
NLSY97 USER'S GUIDE

A Guide to the Rounds 1–5 Data
National Longitudinal Survey of Youth 1997

Prepared for the
U.S. Department of Labor by

Center for Human Resource Research
The Ohio State University
Columbus, Ohio
2003

This publication is prepared under contract #J-9-J-9-0007 with the Bureau of Labor
Statistics, U.S. Department of Labor.

Acknowledgments

The National Longitudinal Survey of Youth 1997 (NLSY97) cohort is an expansive survey sponsored by the Bureau of Labor Statistics, U.S. Department of Labor, and conducted by the National Opinion Research Center (NORC) at the University of Chicago with assistance from the Center for Human Resource Research (CHRR) at The Ohio State University. Additional funding has been provided by the U.S. Department of Defense, the U.S. Department of Justice, the U.S. Department of Education, and the National Institute of Child Health and Human Development.

This document pays tribute to the initiators of the NLS, Herbert S. Parnes and Howard Rosen, who had the vision to see that the collection of longitudinal data on the labor force experiences of men and women in the United States could provide a basis for policies fostering more efficient and equitable labor markets. Additional thanks are given to the NLSY97's current principal investigator, James Walker; project director, Rupa Datta; and to the staff at NORC who have contributed to this guide. Special thanks go to CHRR's director, Randall Olsen; Rosella Gardecki, the original guide compiler and an NLSY97 task leader; the NLSY97 archival task leader, Frank Marino; and the other members of the documentation team, Linda Boyer, Shawn Miller, and Jay Zagorsky.

Jennifer Hering
Amanda McClain
Editors

Help and Information

This manual is one of a set of user-oriented documents available to the NLS research community that contain information on the Original Cohorts (Older Men, Young Men, Mature Women, and Young Women), the NLSY79, the Children of the NLSY79, and the NLSY97. Appendix A in this document contains a quick reference guide to essential information about the NLSY97. Persons needing more information on the NLS in general or the NLS cohort groups are encouraged to obtain copies of the *NLS Handbook*. Summaries of research using the NLS data over the past several decades are presented on-line in the NLS Annotated Bibliography of Research, located at <<http://www.nlsbibliography.org>>. Persons needing detailed information on the other cohorts should read the *NLS of Mature Women User's Guide*, the *NLS of Young Women User's Guide*, the *NLSY79 User's Guide*, and the *NLSY79 Child Handbook 1986–1990*; contact NLS User Services for more information on the men's cohorts. Information on ordering NLS documentation and data sets is available from NLS User Services, 921 Chatham Lane, Suite 100, Columbus, OH 43221-2418; phone (614) 442-7366, fax (614) 442-7329, or e-mail at <usersvc@postoffice.chrr.ohio-state.edu>. Information is also available online at <<http://www.bls.gov/nls>>, and orders can be placed online by following the "Ordering Data" link on that Web site.

Table of Contents

Chapter 1: Introduction	1
1.1 The National Longitudinal Survey of Youth 1997 (NLSY97)	3
1.2 The NLS	3
1.3 NLSY97 Sample	6
1.4 Content of the NLSY97	8
1.5 NLSY97 Data Releases	13
1.6 Organization of this Guide	14
Chapter 2: Sample Design & Fielding Procedures	15
2.1 Sample Design & Screening Process	17
2.2 Interview Methods	21
2.3 Sample Size & Composition	32
2.4 Retention and Reasons for Noninterview	33
2.5 Sample Weights & Design Effects	37
2.6 Confidentiality & Informed Consent	40
Chapter 3: Guide to the NLSY97 Data	53
3.1 Survey Instruments & Other Documentation	55
3.2 Types of Variables	57
3.3 NLSY97 Documentation	66
3.4 NLS Investigator Search Functions	72
Chapter 4: Topical Guide to the NLSY97	75
4.1 Aptitude, Achievement & Intelligence Scores	79
4.1.1 Achievement Tests	80
4.1.2 Administration of the <i>CAT-ASVAB</i>	82
4.1.3 <i>PIAT Math Test</i>	85
4.2 Education	87
4.2.1 College Experience	94
4.2.2 Educational Status & Attainment	97
4.2.3 School-Based Learning Programs	101
4.2.4 School Experience	102
4.2.5 School & Transcript Surveys	104
4.3 Employment	109
4.3.1 Employers & Jobs	124
4.3.2 Fringe Benefits	128
4.3.3 Gaps in Employment	129
4.3.4 Industry	132
4.3.5 Job Search	135
4.3.6 Labor Force Status	137
4.3.7 Occupation	139
4.3.8 Self-Employment Characteristics	143
4.3.9 Tenure	145
4.3.10 Time Spent at Work	147
4.3.11 Wages	150
4.3.12 Work Experience	153
4.4 Event History	155
4.5 Expectations, Attitudes, Behaviors & Time Use	162

4.5.1 Alcohol Use.....	164
4.5.2 Attitudes	166
4.5.3 Autonomy & Parental Control.....	173
4.5.4 Cigarette Use	175
4.5.5 Crime, Delinquency & Arrest	177
4.5.6 Drug Use	181
4.5.7 Expectations.....	184
4.5.8 Sexual Activity & Dating	187
4.5.9 Time Use	189
4.6 Family Background & Demographic Characteristics.....	190
4.6.1 Age of Respondent	192
4.6.2 Gender.....	195
4.6.3 Geographic Indicators	197
4.6.4 Household & Neighborhood Environment.....	200
4.6.5 Household Composition	202
4.6.6 Nonresident Characteristics.....	214
4.6.7 Parent Characteristics	218
4.6.8 Race & Ethnicity.....	224
4.6.9 Youth History	226
4.7 Health	228
4.8 Income, Program Participation & Assets.....	232
4.8.1 Assets & Debts.....	234
4.8.2 Income.....	237
4.8.3 Program Participation	240
4.9 Marital History & Fertility	243
4.9.1 Child Care	244
4.9.2 Fertility, Pregnancy & Children	246
4.9.3 Marital & Marriage-Like Relationships	249
4.10 Training	252
Appendices	255
Appendix A: Quick Reference Guide	257
A.1 Guide to NLS Documentation	257
A.2 About the NLSY97 Data Set.....	258
A.3 How to Get Help	259
A.4 Index to Created Variables, Roster Items, and Symbols.....	260
Appendix B: NLSY97 Glossary of Terms.....	261
B.1 NLS Survey Terms	261
B.2 Terms Used in NLSY97 Interviews.....	267
Index	277

List of Tables and Figures

1.2 Table 1. The NLS: Survey Groups, Sample Sizes, Interview Years & Status	5
1.3 Table 1. NLSY97 Sample Sizes by Subsample, Race/Ethnicity & Gender	7
1.3 Table 2. Round 1 Distribution of NLSY97 Respondents by Household Type.....	8
1.3 Table 3. Round 1 Distribution of NLSY97 Sibling Groups	8
1.4 Table 1. Content of the <i>Youth Questionnaire</i>	11
1.4 Table 2. Content of the <i>Parent Questionnaire</i>	12
1.5 Table 1. Data Releases	13
2.1 Figure 1. Selection of NLSY97 Respondents	17
2.1 Table 1. NLSY97 Round 1 Interview Completion	20
2.1 Table 2. NLSY97 Sampling Status of Youths by Housing Arrangement.....	21
2.2 Table 1. NLSY97 Sample Sizes, Retention Rates, and Fielding Periods	22
2.2 Figure 1. Priority for Choosing Responding Parent	26
2.2 Table 2. NLSY97 Youths by Age and Parent Interview Availability	27
2.3 Table 1. Racial, Ethnic & Gender Composition of NLSY97 Sample.....	33
2.4 Table 1. Retention Rates by Sample Type and Gender.....	34
2.4 Table 2. Reason for Noninterview by Gender	35
2.4 Table 3. Reason for Noninterview by Sample Type.....	35
2.4 Table 4. Reason for Noninterview by Race/Ethnicity	36
2.5 Figure 1. Sampling Weight Variable for All Rounds.....	38
2.6 Figure 1. NLSY97 Round 5 Advance Letter	44
2.6 Figure 2. Privacy Act Statement Used in the NLSY97	46
2.6 Figure 3. BLS Disclosure Notice used in the NLSY97.....	46
3.1 Table 1. Sample Question Names by Youth Questionnaire Section.....	56
3.1 Table 2. Sample Question Names in Screener and Parent Questionnaires	57
3.2 Figure 1. Sample PAPI Roster Grid.....	60
3.2 Figure 2. Example: Variables in CAPI-Generated School Roster for Round 2	61
3.2 Figure 3. How Rosters Are Created	61
3.2 Figure 4. Rosters Included Each Round	63
3.3 Figure 1. NLSY97 Questionnaire Item Codeblock	67
3.3 Figure 2. NLSY97 Created Variable Codeblock	67
4.1 Table 1. Aptitude, Achievement & Intelligence Scores Subtopics and Universe Restrictions.....	79
4.1.2 Figure 1. Contents of the <i>CAT-ASVAB</i>	82
4.2 Table 1. Education Subtopics and Universe Restrictions	88
4.2 Figure 1. Round 1 Schooling Section	90
4.2 Figure 2. Rounds 2–5 Schooling Section.....	91
4.2 Table 2. School Numbers in the NEWSCHOOL Roster	93
4.2.3 Table 1. Definitions of School-Based Learning Programs	101
4.3 Table 1. Employment Subtopics and Universe Restrictions.....	111
4.3 Figure 1. Rounds 1–3 Employment Flowchart.....	113
4.3 Figure 2. Rounds 4–5 Employment Flowchart.....	115
4.3 Figure 3. NLSY97 Unique ID Numbers	118
4.3.2 Figure 1. Fringe Benefits Available through Employer.....	128
4.3.4 Table 1. Industry of NLSY97 Respondents at Current or Most Recent Employee Job	133
4.3.5 Figure 1. Possible Job Search Actions	135
4.3.7 Table 1. Occupation of NLSY97 Respondents at Current or Most Recent Employee Job	140
4.3.12 Table 1. Work Experience Created Variables	153
4.5 Table 1. Expectations, Attitudes, Behaviors & Time Use Subtopics and Universe Restrictions.....	162

