

**Information Collection Request for  
The National Longitudinal Survey of Youth 1997  
OMB # 1220-0157  
Summary and Part A**

**Submitted by the Bureau of Labor Statistics**

## TABLE OF CONTENTS

---

<b>Summary</b> .....	<b>1</b>
<b>Supporting Statement</b> .....	<b>3</b>
A. Justification.....	3
1. Necessity for the Data Collection.....	3
2. Purpose of Survey and Data Collection Procedures.....	3
3. Improved Information Technology to Reduce Burden.....	5
4. Efforts to Identify Duplication.....	5
5. Involvement of Small Organizations.....	6
6. Consequences of Less Frequent Data Collection.....	6
7. Special Circumstances.....	7
8. Federal Register Notice and Consultations.....	7
9. Payment to Respondents.....	9
10. Confidentiality of Data.....	14
a. BLS Confidentiality Policy.....	14
b. Contractor Confidentiality Safeguards.....	15
11. Sensitive Questions.....	16
12. Estimation of Information Collection Burden.....	20
13. Cost Burden to Respondents or Record Keepers.....	23
14. Estimate of Cost to the Federal Government.....	23
15. Change in Burden.....	23
16. Plans and Time Schedule for Information Collection, Tabulation, and Publication.....	23
17. Reasons Not to Display OMB Expiration Date.....	24
18. Exceptions to “Certificate for Paperwork Reduction Act Submissions”.....	24

## SUMMARY

This package requests clearance for the pretest and main fielding of Round 17 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 1979 through 1984 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 BLS has previously contracted with the National Opinion Research Center (NORC) at the University of Chicago to conduct the survey. They are responsible for survey design, interviewing, data preparation, documentation, and the preparation of public-use data files.

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. As has been done in earlier cohort surveys, the NLSY97 has moved to biennial data collection which has resulted in some changes in the existing questionnaire. 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 17 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 external organizations to interview these youths, to study how young people make the transition from full-time schooling to the establishment of their families and careers. Interviews were conducted on a yearly basis through round 15; beginning with Round 16 they are interviewed on a biennial basis. 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 biennial interviews will allow researchers and policymakers to examine 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 release that examines NLSY97 data was published on March 26, 2014, 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**

The NLS program and its contractors, 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. To efficiently reach respondents who do not have telephones or who object to spending expensive minutes to receive phone calls on their cellular telephones, we will purchase inexpensive pre-paid cell phones to give to respondents so they may participate in a telephone interview. These phones can be purchased for approximately ten dollars and have lasting value only if the respondent chooses to pre-pay for additional minutes. The existence of these telephones greatly improves our ability to reach certain respondents who otherwise may require repeated in-person visits before a completed interview is obtained. Finally, the use of computer-assisted recorded interviewing (CARI) reduces respondent burden by introducing recordings of the main interview for data quality assurance to replace post-interview validation calls. We propose to continue our nascent web-based interactions with respondents as part of the Round 17 effort in order to bring forward additional uses of technology into the NLSY97 without jeopardizing representativeness and data quality.

### **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**

Historically, the NLSY97 has been conducted annually since it began, and that frequency has been essential for accurately capturing the educational, training, labor market, and household transitions that young people typically experience. Starting in Round 16 data collection changed to a biennial schedule. As NLSY97 respondents complete their 20s and move into their mid-30s, they experience fewer educational and labor market transitions which make biennial data collection more feasible. Increasing the length of time between rounds could affect the quality of the data in two ways. First, retention rates for the sample would be expected to decrease as it becomes harder to track sample members, especially those with less employment or residential stability. When the NLSY79 went from annual to biennial interviewing after 1994, retention rates, which had always been above 90 percent, began to decline, reaching 80 percent within four rounds. Declining retention rates could introduce some bias into survey estimates.

Second, a longer spacing between interviews could make it more difficult for respondents to recall the details of their experiences. When the NLSY79 went from annual to biennial interviewing, the BLS found that respondents failed to report almost one-third of their jobs that had started and ended in the first year of the new two-year recall period. Also, reports of the timing of unemployment insurance and food stamp receipt and dates of separation and divorce were affected.

