

SUMMARY

This package requests clearance for the pretest and main fielding of Round 13 of the National Longitudinal Survey of Youth 1997 (NLSY97). The main NLSY97 sample includes 8,984 respondents who were born in the years 1980 through 1984 and lived in the United States when the survey began in 1997. Sample selection was based on information provided during the first round of interviews. This cohort is a representative national sample of the target population of young adults. The sample includes an overrepresentation of blacks and Hispanics to facilitate statistically reliable analyses of these racial and ethnic groups. Appropriate weights are provided so that the sample components can be combined in a manner to aggregate to the overall U.S. population of the same ages.

The NLSY97 pretest sample includes 201 respondents who were born in the years 1980 through 1985 and lived in the United States at the time of the initial pretest fielding in 1996. Pretest data are used only for project operational and methodological information and are not released to individuals outside of the project team.

The survey is funded primarily by the U. S. Department of Labor. Additional funding has been provided in some years by the Departments of Health and Human Services, Education, Defense, and Justice, and the National Science Foundation. The Bureau of Labor Statistics has overall responsibility for the project. The project is managed by the National Opinion Research Center (NORC) at the University of Chicago and the Center for Human Resource Research (CHRR) at The Ohio State University. NORC handles the overall design and management, interviewing, initial data preparation, and weighting. Additional data cleanup and preparation, the development of final documentation and the preparation of public-use 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 (NLS). Many of the questions in the NLSY97 also have been designed to reflect the changing nature of institutions and the different problems facing this group of young people. Those data elements of a particularly sensitive nature and those not previously collected are justified in this document.

SUPPORTING STATEMENT

National Longitudinal Survey of Youth 1997 (NLSY97) A Survey of Persons who were Ages 12 to 16 on December 31, 1996 Rationale, Objectives, and Analysis of Content

A. Justification

1. Necessity for the Data Collection

This statement covers the pretest and main fielding of Round 13 of the National Longitudinal Survey of Youth 1997 (NLSY97). The NLSY97 is a nationally representative sample of persons who were ages 12 to 16 on December 31, 1996. The Bureau of Labor Statistics (BLS) contracts with the National Opinion Research Center (NORC) at the University of Chicago and the Center for Human Resource Research (CHRR) at The Ohio State University to interview these youths on a yearly basis to study how young people make the transition from full-time schooling to the establishment of their families and careers. The longitudinal focus of this survey requires information to be collected about the same individuals over many years in order to trace their education, training, work experience, fertility, income, and program participation.

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 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

The major purpose of the data collection is to examine the transition from school to the labor market and into adulthood. The study relates each respondent’s educational, family, and community background to his or her success in finding a job and establishing a career. During Round 1, the study included a testing component sponsored by the Department of Defense that assessed the aptitude and achievement of the youths in the study so that these factors can be related to career outcomes. This study, begun when most participants were in middle school or high school, has followed them as they enter college or training programs and join the labor force. Continued yearly interviews will allow researchers and policymakers to continue examining the transition from school to work. This study will help researchers and policymakers to identify the antecedents and causes for difficulties some youths experience in making the school-to-work transition. By comparing these data to similar data from previous NLS cohorts, researchers and policymakers will be able to identify and understand some of the dynamics of the labor market and whether and how the experiences of this cohort of young people differ from those of earlier cohorts.

The NLSY97 has several characteristics that distinguish it from other data sources and make it uniquely capable of meeting the goals 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 a 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 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, and perceptions of how much control one has over one's environment.

A second major advantage of the NLSY97 is its longitudinal design. This design permits investigations 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 permit. Also, the considerable geographic and environment information available for each respondent for each survey year permits a more careful examination of the impact that local labor market conditions have on the employment, education, and family experiences of this cohort and their families.

Third, the supplemental samples of blacks and Hispanics make possible more detailed statistical analyses of those groups than would otherwise be possible.

The NLSY97 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. In 1979, the National Longitudinal Survey of Youth 1979 (NLSY79), which includes persons who were ages 14-21 on December 31, 1978, began. The NLSY79 was conducted yearly from 1979 to 1994 and has been conducted every two years since 1994. One of the objectives of the National Longitudinal Surveys program is to examine how well the nation is able to incorporate young people into the labor market. These earlier surveys provide comparable data for the NLSY97.

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 releases that examines NLSY97 data was published on January 23, 2009, and is available online at <http://www.bls.gov/news.release/pdf/nlsyth.pdf>. 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/>.

More information about survey applications is provided in attachment 2.

