides respondents with the following confidentiality pledge/informed consent statement:

“We want to reassure you that your confidentiality is protected by law. In accordance with the Confidential Information Protection and Statistical Efficiency Act of 2002, the Privacy Act, and other applicable Federal laws, the Bureau of Labor Statistics, its employees and agents, will, to the full extent permitted by law, use the information you provide for statistical purposes only, will hold your responses in confidence, and will not disclose them in identifiable form without your informed consent. All the employees who work on the survey at the Bureau of Labor Statistics and its contractors must sign a document agreeing to protect the confidentiality of your information. In fact, only a few people have access to information about your identity because they need that information to carry out their job duties.

Some of your answers will be made available to researchers at the Bureau of Labor Statistics and other government agencies, universities, and private research organizations through publicly available data files. These publicly available files contain no personal identifiers, such as names, addresses, Social Security numbers, and places of work, and exclude any information about the States, counties, metropolitan areas, and other, more detailed geographic locations in which survey participants live, making it much more difficult to figure out the identities of participants. Some researchers are granted special access to data files that include geographic information, but only after those researchers go through a thorough application process at the Bureau of Labor Statistics. Those authorized researchers must sign a written agreement making them official agents of the Bureau of Labor Statistics and requiring them to protect the confidentiality of survey participants. Those researchers are never provided with the personal identities of participants. The National Archives and Records Administration and the General Services Administration may receive copies of survey data and materials because those agencies are responsible for storing the Nation’s historical documents.”

BLS policy on the confidential nature of respondent identifiable information (RII) states that “RII acquired or maintained by the BLS for exclusively statistical purposes and under a pledge of confidentiality shall be treated in a manner that ensures the information will be used only for statistical purposes and will be accessible only to authorized individuals with a need-to-know.”

By signing a BLS Agent Agreement, all authorized agents employed by the BLS, contractors at CHRR and their subcontractors pledge to comply with the Privacy Act, CIPSEA, other applicable federal laws, and the BLS confidentiality policy. No interviewer or other staff member is allowed to see any case data until the BLS Agent Agreement, Department of Labor Rules of Conduct, BLS Confidentiality Training certification, and Department of Labor Information Systems Security Awareness training certification are on file. Respondents will be provided a copy of the questions and answers shown in Attachment 6 about uses of the data, confidentiality, and burden. These questions and answers appear on the back of the letter that respondents will receive in advance of the Round 28 interviews.

b CHRR and NORC Confidentiality Safeguards

CHRR and NORC have safeguards to provide for the security of NLS data and the protection of the privacy of individuals in the sampled cohorts. These measures are used for the NLSY79 as well as the other NLS cohorts. Safeguards for the security of data include:

1. Storage of printed survey documents in locked space at NORC.
2. Protection of computer files at CHRR and NORC against access by unauthorized individuals and groups. Procedures include using passwords, high-level “handshakes” across the network, data encryption, and fragmentation of data resources. As an example of fragmentation, should someone intercept data files over the network and defeat the encryption of these files, the meaning of the data files cannot be extracted except by referencing certain cross-walk tables that are neither transmitted

nor stored on the interviewers' laptops. Not only are questionnaire response data encrypted, but the entire contents of interviewers' laptops are now encrypted. Interview data are periodically removed from laptops in the field so that only information that may be needed by the interviewer is retained. If a laptop is lost or stolen it can be wiped clean remotely.

3. Protection of computer files at CHRR and at NORC against access by unauthorized persons and groups. Especially sensitive files are secured through a series of passwords to restricted users. Access to files is strictly on a need-to-know basis. Passwords change every 90 days.
4. To assure the CHRR and its subcontractor NORC are monitoring all cybersecurity risks to NLS assets and data, in addition to regular self-assessments, the NLS system undergoes a full NIST 800-53 audit every 3 years using an outside independent cybersecurity auditing firm.

Protection of the privacy of individuals is accomplished through the following steps:

1. Oral permission for the interview is obtained from all respondents, after the interviewer ensures that the respondent has been provided with a copy of the appropriate BLS confidentiality information and understands that participation is voluntary.
2. Information identifying respondents is separated from the questionnaire and placed into a non-public database. Respondents are then linked to data through identification numbers.
3. The public-use version of the data, available on the Internet, masks or removes data that are of sufficient specificity that individuals could theoretically be identified through some set of unique characteristics.
4. Other data files, which include variables on respondents' State, county, metropolitan statistical area, zip code, and census tract of residence and certain other characteristics, are available only to researchers who undergo a review process established by BLS and sign an agreement with BLS that establishes specific requirements to protect respondent confidentiality. These agreements require that any results or information obtained as a result of research using the NLS data will be published only in summary or statistical form so that individuals who participated in the study cannot be identified. These confidential data are not available on the Internet.
5. In Round 28 we will continue several training and procedural changes to increase protection of respondent confidentiality. These include an enhanced focus on confidentiality in training materials, clearer instructions in the Field Interviewer Manual on what field interviewers may or may not do when working cases, and formal separation procedures when interviewers complete their project assignments. Online and telephone respondent locating activities have been moved from NORC's geographically dispersed field managers to locating staff in NORC's central offices. Respondent social security numbers were removed from NORC and CHRR records during Round 23.