4.5.1 Table 1. Age Respondent First Drank Alcohol.....	164
4.5.2 Table 1. Universe and Reference Parents for Parent Attitude Questions	167
4.5.2 Figure 1. Youth Descriptors of Attitude toward Self	171
4.5.4 Table 1. Age Respondent First Smoked a Cigarette	175
4.5.6 Table 1. Age Respondent First Used Marijuana.....	181
4.5.6 Table 2. Age Respondent First Used Cocaine/Other Drugs	182
4.5.7 Table 1. Topics of Expectations Questions	185
4.5.8 Table 1. Age Respondent First Had Sex	188
4.6 Table 1. Family Background & Demographic Characteristics Subtopics and Universe Restrictions.....	190
4.6.1 Table 1. Age of NLSY97 Respondents as of 12/31/96	192
4.6.1 Table 2. Age of NLSY97 Respondents as of Interview Date.....	192
4.6.1 Table 3. Comparison of Round 1 Variables KEY!AGE and CV_AGE_INT_DATE	194
4.6.2 Table 1. NLSY97 Composition by Gender and Sample Type.....	195
4.6.3 Figure 1. Number of NLSY97 Respondents and States by Census Region	198
4.6.5 Figure 1. Definitions of Relationships in Household Roster	203
4.6.5 Figure 2. Creation of Round 1 Rosters Based on Screener Data.....	207
4.6.5 Table 1. Example Structure of the Round 1 HHI2 Relationship Data.....	209
4.6.5 Table 2. Round 1 ID Variables for Key Relatives of NLSY97 Respondents	211
4.6.5 Table 3. Examples of UID Codes.....	212
4.6.6 Table 1. Data Collected for Nonresident Relatives in Round 1.....	214
4.6.7 Figure 1. Assets Held by the Responding Parent.....	220
4.7 Table 1. Health Subtopics and Universe Restrictions.....	228
4.7 Table 2. Knowledge about Pregnancy and STDs among Youths Born in 1983 (Round 1).....	230
4.8 Table 1. Income, Program Participation & Assets Subtopics and Universe Restrictions	232
4.8 Figure 1. NLSY97 Independence Criteria.....	233
4.8.1 Figure 1. NLSY97 Asset and Debt Categories.....	235
4.8.3 Figure 1. Government Assistance Programs	240
4.8.3 Table 1. Created Program Participation Summary Variables	241
4.9 Table 1. Marital History & Fertility Subtopics and Universe Restrictions.....	243
4.9.3 Table 1. Number of Respondents Ever Married or Never Married and Cohabiting	249
4.10 Table 1. Training Subtopics and Universe Restrictions.....	252
A.1 Table 1. List of NLS Documentation.....	257
A.2 Figure 1. NLSY97 Areas of Interest.....	259

Chapter 1: Introduction

1.1 The National Longitudinal Survey of Youth 1997 (NLSY97)

The NLSY97 is the newest survey in the National Longitudinal Surveys (NLS) program. The survey documents the transition from school to work and is designed to be representative of people living in the United States in 1997 who were born during the years 1980 through 1984. Thus, many of the oldest youths (age 16 as of December 31, 1996) were still in school at the time of the first survey, and the youngest respondents (age 12) had not yet entered the labor market. The NLSY97 cohort includes 8,984 individuals.

Through the NLSY97, researchers will be able to identify characteristics defining the transition that today's youths make from school to the labor market and into adulthood. To achieve this goal, the NLSY97 collects extensive information on youths' labor market behavior and educational experiences. Data on their family and community backgrounds are included in the survey, and in round 1, information was collected from one of the youth's parents. Youths also took the computer adaptive form of the *Armed Services Vocational Aptitude Battery (CAT-ASVAB)*, which measures knowledge and skills in multiple areas including reading and mathematics. Information has been gathered from high schools in the sample areas where the respondents live, and transcripts are gathered as the youths leave high school. This information will help researchers assess the impact of schooling and other environmental factors on these newest labor market entrants. Data from the NLSY97 will also aid in determining how youth experiences relate to establishing a career, participating in government programs, and forming a family. Over time, information from the NLSY97 will allow researchers to compare the progress of this cohort to that of the earlier NLS cohorts.

This detailed guide is designed for researchers who are either working or planning to work with the National Longitudinal Survey of Youth 1997 (NLSY97). Users who are interested in a general overview of the NLS surveys should refer to the *NLS Handbook*. Users may also wish to read recent issues of *NLS News*, a quarterly newsletter, available on-line at <<http://www.bls.gov/nls/nlsnews.htm>>. For those not familiar with the NLS, a brief overview of the NLSY97 and the other NLS cohorts follows.

1.2 The NLS

The National Longitudinal Surveys (NLS) are a set of surveys sponsored by the Bureau of Labor Statistics (BLS), U.S. Department of Labor. These surveys have gathered information at multiple points in time on the labor market experiences of diverse groups of men and women. Each of the NLS samples consists of several thousand individuals, some of whom have been surveyed over several decades. The earliest NLS interviews began in 1966 under the original sponsorship of the Office of Manpower, Automation, and Training (now the Employment and Training Administration). These cohorts were chosen in an effort to understand specific issues pertaining to the U.S. labor market, such as retirement, the return of housewives

to the labor force, and the school-to-work transition. Since that time, however, the surveys have been expanded to provide useful information on an extremely broad range of topics.

The first four cohorts (Older Men, Mature Women, Young Men, and Young Women) were selected in the mid-1960s because each faced important labor market decisions that were of special concern to policy makers. Each of these cohorts contained an oversample of black Americans. Although the initial plan called for only a 5-year period of interviewing, high retention rates and widespread research interest led investigators to continue the surveys. In 1977, a survey of all known NLS data users and the recommendations of a panel of experts convened by the Department of Labor resulted in two decisions. The first was to continue the surveys of the four Original Cohorts for at least an additional five years (as long as attrition did not become a problem). The second decision was to begin a new longitudinal study of a panel of young men and young women.

This new study was initiated to permit a replication of the analysis of the 1960s Young Men and Young Women cohorts and to assist in the evaluation of the expanded employment and training programs for youth legislated by the 1977 amendments to the Comprehensive Employment and Training Act (CETA). To these ends, in 1978 a national probability sample was drawn of young women and young men living in the United States and born between January 1, 1957, and December 31, 1964. This sample included an overrepresentation of blacks, Hispanics, and economically disadvantaged non-black/non-Hispanics. With funding from the U.S. Department of Defense and the Armed Services, an additional group of young persons serving in the military was selected for interviewing. This sample of civilian and military youth, called the National Longitudinal Survey of Youth 1979 (NLSY79), was first interviewed in 1979 and has been re-interviewed regularly since that date.

The Children of the NLSY79 survey, begun in 1986, further enhanced the NLS. With funding from the National Institute of Child Health and Human Development (NICHD) and a number of additional government agencies and organizations, detailed information on the development of children born to NLSY79 women has supplemented the data on mothers and children collected during the regular youth surveys. During these biennial surveys, a battery of child cognitive, socio-emotional, and physiological assessments has been administered to NLSY79 mothers and their children. In addition to these assessments, the Children of the NLSY79 are also asked a number of questions in an interview setting. Beginning in 1994, children age 15 and older, the “Young Adults,” replied to a separate survey with questions similar to those asked of their mothers.

With the aging of the NLSY79 cohort, another longitudinal cohort was started. The National Longitudinal Survey of Youth 1997 (NLSY97) collects information on the circumstances that influence or are influenced by the labor market behaviors of youth ages 12–16 as of December 31, 1996. Data on the

youth’s educational experiences, along with his or her family and community backgrounds, are also included in the survey. Documenting the transition from school to work, this survey is designed to be representative of the U.S. population in 1997 born during the period 1980 to 1984.

Information on sample sizes, interview years, and the current survey status of each NLS respondent group is presented in Table 1 below.

1.2 Table 1. The NLS: Survey Groups, Sample Sizes, Interview Years & Status

Survey Group	Age Cohort	Birth Year Cohort	Initial Sample Size	Initial/Latest Survey Year	# of Surveys to Date	Survey Status
Older Men	45–59 (as of 3/31/66)	4/1/06–3/31/21	5020	1966/1990	13 ¹	Ended
Mature Women	30–44 (as of 3/31/67)	4/1/22–3/31/37	5083	1967/2001	20	Continuing
Young Men	14–24 (as of 3/31/66)	4/1/41–3/31/52	5225	1966/1981	12	Ended
Young Women	14–24 (as of 12/31/67)	1/1/43–12/31/53	5159	1968/2001	21	Continuing
NLSY79	14–21 (as of 12/31/78)	1957–1964	12686	1979/2002	20	Continuing
NLSY79 Children	Birth–14	—	³	1986/2002	9	Continuing
NLSY79 Young Adults ²	15 & older	—	³	1994/2002	5	Continuing
NLSY97	12–16 (as of 12/31/96)	1980–1984	8984	1997/2002	5	Continuing

¹ Twelve interviews occurred from 1966–83. The 1990 interview surveyed living respondents and next-of-kin of deceased respondents.

² NLSY79 Young Adult respondents were initially interviewed as part of the NLSY79 Children sample. Beginning in 1994, those 15 and older were surveyed separately. In 1998, youths age 21 and older were not interviewed.

³ The sizes of the NLSY79 Children and Young Adult samples are dependent on the number of children born to NLSY79 respondents. Since this number is still increasing, original sample sizes are omitted.

As the table indicates, five rounds of the NLSY97 have been completed (the sixth began in late 2002). Since only five rounds of data are available to the public, this guide discusses only those rounds. Information on the content of the round 6 survey is available in the current *NLS Handbook*. A new edition of this guide will be published when round 6 data processing is complete.

Administration of the NLS: Responsibility for the administration of the NLS resides with the Bureau of Labor Statistics (BLS), an agency of the U.S. Department of Labor. Established in 1884, BLS is responsible for the analysis and publication of data series on employment and unemployment, prices and living conditions, compensation and working conditions, productivity, occupational safety and health, and economic growth and employment projections. Its mission is to promote development of the U.S. labor force by gathering information about the labor force and disseminating it to policy makers and the public so that participants in those markets can make more informed, and thus more efficient, choices.

The NLS program supports BLS in this mission. The surveys are part of a longitudinal research program that includes in-house analyses, extramural grants, and other special projects. The NLS program is housed within the Office of Employment and Unemployment Statistics at BLS.

BLS contracts with the Center for Human Resource Research (CHRR) at The Ohio State University to manage the NLS Original Cohorts and the NLSY79 survey, to share in the design of the survey instruments, to disseminate the data, and to provide user services. The Census Bureau collects data for the Original Cohorts through an interagency agreement with BLS. Data collection for the NLSY79 and NLSY79 Children samples is subcontracted to the National Opinion Research Center (NORC) at the University of Chicago. BLS contracts with NORC to manage the NLSY97 cohort and to collect the data for this survey; CHRR is subcontracted for data dissemination, documentation, and variable creation.

The project is assisted in its efforts by the NLS Technical Review Committee. Meeting twice each year, committee members provide recommendations regarding questionnaire design, additional survey topics, potential research uses, methodological issues, data distribution, and user services. The committee is multidisciplinary, reflecting the wide range of social scientists utilizing NLS data.

Ultimate responsibility rests with BLS for overseeing all aspects of the work undertaken by other organizations. For more information about the administration of the NLS program, contact:

The National Longitudinal Surveys
Bureau of Labor Statistics
2 Massachusetts Avenue NE, Room 4945
Washington, D.C. 20212-0001
(202) 691-7410; fax: (202) 691-7425
e-mail: NLS_Info@bls.gov
Web site: <http://www.bls.gov/nls>

1.3 NLSY97 Sample

The design of the NLSY97 sample allows researchers to analyze the experiences both of the population as a whole and also of groups such as Hispanics and blacks. Chapter 2 of this guide further details the NLSY97 sample design and fielding procedures; a brief overview of the sample composition is provided here. Two subsamples comprise the NLSY97 cohort:

- (1) a cross-sectional sample of 6,748 respondents, which is designed to be representative of people living in the United States during the initial survey round and born between January 1, 1980, and December 31, 1984; and
- (2) a supplemental sample of 2,236 respondents, which is designed to oversample Hispanic and black people living in the United States during the initial survey round and born during the same period as the cross-sectional sample.