3. Improved Information Technology to Reduce Burden

CHRR and NORC have led the industry in survey automation and continue to use state-of-the-art methods for the NLSY97. This includes the continued use of computer-assisted personal interviewing (CAPI) for the survey. For sensitive questions, such as those about drug or alcohol use, the NLSY97 uses an audio computer assisted self-interview (ACASI) that allows the respondent to see the questions on the screen and listen to them through earphones and record the answers on the keyboard. This method helps to make the respondent more comfortable with these questions and encourages more truthful and complete responses. CAPI interviews reduce respondent burden and produce data that can be prepared for release and analysis faster and more accurately than is the case with pencil-and-paper interviews. Mode experiments on another

NLS cohort showed that the same interview took 10 percent less time to administer using a computer. For Round 13, we propose to collect selected locating information via web.

4. Efforts to Identify Duplication

We do not know of a national longitudinal survey that samples this age bracket and explores an equivalent breadth of substantive topics including labor market status and characteristics of jobs, education, training, aptitudes, health, fertility, marital history, income and assets, participation in government programs, attitudes, sexual activity, criminal and delinquent behavior, household environment, and military experiences. Data collection for the National Longitudinal Study of Adolescent Health (Add Health) is less frequent and addresses physical and social health-related behaviors rather than focusing on labor market experiences. The studies sponsored by the National Center for Education Statistics do not include the birth cohorts 1980 through 1984. The Children of the NLSY79, also part of the NLS program, spans the NLSY97 age range and touches on many of the same subjects but does not yield nationally representative estimates for these birth cohorts. Further, the NLSY97 is a valuable part of the NLS program as a whole, and other surveys would not permit the kinds of cross-cohort analyses that are possible using the various cohorts of the NLS program.

The repeated collection of NLSY97 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) a focus on youths and young adults; (3) national representation; (4) large minority samples; and (5) 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 research on youth issues unique.

5. Involvement of Small Organizations

The NLSY97 is a survey of individuals in household and family units and therefore does not involve small organizations.

6. Consequences of Less Frequent Data Collection

BLS believes that annual interviews are the lowest frequency that can be conducted at this stage of respondents' lives without seriously sacrificing data quality. The core questions of the NLSY97 focus on labor force behavior, training, and education. It is very difficult to reconstruct behavior retrospectively in the pattern of being with a job and at work or not at work, being unemployed, and being out of the labor force, and relating those different labor force states to experiences with schooling and training programs. The evaluation of schooling and training as they relate to labor force behavior depends upon having an accurate temporal record of these activities. Less frequent data collection can lead to such crucial errors as whether a particular job began before or after a training program. This is the single most important reason BLS strives to maintain yearly interviews with these young adults, who on average have frequent transitions in employment, income and earnings, schooling, training, 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 recalled easily. If these data are not collected, the ability of BLS to provide information on schooling and employment transitions would be severely compromised.

Information about the vocational aptitudes of the NLSY97 sample will enable BLS to map current trends and the effect of these variables on labor market attachment in the United States; excellent scales are available in the Round 1 collection of Armed Services Vocational Aptitude Battery (ASVAB) data from which new national norms have been created. No other survey has this depth of information, and a comprehensive longitudinal record will support high quality research into the dynamics of the school-to-work transition. Less frequent data collection would impede the ability of researchers and policymakers to exploit these data.

As with employment-related information, data about a respondent's education and training history are also extremely difficult to recall retrospectively. Completion dates of training and education programs are subject to memory biases. Thus, causal analyses that require a sequencing of education, training, and work experiences cannot be easily or accurately accomplished with data collected retrospectively with a long recall period. Not only are completion dates of educational and training experiences frequently difficult to recall, but there is evidence that misreporting of program completion is not unusual. For younger respondents, such as those in the NLSY97, events come rapidly and recall can be more problematic. Less frequent data collection would increase the risk of schooling and training experiences becoming confused with one another.

The precise timing and dating of demographic, socio-economic, and employment-related 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, there is evidence 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. Some respondents even 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 be reported accurately only when it is collected as contemporaneously as possible. 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 describes 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 further in the past one tries to recall events, objective or subjective, 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.

7. Special Circumstances

None of the listed special circumstances apply.

8. Federal Register Notice and Consultations

One comment was received as a result of the Federal Register notice published in 74 FR 8813, on February 26, 2009. The comment, which was e-mailed to BLS on February 28, 2009, expressed the opinion that the survey does not benefit the citizens of the country.

There have been numerous consultations regarding the NLSY97. 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. The participants endorsed the notion of conducting a new youth survey. The NLSY97 incorporates many of the major recommendations that came out of that conference.

Individuals consulted in federal agencies other than BLS regarding the content of the NLSY97 include:

Dr. Christine Bachrach
Chief, Demographic and Behavioral Sciences Branch
National Institute of Child Health and Human Development
6100 Executive Blvd. Room 8B13
Bethesda, MD 20892

Sharon Miller

U.S. Department of Education
Office of Vocational and Adult Education
Potomac Center Plaza
550 12th St SW
Washington, DC 20065

David Goodwin
U.S. Department of Education
Planning and Evaluation Service
400 Maryland Ave., SW
Washington, D.C. 20202

Marsha Silverberg
U.S. Department of Education
Planning and Evaluation Service
400 Maryland Ave., SW
Washington, D.C. 20202

Barbara Allen-Hagen
U.S. Department of Justice
Office of Juvenile Justice and Delinquency Prevention
810 7th Street NW
Washington, D.C. 20531

The NLS program also has a technical review committee that provides advice on interview content and long-term objectives. That group meets twice each year. Table 1 below shows the current members of that committee.