11. Sensitive Questions

Several sets of questions in the NLSY79 and Children of the NLSY79 might be considered sensitive. This section describes these questions and explains why they are a crucial part of the data collection. All of these topics have been addressed in previous rounds of the surveys, and respondents generally have been willing to answer the questions. Respondents are always free to refuse to answer any question that makes them feel uncomfortable.

NLSY79 main survey

Income, Assets, and Program Participation. One major set of sensitive questions collects information about respondents' income and assets. The interviews record information about the sources and amounts of income received during the past calendar year by the respondent, his or her spouse or partner, or other family members. Income sources identified include the respondents' and their spouses' or partners' wages and salaries, income from military service, profits from a farm or business, Social Security, pensions and annuities, and alimony/child support. These questions, or variants, have appeared in the NLSY79 since 1979. The interviews collect information about what assets and debts the respondent and his or her spouse or partner hold, as well as the values of these assets and debts. Variants of these questions have appeared in the NLSY79 since 1979. While some respondents refuse to answer these questions, our item nonresponse rate is lower than for most surveys. The survey also asks about income received by the respondent and spouse or partner from unemployment compensation, Temporary Assistance for Needy Families (TANF), food stamps, Supplemental Security Income, and other public assistance programs. While questions on program participation have changed over the years, they have been in the survey for the past 30 years.

As part of the financial section, we include a brief series of questions on personal finance. These questions, which explore issues closely related to income, ask respondents to report problems that may affect their ability to obtain credit. Issues explored include missed payments on bills, credit cards on which the respondent carries the maximum balance, and whether the respondent has been turned down for credit in the past 5 years. These questions were asked in Rounds 21, 23, 25, 26, and 27 of the NLSY79, and similar questions are asked in the Survey of Consumer Finances. We did not experience nonresponse problems in previous rounds.

Income and assets questions are central to the usefulness of the NLSY79 data collection. Most analyses based on the survey include some measure of the financial resources available to the respondent, either as an input variable, outcome variable, or control. It is very difficult to conceive of a replacement measure that would function in the same way in research about the returns to education, participation in the labor market, child development, and so on. The public assistance questions additionally permit research on the effects of the welfare reforms enacted in 1997, providing important information to public officials charged with overseeing the country's public assistance programs.

In Round 28, we plan to continue the use of follow-up questions for respondents who answer "don't know" or "refuse" to income and asset questions. As described in a report submitted to OMB with the Round 22 clearance package, we have tested these follow-up questions to ensure that we are getting the best information possible. Based on this analysis, we implemented a hybrid approach beginning in Round 22 and plan to continue using this approach in Round 28. Briefly, respondents who answer an income or asset question with don't know or refuse are first asked to provide a self-generated range. Respondents who are unable to answer the self-generated intervals question then are asked unfolding bracket questions, where they are asked if the amount is more or less than a specified amount, and then more or less than a second specified amount. These unfolding brackets are a common follow-up approach in major surveys including the Health and Retirement Study. To limit any negative effects from using a hybrid follow-up, we skip potentially uncooperative respondents (for example, those giving at least one refusal in the income and asset section) past the hybrid follow-up to subsequent refusals.

Menopause: The NLSY79 will ask a brief set of questions in order to date the onset of menopause. These questions are useful in research on fertility. Similar questions were asked in the NLS Young Women.

Cigarette and Alcohol Use. Questions on smoking ask whether the respondent has smoked more than 100 cigarettes in his or her life, the age the respondent began smoking daily, whether the respondent now smokes daily, the age when the respondent quit smoking, and the number of cigarettes smoked per day. These questions have been asked in several previous rounds of the survey with very low nonresponse rates.

The survey also will include a series of four questions asking whether the respondent drank alcohol in the past 30 days, the number of days on which alcohol was consumed, the average number of drinks per day, and the number of days on which the respondent consumed 6 or more drinks (an indication of alcohol abuse). These questions are important for both economic and social research. Economists are interested in the impact that alcohol use and abuse may have on employment and earnings (for example, Dooley and Prause 1998; Kenkel and Wang 1998). Sociologists and public health researchers can use alcohol data, along with the other information collected, to examine the social and psychological impact of alcohol use and abuse.