Table 1 lists the sample size in rounds 1–5 by the subsample, race/ethnicity, and gender of respondents.

1.3 Table 1. NLSY97 Sample Sizes by Subsample, Race/Ethnicity & Gender

	Total sample	Cross-sectional sample					Supplemental Sample			
		Cross-sect. total	Non-black, non-Hisp.	Black, non-Hisp.	Hispanic	Mixed race	Supp. total	Black, non-Hisp.	Hispanic	Mixed race
Round 1										
Male	4599	3459	2413	537	469	40	1140	632	508	—
Female	4385	3289	2252	544	452	41	1096	622	472	2
Total	8984	6748	4665	1081	921	81	2236	1254	980	2
Round 2										
Male	4283	3213	2238	504	433	38	1070	599	471	—
Female	4103	3066	2095	517	417	37	1037	584	451	2
Total	8386	6279	4333	1021	850	75	2107	1183	922	2
Round 3										
Male	4170	3144	2193	490	422	39	1026	572	454	—
Female	4039	3029	2076	503	412	38	1010	568	441	1
Total	8209	6173	4269	993	834	77	2036	1140	895	1
Round 4										
Male	4117	3098	2153	485	423	37	1019	580	439	—
Female	3964	2957	2027	489	402	39	1007	570	435	2
Total	8081	6055	4180	974	825	76	2026	1150	874	2
Round 5										
Male	3989	3012	2110	455	411	36	977	541	436	—
Female	3894	2907	1991	478	401	37	987	558	427	2
Total	7883	5919	4101	933	812	73	1964	1099	863	2

Note: This table was created using the following variables: CV_SAMPLE_TYPE (R12358.), KEYIRACE_ETHNICITY (R14826.), KEYISEX (R05363.), and RNI (R25102. and R38297.).

Data hint →

The sample type of a respondent is identified in variable CV_SAMPLE_TYPE (R12358.).

Multiple Respondent Households

In the NLSY97 cohort, 8,984 respondents originated from 6,819 unique households. Because the sample design selected all household residents in the appropriate age range, 1,862 households included more than one NLSY97 respondent. Table 2 lists the numbers of respondents living in multiple respondent households during the initial survey round. The most common relationship between multiple respondents living in the same household during the first round was that of siblings. Table 3 provides the subset of respondents living with at least one sibling. The NLSY97 does not contain nationally representative samples of siblings of all ages and living arrangements; users should exercise caution when generalizing from the findings of NLSY97 sibling studies.

1.3 Table 2. Round 1 Distribution of NLSY97 Respondents by Household Type

	Respondents	Households
1 Respondent	4957	4957
Total Multiple Respondents	4027	1862
2 Respondents	3192	1596
3 Respondents	705	235
4 Respondents	100	25
5 Respondents	30	6
Total	8984	6819

1.3 Table 3. Round 1 Distribution of NLSY97 Sibling Groups

Type	Respondents
No Siblings	5129
Total Multiple Siblings	3855
2 Siblings	3134
3 Siblings	627
4 Siblings	84
5 Siblings	10
Total	8984

Notes: Table 2 is based on the household ID code (R11930.) and reports multiple respondent households regardless of the relationship between respondents. Table 3 is based on the household ID code (R11930.) and the relationship variables from the round 1 household roster (HHI2_RELx.xx); it reports multiple respondent households where respondents are siblings. In this table, siblings include biological, adoptive, half-, and step- relationships but not foster relationships.

1.4 Content of the NLSY97

The primary purpose of the NLSY97 is to collect data on youth labor force experiences and investments in education. However, the actual content of the NLSY97 is much broader due to the interests of other governmental agencies besides BLS. The following agencies have funded special sets of questions:

- (1) The departments of Defense and Labor jointly sponsored the round 1 administration of the *Armed Services Vocational Aptitude Battery (ASVAB)*, a military enlistment test, and the *Interest-Finder (I-F)*, an occupational interest inventory, to the NLSY97 respondents. See section 4.1.2, “Administration of the *CAT-ASVAB*,” for further details.
- (2) The National School-to-Work Office, a joint agency of the departments of Education and Labor, funded the 1996 school survey. This survey, which collected information from all high schools with a 12th grade in the 147 NLSY97 primary sampling units, recorded the characteristics of each school and its staff. See section 4.2.5, “School & Transcript Surveys,” for details and availability information.
- (3) The Department of Justice, Office of Juvenile Justice and Delinquency, sponsors the set of questions on crime and criminal activities found in the self-administered portion of the *Youth Questionnaire*. These questions are described in section 4.5.5, “Crime, Delinquency & Arrest.”
- (4) The National Institute of Child Health and Human Development (NICHD) provides funding for various questions in the self-administered portion of the *Youth Questionnaire*.

The wide range of employment, education, and background information in the NLSY97 has been gathered in several different survey instruments. Briefly, the *Screener, Household Roster, and Nonresident Roster Questionnaire* identified youths eligible for the survey and collected demographic information on household occupants and key nonresident relatives during the initial survey round. The *Youth Questionnaire*, administered each round, asks the youth about a number of topics, including employment, schooling, family background, health, and attitudes and behaviors. The round 1 *Parent Questionnaire* interviewed one of the youth’s parents about the youth’s history and status and about key aspects of the

parent's own life. School and transcript surveys collected information in several rounds about schools in the NLSY97 sample areas and the academic records of NLSY97 respondents. The administration of the computer adaptive form of the *Armed Services Vocational Aptitude Battery* during round 1 to most NLSY97 respondents will provide researchers with a common aptitude score when the data are publicly available. Finally, in rounds 2–5, data about family income were collected from one of the respondent's parents using the *Household Income Update*. The contents of these survey instruments are described in the following paragraphs.

Screener, Household Roster, and Nonresident Roster Questionnaire

This instrument was initially administered to a member of each household selected for sampling in the NLSY97 survey areas. It was completed by a household resident age 18 or older, referred to as the household informant. This questionnaire was used to identify youths potentially eligible for the NLSY97 survey and/or the administration of the computer adaptive version of the *Armed Services Vocational Aptitude Battery (CAT-ASVAB)*, a military enlistment test (see section 4.1.2 for more information). In general, eligibility for either group was dependent on the youth's age and, in some sample areas, on the youth's race or ethnicity. Section 2.1, "Sample Design & Screening Process," provides more detailed information about the precise age and race/ethnicity requirements. This instrument actually comprised four sections: the simple screener, extended screener, household roster, and nonresident roster, administered in that order and described below. Question names for the simple and extended screeners begin with "SE," household roster questions begin with "SH," and nonresident roster questions begin with "SN."

The simple screener section was first administered at each household in the sample areas to identify potentially eligible sample members. The simple screener collected the name and birth date or age of each person who could be linked to the household (see section 2.1, "Sample Design & Screening Process," for more information). If the household contained a youth potentially eligible for the NLSY97 or the administration of the *CAT-ASVAB*, the household informant completed the extended screener. This section collected the gender and race/ethnicity of each person in the household and the grade in school the potentially eligible youth was currently attending, if any. Race and ethnicity were used in supplemental sample areas as further criteria for NLSY97 eligibility; in certain cases grade in school affected eligibility for the administration of the *CAT-ASVAB*. If the household contained a youth eligible only for the administration of the *CAT-ASVAB*, the interviewer solicited his or her participation. If the household included one or more youths potentially eligible for the NLSY97, the interview continued with an extensive two-part roster.

The household informant first answered the questions in the household roster section of the *Screener, Household Roster, and Nonresident Roster Questionnaire*. This section established the relationships between household occupants and collected basic demographic information (e.g., marital status, highest grade of schooling completed, employment status) for all household members, including any NLSY97-eligible youths. After the roster was created, one parent of the sample youth was chosen from the list of household occupants and asked to participate in the *Parent Questionnaire* (see “Choice of Parent” in section 2.2, “Interview Methods,” for a description of this process). Finally, the household informant was administered the nonresident roster section, which gathered data on members of the youth’s immediate family (e.g., biological, adoptive, or stepparents; biological or adoptive siblings; spouse; biological children; parent of the youth’s biological children) who lived elsewhere at the survey date.

Youth Questionnaire

In each survey round, the *Youth Questionnaire* is administered to every respondent. This instrument focuses in detail on the youth’s schooling and employment activities; the hour-long interview also collects extensive data on the respondent’s family background, social behavior, and health status. Table 1 lists the name given to each of the major sections included in the *Youth Questionnaire* and provides a brief, general description of the kinds of information collected in each section. The abbreviations for each section are incorporated into the question names for this survey; they are included here for reference.

1.4 Table 1. Content of the Youth Questionnaire

Section		Description of Data Collected
Information	YINF	In rd. 1, verified youth data in the <i> Screener, Household Roster, and Nonresident Roster Questionnaire</i> .
Household Information	YHHI	Confirms and updates information on members of the youth's household after rd. 1.
CPS	YCPS	In rds. 1 and 4, established employment status using questions from the Current Population Survey.
Schooling	YSCH	Gathers information about current schooling and school environment.
Peers/Opportunity Sets	YPRS	In rd. 1, provided the youth with a list of activities and asked him or her to estimate the percentage of peers who participate in each.
Time Use	YTIM	In rds. 1–3, asks the youth about time spent during the day and week on various activities.
Employment	YEMP	Collects data about each employer for whom the youth worked since age 14; also includes data on freelance employment.
Training	YTRN	Asks about training programs the youth has participated in outside of regular schooling.
Health	YHEA	Asks about general state of youth's health and long-standing problems he or she has.
Self-Administered	YSAQ	Completed by youth; asks sensitive questions. Subsections include household and neighborhood environment, relationship with parents, puberty, dating and sexual activity, pregnancy and abortion, attitudes toward self, substance use, and criminal and delinquent activities. (This section is divided into SAQ1 and SAQ2 sections in round 5 and beyond.)
Marriage	YMAR	Asks questions about any marriages or marriage-like relationships that the youth may have had.
Fertility	YFER	Gathers information about any biological children of the youth and the parentage of each.
Child Care	YCCA	In rd. 5, collects details about child care arrangements or child care availability.
Program Participation	YPRG	Gathers data about any assistance programs in which the youth and the youth's spouse/partner may have participated.
Income / Assets	YINC/ YAST	Collects data on the income and assets of youth and the youth's spouse/partner. These questions were combined in the YINC section in rd. 1 but split into two sections for subsequent rounds.
Expectations	YEXP	In rds. 1, 4, and 5, asked youths to predict characteristics of their lives at certain points in the future.
PIAT Math	YPIA	Administers the <i>PIAT Math Assessment</i> to eligible respondents.

Parent Questionnaire (round 1)

The *Parent Questionnaire* collected extensive background information from one of the youth's biological parents. If no biological parent lived in the respondent's primary household, another adult household member was selected for the *Parent Questionnaire* according to predetermined criteria (see "Choice of Parent" in section 2.2, "Interview Methods," for a detailed description of this process). In the *Parent Questionnaire*, the resident adult was asked to provide information about his or her activities and the activities of the NLSY97 youth. General questions also determined the family's nationality, religious orientation, and community environment. The names and abbreviations given in Table 2 refer to the major sections included in the round 1 *Parent Questionnaire*. This table also includes a brief, general description of the subsets of questions asked within each major section.