Table 1. National Longitudinal Surveys Technical Review Committee (2008)

David Autor
Massachusetts Institute of Technology
Department of Economics
50 Memorial Drive, E52-371
Cambridge, MA 02142
Email: dautor@mit.edu
Phone: 617-258-7698

Janet Currie
Professor, Department of
Economics
Columbia University
Room 1038 IAB
420 West 118th Street
New York, NY 10027
Email: jc2663@columbia.edu
Phone: 212-854-4520

Jeff Grogger
Harris School of Public Policy
University of Chicago
Suite 139
1155 E. 60th Street
Chicago, IL 60637
E-mail: jgrogger@uchicago.edu
Phone: 773-834-0973

Arie Kapteyn
Senior Economist
RAND Corporation
1776 Main Street
P.O. Box 2138
Santa Monica, CA 90407
Email: kapteyn@rand.org
Phone: 310-393-0411 x7973

Annamaria Lusardi
Dartmouth College
Economics Department
301 Rockefeller Hall
Hanover, NH 03755
Email: Annamaria.Lusardi@Dartmouth.edu
Phone: 603-646-2099

Derek Neal
Professor, Department of
Economics
University of Chicago
1126 E. 59th Street
Chicago, IL 60637
Email: d-neal@uchicago.edu
Phone: 773-702-8166

Seth Sanders
Department of Economics
Duke University
213H Social Sciences Building
Durham, NC 27708
Email: seth.sanders@duke.edu
Phone: 919-660-1800

Chris Taber
Professor of Economics
University of Wisconsin
1180 Observatory Drive
Madison, WI 53706-1391
Email: ctaber@ssc.wisc.edu
Phone: 608-263-7791

Bruce Western
Professor of Sociology
Harvard University
430 William James Hall
33 Kirkland Street
Cambridge, MA 02138
Email: western@wjh.harvard.edu
Phone: (617) 495-3879

9. Payment to Respondents

The NLSY97 is a long-term study in which the same subjects are interviewed annually. For that reason, respondents in all prior rounds have been offered financial and in-kind incentives as a means of securing their long-term cooperation and arresting the decline of response rates. For Round 13 we propose to continue the same incentive structure as Round 12. The base incentive will consist of \$30 cash and an in-kind payment of average value \$10 and a maximum value of \$30. In addition, we would offer an additional \$10 per missed round to those respondents returning to the survey after having missed Round 12 and possibly other prior rounds. Respondents would receive at most \$30 additional dollars for prior missed rounds. This proposal is based on data from a two-round incentive experiment approved by OMB.

We also plan to administer a noninterview respondent (NIR) questionnaire in Round 13 to sample members who have missed at least 5 consecutive rounds and who do not complete the Round 13 interview on first approach. Responding to this questionnaire will still classify sample members as noninterviews for Round 13, but it will enable us to learn more about our long-term nonrespondents and therefore help us to understand attrition patterns and any nonresponse bias. We expect about 600 respondents to be eligible for the NIR questionnaire, with about 120 individuals agreeing to complete the NIR questionnaire. Those who complete the NIR questionnaire will receive \$13 for their effort.

Contingent on outside funding, we propose to collect permission forms to obtain postsecondary educational records (that is, college transcripts) from our respondents. Releases will be sought from all respondents who have received a high school diploma or GED or completed coursework in a postsecondary degree program. Releases would be sought from respondents first by field interviewers at the time of the Round 13 in-person interview, but no incentive payments would be offered at that time. A follow-up mail effort would take place after the close of the Round 13 data-collection period and would request return of signed releases from sample members completing the Round 13 interview by phone or not completing the Round 13 interview at all. Respondents who are asked to mail a signed release would be offered \$10 for their efforts. We propose to offer \$20 to sample members from disadvantaged backgrounds who have historically not cooperated well in past survey activities if those sample members send us a signed release. We plan to offer this added incentive to help ensure a sufficiently large sample representation for this important population group. Measures of disadvantage will include growing up in a single-parent or no-parent family, poverty status, and low ASVAB scores, family income, and parental education. Measures of survey cooperation will include whether the sample member granted permission for the NLSY97 high school transcript study and participated consistently in previous NLSY97 interviews. Of the approximately 6,300 people we expect will grant their permission to obtain college transcripts, about 20 percent are expected to receive the \$20 incentive.