The set of alcohol questions has been asked previously in identical or similar form (the time reference period was different in the early surveys) in 1982–85, 1988, 1989, 1992, 1994, 2002, 2006, 2008, 2010, 2012, 2014, and 2016. In the earlier surveys, the questions were generally part of a longer series on alcohol use and abuse. No difficulties were encountered with these longer series in past rounds, and nonresponse is very low. For example, for the set of four questions being included this round, the largest number of respondents who refused or answered “don’t know” in 2010 was 18. No problems were experienced in Rounds 22–26 with the shorter and less intrusive set of questions, and none are expected in Round 28.

NLSY79 young adults

Income. Young adults (those age 14 and older) are asked a series of income questions similar to, but somewhat less detailed than, those asked of their mothers in the main interview. As described above, income data are crucial to many kinds of analysis in a variety of research fields. The young adult data additionally allow researchers to examine similarities or differences in the income sources of mothers and their children, providing information about the transmission of the ability to earn income across generations.

Smoking, Drug and Alcohol Use, and Criminal Activity. Young adults age 12 and older are asked about smoking, drug use, and alcohol use. These questions record whether the respondent has ever used a number of substances, including alcohol, cigarettes, marijuana, cocaine, and other drugs, and ask about the extent of the respondent’s use in the past 30 days. Additional delinquent and criminal behavior questions record whether the young adult has run away from home or been convicted of criminal activities such as selling drugs, possessing drugs, theft, assault, and so on. If the respondent reports convictions, he or she is asked to describe the type of crime committed and the punishment received.

Questions about substance use and criminal behavior are crucial in understanding the education and employment outcomes of this group of young adults. To quote a report based on data from the 1990 Youth Risk Behavior Surveillance System (U.S. Department of Health and Human Services), “Patterns of tobacco, alcohol and other drug use usually are established during youth, often persist into adulthood, contribute substantially to the leading causes of mortality and morbidity, and are associated with lower educational achievement and school dropout.” One concern with long-term drug and alcohol use is the

gateway effect that can occur, leading to use of heavier drugs and an increase in other risky behaviors (for example, sexual activity or criminal acts). For examples of such research, see Pacula (1997); Desimone (1998); and Parker, Harford, and Rosenstock (1994). The negative relationship between drug and alcohol use and educational attainment has been investigated by authors such as Yamada, Kendix, and Yamada (1996). Finally, as mentioned above, substance use may have a negative effect on the probability of employment and the wages and benefits received.

These sensitive questions have been asked in a nearly identical form since 1994 without difficulty. Refusals and don't knows often have been less than 1 percent. Prior to computer-assisted administration, some respondents did not fill out the self-report booklets correctly or completely. Because these instruments are now computer-administered, starting in 2000 for young adults 15 and older and in 2002 for those under age 15, this problem has been ameliorated.

School Safety. Reflecting the growing national concern with weapons in schools, questions on this topic have been administered to adolescents in several other national surveys. Questions on whether a respondent has carried a weapon or been threatened by a weapon have been directed toward adolescents ranging in age from 12-18 in the following surveys:

- The National Youth Study (Tulane, 1998)
- Welfare, Children, and Families: A Three-City Study (Johns Hopkins, 2001)
- NLSY97 (BLS, 1997-2004)
- Youth Risk Behavior Surveys (CDC)

Other surveys have asked questions about whether young respondents carry a weapon to school for protection (Josephson Institute, 1999).

For many previous rounds, the NLSY79 Child and Young Adult surveys have included questions about the child's attitudes and opinions regarding school, including whether the child feels safe at school. In 2002, we added two related questions that ask whether a respondent has ever seen a student with a gun, knife, or other weapon on school property and, if so, the number of times in the last year that the respondent has seen a student with a gun, knife, or other weapon on school property. These questions will continue to be asked of respondents aged 12–18 who are attending school. These questions will aid researchers in investigating the presence of weapons in schools as it relates to school characteristics, neighborhood environment, child behavior, child success in school, subsequent criminal behavior, and so on.

Sexual Activity. Young adults aged 14 and older are also asked about sexual intercourse. Because puberty and the initiation of sexual activity occur during the teenage years for many youths, and because this information may not be recalled accurately if collected retrospectively, it is important to ask these questions of respondents in this age range in each survey round. Results from a number of different surveys, including early rounds of the NLSY97, indicate that a significant proportion of adolescents report that they are sexually active. It is vital that we continue to trace the progression of sexual activity in relation to the realization of educational and occupational goals and with respect to the promotion of good health practices. The level of sexual activity and contraceptive use are important indicators of how serious young people are about reaching higher levels of educational and occupational attainment, and there should be significant congruence between anticipated life goals, sexual activity, and its associated outcomes. These questions also provide information important for analyses of programs and policies related to adolescent health and well-being.

Further, age at first intercourse is important to understanding labor market behavior because of the central role that adolescent fertility plays in the future life course of women. Early childbearing not only retards

the education of the mother and hence is deleterious to her labor market opportunities but also tends to play a powerful role in the intergenerational transmission of poverty. AIDS and other sexually transmitted diseases also make sexual behavior a significant public health issue. For these reasons, this line of questioning is important to the central point of the survey.