1.4 Table 2. Content of the Parent Questionnaire

Section		Description of Data Collected
Information	PINF	Verified own and youth's data collected in the <i> Screener, Household Roster, and Nonresident Roster Questionnaire</i> .
Family Background	P2	Gathered data about parents' family background (e.g., birthplace, siblings).
Calendars	P3	Collected dates of significant events in parent's life (e.g., marital history, employment history).
Parent Health	P4	Gathered information about the general state of the parents' health (e.g., long-term problems).
Income and Assets	P5	Asked about 1996 earnings of all household members and other parental income and assets.
Self-Administered	P6	Parent answered questions about self-esteem, religion, health, behavior/beliefs, and spouse or partner relations.
Child Calendar	PC8	Gathered data about each eligible youth in the household regarding places lived, school history, paternity, adoption or custody, poverty, and child care.
Child Health	PC9	Collected data about the health and health insurance of each eligible youth.
Child Income	PC10	Inquired about the 1996 financial support and earnings of each eligible youth.
Expectations	PC11	Asked parent to make predictions about each eligible youth's life in the next year and at age 20 and 30.
Family	PC12	Collected data about the family situation by asking questions about the youth's social skills, decision-making, positive behavior, and relationships with parents.

School and Transcript Surveys

Designed with an emphasis on the school-to-work transition, the NLSY97 also gathered information from two school surveys conducted by mail with telephone follow-up. In the winters of 1996–97 and 2000–01, all high schools with a 12th grade in the 147 NLSY97 primary sampling units were asked to participate. The 2000 sample also included vocational schools. Participating school administrators provided detailed data on the characteristics of the school, the staff, and the student body. Additional information included the school's general practices, graduation policies, and school-to-work programs. Due to confidentiality restrictions, access to these data is limited; interested researchers should contact BLS at NLS_Info@bls.gov for more information.

The NLSY97 program has also collected specific educational information about respondents directly from their high schools. In winter 1999–2000, the NLSY97 transcript survey sought high school transcripts for all respondents who either had graduated from high school or were age 18 or older and no longer enrolled in high school. Transcripts were received for 1,417 respondents. Based on these transcripts, survey staff constructed a series of variables describing each respondent's high school experience.

Administration of the CAT-ASVAB

The Department of Defense (DOD) used the NLSY97 sample as part of a larger effort to establish new norms for the computer adaptive form of the *Armed Services Vocational Aptitude Battery (CAT-ASVAB)*, a military enlistment test. The scores from this group will also be instrumental in creating norms for the DOD *Interest-Finder (I-F)*, an occupational interest inventory. Administered from summer 1997 to spring

1998, the *CAT-ASVAB* and *I-F* collected baseline information on the aptitudes and vocational interests of NLSY97 respondents. Additionally, the *CAT-ASVAB* and *Interest-Finder* were administered to a group of participants age 17–23 who were sampled at the same time as NLSY97 respondents (see section 4.1.2, “Administration of the *CAT-ASVAB*,” for more information). A number of *ASVAB* variables are included in the round 4 data; section 4.1.2 contains details about the types of information currently available and the additional variables planned for future release.

Household Income Update

Beginning in round 2, the *Household Income Update* gathers basic income information concerning the respondent’s parent and his or her spouse/partner in the absence of a detailed parent questionnaire. If the youth respondent lives with a parent, this questionnaire collects the parent’s total pre-tax income from wages, salaries, commissions, and tips during the past calendar year; the same data for the parent’s spouse or partner; and the total pre-tax amount of any other income (i.e., farm or business income, inheritances, child support, government programs) received by the parent and spouse/partner together. These data are collected for the respondent’s current household regardless of his or her residence during the past calendar year. Question names for the variables based on the *Household Income Update* begin with “HIU.”

1.5 NLSY97 Data Releases

NLSY97 data are available to the public via download from <<http://www.bls.gov/nls>> (no charge). The NLS cross-cohort CD also includes NLSY97 data. Each data set contains the data record of each youth, including information from the *Youth Questionnaire*, the *Parent Questionnaire*, and the *Screener, Household Roster, and Nonresident Roster Questionnaire*. Geocode files containing detailed geographic information are also available; this restricted-use data is only available on CD. Table 1 presents brief descriptions of these NLSY97 data files. Subsequent data releases will be announced in *NLS News*, the quarterly NLS newsletter.

1.5 Table 1. Data Releases

Data set	Content
NLSY97 Main File	The record of each youth includes information from the <i>Youth Questionnaire</i> ; the round 1 <i>Screener, Household Roster, and Nonresident Roster Questionnaire</i> ; the round 1 <i>Parent Questionnaire</i> ; and the rounds 2–5 <i>Household Income Update</i> .
NLSY97 Event History File	The event history variables provide information on the youth’s education, employment, marriages, and program participation in a status array format. This data set also includes all main file variables.
NLSY97 Geocode File	Information on the geographic area of the youth’s current residence and select environmental variables is available to users satisfactorily completing BLS’s geocode accessing agreement procedure. This data set also includes all main file variables.
NLSY97 Public Screener Data	Data collected in the round 1 screener that are appropriate for public use will be included in a future data release.

1.6 Organization of this Guide

The *NLSY97 User's Guide* is designed for researchers who are either working or planning to work with the NLSY97 data. Users interested in a general overview of NLS cohorts should refer to the *NLS Handbook*. The remainder of this guide is organized as follows:

Chapter 2 presents a description of the procedures used to select the cohort members and the interview methods used during each survey round. This chapter also provides information about the size and composition of the sample, attrition after round 1, and the creation and use of sampling weights.

Chapter 3 describes the various survey instruments, types of variables, and available documentation, as well as the search options included in the data set.

Chapter 4, the topical portion of this guide, presents summary discussions of sets of variables. Each section details the NLSY97 variables and refers the user to the portion of the survey containing questions in the topical area, as well as to other related supporting documentation. The user is also directed to related sections within this guide and to the relevant areas of interest in the data set. The table of contents lists the topics and corresponding subtopics found in this chapter. The topical discussions are arranged alphabetically by topic (e.g., “Employment,” then “Event History”); the subtopics are also in alphabetical order (e.g., “Fringe Benefits,” then “Gaps in Employment”) within each topic. Persons interested in reviewing, for example, variables containing data on a respondent's rate of pay should turn to the “Employment” section and examine the subtopic “Wages.”

Appendices make up the final section of this guide. The first, a quick reference guide, includes a list and description of NLS documentation; details concerning the NLSY97 data releases; an index to created variables, roster items, and symbols; and contact information for assistance in using the data. The second appendix is a glossary of NLSY97 terms and their definitions.

Throughout the guide, the sections may include “user notes,” which contain important information about some characteristic of the surveys or data file, and “data hints,” which direct researchers to a particular variable or variables that should be used for a specific function. These boxed paragraphs are valuable sources of advice about idiosyncratic aspects of the data set.

Chapter 2: Sample Design & Fielding Procedures

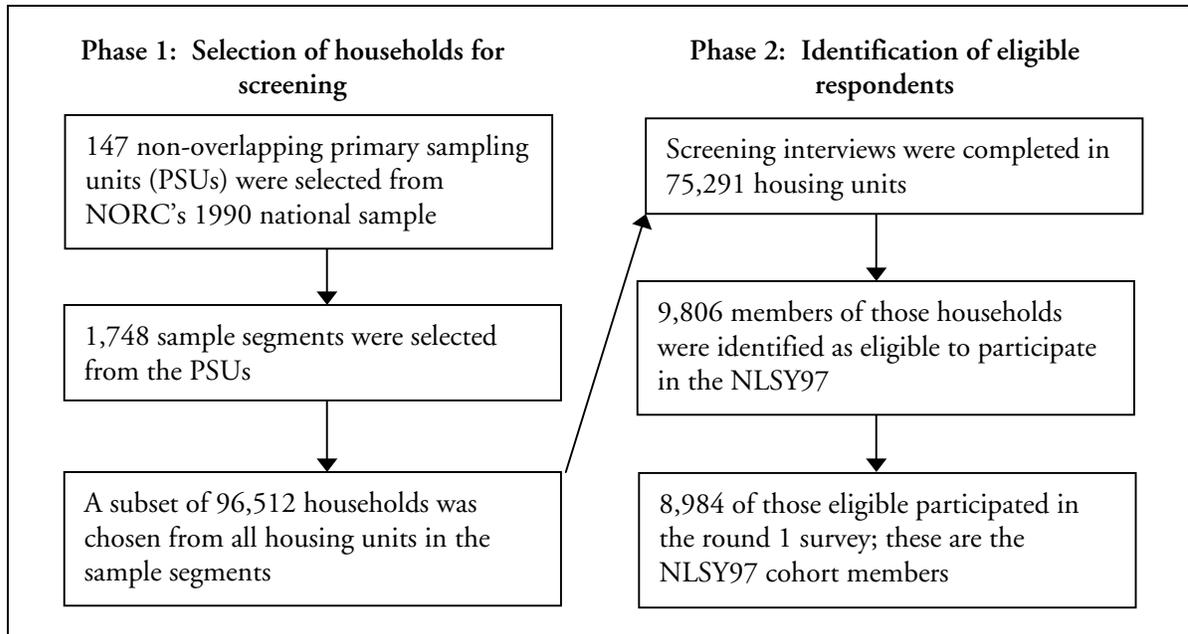
2.1 Sample Design & Screening Process

Sampling Procedures

The NLSY97 cohort comprises two independent probability samples: a cross-sectional sample and an oversample of black and/or Hispanic respondents. The cohort was selected using these two samples to meet the survey design requirement of providing sufficient numbers of black and Hispanic respondents for statistical analysis. Information on the actual size and racial/ethnic composition of the cohort is presented in section 1.3, “NLSY97 Sample.”

The NLSY97 cohort was selected in two phases, as pictured in Figure 1. In the first phase, a list of housing units for the cross-sectional sample and the oversample was derived from two independently selected, stratified multistage area probability samples. This ensured an accurate representation of different sections of the population defined by race, income, region, and other factors. In the second phase, subsamples of the eligible persons identified in the first phase were selected for interview.

2.1 Figure 1. Selection of NLSY97 Respondents



The listing of eligible housing units was composed of 96,512 households, defined as a single room or group of rooms intended as separate living quarters for a family, for a group of unrelated persons living together, or for a person living alone. The list of housing units for each sample was selected in the following manner: First, 100 primary sampling units (PSUs)¹ for each sample were chosen from NORC’s

¹ There are 100 PSUs in the cross-sectional sample and 100 PSUs in the oversample; however, some PSUs were selected in both samples. Thus, there are a total of 147 non-overlapping PSUs included in the NLSY97.

1990 national sample. In the cross-sectional sample, each PSU represented either a metropolitan area or one or more non-metropolitan counties with a minimum of 2,000 housing units. The supplemental sample defined PSUs differently from the cross-sectional sample; counties containing large percentages of minorities were merged to create areas containing a minimum of 2,000 housing units. Second, regardless of sample, segments containing one or more adjoining blocks—and at least 75 housing units—were selected from each PSU. Finally, a subset of housing units within the segment comprised the NORC listing of households eligible for interview.