In Round 16 (the first biennial interview), we continued to collect event histories since the date of last interview for employment, marriage, fertility, and schooling in order to maintain the detailed event history data we have collected since Round 1. However, in other domains the reference period for the interview questions was the last 12 months, which would not increase the length of the recall period for most respondents. As in prior rounds, we continued to use memory aides and bounding techniques in our interviewing to elicit accurate recall of events and dates.

We did experience a decline in retention rate in Round 16, much like the decline experienced by the NLSY79 in its transition to biennial interviewing. The likely causes of the decline were manifold, including a pause in data collection related to a partial shutdown of the federal government, as well as an extremely severe winter that hampered travel, staff and respondent availability and communications. We are still investigating any bias resulting from this decline in retention rate. For Round 17, we do not propose any further changes to the instrument to accommodate biennial interviewing.

Table 1 shows the retention rates experienced on the NLSY79 and the NLSY97 just before and after the transition from annual to biennial interviewing. Although there were important extenuating circumstances affecting the Round 16 NLSY97 fielding, the drop in response rates in the first biennial year was 2.3 percent for the NLSY79 and 3.1 percent for the NLSY97.

**Table 1. NLSY79 and NLSY97 Retention Rates Surrounding the Transition to Biennial Interviewing**

	NLSY79 Retention Rate* (Year)	NLSY97 Retention Rate* (Year) (Round)
2 rounds prior to transition to biennial	92.1 (1993) (R15)	83.7 (2010) (R14)
Round prior to transition to biennial	91.1 (1994) (R16)	83.9 (2011) (R15)
First round after 2-year gap	88.8 (1996) (R17)	80.8 (2013) (R16)
2 <sup>nd</sup> round after transition to biennial	86.7 (1998) (R18)	<i>Projected 81.6 (2015) (R17)</i>

\* Retention rates exclude deceased and out of sample cases.

## 7. Special Circumstances

None of the listed special circumstances apply.

## 8. Federal Register Notice and Consultations

No comments were received as a result of the Federal Register notice published in 79 FR 72704 on December 8, 2014.

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.



## 9. Payment to Respondents

The NLSY97 is a long-term study in which the same respondents have been interviewed on an annual basis. Beginning with Round 16, the interval period between interviews moved to two years. Because minimizing sample attrition is critical to sustaining this type of longitudinal study, respondents in all prior rounds have been offered financial and in-kind incentives as a means of securing their long-term cooperation and slowing the decline in response rates.

Evidence shows that incentives can have effects on both respondent and interviewer behavior, and therefore are indispensable in reaching the project response rates. Incentives result in conversions among those least likely to participate and quicker cooperation among the more likely. Small increases or one-time bonuses can have a halo effect that results in future participation. Many interviewers find that having something to offer respondents, such as a monetary incentive, additional in-kind offering, or new conversion materials, allows them to open the dialogue with formerly reluctant respondents. Interviewers want a variety of options to respond to the particular needs, issues and objections of the respondent. Fully loaded, interviewer costs come to around \$30 / hour, and interviewers spend an average of almost 7 hours to obtain each interview. If incentives reduce this time by only 5%, nearly \$100,000 in interviewer costs are saved.