Suicidal Ideation. In round 25, we added two questions on suicidal ideation which will continue in Round 28 of respondents aged 14 and up. We believe this is an important topic to address and one that can be captured relatively simply. As diagnoses of depression have risen in recent years among adolescents and young adults, the use of anti-depressants also has risen. A supposed side effect of these medications is an increase in suicidal ideation. No large scale surveys currently exist, however, that enable one to estimate the extent of suicidal ideation in the general population across a range of youthful ages. The NLSY79 Young Adults, with repeated CES-D scores, represent an ideal sample to use. These data also will provide researchers with an opportunity to answer important questions about socioeconomic determinants of suicidal risk. Both questions are taken from the CDC's Youth Risk Behavior Survey. If a respondent answers that he or she has seriously considered suicide during the last 12 months, we will instruct interviewers to provide the respondent the National Suicide Prevention Lifeline number which is staffed through the U.S. Department of Health and Human Services' Substance Abuse and Mental Health Services Administration (SAMHSA). This number takes in all calls and then funnels them to an available crisis center.

Gender Identity and Sexual Orientation. With growing openness about and awareness of these issues and how they might affect the lives of our respondents across multiple domains, and in direct response to requests from researchers, we are adding items on this topic beginning in 2018. Sexual and gender minorities, particularly those transitioning into adulthood, are at a significantly increased risk of mortality and morbidity (e.g. substance use, mental health conditions), and thus understanding the key determinants of health among sexual and gender minorities is a fundamental and imperative research aim for ameliorating health disparities at large. Although gender identity and sexual orientation have sometimes been combined into a single survey question, after reviewing question wording in several large national studies and in published research (Betts, 2009; GenIUSS, 2014; Miller and Ryan, 2011; Redford and Van Wagen, 2012), we plan to ask about gender identity separately from sexual orientation. Several large studies, such as the National Survey of Family Growth (NSFG), the National Health and Nutrition Examination Study (NHANES), and the General Social Survey (GSS), have included sexual orientation and/or gender identity questions in their surveys. The Center for Disease Control (CDC) has done cognitive testing on gender identity and sexual orientation questions for inclusion in the National Health Interview Survey (NHIS) and the National Center for Health Statistics (NCHS) surveys.

The sensitive questions on substance use, criminal behavior, and sexual activity are only asked with the consent of a parent or guardian of a young adult under age 18. We inform parents about the questions we ask, and both parents and teenagers are free to refuse to answer these questions. Our experience has been that participants recognize the importance of these questions and rarely refuse to answer. To further protect respondents and encourage honest reporting, these questions are contained in a self-administered section of the interview for any young adults interviewed in person. Because most young adults will be interviewed on the telephone, the sensitive questions have been written in such a way that the respondent can answer the questions without revealing personal information to anyone (such as a parent) who might overhear the conversation. Although we now ask about the age of the respondent's most recent sexual partner and his or her relationship with the respondent, no identifying information is collected about sexual partners.

Informed Consent. At OMB's request, we conducted cognitive testing before Round 20 to determine whether children and young adults understand the informed consent statement. A report summarizing this

research was submitted with the Round 20 OMB clearance package. We will continue to use the consent statement developed as a result of that research and used for the first time in Round 20.

References

- Betts, Peter. "Developing survey questions on sexual identity: Cognitive/in-depth interviews." Office for National Statistics, July 2009.
- Desimone, Jeffrey. "Is Marijuana a Gateway Drug?" *Eastern Economic Journal* 24,2 (Spring 1998): 149-163.
- Dooley, David and Prause, Joann. "Underemployment and Alcohol Misuse in the National Longitudinal Survey of Youth." *Journal of Studies on Alcohol* 59,6 (November 1998): 669-80.
- The GenIUSS Group. "Best Practices for Asking Questions to Identify Transgender and Other Gender Minority Respondents on Population-Based Surveys." J.L. Herman (Ed.). Los Angeles, CA: The Williams Institute, 2014.
- Harford, Thomas C. and Muthen, Bengt O. "Adolescent and Young Adult Antisocial Behavior and Adult Alcohol Use Disorders: A Fourteen-Year Prospective Follow-Up in a National Survey." *Journal of Studies on Alcohol* 61,4 (July 2000): 524-528.
- Kenkel, Donald S. and Wang, Ping. "Are Alcoholics in Bad Jobs?" NBER Working Paper No. 6401, National Bureau of Economic Research, March 1998.
- Miller, Kristen and J. Michael Ryan. "Design, Development and Testing of the NHIS Sexual Identity Question." National Center for Health Statistics, October 2011.
- Pacula, Rosalie Liccardo. "Adolescent Alcohol and Marijuana Consumption: Is There Really a Gateway Effect?" NBER Working Paper No. 6348, National Bureau of Economic Research, January 1997.
- Parker, Douglas A.; Harford, Thomas C.; and Rosenstock, Irwin M. "Alcohol, Other Drugs, and Sexual Risk-Taking among Young Adults." *Journal of Substance Abuse* 6,1 (1994): 87-93.
- Redford, Jeremy and Aimee Van Wagen. "Measuring Sexual Orientation Identity and Gender Identity in a Self-Administered Survey: Results from Cognitive Research with Older Adults." Presented: San Francisco, CA, Population Association of America Meetings, May 2012.
- Yamada, Tetsuji; Kendix, Michael; and Yamada, Tadashi. "The Impact of Alcohol Consumption and Marijuana Use on High School Graduation." *Health Economics* 5,1 (January-February 1996): 77-92.