In Rounds 5 through 9, the base incentive payment for respondents was \$20 cash. Beginning in Round 7, prior-round nonrespondents received \$5 for each consecutive prior missed round up to 3 rounds. In Round 10, the base incentive increased to \$30 cash for all respondents. Prior-round nonrespondents who resumed their participation in Round 10 were also paid \$10 extra per missed round, with a cap of \$30 extra.

In October, 2006, BLS proposed conducting an experiment in Rounds 10 and 11 to study two questions that would help us craft an effective long-term incentive policy:

1. Can targeted in-kind incentives result in higher response than cash incentives?
2. Will large increases in incentives have significantly greater impact than smaller increases in the current round and in subsequent rounds of data collection?

With OMB approval, the Rounds 10-11 incentive experiment began in December of 2006. At that time, interviews had been completed with 5,000 respondents who were regarded as very cooperative. The remaining 3,825 sample members were randomized into three equal-sized groups: the control group, the discretionary in-kind treatment group, and the cash payment group. The control group continued to receive the basic Round 10 respondent incentive of \$30 in cash. In addition, respondents who had missed previous rounds continued to receive the increased payments previously approved (\$10 per missed round, up to \$30). This regime continued for Round 11 as well.

The first treatment group (discretionary in-kind payments) received the previously approved \$30, and respondents who had missed previous rounds were eligible for the increased payments of \$10 per missed round. Respondents in this group were also eligible to receive in-kind payments that averaged \$20 in value with a maximum value of \$30. All respondents in this treatment group received some form of in-kind incentive. BLS allowed field managers and field interviewers the ability to determine the in-kind incentive that they judged would be most effective at securing cooperation. This regime continued for Round 11 as well. For the second treatment group (cash incentive), BLS proposed increasing the respondent payment to \$50 in cash. As with the control group and the in-kind treatment group, respondents in the cash treatment group who had missed previous rounds received the payments of \$10 per missed round (up to \$30 maximum). This regime continued for Round 11 as well.

The purpose of this large increase in payments from Round 9 to Round 10 was to examine whether such a large increase would have any impact on response rates. Because the respondents who had already completed the survey at this point could not be affected by the increased payments, we effectively stratified our respondents by their willingness to participate. The failure to detect an impact of the increased respondent incentives for this group of sample members would provide strong evidence that respondent incentives are not effective at increasing response rates.

In Round 11, respondents who had completed their survey prior to random assignment and who are not siblings of participants included in the experiment continued to receive the \$30 respondent payment authorized for Round 10.

This experimental design offered the NLS program an opportunity to test the effectiveness of discretionary in-kind incentives and large increases in cash incentives. Results from rounds 10 and 11 indicate that response rates have increased for respondents in both the two treatment groups. Our results also suggest that an in-kind incentive supplement makes the best use of project dollars to improve respondent cooperation, invest in the unique relationship that each respondent has with the NLSY97, and permit some flexibility in aggregate cost.

The design of this experiment has a strong dynamic component. An advantage of this design is that we can study the effects not only of incentive amounts, but of sequences of incentive amounts on respondent behavior. In the Round 7 experiment, in which we introduced additional payments for missed rounds, we saw the perhaps surprising result that respondents who received increased incentives were more cooperative in subsequent rounds, even though they did not receive any supplemental incentives in those later rounds. We saw similar patterns of improved longer-term cooperation among sample members who received additional incentives in rounds 8-10. When more information is available on Round 12, we can further analyze whether Round 10 and Round 11 incentives affected the Round 12 completion rates.

When we mail the Round 13 advance letter and locating card shown in attachment 6, the envelope also will include either a Presidential dollar coin or a reusable “eco bag” that respondents can use at the grocery store or for other purposes. The purpose of enclosing these items is to give the envelope some heft and make it less likely that recipients of the letter will simply throw it away without opening the envelope. The items also serve as a low-cost gesture of goodwill that might encourage sample members to update their contact information or call to schedule an interview.

10. Confidentiality of Data

a. BLS Confidentiality Policy

The information that NLSY97 respondents provide is protected by the Privacy Act of 1974 and the Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA). CIPSEA is shown in attachment 3.

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.

The Bureau of Labor Statistics Commissioner's Order No. 1-06, "Confidential Nature of BLS Statistical Data," explains the Bureau's policy on confidentiality: "In conformance with existing law and Departmental regulations, it is the policy of the BLS that respondent identifiable information collected or maintained by, or under the auspices of, the BLS for exclusively statistical purposes and under a pledge of confidentiality shall be treated in a manner that will ensure that the information will be used only for statistical purposes and will be accessible only to authorized persons." Commissioner's Order 1-06 is shown in attachment 4.

By signing a BLS Agent Agreement, all authorized agents employed by the BLS contractors at NORC, 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, 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 5 about uses of the data, confidentiality, and burden. These questions and answers will appear on the back of the letter that respondents will receive in advance of the Round 13 interviews. Attachment 6 shows the combination advance letter and locating card that was we plan to use in Round 13.