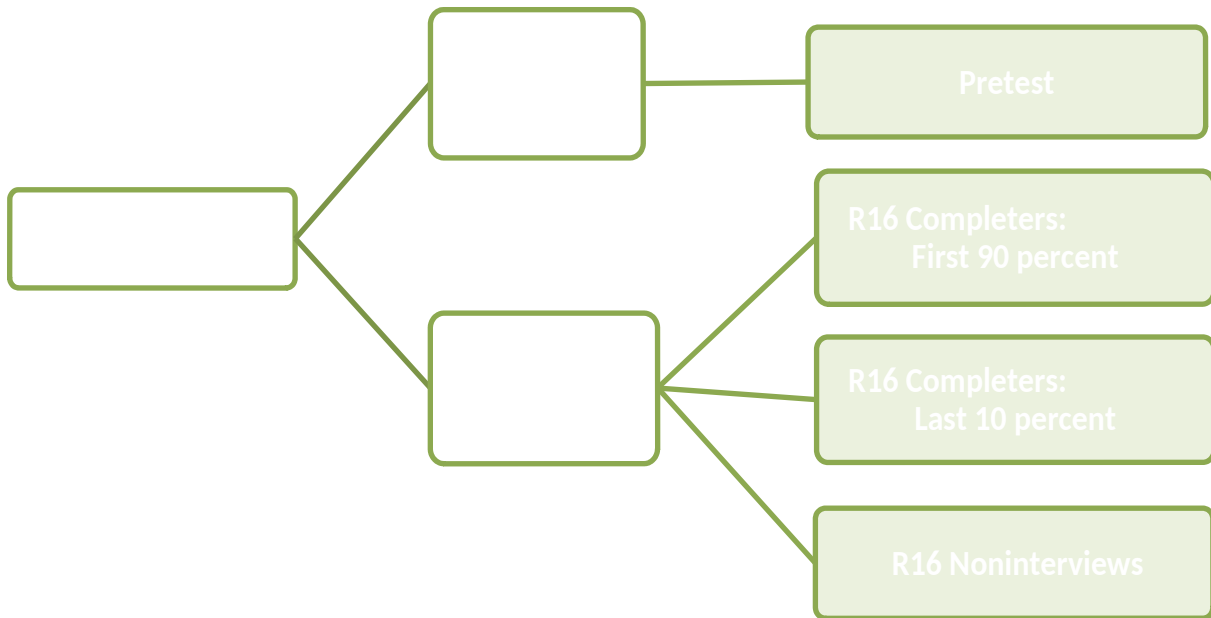
Round 17 will utilize several incentive strategies similar to those used in Rounds 13-16, including base, small in-kind, and missed round incentives. In addition, we propose two incentives that have proven effective with the NLSY79 cohort and other large scale studies. For the purpose of this discussion, a brief summary of the respondent pools and incentive types is provided below.

### ***Respondent Pools***

For the Round 17 incentive plan, respondents are grouped into one of four major categories:

1. Respondents in the pretest sample,
  2. Respondents in the main sample who *completed* Round 16, and were among the first 90 percent of completed interviews,
  3. Respondents in the main sample who *completed* Round 16, and were among the last 10 percent of completed interviews,
  4. Respondents in the main sample who *missed* Round 16 (and are not known to be deceased)
-

**Table 3. Round 17 Respondent Pools**



**Incentive Types**

As in past rounds, the incentive strategy includes a base incentive, small discretionary supplements, and payments for missed interviews. Two additional monetary incentives are proposed for Round 17: an early completion incentive targeting the last 10 percent of interview completers in Round 16 if they complete their Round 17 interviews within the first four months of data collection, and a final push incentive for individuals for whom the standard protocol has been unsuccessful in securing cooperation.

**Table 4. Round 17 Incentive Structure**

Incentive Type	Pretest	R16 Completers: First 90%	R16 Completers: Last 10%	Missed R16 (or more)
Base	\$50	\$40	\$40	\$40
Missed Round(s)	\$15 per round	-	-	\$15 per round (up to 3 rounds missed)
Early Completion			\$20	
Final push	-	<= \$20	<=\$20	<= \$20
Min	\$50	\$40	\$40	\$55
Max	\$95	\$60	\$60	\$105
Typical	\$50	\$40	\$40	\$55

Token In-Kind	-	< \$10 per respondent for up to 200 respondents may be spent on in-kind token gestures. The distribution of these respondents across the almost 9,000 respondents in these three columns is not known ex ante.
---------------	---	--

### *Base Incentive*

A base incentive will be given to all respondents who complete the Round 17 interview. For the pretest sample, the base incentive will remain \$50. For main study respondents, we propose to increase the base incentive to \$40. The \$40 incorporates the \$30 base fee paid since Round 10 [2006] as well as the \$10 transition-to-biennial premium paid in Round 16. We will also continue to offer payments of up to \$15 for each missed round up to 3 for those who have missed one or more consecutive rounds.

### *Missed round incentive*

Payments for missed rounds are designed to encourage attriters to return to the study. For individuals who missed Round 16 (and possibly earlier consecutive rounds), we would offer an additional incentive for each missed round if those respondents complete Round 17. As in previous rounds, we propose to pay \$15 for one or more consecutive rounds (not to exceed 3). Respondents would receive at most an additional \$45 for prior missed rounds. Our objective is always to keep the missed-round payment lower than the payment for a timely interview; the three rounds prior to Round 17 all offered incentives of at least \$30 per interview.