12. Estimation of Information Collection Burden

The main NLSY79 interview will be administered to approximately 6,900 respondents, and the average response time is about 75 minutes per respondent. Note that with the elimination of the separate child survey, this time estimate includes the time for administration of the set of questions previously called the "Mother Supplement". These questions only apply to about 95 younger children and will be addressed to about 80 mothers.

The Young Adult survey will be administered to approximately 5,060 respondents ages 12 and older. These respondents will be contacted for an interview regardless of whether they reside with their mothers. The expected average response time for respondents without children is 50 minutes per respondent in the 12-13 age group, 66 minutes per respondent in the 14-18 age group, 63 minutes per respondents in the 19-

22 age group, 60 minutes per respondent in the 23-28 age group, and 70 minutes per respondent age 29 and older (the longer interview time is due to the age 29 health module and the longer gap between interviews starting at age 31). Respondents across all age groups with children in their household take an average of 20 additional minutes of survey time; total time including this additional child section is reflected in the table below. Members of the Young Adult grant sample are contacted for interviews every other round once they reach age 31.

The time required to finish an interview varies in the sample. While women with young children take longer to answer the fertility and childcare questions, men are asked more questions about employment because they tend to hold more jobs. The data show the standard deviation of interview time for the main survey is around 25 minutes.

During the field period, less than 10 interviews may be validated to ascertain whether the interview took place as the interviewer reported and whether the interview was done in a polite and professional manner. These reinterviews average about 6 minutes each. As mentioned earlier, we employ a number of analytical techniques in order to limit the number of validation interviews required and to target these interviews to the respondents most likely to have had a problematic experience.

Table 4 below summarizes the estimated respondent burden for Round 28 of the NLSY79.

Table 4. Number of respondents and average response time by survey questionnaire, Round 28

Instrument	Total Respondents	Total Responses	Average Time per Response	Estimated Total Burden
NLSY79 Round 28 Main Survey	6,900	6,900	75 minutes	8,625 hours
Round 28 Validation Interviews	10	10	6 minutes	1 hour
Young Adult Survey (Ages 12 to 13)	20	20	50 minutes	17 hours
Young Adult Survey (Ages 14 to 18)	294	294	66 minutes	323 hours
Young Adult Survey (Ages 19 to 24)	962	962	63 minutes	1,010 hours
Young Adult Survey, Grant component (Age 25 to 28), interview	1,210	1,210	60 minutes	1,210 hours
Young Adult Survey, Grant component (Age 29 and older), interview	2,574	2,574	70 minutes	3,003 hours
TOTALS¹	11,960	11,970	—	14,189 hours

¹The total number of 11,960 respondents across all the survey instruments is a mutually exclusive count that does not include the 10 reinterview respondents, who were previously counted among the main and young adult survey respondents.

13. Cost Burden to Respondents or Record Keepers

Respondents for this survey will not incur any capital and start-up costs; respondents also will not incur any operation, maintenance, or purchase-of-service costs.

14. Estimate of Cost to the Federal Government

The total estimated cost for the main NLSY79 and the associated Young Adult surveys in Round 28 is \$13.5 million. This cost includes survey management, questionnaire design, instrument development, pretest and main data collection including incentive payments, cleaning and preparation of data files for users, and services to users of the data files.

The BLS bears approximately \$11.5 million of this cost. The National Institute of Child Health and Human Development is expected to provide approximately \$2 million in funding for Round 28 through interagency agreements with the BLS and through a grant awarded to researchers at the Ohio State University Center for Human Resource Research (CHRR). The interagency agreements funds interviews with children and young adults up to age 21. The grant funds interviews of young adults age 22 and older.

15. Change in Burden

The estimated total respondent burden of 14,189 hours for Round 28 is higher than the estimated burden of 13,964 hours for Round 27. The increase is due to an increase in the youth interview length as well as a larger number of older Young Adult respondents and Young Adult respondents with children, who tend to have longer interview times.