The following confidentiality pledge is given to NLS respondents:

"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."

b. NORC and CHRR Confidentiality Safeguards

NORC and CHRR 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 NLSY97 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 its subcontractors 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.
3. Protection of computer files at NORC and at CHRR against access by unauthorized persons and groups. Especially sensitive files are secured via a series of passwords to restricted users. Access to files is strictly on a need-to-know basis. Passwords change every 30 days.

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. After the final interview round, respondent identifier computer files will be destroyed.
4. The public-use version of the data, available on the Internet, masks data that are of sufficient specificity that individuals could theoretically be identified through some set of unique characteristics.
5. 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.
6. Questions of a more private nature are contained in self-administered portions of the survey so the respondents’ answers are concealed both from the interviewer and anyone in the household who might overhear the interview.
7. In Round 13 we will continue several training and procedural changes that were started in Round 11 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 the introduction of 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 field interviewer laptops in Round 10 and from NORC and CHRR records during Round 11.

11. Sensitive Questions

Continuing the practice of the last few rounds of the NLSY97, the Round 13 questionnaire includes a variety of items that permit the respondents to provide more qualitative information about themselves. Informal feedback from the interviewers and respondents indicates that this type of subjective data carries greater resonance with respondents as being informative about who they are, rather than the

behavioral data that are the mainstay of the NLSY97 questionnaire. The items selected for self-description are all hypothesized in the research literature to be predictive of or correlated with labor market outcomes. In Round 13, these items include questions about the respondent's favorite person and respondents' perceptions about their life so far. There are several broad sets of questions in the NLSY97 data-collection instruments that may be considered sensitive. We address each of these categories separately below.

a.) Sexual Activity

Because puberty and the initiation of sexual activity occurred for many of the sample members during the first few survey rounds, this information has been carefully collected. Results from a number of different surveys, including early rounds of the NLSY97, indicate that a significant proportion of adolescents between the ages of 13 and 17 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.

The survey will continue to collect information on the number of times male respondents have made a woman pregnant, as well as information on the live births from those pregnancies. Few studies have examined the linkages between early childbearing for men and subsequent education and employment outcomes in relation to family commitments. However, there is now some research indicating a modest connection between male labor supply and how many children they have. In an age where social responsibility is a salient public issue, longitudinal collection of data on the number of children men have fathered is essential to guarantee adequate representation of their children. At a minimum, collection of information about the offspring of male respondents is necessary for linking economic outlays of child support or lack of outlays of child support with potential determinants of both men's and women's labor supply behavior. Cross-sectional estimates of childbearing data for men can underestimate the number of children ever born to males.

Questions related to puberty, dating, sexual activity, birth control, and pregnancy outcomes (including outcomes other than live births) will be asked of all respondents in a self-administered portion of the questionnaire. Respondents will be eligible for the puberty questions only if they had not been interviewed in Round 5 or later. During administration of these questionnaire segments, respondents first will be instructed how to use the computer to enter their responses. They also will be instructed on the use of the audio headset that will allow them to hear a question read to them at the same time that the question text appears on the screen. Question response sets will also be audio as well as visual. The audio portion will help to improve response in situations where literacy or visual impairment is a problem. During self-administration, the computer screen is not visible to the interviewer and the program automatically directs the respondent through the appropriate universe of questions. Upon ending the self-administered section, the program automatically saves the data and the interview reverts back to the next interviewer-administered module; the self-administered section cannot be re-entered during the interview after the respondent exits the section by entering a password. The respondent will be reassured by the interviewer that his or her responses, once entered into the computer, are not available for retrieval by non-survey personnel. There is evidence that using an audio computer-assisted self-interview approach favorably affects data quality (O'Reilly, J.; Hubbard, M., Lessler, J., and Biemer, P., 1992; Johnston, J. and Walton, C., 1992; Kinsey, S., Thornberry, J., Carson, C., and Duffer, A., 1995; Tourangeau, R. and Smith, T., 1995). No respondent will be pressured to answer the questions, and interviewers will be instructed to accept refusals without attempting to encourage response. Previous experience indicates that respondents usually recognize the importance of these questions, and field interviewers generally have not reported difficulties with these types of questions.

b.) Anti-Social Behavior

The educational and labor force trajectory of individuals is strongly affected by their involvement in delinquent and risk-taking behaviors, criminal activity, and alcohol and drug use. There is widespread interest in collecting data on such behaviors. The challenge, of course, is to obtain accurate information on activities that are socially unacceptable or even illegal. Questions on these activities are asked in the self-administered portions of the NLSY97.