### *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.

### *Early Completion Incentive*

The early completion incentive acknowledges that the last cases to be completed each round differ on observable characteristics from those completed earlier. While these cases are important for preserving the representativeness of the sample, the extension of the field period is costly to the study and to data quality. By offering an incentive for earlier Round 17 completion to Round 16 late completers, we hope to achieve some cost savings in data collection; the incentive is a mechanism for sharing these savings with the respondents. The strategy is somewhat similar to the Early Bird incentive used for many years on the NLSY79. In the NLSY79 Early Bird, respondents are asked to take a different type of action to participate (proactively contact the project for participation). In the NLSY97 case, respondents who completed very late one round are at higher risk of not completing at all, or completing very late the subsequent round, thus requiring considerable fielding effort. The purpose of this incentive is to try to encourage respondents at risk of delaying completion to complete their interviews in a timely fashion. In excess of 60 percent of the sample already completes their interviews in a timely fashion; it is not necessary to offer them an incentive to do so (nor would any benefit derive to the project for offering them such an incentive). Respondents eligible for this payment but who did not complete within the early window will receive the base incentive and possibly a final push incentive but not the earlier completion inducement.

The motivation for an early completion incentive is based on three concepts. First, when attrition in the NLSY97 sample is examined in the context of prior participation, the highest rate is among those who complete a given interview, but have previously missed a round. These individuals are much less likely to complete the subsequent interview than are people who have never missed a round.

The second motivation is the desire to constrain the length of the field period. An extended field period is costly to implement, and the longer field period may compromise the comparability of data within the round. Completing a higher percentage of interviews earlier in the field period and with fewer contact attempts would be an important benefit to the project.

Finally, the Early Bird effort on the NLSY79 has demonstrated the value of offering incentives to respondents to help reduce the effort required for us to seek them out. The Early Bird program has been successful in getting large fractions of the NLSY79 sample to come forward for their interview rather than waiting for our field staff to reach them. The design of the early completion incentive utilizes this key feature of the NLSY79 Early Bird program.

The table below shows the completion rates for individuals by their prior round completion and when they completed the interview. We see that being among the first 90 percent of completed interviews in one round is linked to extremely high rates of completion in the following round. In contrast, individuals who were among the last 10 percent of interviews to be completed had more than 12 percent less likelihood of completing the Round 15 interview, and almost 16 percent less likelihood of completing the Round 16 interview.

**Table 5: R16 Respondent Completion Rates by their R15 Completion Status**

Prior round completion status	No. R15 Respondents	R15 completion %	No. R16 Respondents	R16 completion %
First 90 %	6650	97.0 %	6529	93.9 %
Last 10 %	831	84.8 %	894	77.6 %
NIR*	1503	18.0 %	1561	20.1%

\*Non-Interview Respondents

*Final Push Incentive*

In past rounds, we have had permission to spend no more than \$5 on average, with a limit of \$20 per case, on in-kind incentives to encourage the cooperation of the most difficult respondents. The administrative burden of obtaining and tracking these in-kind incentives continues to increase, especially with many interviews being completed on the phone. To make the field period more efficient and simplify the administrative issues, we request permission to offer this extra payment in cash along with the regular incentive. Starting after the first 12 weeks of the round, cases that have had at least 7 contact attempts will be eligible for a final push incentive of up to \$20 (bringing the total incentive to \$60). 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.

### *Incentive Costs*

A table listing the total cost of incentives by respondent pool is provided below.

**Table 6: Round 17 Incentive Costs by Respondent Pool**

Incentive Type	Pretest	R16 Completers: First 90%	R16 Completers: Last 10%	Missed R16 (or more)	Total
Sample size*	196	6,427	714	1,692	9,029
Expected completes	156	6,272	500	428	7,356
Base	\$7,800	\$250,880	\$10,000	\$17,120	\$278,000
Final Push Incentive	-	\$31,360	\$2,500	\$2,140	\$36,000
Missed Round(s)	\$270	-		\$8,560	\$8,830
In-kind	-	\$500	\$200	\$300	\$1,000
Earlier Completion Inducement	-	-	\$10,000		\$10,000
Totals	\$8,070	\$282,240	\$22,500	\$27,820	\$340,630

\* Note: Sample sizes exclude deceased and blocked cases which are not fielded and therefore have no potential to receive incentive.