16. Plans and Time Schedule for Information Collection, Tabulation, and Publication

Following transmittal of the survey data from NORC to CHRR, staff at CHRR spend approximately 14 months cleaning the data and preparing a main NLSY79 public-use data file. A month later, the Child and Young Adult data files are released and reports are written for the National Institute of Child Health and Human Development. The timing of these events is as follows:

NLSY79 Young Adult Interviews	October 2018– October 2019
Main NLSY79 Interviews	October 2018– October 2019
Data Cleaning and Coding	October 2019–January 2020
Public-Use Data File Preparation	February–November 2020
Release of Main NLSY79 Public-Use Data File	December 2020
Release of Young Adult File	January 2021

17. Reasons Not to Display OMB Expiration Date

The OMB number and expiration date will be provided in the advance letter.

18. Exceptions to “Certificate for Paperwork Reduction Act Submissions”

We do not have any exceptions to the “Certificate for Paperwork Reduction Act Submissions” statement.

National Longitudinal Survey of Youth 1979
1220-0109
November 2017

**Information Collection Request for
The National Longitudinal Survey of Youth 1979
OMB # 1220-0109
Part B**

Submitted by the Bureau of Labor Statistics

TABLE OF CONTENTS

<i>B. Collections of Information Employing Statistical Methods.....</i>	<i>3</i>
1. Respondent Universe and Respondent Selection Method.....	3
2. Design and Procedures for the Information Collection.....	4
3. Maximizing Response Rates.....	5
4. Testing of Questionnaire Items	7
5. Statistical Consultant	8

B. Collections of Information Employing Statistical Methods

1. Respondent Universe and Respondent Selection Method

The initial NLSY79 sample was selected to represent (after appropriate weighting) the U.S. civilian and military population of 33,570,000 persons who were ages 14 to 21 as of December 31, 1978. The sample selection procedure included an overrepresentation of Hispanic and black youths so as to include sufficient sample cases to permit racial and ethnic analytical comparisons. Economically disadvantaged nonblack/non-Hispanic youths also were oversampled.

The NLSY79 originally included a supplemental sample of youths in the military. In 1985, the military supplemental sample was discontinued. In 1991, the economically disadvantaged nonblack/non-Hispanic oversample was discontinued. Appropriate weights have been developed so that the sample components can be combined in a manner to aggregate to the overall U.S. population born in the years 1957-64 and living in the United States when the sample was selected in 1978. The number of sample cases in 1979, excluding the discontinued military and nonblack/non-Hispanic samples, was 9,964. A breakdown by sex and race is depicted in table 5 below. We anticipate a response rate in Round 28 that is similar to the Round 27 experience.

Table 5. Interviews Completed in 1979 and 2016 (Preliminary) by Race and Sex

Race and Hispanic origin	1979			2016					
	Number of men	Number of women	Total number	Number of men	Retention rate for men	Number of women	Retention rate for women	Total number	Total retention rate
Nonblack, non-Hispanic	2,518	2,484	5,002	1,649	65.5	1,758	70.1	3,407	68.1
Hispanic	981	980	1,961	628	64.0	684	69.8	1,312	66.9
Black	1,524	1,477	3,001	1,048	68.8	1,144	77.5	2,192	73.0
Total	5,023	4,941	9,964	3,325	66.2	3,586	72.6	6,911	69.4

Retention rates for the NLSY79 are significantly affected by attrition due to death. Approximately 9.1% of the 9,964 NLSY79 respondents still eligible for interviewing were deceased after the 2016 survey; we are investigating ways to determine whether some respondents we cannot locate may be deceased. Table 6 provides information about retention (percent of all respondents interviewed) and response (percent of living respondents interviewed) rates for each year of the NLSY79.

Table 6. NLSY79 retention and response rates by sample type

Year	Number Interviewed	Retention Rate ¹	Number of Deceased Respondents	Response Rate ¹
1979	12,686	—	—	—
1980	12,141	95.7	9	95.8
1981	12,195	96.1	29	96.3
1982	12,123	95.6	44	95.9
1983	12,221	96.3	57	96.8
1984	12,069	95.1	67	95.6
1985	² 10,894	93.9	79	94.5
1986	10,655	91.8	95	92.6
1987	10,485	90.3	110	91.2
1988	10,465	90.2	127	91.2
1989	10,605	91.4	141	92.5
1990	10,436	89.9	152	91.1
1991	³ 9,018	90.5	144	91.8
1992	9,016	90.5	156	91.9
1993	9,011	90.4	177	92.1
1994	8,891	89.2	204	91.1
1996	8,636	86.7	243	88.8
1998	8,399	84.3	275	86.7
2000	8,033	80.6	313	83.2
2002	7,724	77.5	346	80.3
2004	7,661	76.9	399	80.1
2006	7,654	76.8	456	80.5
2008	7,757	77.9	503	82.0
2010	7,565	75.9	573	80.6
2012	7,301	73.3	689	78.7
2014	7,066	70.9	790	77.0
2016	⁴ 6,911	69.4	⁴ 909	76.3

¹ Retention rate is defined as the percentage of base-year respondents remaining eligible who were interviewed in a given survey year; deceased respondents are included in the calculations. Response rate is defined as the percentage of base-year respondents remaining eligible and not known to be deceased who were interviewed in a given survey year.