The second phase identified all NLSY97-eligible individuals age 12 to 16 as of December 31, 1996, in each household. NORC interviewers went to the households and administered a short interview called the simple screener, a portion of the *Screener, Household Roster, and Nonresident Roster Questionnaire*, which collected the age or date of birth of every person linked to a particular household. The survey collected these data for more than 150,000 people. In cross-sectional sampling units, if the household included one or more occupants in the eligible age range, NORC interviewers asked those individuals to participate in the first NLSY97 interview. In supplemental sampling units, the interviewer continued with the extended screener, which established the race and ethnicity of household members. If a person of the correct age and of black or Hispanic race/ethnicity resided in the household, he or she was asked to participate in the survey. Any person in the above age range who completed the first round interview is considered a member of the NLSY97 cohort. Base year interviews were conducted between January and early October 1997 and between March and May 1998 (see section 2.2 for details). Of the 9,806 individuals selected for interview during household screenings, a total of 8,984 (91.6 percent) were interviewed.

During the NLSY97 screening process, two additional nationally representative samples were identified to participate in the administration of the *CAT-ASVAB*. The first group, the Student Testing Program (STP), consisted of students who expected to be in the 10th through 12th grades in the fall of 1997. Included were many respondents who also participated in the main NLSY97 survey, as well as youths who refused to participate in or were not eligible for the NLSY97. The second sample, the Enlistment Testing Program (ETP), was a nationally representative sample of youths 18 to 23 years old as of June 1, 1997. This group provided the normative information that will be used by the Department of Defense to determine the score distribution of military-eligible youths and will help to assess the impact of these tests on minority and female military eligibility.

Cross-Sectional Sample

For the cross-sectional sample, 54,179 screening interviews were carried out among 1,149 sample segments in 100 primary sampling units (PSUs), drawn from the NORC master probability sample of the United States. The cross-sectional screening established three samples:

- (1) **Main NLSY97 Sample:** A cross-sectional sample designed to be representative of young people living in the United States during round 1 and born January 1, 1980, through December 31, 1984. This sample is designed to maximize the statistical efficiency of samples through the several stages of sample selection (counties, enumeration districts, blocks, sample listing units). Probabilities of selection are based upon total housing units in a geographic area.

Following the initial screening process, 7,327 individuals from the cross-sectional sample were designated to be interviewed in the NLSY97 survey; of those, 92.1 percent, or 6,748 respondents, completed the round 1 interview.

- (2) **Department of Defense Student Testing Program (STP) Sample:** A nationally representative sample of students living in the United States during round 1 and born June 2, 1973, through December 31, 1984, who—depending on the time of the household screening—were in grades 9–11 in the spring or summer of 1997, were not enrolled during the spring and summer but expected to be in grades 10–12 in the fall of 1997, or were enrolled in grades 10–12 during the fall of 1997. (See the “Administration of the *CAT-ASVAB*” section of this guide for more information.) Some NLSY97 respondents were also eligible for the STP sample.
- (3) **Department of Defense Enlistment Testing Program (ETP) Sample:** A cross-sectional sample designed to be representative of the noninstitutionalized segment of young people living in the United States during round 1 and born June 2, 1973, through June 1, 1979.

Supplemental Sample

Statistically efficient samples of black and Hispanic respondents were created by oversampling these minorities in 100 PSUs in NORC’s national sample. For the supplemental sample, 21,112 screening interviews were conducted in 599 sample segments. The supplemental screening produced three samples:

- (1) **NLSY97 black and Hispanic Oversample:** A supplemental sample designed to oversample Hispanic and black respondents living in the United States during round 1 and born January 1, 1980, through December 31, 1984. Stratification specifically relevant for Hispanics and blacks was used. Oversample respondents were chosen with a probability based on size measures for these groups rather than for the general population. This should make it possible to equalize the distribution of the targeted groups among the various sampling units more than would otherwise be the case.

After screening, 2,479 individuals from the supplemental sample were designated for interview in the NLSY97, and of these, 90.2 percent, or 2,236 respondents, completed the round 1 interview.

- (2) **Department of Defense STP Sample:** A nationally representative sample of students, selected regardless of race and/or ethnicity, living in the United States during round 1 and born June 2, 1973, through December 31, 1984. Members of this sample are those who—depending on the time of the household screening—were in grades 9–11 in the spring or summer of 1997, were not enrolled during the spring and summer but expected to be in grades 10–12 in the fall of 1997, or were enrolled in grades 10–12 during the fall of 1997.
- (3) **Department of Defense ETP black and Hispanic Oversample:** A sample of black or Hispanic youths living in the United States during round 1 and born June 2, 1973, through June 1, 1979.

Data hint ➔

Users can identify the cross-sectional or supplemental sample type of each respondent by referring to the sample type variable (CV_SAMPLE_TYPE—R12358.).

2.1 Table 1. NLSY97 Round 1 Interview Completion

Sample	Eligible for interviewing	Interviewed round 1	
Total Cohort	9806	8984	91.6%
Cross-Sectional Sample	7327	6748	92.1%
Supplemental Sample	2479	2236	90.2%

Screening Procedures

The screening interview was completed by NORC in 75,291 housing units. These interviews occurred in 1,748 sample segments of 147 non-overlapping PSUs, including most of the fifty states and the District of Columbia.² The screening interview was designed to elicit information allowing identification of household occupants eligible for inclusion in the NLSY97 sample. The NLSY97 screening interviews were completed within 94.1 percent of the cross-sectional and 93.1 percent of the supplemental occupied housing units selected for screening. Table 1 presents a summary of completed interviews in round 1.

Sampling procedures were developed to establish links between housing units in the sample PSUs and individuals who might be temporarily absent. As part of the screening process, household informants were asked if there were any persons for whom the housing unit was the usual place of residence, but who were away from the housing unit at the time of the survey. Included in this group were college students, persons in the military, and persons in prisons or other institutions. Sampling procedures were also established for those residing in a selected housing unit whose usual place of residence was elsewhere. Table 2 lists the NLSY97 status (e.g., included in the sample, excluded, or restricted) for youths not in their usual place of residence at the time of the survey.

² There are 100 PSUs in the cross-sectional sample and 100 PSUs in the oversample; however, some PSUs were selected in both samples. Thus, there are a total of 147 non-overlapping PSUs included in the NLSY97.

2.1 Table 2. NLSY97 Sampling Status of Youths by Housing Arrangement

Housing arrangement	Status
Exchange students	Included if the youth lived in the sample housing unit for at least six months during 1997.
Youths whose temporary residence was a group quarters structure (e.g., prisons, boarding school, college dormitories)	Included if their usual place of residence was in a selected PSU. Excluded otherwise.
Youths whose usual place of residence was not in a selected PSU, but whose temporary residence was within a PSU	Excluded.
Youths in a foreign school	Included.
Youths linked to two or more housing units	If the respondent's mother is alive and her housing unit is in a sample housing unit, the youth is linked there. Otherwise, the youth is linked to the father's housing unit. If neither the mother nor the father is living in a sample housing unit, the youth is linked to one of the sample housing units at random.
Youths who cannot be linked to any other housing unit	Included if the youth is residing at a sample housing unit when the screening interview is conducted.

Siblings: The NLS sample design, which selected every eligible person connected to the housing unit, generated a sample of siblings living in the same housing unit and satisfying the NLSY97 age restrictions. However, the NLSY97 samples do not contain nationally representative samples of siblings of all ages and living arrangements. Care should be used in generalizing from the findings of sibling studies based on the NLSY97. See Table 3 in section 1.3 for the numbers of sibling groups in the NLSY97.

Other technical information on the sample assignment process can be found in (1) the *Field Interviewer Reference Manual*, which includes a copy of the screening instrument, and (2) the *Technical Sampling Report*, which describes the NLSY97 sample selection procedures for both subsamples. Both of these documents are available on the NLS Web site at <<http://www.bls.gov/nls>>.

2.2 Interview Methods

This section first discusses the data collection methods used for the five round 1 survey instruments: the *Screener, Household Roster, and Nonresident Roster Questionnaire*; the *Youth Questionnaire*; the *Parent Questionnaire*; the *School Survey*; and the *CAT-ASVAB*. Following this overview, the section briefly describes interview administration in subsequent survey rounds. The content of these instruments is described in section 1.4, “Content of the NLSY97.”

Users should note that respondents have received \$10 for their participation in rounds 1–3, and responding parents received \$10 when they completed the round 1 interview. In round 4, survey administrators offered different levels of incentives to respondents in an effort to study the effects of incentive level on survey participation. Three levels of compensation were offered: \$10, \$15, and \$20. In

In addition, half of the respondents at each level were paid in advance and half were paid upon completion of the interview. Both the level and the timing of the compensation are included in the variable PAYINCENT, found in the round 4 data. In round 5 all respondents received \$20.

The field periods have differed somewhat across rounds. Table 1 indicates when the first several rounds were fielded, along with the total response rate.

2.2 Table 1. NLSY97 Sample Sizes, Retention Rates, and Fielding Periods

Round	Fielding period	Cross-sectional sample		Supplemental sample		Total sample	
		Total	Retention rate	Total	Retention rate	Total	Retention rate
1	February–October 1997 and March–May 1998	6748	—	2236	—	8984	—
2	October 1998–April 1999	6279	93.0	2107	94.2	8386	93.3
3	October 1999–April 2000	6173	91.5	2036	91.1	8209	91.4
4	November 2000–May 2001	6055	89.7	2026	90.6	8081	89.9
5	November 2001–May 2002	5919	87.7	1964	87.8	7883	87.7
6	November 2002–May 2003 ¹	NA	NA	NA	NA	NA	NA

Note: Retention rate is defined as the percentage of base year respondents remaining eligible who were interviewed in a given survey year; deceased respondents are included in the calculations.

¹ Round 6 was fielded in 2002–03 but is not discussed in this guide.

Round 1 Interview Methods

Fielding Period: Most round 1 NLSY97 interviews were conducted between January and early October 1997. Due to concerns about the number of eligible youths found during the initial field period, investigators decided to conduct a refielding between March and May 1998. During this second part of the initial survey round, 395 additional respondents were interviewed. These respondents were administered the same instrument as those initially interviewed in 1997. See section 2.3 for more information about the composition of the NLSY97 sample.

Data hint →

Respondents selected for the NLSY97 sample during the refielding are identified by the refielding symbol (CV_REFIELD_YOUTH).

Researchers analyzing topics where time periods are critical should carefully examine the reference period of the questions, as well as the actual interview date for individual respondents. In particular, the round 1 fielding period has implications for questions on education; see section 4.2.2, “Educational Status & Attainment,” for more information.

Researchers should also pay close attention to the elapsed time between interviews for each respondent. While the time between the first and second interviews was about 18 months for most respondents, it may be somewhat less for those first interviewed during the refielding period.

Data hint →

The respondent's interview date for each round can be identified by using three created variables: CV_INTERVIEW_DATE_D, CV_INTERVIEW_DATE_M, and CV_INTERVIEW_DATE_Y.

Screener, Household Roster, and Nonresident Roster Questionnaire

Choice of household informant: To identify youths potentially eligible for the NLSY97, the screener collected data from selected households within a sample area. A single member of the household, designated as the household informant, was asked to provide certain information on persons who usually resided in the household. To ensure more accurate reporting of these data, the NLSY97 required the household informant to be age 18 or older and to consider the selected household his or her usual place of residence.