### *Electronic payments*

In the more than 80 percent of NLSY97 interviews that are conducted in person, the respondent receives a cash incentive payment at the close of the interview. For the approximately one thousand interviews to be done by phone, the payment process is less streamlined. The case management system identifies a telephone interview as having occurred. Then either cash or a money order must be secured and mailed to the respondent, who receives it at some delay. In the case of loss or neglect by the respondent, the project has no ability to track or restore the payment. On some other OMB-approved data collections such as the National Survey of Early Care and Education for the Administration for Children and Families in the U.S. Department of Health and Human Services, incentives have been paid through electronic means such as e-mailed gift cards. The feasible technologies are rapidly increasing, with Google Wallet and iPhone 6 features as well as PayPal and retailers and financial institutions like Walmart, Amazon and Visa all offering tools for direct electronic payments without cash, check or credit cards. This type of payment system offers the opportunity to achieve project savings in the payment and administration of fees, improve the speed of payment, get more and more accurate contact information for respondents (such as mobile phone number or e-mail address to which payment can be made), and simplify the process if respondents need help tracking down their payments. For Round 17, we request permission to offer electronic payments on the pretest to telephone

respondents. Payments would only be associated with the contractor and not the NLS program, and would be issued from the central office rather than individual interviewers. Telephone respondents would be given the option to choose electronic payment, but all respondents could continue to choose the traditional payment mechanisms if desired. If successful in the pretest, we would propose to also offer the option for electronic payment to Round 17 main survey respondents completing telephone interviews.

## **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.

The Confidential Information Protection and Statistical Efficiency Act of 2002 (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, the prime contractor and associated 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 4 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 17 interviews. Attachment 5 shows the combination advance letter and locating card for Round 16.

#### **b. Contractor Confidentiality Safeguards**

NLS contractors 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.
2. Protection of computer files 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 frequently removed from laptops in the field so that only information that may be needed by the interviewer is retained.
3. Protection of computer files 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 90 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 nonpublic 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 17 the project team will continue several training and procedural changes that were begun 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, continued reminders of respondent confidentiality in field communications throughout the field period and formal separation procedures when interviewers complete their project assignments. Online and telephone respondent locating activities have been moved from geographically dispersed field managers to locating staff in central offices.

## **11. Sensitive Questions**

Continuing the practice of the last few rounds of the NLSY97, the Round 17 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. 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 sexual activity, birth control and pregnancy outcomes (including outcomes other than live births) will be asked of in-person respondents in a self-administered portion of the questionnaire. 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 anyone who does not work on the survey. 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). In the small portion of interviews that are conducted by telephone, the interviewer will administer these questions directly. Questions are worded so that respondents may answer without having to speak aloud sensitive terms such as contraception methods which could be overheard in their immediate environment. 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 17 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 17 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. In Round 17, we are also including questions on religious attendance as have been asked in several prior rounds.

e.) Income, Assets, and Program Participation

The questionnaire 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 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, 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 have been asked in the first interview after the respondent turns 18, and the first interviews after the respondent’s 20<sup>th</sup>, 25<sup>th</sup>, and 30<sup>th</sup> birthdays. In Round 17, we will begin asking respondents first interviewed after their 34<sup>th</sup> birthdays. Because asset accumulation is slow at these young ages, this periodic collection is sufficient to capture changes in asset holdings. Round 14 (begun in Fall 2010) was the first time any respondents were asked to report their assets as of age 30. Five years later, we will begin administering questions about assets head as of age 35. No substantive changes were necessary for the age 30 section to be made appropriate for administration at age 35.

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 14 we added a question on present value of the home. We also ask respondents who owned a house or other dwelling previously and no longer live there about what happened to

their house or dwelling. To allow for respondent-interviewer rapport to build before these potentially sensitive items are addressed, the housing value questions were moved from the (first) Household Information section to the Assets section, which occurs late in the interview. This question series also permits the respondent to report the loss of a house due to foreclosure.