Crime and delinquency. The longitudinal collection of self-reported criminal behavior permits examination of the effects of these deviant behaviors on employment activity. This includes the ability to study whether there is a sustained pattern of criminal activities through the life cycle and how these patterns are related to employment difficulties. An additional area of study is the ways in which deviant behaviors may be causally associated with a disposition towards other aberrant behavior such as excessive alcohol and drug use. Use of both self-reports of behavior and of official disciplinary and court actions allows the NLSY97 to separate the effects of criminal activity that lead to an arrest or other legal action versus criminal activity that remains unpunished.

The design of the crime and delinquency module for the NLSY97 has taken great care to avoid weaknesses contained in other surveys' instruments. As a result, unlike other surveys, the NLSY97 elicits information on a wider scope of activities and experiences related to crime. This includes questions concerning the type and frequency of criminal activity as well as self-reports about convictions, time served, and income received as a result of criminal activity.

Experiences with the correctional system. The Round 13 questionnaire continues to ask several questions on incarceration and parole that were added in Round 12. These include finer detail on parole and probation status as well as violations of that status, questions about experiences and services received while incarcerated, and questions about experiences and behaviors since release from incarceration. The questions on experiences and services received while incarcerated will be asked of currently incarcerated respondents and those who were incarcerated and released since the last interview. Respondents who were released from incarceration since the last interview also will be asked about their experiences and behaviors since they were released. These questions appear in the self-administered section, as do all other questions pertaining to arrest and incarceration.

Substance use. 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." It is important that the NLSY97 continue to collect this information because of the potential impact of substance use on education and employment outcomes.

c.) Mental health

The literature linking mental health with various outcomes of interest to the NLSY97, including labor force participation, is fairly well-established. The Round 13 questionnaire includes questions on how many times the respondent has been treated for emotional, mental or psychiatric problems and how many times the respondent missed work or activities because of such problems in the past year.

d.) Religion

The NLSY97 has included questions about religious identification and attendance in most rounds. The Round 13 survey will include questions about religious attendance. Religion and spirituality are an important part of life for a majority of Americans. Belief systems affect a wide variety of outcomes relevant to labor market participation, ranging from the type and intensity of work and career orientations, to labor force participation and other economic outcomes that influence social and economic mobility. Religious denomination and frequency of attendance at worship services

also indirectly affect labor force participation through their impact on other dimensions of individual lives.

e.) Income, Assets, and Program Participation

The survey asks all respondents about their income from wages, salaries, and other income received in the last calendar year. Other income is collected using a detailed list of income sources such as self-employment income, receipt of child support, interest or dividend payments, or income from rental properties. Respondents also are asked about their participation in government programs. Included are specific questions (number of spells, duration of each spell, amount of benefit, and so forth) regarding a number of government assistance programs such as Unemployment Compensation, AFDC/TANF/ADC, and food stamps.

In addition to income, respondents are periodically asked about current asset holdings. Questions include the market value of any residence or business, whether the respondent paid property taxes in the previous year, the average amount spent on utilities per month, and the amount owed on motor vehicles. Other questions ask about the respondent's current checking and savings account balances, the value of various assets such as stocks or certificates of deposit, and the amount of any loans of at least \$200 that the respondent received in the last calendar year. To reduce respondent burden, the asset questions are not asked of each respondent in every round. These questions are asked in the first interview after the respondent turns 18, the first interview after the respondent turns 20, and the first interview after the respondent turns 25. Because asset accumulation is slow at this young age, this periodic collection is sufficient to capture changes in asset holdings.

Given the high fraction of household wealth associated with home ownership, the NLSY97 questionnaire collects home ownership status and (net) equity in the home from respondents each year that they are not scheduled for the full assets module. In Round 13 we have added a question on present value of the home. We also will ask respondents who owned a house or other dwelling previously and no longer live there about what happened to their house or dwelling.

f.) Financial Health

To get a better understanding of the financial well-being of the respondents, we continue in Round 13 questions we introduced in Round 10 about a respondent's financial condition. We ask a set of questions to measure the financial distress of the respondents in the past 12 months. In particular, we ask whether respondents have had a "payday loan" (a loan to be repaid on the next payday), whether respondents have been 60 days late in paying their mortgage or rent, and whether they have been pressured to pay bills by stores, creditors, or bill collectors. In addition, we ask respondents to pick the response that best describe their financial condition from the following list:

1. very comfortable and secure
2. able to make ends meet without much difficulty
3. occasionally have some difficulty making ends meet
4. tough to make ends meet but keeping your head above water
5. in over your head

The goal of these questions is to understand better the financial status of these youths and how this status affects and is affected by their labor market activities.

g.) Skin Color

The Round 12 questionnaire included an item in the interviewer remarks section asking the interviewer, but not the respondent, to code the respondent's skin color on a scale from 0 to 10. This item had previously appeared in the New Immigrant Survey (<http://nis.princeton.edu/>) in the U.S. and in data collections in Latin American countries. The substantive interest in skin color is multi-fold. In the health area, the American Academy of Dermatology (1998) describes skin color as the most important factor determining an individual's risk for skin cancer. Additional health interest focuses on skin color as a risk factor for coronary heart disease (Costas, R. Jr., Garcia-

Palmieri, M.R., Sorlie, P. and Hertzmark, E, 1981) and in association with periodontal disease (Peres, M.A., Antunes, J.L.F., Boing, A.F., Peres, K.G., and Bastos, J.L.D, 2007). In the social science literature, skin color is studied in the context of workplace discrimination (Hersch, J, 2007), social stratification (Keith, V.M, and Herring, C, 1991), and within-household resource allocation (Rangel, M, 2007), especially among American Latinos and African-Americans.