Computer-Assisted Personal Interview (CAPI): After a household informant was chosen to complete the *Screener, Household Roster, and Nonresident Roster Questionnaire*, interviewers used a CAPI system to collect data. Computer software automatically guided interviewers through an electronic questionnaire, selecting the next question based on a respondent's answers. The program also prevented interviewers from entering invalid values and warned interviewers about implausible answers. A set of checks within the CAPI system lowered the probability of inconsistent data both during an interview and over time. To ensure that accurate data were collected from Spanish-speaking respondents, CHRR prepared both English and Spanish versions of all survey instruments, and NORC employed bilingual Spanish-speaking interviewers to administer the Spanish version to those requesting it. During the initial round, the Spanish version of the questionnaire was requested by 297 responding parents and 96 NLSY97 youths.

Screen and Go: In round 1, use of the computer-assisted personal interviewing system (CAPI) allowed for a screen and go method of screening households. When an NLSY97-eligible youth was identified in the simple screener portion of the interview, information from the remainder of the *Screener, Household Roster, and Nonresident Roster Questionnaire* was collected. Selected data (e.g., basic demographic information, a roster of household members) were then transferred automatically into the *Parent and Youth Questionnaires* for verification and use during the interview. Therefore, the interviewer could administer the parent or the youth portion of the NLSY97 immediately. It was expected that this would increase the likelihood that eligible youths participated in the survey since the number of visits interviewers had to make to a household decreased.

However, in some cases, the respondents (parent and youth) were not available to participate in the parent and youth interviews immediately after screening. In these cases, a screen and come back method was utilized, in which the interviewer made an appointment to return to the household to administer the *Youth* and *Parent Questionnaires* at a convenient time.

Paper Screener: During round 1, the interviewers had the option of using a paper screener to perform the initial screening of the household. The paper screener collected the same basic information as the initial CAPI screener. This was useful in cases where the simple screener information could not be collected using CAPI (e.g., weather conditions, computer battery life, dangerous neighborhood) and also gave the interviewer an alternative medium for collecting the initial screener data. Like the screen and go model, the paper screener was designed to determine if anyone residing in the housing unit was eligible for either the NLSY97 or the administration of the *CAT-ASVAB*. If a youth was identified as being potentially eligible for the NLSY97, the information from the paper screener was entered into CAPI. The interviewer could then continue in CAPI with the *Screener, Household Roster, and Nonresident Roster Questionnaire* and the *Youth* and *Parent Questionnaires*. Approximately 28,000 paper screeners were administered, including those used for the screen and come back method described above.

Proxy Screener: In cases where a round 1 interviewer made several visits to a household and still had difficulty contacting household members to administer the initial screener, a proxy screener was administered to an adult living either next door to or directly across from the selected housing unit. Before the interviewer could administer a proxy screener, at least three attempts were made by the interviewer, on different days and at different times, to contact anyone in the selected housing unit.

The purpose of the proxy screener, a paper questionnaire, was to assess whether a person eligible for the NLSY97 resided in the household. In particular, the proxy screener was designed to determine the best time to establish contact with a household member, whether or not a person between the ages of 8 and 28 currently lived in the household, and the steps required to contact a household member. The broad 8–28 age range was intended to ensure that youths close to the endpoints of the actual age range were not missed due to inaccurate reporting. If the proxy screener indicated that none of the household members were in the age range of 8 to 28, the screener was coded as a proxy screener and no more attempts were made to contact the household. However, if the proxy informant was unable to definitively deny the presence of residents ages 8–28, the interviewer was instructed to return as many times as reasonable and necessary to administer the simple screener and, if appropriate, the remainder of the survey instruments. A total of 5,175 proxy screeners determined that no one between ages 8 and 28 lived in the household.

Gatekeepers: The gatekeeper disposition code was used in cases where the interviewer could not gain direct access to the sample household, such as a high-rise building with a locked door where access was

denied by a building manager or a gated housing community where the entry guard refused entrance. In these cases, the interviewer asked the gatekeeper or other community official whether anyone between the ages of 8 and 28 lived in the sample households. If the gatekeeper was unable to definitively deny the presence of household members ages 8–28, the interviewer then attempted to gain access to the household in order to complete the *Screener, Household Roster, and Nonresident Roster Questionnaire* and was not permitted to use this disposition code. A total of 4,055 cases were closed with a gatekeeper disposition code after the interviewer determined that no one between ages 8 and 28 lived in the household. This code was mainly used in gated housing communities for senior citizens.

Telephone Screener: In rare cases at the conclusion of the field period, the simple screener was conducted by telephone. A total of 931 telephone screeners were administered. Instances in which the housing unit was contacted by telephone include:

- (1) The proxy screener revealed a person between the ages of 8 and 28 living in the household and the interviewer was unable to contact anyone in the housing unit on three subsequent in-person visits; or
- (2) The interviewer made three in-person visits but was unable to find a neighbor to whom he or she could administer the proxy screener.

The full *Screener, Household Roster, and Nonresident Roster Questionnaire* was also administered by telephone in rare instances. Situations in which the full instrument was conducted by telephone include:

- (1) After completing the paper screener, the interviewer was unable to contact anyone in the housing unit to complete the full extended screener. At least three in-person contacts must have been attempted before the telephone contact was approved.
- (2) The sample housing unit was inside a residential community to which the interviewer was barred access by the community (e.g., housing board authority). Prior to the telephone interview, the correct person must have been contacted about gaining access at least three times (in person, by telephone, or by letter).

NLSY97 Parent Questionnaire and Youth Questionnaire

When the *Screener, Household Roster, and Nonresident Roster Questionnaire* was complete, any NLSY97-eligible youth(s) and one of the youth's parents (the responding parent) were interviewed using CAPI. Prior to these interviews, selected data (e.g., basic demographic information, a roster of household members) were automatically transferred into the *Parent Questionnaire* and the *Youth Questionnaire* for verification and use during the interviews. Consequently, the interviewer was able to administer the parent or the youth portion of the NLSY97 immediately. CAPI interviews were conducted in either English or Spanish; parent and youth respondents could choose either version.

Data hint →

In round 1, the NLSY97 youth respondent(s) and responding parent(s) in the household are listed on the household roster, but they are referred to as “Household Member #” in the same way as noninterviewed household members. The youth respondent’s position on the household roster can be identified by using the variable YOUTH_HHID.01. The responding parent’s position on the roster is provided in PARYOUTH_PARENTID. See section 4.6.5, “Household Composition,” for further discussion of the structure and use of the household roster.

Choice of Parent: One parent of each respondent was asked to participate in the parent interview. This parent was identified during the household roster portion of the survey. The responding parent (or guardian) was asked for extensive background information, including marital and employment histories. He or she was also asked to answer questions about the family in general, as well as to provide information about aspects of his or her (NLSY97-eligible) children’s lives.

The choice of the preferred responding parent was based on the pre-ordered list in Figure 1. For example, a biological mother was chosen before a biological father, and so forth. However, in some cases a parent figure lower on the list was chosen if a parent higher on the list was in the household but was not available at the time of the interview. If the youth did not live with a parent-type figure, or lived with a guardian or parent not listed, no parent was interviewed; the youth’s record will not contain any data from the *Parent Questionnaire*. Users should note that the records of some youths who do live with a listed parent or parent-figure do not contain any data from the *Parent Questionnaire* due to nonresponse.

2.2 Figure 1. Priority for Choosing Responding Parent

1	Biological mother
2	Biological father
3	Adoptive mother
4	Adoptive father
5	Stepmother
6	Stepfather
7	Guardian, relative
8	Foster parent, youth lived with for 2 or more years
9	Other non-relative, youth lived with for 2 or more years
10	Mother-figure, relative
11	Father-figure, relative
12	Mother-figure, non-relative youth lived with for 2 or more years
13	Father-figure, non-relative youth lived with for 2 or more years

Interviews are available with 6,124 parents; 7,942 youth respondents have information available from a parent interview. Table 2 shows the number of respondents by age who had a parent participate in the round 1 survey.

2.2 Table 2. NLSY97 Youths by Age and Parent Interview Availability

Age (birth year)	Total number of youths	Youths with a parent interview
12 (1984)	1771	1583 (89.4%)
13 (1983)	1807	1615 (89.4%)
14 (1982)	1841	1595 (86.6%)
15 (1981)	1874	1668 (89.0%)
16 (1980)	1691	1481 (87.6%)
Total	8984	7942 (88.4%)

Note: Table based on R05367. and R07359.

In multiple respondent households, more than one parent may have been interviewed during round 1 if the selection criteria above indicated different parents for different NLSY97-eligible youths in the household. For example, if a couple residing in a sample household each had an NLSY97-eligible youth from a previous marriage, the biological parent of each youth would be interviewed. The survey first collected parent-specific information from each parent and then asked for information about the NLSY97-eligible youth matched to that parent. In this example, each parent would be asked to provide youth-specific information only for his or her NLSY97-eligible biological child.

Due to a computer programming error, however, both parents in some multiple respondent households were asked to provide youth-specific information only for the oldest NLSY97-eligible youth(s) living in the household. In the example above, both parents would be asked to give information about the older of the two children. In these infrequent instances, the correct parent-specific information is matched to each youth, but one or more youths in the household do not have any youth-specific information. This programming error was corrected during the survey period and affected only 33 youth cases.

Audio Computer-Assisted Self-Interview (ACASI): The parent and youth portions of the NLSY97 survey used an audio computer-assisted self-interview (ACASI) to obtain potentially sensitive information. The respondent was able to listen to the questions with earphones or turn off the audio and read the questionnaire from the computer screen. Compared to traditional paper-and-pencil self-administered sections, the computerized version permits more complex questionnaire structuring, and the audio component theoretically improves response quality when the respondent’s literacy is in question. As with the interviewer-administered instruments, the ACASI was available in Spanish or English.

User Notes: Each NLSY97 questionnaire includes an interviewer remarks section, which interviewers complete after finishing the interview with the respondent. This section records objective information about the interview, such as the presence of another person during the survey, where the interview took place, and the language in which the questionnaire was administered. Interviewers are also asked to provide an assessment of the interview, stating how cooperative the youth was, how well the youth appeared to understand the questions, whether the youth seemed to be candid and honest, and whether there were any special circumstances that might affect the quality of the data (e.g., respondent lacks social skills, has a mental impairment, has a physical disability, is under the influence of alcohol or drugs). Finally, the interviewer observes the youth’s home and neighborhood environment, describing the interior and exterior condition of the youth’s home, the type of neighborhood (rural and agricultural, suburban residential, urban residential, urban mixed residential and commercial, etc.), the type of residence most common on the youth’s street, and whether the interviewer felt safe in the youth’s neighborhood. These questions help survey staff to plan for future interviews by anticipating potential problems and provide researchers with a general idea of the quality of the respondent’s answers. Questions found in the interviewer remarks section have the prefix “YIR” in their question name.

Supplemental NLSY97 Studies

School Survey (1996). Designed with an emphasis on the school-to-work transition, round 1 of the NLSY97 also included a mail survey of schools. Principals (or their proxies) were asked to complete a self-administered instrument that focused on institutional-level attributes such as school policies and management as well as student-level “experience” data. See section 4.2.5, “School & Transcript Surveys,” for more detail about the content of the survey.