The Rounds 1-13 NLSY97 questionnaires collected month-by-month participation status information for several government programs. These questions have been substantially reduced so that monthly data were no longer collected beginning with the Round 14 interview except from respondents who did not complete the Round 13 interview. The relatively low levels of participation reported did not seem to justify the respondent burden imposed by these questions. More detailed questions about receipt of income from government programs are now asked in the Income section with other sources of income.

f.) Financial Health

To get a better understanding of the financial well-being of the respondents, we continue in Round 17 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 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.

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 17 field effort will seek to interview each respondent identified when the sample was selected in 1997. We will attempt to contact approximately 8,800 sample members who are not known to be deceased. BLS expects that interviews with approximately 7,200 of those sample members will be completed. The content of the interview will be similar to the interviews in Round 16. Based upon interview length in past rounds, we estimate the interview will require about 61 minutes. Changes from round 16 to round 17 are discussed in detail in Attachment 6.

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 17 includes an allowance for attrition that takes place during the course of longitudinal surveys. To minimize the effects of attrition, we 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 17 many respondents 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 17 field period, we will conduct validation interviews with no more than 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. In Round 16, the validation percentage was just over 1 percent of all completed cases. In that round, we had a 12 percent CARI review rate overall. We contacted 53 cases for validation in R16 of which we were able to complete 68 percent and confirmed proper administration on all such cases. Reasons for validation interview included missing audio on CARI files and High CARI recording refusal rate.

Based on our experience of recording segments of the main interview in rounds 11 through 13 and use of these segments for data quality assurance in Rounds 14 through 16, respondents will be asked to provide their consent for the recording of segments of the main interview in Round 17. Recording these interviews will enable BLS to improve data quality while reducing respondent burden.

Recording segments of the main interviews can help 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. Our experiences with OMB-approved

interview recordings since Round 10 indicate that respondents are generally quite willing to consent to be recorded, and the quality of recordings is sufficient for meaningful data-quality assurance.

Interviewers in Round 17 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, research, testing and training purposes. As always your confidentiality is protected by Federal law and the policies of the Bureau of Labor Statistics and (Name of Contractor). 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.

**Table 7: No. of Respondents and Average Response Time, NLSY97 Round 17**

<b>Form</b>	<b>Total Respondents</b>	<b>Frequency</b>	<b>Total Responses</b>	<b>Average Time per Response</b>	<b>Estimated Total Burden</b>
NLSY97 Pretest May/June 2015	150	One-time	150	61 minutes	152.5 hours
NLSY97 R17 web update June/July 2015	1,200	One-time	1,200	10 minutes	200 hours
Main NLSY97: September 2015-May 2016	7,200	One-time	7,200	61 minutes	7320 hours
Validation interview: October 2015 – June 2016	144	One-time	144	4 minutes	9.6 hours
<b>TOTALS*</b>	<b>7,400</b>	<b>—</b>	<b>8,694</b>	<b>—</b>	<b>7682 hours</b>

\* The difference between the total number of respondents and the total number of responses reflects the fact that about 7200 are expected to complete the main interview. In addition, about 144 respondents will be interviewed twice, once in the main survey and a second time in the 4-minute validation interview. We estimate achieving about 1,200 web updates; most of these will be from respondents who also complete the main interview, but

a small number (perhaps 50) may complete only the web update and not also the main interview.

The total response burden for the survey is 7,682 hours. The total annualized cost to respondents, based on burden hours and the federal minimum wage of \$7.25 per hour, is \$55,694.50.

### **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 17 survey is approximately \$13 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, services to users of the data files, and BLS expenses.

### **15. Change in Burden**

The burden of 7,682 hours requested for Round 17 of the NLSY97, which will be conducted during parts of fiscal years 2015 and 2016, is 170 hours lower than the burden of 7,852 hours approved for Round 16. The difference results from projections of a reduced number of completed main interviews relative to what was requested in the clearance package for Round 16. Note that we are projecting more completed interviews than were obtained in the Round 16 main field period.

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

The following is the planned schedule for the data collection for Round 17.

Questionnaire Development	September 2013 – April 2014
Respondent Materials Development	January 2015 – September 2015
Pretest Data Collection	May 2015 – June 2015
Main Data Collection	September 2015 – May 2016
Data Processing	June 2015 – October 2016
Mail advance materials for Round 18	February 2017
Publication of BLS News Release	September 2017
Release of Public-Use Main Data Files	November 2017

### **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.