In Round 13, the interviewers will be asked to code the respondent's skin color for all those respondents who either did not complete a Round 12 interview or completed a Round 12 interview by telephone and are completing the Round 13 interview in person.

Respondents are free to refuse to answer any survey question, including the sensitive questions described above. Our experience has been that participants recognize the importance of these questions and rarely refuse to answer.

12. Estimation of Information Collection Burden

The Round 13 field effort will seek to interview each respondent identified when the sample was selected in 1997. NORC will attempt to contact approximately 8,800 sample members who are not known to be deceased. BLS expects that NORC will complete interviews with approximately 7,350 of those sample members. The content of the interview will be similar to the interviews in Round 12. Based upon interview length in past rounds, we estimate the interview will require about 65 minutes. One purpose of the Round 13 pretest is to improve our estimates of the administration time for the Round 13 questionnaire.

Interview length will vary across respondents. For example, the core of the interview covers schooling and early labor market experience. Naturally, respondents vary in the number of jobs they have held, the number of schools they have attended, and their experiences at work and at school. Our aim is to be comprehensive in the data we collect, and this leads to variation in the time required for the respondent to remember and relate the necessary information to the interviewer. In addition, the audio self-administered component of the interview is, in some cases, very engaging, and the respondents sometimes take longer to complete the task than one might expect. For these reasons, the timing estimate is more accurate on average than for each individual case.

The estimated burden in Round 13 includes an allowance for attrition that takes place during the course of longitudinal surveys. To minimize the effects of attrition, NORC will seek to complete interviews with living respondents from Round 1 regardless of whether the sample member completed an interview in intervening rounds. We anticipate that an increased fraction of completed cases will be respondents who are returning to the survey after missing one or more rounds and therefore will have longer interviews than the typical respondent who has been in the survey every year.

Household burden will vary with the number of in-scope sample members present, so households with three sample members may require three hours, and so forth. Although more than 1,800 households included multiple respondents at the time of the initial interview, by Round 13 many respondents will have established their own households, and very few multiple respondent households remain. We are sensitive to the fact that the interviews in households with several sample members theoretically can pose interviewing problems, but that has not been our experience in previous rounds.

During the Round 13 field period, NORC will conduct validation interviews with about 2 percent of respondents to ascertain that the interview took place as the interviewer reported and to assess the quality of the data collected. These cases will be selected purposefully, based on data and assessments by survey management that indicate a field interviewer's caseload merits further scrutiny. These validation interviews average about four minutes each and will be conducted only for the main fielding

For Rounds 10 and 11, BLS requested and was granted OMB approval to record the validation interviews. With this approval, validation interviews were recorded, and standardized abstracts from these recordings were prepared for rating by listeners. Audio recordings were obtained for 520 respondents to Round 10 validation interviews. U.S.-born raters listened to the standardized audio clips and provided their perceptions of the speakers' characteristics, including sex, race, education level, and region of origin.

Raters listened to audio clips in batches of 20. Five listeners were recruited for each batch, resulting in 2,600 (520 x 5) speaker-listener observations.

BLS had three objectives for recording the validation interviews in Rounds 10 and 11.

1. Assess any technical challenges to recording the main interviews in future rounds. Although technical issues prevented us from recording all validation interviews in Round 10, no technical challenges were identified that would pertain to future rounds. In particular, telephone access and instrumentation variation among respondents did not materially affect recording quality.
2. Identify any resistance from respondents when NORC asked for their consent to record the interviews. We did not experience noteworthy resistance from respondents about recording the interviews.
3. Determine whether it is feasible to develop a taxonomy and code respondents' speaking skills. Listeners were able to code perceived demographic characteristics of respondents that were highly correlated with educational attainment and AFQT score, as well as actual demographic characteristics.

With OMB approval, NORC recorded randomly selected segments of the main Round 11 and Round 12 interviews. Recording segments of the main interviews can help BLS and NORC to ensure that the interviews actually took place and that the interviewers did not fabricate the data. Recording can help to ensure that interviewers are reading the questions exactly as worded and entering the responses properly. Recording also can help to identify parts of the interview that might be causing problems or misunderstanding for interviewers or respondents. Early evaluation of these recordings indicates that a large fraction are of acceptable quality for listening and coding.