Schools in the NLSY97 sample areas that had a 12th grade comprised the sample for this survey. As depicted in Figure 1 in section 2.1 of this chapter, the NLSY97 sample was drawn from 147 primary sampling units (PSUs).³ The PSUs were further divided into sample segments. All schools in any county with a segment selected for NLSY97 sampling were included in the survey. There were some counties in the PSUs from which no sample segments were selected. The 1996 survey did not include schools in these counties. Schools were identified using the Quality Education Data (QED) file, a proprietary national database of primary and secondary schools in the United States.

³ There are 100 PSUs in the cross-sectional sample and 100 PSUs in the oversample; however, some PSUs were selected in both samples. Thus, a total of 147 non-overlapping PSUs are included in the NLSY97.

The original school survey form was mailed in September 1996; in-scope schools that did not respond by December 1996 were sent a shorter version of the survey, the “critical items” questionnaire. Of the 7,390 in-scope schools that received the survey, 5,295 responded to either the original school survey or the critical items questionnaire. The response rate by the end of the field period, April 5, 1997, was 71.6 percent.

Answer forms for the original school survey were electronically scanned by NORC. However, some hand editing was necessary. The majority of the edited questions were in decimal format. To ensure clean data, the answers were verified by randomly selecting cases, keying the data, and comparing the keyed data files against the scanned data files. The critical items questionnaire did not use a scannable format; the data were keyed using Computer Assisted Data Entry (CADE) and verified twice.

CAT-ASVAB: From summer 1997 through spring 1998, most NLSY97 respondents were administered the computer adaptive version of the *Armed Services Vocational Aptitude Battery (CAT-ASVAB)*, as well as the *Interest-Finder*. See section 4.1.2, “Administration of the *CAT-ASVAB*,” for more information.

Rounds 2–5 Interview Methods

Fielding Periods: The round 2 survey was conducted from October 1998 through April 1999. Most respondents were surveyed approximately 18 months after their first interview, although the elapsed time between interviews is substantially less for some respondents. The round 3 survey was conducted from October 1999 through April 2000. Round 4 surveys were administered from November 2000 through May 2001, and round 5 surveys were conducted from November 2001 through May 2002.

Locating respondents is a coordinated effort of NORC’s central office, locating shop, and local-level field staff. Prior to fielding, NORC’s central office sends a short, informative “locator letter” to each respondent reminding him or her of the upcoming interview and confirming the respondent’s current address and phone number.

Youth Questionnaire: As in round 1, the interviews were conducted using a CAPI instrument, administered in person by an interviewer with a laptop computer. During sensitive portions of the interview, the respondents entered their answers directly into the laptop rather than interacting with the interviewer. This self-administered portion, called ACASI, included an audio option so that the respondents could listen to the questions and answers being read via headphones if they preferred.

Household Income Update: This brief questionnaire collected basic income information from one of the respondent’s parents (usually the parent who signed the youth’s interview consent form). All respondents who live with a parent are eligible for this questionnaire, regardless of age or other criteria for

independence. The parent answered these questions on a self-administered paper instrument. Interviewers then entered the data into a computer-assisted questionnaire on their laptops and attached the information to the records of all NLSY97 youths in the household. Additional quality control checks were performed in the central office, where hard copy questionnaires were reviewed against the coded data. In round 2, parents of 7,601 respondents answered at least one question from the *Household Income Update*. Parents of 5,488 respondents answered at least one question in round 3, and 5,225 parents of respondents answered at least one question in round 4. Parents of 4,090 respondents answered at least one question in round 5.

Transcript Survey. In winter 1999–2000, the 2000 NLSY97 transcript survey sought high school transcripts for all sample respondents who were no longer enrolled in high school and for whom field interviewers had secured parent and respondent consent for transcript release. Eligible respondents were those who either had graduated from high school or who were age 18 or older and no longer enrolled in high school. Transcripts were received and processed for 1,417 respondents. Using course catalogs, transcript data, and clarification calls to school administrators, survey staff constructed histories of courses taken and term enrollment calendars for each youth. Data files also include information on absences, standardized test scores, and indicators of special education, gifted/talented, and high school graduation status. Courses were coded into the Revised Secondary School Taxonomy (SST-R). Public use data are available on the round 4 Event History data release.

School Survey (2000). Round 3 of the NLSY97 also included a repeat survey of schools. Principals (or their proxies) were asked to complete a self-administered instrument similar to that used in 1996. To reduce the time burden, questionnaire items from the 1996 instrument were modified to encourage respondents to provide approximate values rather than requiring them to consult administrative records for exact figures. See section 4.2.5, “School & Transcript Surveys,” for more detail about the content of the survey.

As in 1996, schools in the NLSY97 PSUs that had a 12th grade were mailed survey instruments. However, the 2000 sample was expanded to include vocational schools. The sample also included schools in the counties that were in NLSY97 PSUs but did not include any sample segments. Schools in these counties had been omitted from the 1996 survey but were included for limited data collection in 2000. No telephone follow-up was done for schools in these “omitted counties.” Finally, in addition to the geographically based sample, other schools were included if an NLSY97 respondent was enrolled during round 2 and that school met the grade and program requirements for eligibility. Schools were identified using the 1998 Quality Education Data (QED) file.

By January 2000, survey staff had secured cooperation from state school officers and local school districts. In February 2000, questionnaires were mailed to 9,632 sampled schools, including 8,925 schools in a longitudinal sample (comparable to the 1996 school survey), 492 in the omitted counties sample, and 215 eligible only due to round 2 youth enrollment. After mail and telephone follow-up, 5,955 schools (71.6 percent) in the longitudinal sample (comparable to the 1996 school survey) completed questionnaires. The overall response rate for all schools in the 2000 survey was 71 percent.

Due to “births” and “deaths” of schools between 1996 and 2000 and nonresponse in 1996, not all schools in the longitudinal sample are present in the 1996 data. The retention rate of 1996 schools into the 2000 survey was 74.2 percent (3,900 of 5,253).

Validation Reinterviews. After each round of the NLSY97, validation reinterviews are conducted with randomly selected respondents in order to confirm that their interviews were administered as reported by the interviewer and to solicit feedback on interviewers’ conduct. Most validations are conducted over the telephone by the NORC phone center, with a small number conducted in person or by mail. These data offer opportunities for studying response variance, item reliability, and other methodological issues. Though these reinterviews have been administered each year since round 2, only the rounds 4 and 5 data have been released for public users. These variables have “VALIDR4” or “VALIDR5” as the beginning of each question name and are found in the main file data set.

Round 4: Between November 2000 and July 2001, 989 respondents completed validation reinterviews for round 4. This produced an overall project validation rate of 12.2% of completed interviews. The short telephone questionnaire included a validation component that asked for details about the respondents’ original round 4 interview (e.g., duration, mode) and information on whether or not they were paid for their participation. The reinterview component involved re-asking questions that were drawn directly from the youth interview. This component included some characteristics of their current residence, several expectations questions, a question about weekly family activities, and two questions concerning the respondent’s income from the previous year. Comparable to similar questions from the main interview data, these re-interview data are chosen to represent a variety of question types with different response variance characteristics. Finally, respondents are asked whether the interviewer they had in round 4 was the same one who conducted their interview in round 3.

Round 5: Between November 2001 and June 2002, 1,036 respondents completed validation reinterviews for round 5. This produced an overall project validation rate of 13.0% of completed interviews. The short telephone questionnaire included a validation component similar to the round 4 validation component described above. The reinterview component involved re-asking questions that were drawn directly from the youth interview. This component included their marital status, current employer name

and start date, three child care questions (if female and born in 1983–84), and the respondent’s total income from the previous year. Comparable to similar questions from the main interview data, these re-interview data are chosen to represent a variety of question types with different response variance characteristics. Respondents also are asked whether the interviewer they had in round 5 was the same one who conducted their interview in round 4. Finally, new in round 5 was the addition of a question that asked respondents why they completed the interview. Verbatim answers to this question were recorded and later coded into reason categories.

2.3 Sample Size & Composition

For more information about the representativeness of the sample members, users should consult the *NLSY97 Technical Sampling Report* (2000). Although fewer age-eligible youths than expected were found during the household screenings, no correlation has been identified between education, income, area of residence, etc., and participation in the survey.

Of the youths eligible for interview in the first round, 8,984 were actually interviewed. Table 1 illustrates the racial, ethnic, and gender composition of the initial sample and the respondents participating in subsequent rounds.

2.3 Table 1. Racial, Ethnic & Gender Composition of NLSY97 Sample

Gender	Race/Ethnicity				Total
	Black	Hispanic	Non-black/ non-Hispanic	Mixed	
Round 1					
Male	1169	977	2413	40	4599
Female	1166	924	2252	43	4385
Total	2335	1901	4665	83	8984
Round 2					
Male	1103	904	2238	38	4283
Female	1101	868	2095	39	4103
Total	2204	1772	4333	77	8386
Round 3					
Male	1062	876	2193	39	4170
Female	1071	853	2076	39	4039
Total	2133	1729	4269	78	8209
Round 4					
Male	1065	862	2153	37	4117
Female	1059	837	2027	41	3964
Total	2124	1699	4180	78	8081
Round 5					
Male	996	847	2110	36	3989
Female	1036	828	1991	39	3894
Total	2032	1675	4101	75	7883

Note: Table based on KEY!RACE_ETHNICITY (R14826.), KEY!SEX (R05363.), and RNI (R25102. and R38297.).

User Notes: The initial NLSY97 data release contained records for 9,022 respondents. However, an evaluation of the round 1 data revealed that 38 of these respondents either were not age-eligible for the cohort or were duplicates. The records of these out-of-scope respondents have been removed from the data, and numbers in this guide have been updated to reflect the new sample size of 8,984 respondents. Identification numbers of dropped respondents are included in the round 1 *NLSY97 Codebook Supplement* and are available from NLS User Services.

2.4 Retention and Reasons for Noninterview

After the initial survey round, some sample members do not respond to one or more subsequent interviews. Table 1 shows the retention rates by sample type for rounds 2–5 of the NLSY97.

2.4 Table 1. Retention Rates by Sample Type and Gender

	Cross-sectional		Supplemental		Sample Total	
	Interviewed	Retention Rate	Interviewed	Retention Rate	Interviewed	Retention Rate
<i>Round 2</i>						
Male	3213	92.9%	1070	93.9%	4283	93.1%
Female	3066	93.2	1037	94.6	4103	93.6
Total	6279	93.0	2107	94.2	8386	93.3
<i>Round 3</i>						
Male	3144	90.9	1026	90.0	4170	90.7
Female	3029	92.1	1010	92.2	4039	92.1
Total	6173	91.5	2036	91.1	8209	91.4
<i>Round 4</i>						
Male	3098	89.6	1019	89.4	4117	89.5
Female	2957	89.9	1007	91.9	3964	90.4
Total	6055	89.7	2026	90.6	8081	89.9
<i>Round 5</i>						
Male	3012	87.1	977	85.7	3989	86.7
Female	2907	88.4	987	90.1	3894	88.8
Total	5919	87.7	1964	87.8	7883	87.7

Note: Table based on RNI (R25102. and R38297.), KEY!SEX (R05363.), and CV_SAMPLE_TYPE (R12358.). Retention rate is defined as the percentage of all base-year respondents participating in a given survey. Deceased respondents are included in the calculations.