² A total of 201 military respondents were retained from the original sample of 1,280; 186 of the 201 participated in the 1985 interview. The total number of NLSY79 civilian and military respondents eligible for interview beginning in 1985 was 11,607.

³ The 1,643 economically disadvantaged nonblack/non-Hispanic male and female members of the supplemental subsample were not eligible for interview as of the 1991 survey year. The total number of NLSY79 civilian and military respondents eligible for interview beginning in 1991 was 9,964.

⁴ Preliminary.

2. Design and Procedures for the Information Collection

The survey includes telephone or personal-visit interviews with all the respondents, regardless of their place of residence. At each interview, detailed information is gathered about relatives and friends who could be of assistance in locating the sample member if he or she cannot be readily located in the subsequent survey round. Most interviews in Round 28 will be carried out between

October 2018 and October 2019, with the field period extending later into 2019 if necessary. Every effort is made to locate respondents, as the attrition information above suggests. Interviewers are encouraged to attempt contacting respondents until they are located. There is no arbitrary limit on the number of callbacks. The success of NORC interviewers in this regard is indicated by a very low rate of attrition over the first 27 rounds of the survey. More than 76 percent of the living, in-scope, original respondents were surveyed in Round 27.

Preceding the data collection, interviewers are carefully trained, with particular emphasis placed on resolving sensitive issues. Most of the interviewers have lengthy experience in the field from participation in earlier NLSY79 interview rounds, as well as from involvement with other NORC surveys. Experienced interviewers receive self-study training consisting of over 8 hours spent on specially designed materials requiring study of the questionnaire and question-by-question and procedural specifications, with exercises on new or difficult sections and procedures. Experienced interviewers working on the Early Bird phone phase also receive several hours of remote training (by computer and conference call), mainly focused on using the call management system. All interviewers must successfully complete a practice interview with their supervisor before they are permitted to begin field work.

Efforts to assure quality data from the field are instigated at several points. The first 100 cases completed are reviewed, answer by answer, to determine whether there are any problems with the instrument. After this, every case identified by the interviewer as having a problem during the interview is reviewed in detail. Throughout the field period, individual cases are checked for problems, and rapid feedback is given to the interviewers so they can improve interviewing methods.

We will reduce burden by employing targeted validation. Cases that have unusual patterns in terms of length, time of day, break-offs, an incorrect entry to the question on the respondent's date of birth, or height and weight entries that are inconsistent with previous rounds will be validated. We also use review of recordings for validation purposes. If a case fails to validate, the entire caseload of the interviewer will be validated through review of recordings or validation re-interviews.

The questionnaires are prepared by professional staff at CHRR. When new materials are incorporated into the schedule, special assistance is generally sought from appropriate experts in the specific substantive area. The technical expertise of staff at NORC is also used in this regard.

Because sample selection took place in 1978 in preparation for the 1979 baseline interview, sample composition has remained unchanged except for the discontinuation of some of the oversamples as previously mentioned. A more detailed discussion of sampling methodology is available from the NLSY79 Technical Sampling Report at:
<http://www.nlsinfo.org/content/cohorts/nlsy79/other-documentation/technical-sampling-report>.

In an effort to reduce respondent burden while still providing a broad spectrum of variables for researchers and policymakers to use, certain topical modules are cycled in and out of the survey from one round to the next. Although the data from these modules are important, it is not necessary to collect data on all topics in every round. An example of such a topical module is the assets module, now asked every other round.

3. Maximizing Response Rates

A number of the procedures used to maximize the response rate already have been described in items 1 and 2 above. The success of the procedures is demonstrated by the low attrition rates

indicated in tables 5 and 6. Hispanic attrition has been slightly higher than among blacks and nonblack/non-Hispanics, and attrition for men is higher than for women. Attachment 7 analyzes the patterns of non-response and its implications on the estimates of labor market outcomes. Clearly, it is more difficult to gain the cooperation of some respondents. Past non-response indicates a lower probability of participation in subsequent rounds. Still, there is little evidence that any selective response biases the analytical results. To the best of our knowledge, the NLSY79 has the best retention rate of any longitudinal survey in the U.S. We note, however, that interviewing becomes a little more difficult each round.

The other component of missing data is item nonresponse. The rate of item nonresponse due to the refusal of a respondent to answer a particular question in the NLSY79 is between 1 and 2 percent per question, depending on how questions are counted. The highest nonresponse rates occur for income and asset items.