Based on our experience of recording segments of the main interview in rounds 11 and 12 and favorable results from recording validation interviews in prior rounds, respondents will be asked to provide their consent for the recording of segments of the main interview in Round 13. Recording these interviews will enable BLS to improve data quality while reducing respondent burden.

NORC interviewers in Round 13 will read the following script to ask respondents for their consent to record the main interviews.

“My computer is equipped to record this interview for quality control, training and research purposes. As always your confidentiality is protected by Federal law and the policies of the Bureau of Labor Statistics and NORC. May I continue with the recording?”

YES
NO

If the respondent objects to the recording of the interview, the interviewer will confirm to the respondent that the interview will not be recorded and then proceed with the interview.

Specific questions also will be recorded as part of a proposed research project undertaken by Professor Jeffrey Grogger of the University of Chicago. The data from recordings will be used to attempt to expand our understanding of the role that speech patterns play in explaining labor market differences between blacks and whites. To collect speech data, respondents in Round 13 will be asked to respond to two sets of stimulus questions. One set is designed to capture formal speech, the other casual speech. By priming respondents for formal and informal speech, we plan to construct a measure of code-switching among bi-dialectical respondents. To analyze the effects of speech patterns, we plan to transform the audio responses to our stimulus questions into measures that can be incorporated into econometric models. We propose to do this through two approaches. First, we plan to collect listener perception data of the type employed in Grogger (2008). Second, we plan to construct quantitative dialect density measures (DDM's) based on linguistic analyses of the respondents' speech. The construction of these measures is discussed in further detail in the section on testing of questionnaire items.

To provide perceptions data, listeners will be recruited from NORC's staff of telephone interviewers. Using NORC staff members provides for much greater security than could be achieved through other means. NORC interviewers are trained to maintain the confidentiality of respondent data. Indeed their continued employment depends on their doing so. Interviewers are required to sign the BLS agent agreement, by which they pledge to abide by the terms of the Confidential Information Protection and Statistical Efficiency Act of 2002, the Privacy Act, and BLS policies pertaining to respondent confidentiality and data security.

Although the audio recordings themselves are confidential due to their potential to identify respondents, the measures we propose to construct and make available on public-use files need not be. Perceptions data take the form of the number of anonymous listeners who perceived the speaker to be white, for example. The DDM is a number between 0 and 1, which seems to pose little more confidentiality risk than ASVAB scores and other measures released on NLSY97 public-use files.

Table 2 provides additional details on the estimated respondent burden in Round 13. The estimated total respondent burden for the Round 13 main and pretest interviews is expected to be 8,314 hours.

Table 2. Number of Respondents and Average Response Time, NLSY97 Round 13

Form	Total Respondents	Frequency	Total Responses	Average Time per Response	Estimated Total Burden
NLSY97 Pretest: June-July 2009	150	Annually	150	65 minutes	163 hours
Main NLSY97: September 2009-May 2010	7,350	Annually	7,350	65 minutes	7,963 hours
Validation interview	147	Annually	147	4 minutes	10 hours
NIR questionnaire	120	Annually	120	10 minutes	20 hours
College Transcript Release Form	6,311	Once	6,311	1.5 minutes	158 hours
<i>TOTALS*</i>	<i>7,620</i>	<i>—</i>	<i>14,078</i>	<i>—</i>	<i>8,314 hours</i>

* The difference between the total number of respondents and the total number of responses reflects the fact that about 6,311 are expected to complete the main interview and the college transcript release form. In addition, about 147 respondents will be interviewed twice, once in the main survey and a second time in the 4-minute validation interview.

13. Cost Burden to Respondents or Record Keepers

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

14. Estimate of Cost to the Federal Government

The total estimated cost of the Round 13 survey is \$9,200,000. 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.

15. Change in Burden

The burden of 8,314 hours requested for Round 13 of the NLSY97, which will be conducted during parts of fiscal years 2009 and 2010, is 954 hours higher than the burden of 7,360 hours approved for the pretest and main fielding of Round 12. The increase reflects the expectation that the average interview will be 5

minutes longer in Round 13 than in Round 12. The increase also reflects the planned administration of the NIR questionnaire and the possible collection of permission forms to obtain college transcripts if funding becomes available for that activity.

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

The following is the planned schedule for the data collection for Round 13.

Questionnaire Development	September 2007 – April 2009
Respondent Materials Development	June 2009 – September 2009
Pretest Data Collection	June 2009 – July 2009
Main Data Collection	September 2009 – May 2010
Data Processing	May 2010 – October 2011
Release of Public-Use Main Data Files	August 2011
Publication of BLS News Release	January 2012

The data collection schedule has been placed earlier in the calendar year so that we can minimize the overlap of NLSY97 Round 13 main fielding with the decennial census. This schedule change also moves the NLSY97 pretest into late spring.

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.