One natural issue for longitudinal surveys is to determine whether the sample still represents its portion of the U.S. population. The NLSY79 originally was weighted to represent the 1978 population of 14-21 year-olds and closely matches the official statistics for that year. Sampling weights are prepared each year to adjust the remaining sample to representative proportions. These sampling weights are released with the other data on the public-use data file.

To investigate the issue of continued sample representation, table 7 compares numbers from the 2010 decennial census with NLSY79 population estimates. Census data are taken from the [American FactFinder](#) page of the Census Bureau website.^[1] The American FactFinder website provides the number of people living in the United States, who were ages 46 to 53 on April 1, 2010, which is the same age group that the NLSY79 sample represents. NLSY79 population estimates are from the weighted results of the Round 24 (year 2010) survey.

Table 7 shows the percentage of the NLSY79 sample and U.S. population by sex and race. Overall, the table has two significant features. First, the 2010 NLSY79 sample slightly overrepresents men, since there is a larger percentage of men in the NLSY79 sample than in the U.S. population. For comparison, the original NLSY79 sample in 1979 was composed of 50.8 percent men and 49.2 percent women, while the Census Bureau reported a 1979 population for the same age group that was 50.6 percent men and 49.4 percent women. (Source: U.S. Bureau of the Census, publication P25-917, *Preliminary Estimates of the Population of the United States, by Age, Sex, and Race: 1970 to 1981*; and 1979 NLSY79 data) Hence, the current male-female composition of the NLSY79 does not suggest any gender-biased attrition. Rather, the composition is similar to the sex ratio of the original panel. Nevertheless, the comparison with the recent Census Bureau estimates may suggest that the NLSY79 sample is experiencing lower male mortality than the overall U.S. population of the same age.

Second, the NLSY79 sample underrepresents the current U.S. population of Hispanics. The NLSY79 sample does not include persons who entered the United States after 1978, and the rate of immigration among Hispanics has been very high since the NLSY79 sample was selected. (The differences between the NLSY79 sample and the recent U.S. population estimates in the

^[1] Three series were extracted from the American FactFinder website; PCT12, PCT12J, and PCT12H which track the U.S. population's age by sex and race/ethnicity. PCT12 tracks the total population, PCT12J tracks black non-Hispanics and PCT12H tracks the Hispanic population. Nonblack, non-Hispanic figures are computed by subtracting PCT12J and PCT12H from PCT12. All data come from the Census 2000 Summary File 1 (SF 1), which combines both short and long form answers and does not subsample any information.

percentages of nonblack/non-Hispanics and black non-Hispanics can be explained largely by the shortfall of Hispanics in the NLSY79 sample.) Comparing the NLSY79 sample with the U.S. population estimates for 1978, the NLSY79 sample correctly represents the Hispanic population on a weighted basis, and as described earlier in this document, the NLSY79 sample intentionally *overrepresents* the 1978 Hispanic population on an unweighted basis.

Overall, table 7 shows that, except for the Hispanic population, the NLSY79 sample still is similar to the U.S. population estimates for the same age group. If one accounts for the large amount of Hispanic immigration since the survey began, the remaining differences are not large. Moreover, the weights that are produced after each round compensate for the modestly different rates of attrition and mortality across demographic groups.

Table 7. NLSY79 Weighted Sample Composition in 2010 versus U.S. Census Data for Persons Ages 43 to 56 as of April 1, 2010

	2010 NLSY79	Census Data
Total	100.0%	100.0%
Men	50.9%	49.2%
Women	49.1%	50.8%
Nonblack, non-Hispanic	79.3%	75.9%
Men	40.2%	37.5%
Women	39.1%	38.4%
Black, non-Hispanic	14.2%	12.0%
Men	7.3%	5.7%
Women	6.9%	6.4%
Hispanic	6.5%	12.0%
Men	3.4%	6.0%
Women	3.1%	6.0%

4. Testing of Questionnaire Items

BLS is cautious about adding items to the NLSY79 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 NLSY79 design process employs a multi-tiered approach to the testing and review of questionnaire items.

When new items are proposed for the NLSY79 questionnaire, we often adopt questions that have been used previously in probability sample surveys. 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, the National Institute of Aging-funded Health and Retirement Study and Midlife in the United States, and other Federally funded surveys.

New questions are reviewed in their proposed NLSY79 context by survey methodologists who consider the appropriateness of questions (reference period, terms and definitions used, sensitivity, and so forth). Questions that are not well-tested with NLSY79-type respondents undergo cognitive testing with convenience samples of respondents similar to the NLSY79 sample members.

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, and comparison of NLSY79 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, to get a divorce, or to retire from the labor force. These respondents are often atypical, but their interviews can reveal problems in question functionality or comprehensibility.

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.

5. Statistical Consultant

Kirk M. Wolter

Senior Fellow and Director

Center for Excellence in Survey Research

NORC