For each respondent who is not interviewed in a given round, NORC personnel assign a reason for noninterview code, contained in the variable RNI. Tables 2–4 summarize the reasons for noninterview among NLSY97 respondents during rounds 2–5.

2.4 Table 2. Reason for Noninterview by Gender

Reason for noninterview	Deceased	Not locatable	Technical problem	R too ill	R unavailable	Refused interview	Other	Total
Round 2 total	7	104	6	6	42	428	5	598
Male	3	52	3	3	22	229	4	316
Female	4	52	3	3	20	199	1	282
Round 3 total	16	192	2	1	51	510	3	775
Male	7	107	2	1	34	275	3	429
Female	9	85	—	—	17	235	—	346
Round 4 total	15	172	6	6	80	612	12	903
Male	6	87	—	2	53	326	8	482
Female	9	85	6	4	27	286	4	421
Round 5 total	25	278	—	1	77	718	2	1101
Male	14	152	—	—	57	386	1	610
Female	11	126	—	1	20	332	1	491

Note: Table based on RNI (R25102. and R38297.) and KEY!SEX (R05363.).

2.4 Table 3. Reason for Noninterview by Sample Type

Reason for noninterview	Deceased	Not locatable	Technical problem	R too ill	R unavailable	Refused interview	Other	Total
Round 2 total	7	104	6	6	42	428	5	598
Cross-sectional	6	63	3	6	37	350	4	469
Supplemental	1	41	3	—	5	78	1	129
Round 3 total	16	192	2	1	51	510	3	775
Cross-sectional	13	121	2	1	35	400	3	575
Supplemental	3	71	—	—	16	110	—	200
Round 4 total	15	172	6	6	80	612	12	903
Cross-sectional	12	106	5	5	61	496	8	693
Supplemental	3	66	1	1	19	116	4	210
Round 5 total	25	278	—	1	77	718	2	1101
Cross-sectional	19	171	—	1	53	583	2	829
Supplemental	6	107	—	—	24	135	—	272

Note: Table based on RNI (R25102. and R38297.) and CV_SAMPLE_TYPE (R12358.).

2.4 Table 4. Reason for Noninterview by Race/Ethnicity

Reason for noninterview	Deceased	Not locatable	Technical problem	R too ill	R unavailable	Refused interview	Other	Total
Round 2 total	7	104	6	6	42	428	5	598
Non-black/non-Hisp.	2	22	2	3	22	278	3	332
Black	4	39	—	1	8	79	—	131
Hispanic	1	40	4	2	11	69	2	129
Mixed	—	3	—	—	1	2	—	6
Round 3 total	16	192	2	1	51	510	3	775
Non-black/non-Hisp.	8	65	1	1	23	297	1	396
Black	6	59	—	—	13	123	1	202
Hispanic	2	67	1	—	14	87	1	172
Mixed	—	1	—	—	1	3	—	5
Round 4 total	15	172	6	6	80	612	12	903
Non-black/non-Hisp.	6	61	1	5	33	375	4	485
Black	8	44	3	1	21	128	6	211
Hispanic	1	66	1	—	26	106	2	202
Mixed	—	1	1	—	—	3	—	5
Round 5 total	25	278	—	1	77	718	2	1101
Non-black/non-Hisp.	9	100	—	1	35	417	2	564
Black	14	77	—	—	23	189	—	303
Hispanic	2	96	—	—	19	109	—	226
Mixed	—	5	—	—	—	3	—	8

Note: Table based on RNI (R25102. and R38297.) and KEY!RACE_ETHNICITY (R14826.).

2.5 Sample Weights & Design Effects

Sample Weights

The sampling weights, which are constructed in each survey year, provide the researcher with an estimate of how many individuals in the United States are represented by each respondent. Weighting decisions for the round 1 NLSY97 data were guided by the following principles. Individual case weights were assigned to produce group population estimates when used in tabulations. The assignment of individual respondent weights involved at least three types of adjustment. Interested users should consult the *NLSY97 Technical Sampling Report* for a step-by-step description of the following adjustment process.

Adjustment One: The first weighting adjustment involves the reciprocal of the probability of selection. Specifically, this probability of selection is a function of the probability of selection associated with the housing unit in which the respondent was located as well as the subsampling (if any) applied to individuals identified in screening.

Adjustment Two: This process adjusts for differential response (cooperation) rates in the screening phase. Differential cooperation rates are computed (and adjusted) based on geographic location, group membership, and within-group subclassification.

Adjustment Three: This weighting adjustment attempts to correct for certain types of random variation associated with sampling as well as sample “undercoverage.” These ratio estimations are used to conform the sample to Census Bureau estimates of population totals.

Sampling Weights and Readjustments: NORC recalculates the sampling weights for all interviewed respondents after each survey round. These readjustments correct for differential nonresponse. The weights are created using base year sample parameters in a procedure similar to that described above. However, in the final stage of post-stratification, the weights are computed on the basis of completed cases in that survey year rather than on the number of respondents in the entire sample.

User Notes: Various sampling weights have been created in different survey years. The figure below shows the variables created in each round and the question name of each variable. Cross-sectional weights refer to the cross-sectional sample. Panel weights include only those respondents who have been interviewed in every round up to that round's interview date. The cumulative cases method refers to a new method for creating more accurate sampling weights.

2.5 Figure 1. Sampling Weight Variable for All Rounds

Sampling Weight Variables	Round 1	Round 2	Round 3	Round 4	Round 5
Sampling Weight (includes round 4 "old method")	SAMPLING_WEIGHT	SAMPLING_WEIGHT	SAMPLING_WEIGHT	SAMPLING_WEIGHT	--
Cross-Sectional Sampling Weight	CS_SAMPLING_WEIGHT	CS_SAMPLING_WEIGHT	CS_SAMPLING_WEIGHT	--	--
Sampling Weight Cumulative Cases Method	--	--	--	SAMPLING_WEIGHT_CC	SAMPLING_WEIGHT_CC
Sampling Weight Panel Method	--	--	SAMPLING_PANEL_WEIGHT_R3	SAMPLING_PANEL_WEIGHT	SAMPLING_PANEL_WEIGHT
Cross-Sectional Panel Weight	--	--	CS_PANEL_WEIGHT	--	--

Practical Usage

Researchers should weight the observations using the weights provided if tabulating sample characteristics in order to describe the population represented (i.e., computing sample means, totals, or proportions). The use of weights may not be appropriate without other adjustments for the following applications:

Samples Generated by Dropping Observations with Item Nonresponses: Often users confine their analysis to subsamples of respondents who provided valid answers to certain questions. In this case, a weighted mean will not represent the entire population, but rather those persons in the population who would have given a valid response to the specified questions. Item nonresponse due to refusals, don't knows, or invalid skips is usually quite small, so the degree to which the weights are incorrect is probably quite small. In the event that item nonresponse constitutes a small proportion of the variables under analysis, population estimates (i.e., weighted sample means, medians, and proportions) would be reasonably accurate. However, population estimates based on data items that have relatively high nonresponse rates—such as family income—may not necessarily be representative of the underlying population of the cohort under analysis.

Data from Multiple Waves: Because the weights are specific to a single wave of the study, and because respondents occasionally miss an interview but are contacted in a subsequent wave, a problem similar to item nonresponse arises when the data are used longitudinally. In addition, occasionally the weights for a respondent in different years may be quite dissimilar, leaving the user uncertain as to which weight is appropriate. In principle, if a user wished to apply weights to multiple wave data, weights would have to be recomputed based upon the persons for whom complete data are available. In practice, if the sample is limited to respondents interviewed in a terminal or end point year, the weight for that year can be used.

Users may also create longitudinal weights for multiple survey years by going to the Bureau of Labor Statistics Web site at <http://www.bls.gov/nls> and choosing the NLSY97 link. Picking "Create a set of

custom weights” on this page brings up the NLSY97 Custom Weighting program. To create a set of custom weights, first type in your e-mail address, then select the survey years corresponding to your research and pick the “Create Custom Weights” button. The NLS server will generate a set of longitudinal weights and e-mail you a compressed file in Winzip format.

Regression Analysis: A common question is whether one should use the provided weights to perform weighted least squares when doing regression analysis. Such a course of action may lead to incorrect estimates. If particular groups follow significantly different regression specifications, the preferred method of analysis is to estimate a separate regression for each group or to use indicator variables to specify group membership; regression on a random sample of the population would be misspecified. Users uncertain about the appropriate method should consult an econometrician, statistician, or other person knowledgeable about the data before specifying the regression model.

User Notes: The NLSY97 data set contains two sampling weight variables for survey rounds 1 through 3: `SAMPLING_WEIGHT` and `CS_SAMPLING_WEIGHT`. The first set includes all NLSY97 respondents. These weights (when divided by 100) will add up to an estimate of the number of U.S. residents in the sample age range in 1997. The second set contains weights only for respondents in the cross-sectional sample; all oversample cases have a zero weight. These weights are also designed to produce an estimate of the number of U.S. residents in the sample age range. Since there are fewer respondents if the oversample is omitted, however, each black or Hispanic respondent in the cross-sectional sample has a larger value.

For research that includes analysis by race, using the regular sampling weights rather than the cross-sectional weights will produce results with higher precision for black and Hispanic youths. For research that focuses only on non-black, non-Hispanic youths or that does not include any analysis by race/ethnicity, using the cross-sectional weights will save processing time.

Design Effects

Because the samples are multi-stage stratified random samples instead of simple random samples, respondents tend to be clustered in geographic areas (for more information on the sample design and screening process, see section 2.1). In general, these clusters tend to be alike in a variety of ways for a variety of reasons. For example, there may be cultural differences by locality or ecological differences in labor market conditions. Depending upon the degree of this homogeneity, the conventionally computed standard deviations for the variables, which assume a simple random sample, may be too small. However, by controlling the rate at which particular strata are sampled, multi-stage stratified random samples can improve upon simple random samples. The ratio of the correct standard error to the standard error

computed under the assumption of a simple random sample is known as the design effect. The *NLSY97 Technical Sampling Report* provides design effects for the various strata.

As respondents in the cohort get older, mobility may mix the respondents more uniformly through the country, reducing the clustering of the sample as well as the design effects. Many of the persons who started out in the same PSU will have moved to different areas and may no longer be affected by similar unobservable labor market conditions. As this occurs, the error terms in a regression will more closely approximate the standard error computed for a completely random sample. However, some correlation due to respondents coming from the same household or neighborhood will, almost surely, remain.

By examining the geocode data for the NLSY97, it may be possible to control for some of the environmental factors generating design effects or, if desired, to compute design effects based upon county or metropolitan area clusters.

Reference

Moore, Whitney; Pedlow, Steven; Krishnamurty, Parvati; and Wolter, Kirk. *National Longitudinal Survey of Youth 1997 (NLSY97) Technical Sampling Report*. Chicago: NORC, 2000.

2.6 Confidentiality & Informed Consent

Like all surveys, the NLSY97 relies on the cooperation of respondents to provide relevant, accurate, and timely information that researchers and policy makers can use to understand economic and social phenomena. At a minimum, securing cooperation requires survey administrators to explain clearly to potential respondents the uses of the survey and the importance of each respondent's information to the success of the survey. Ethical survey practice entails more than just convincing members of the survey sample to participate, however. Sample members need to be informed about how their confidentiality will be protected, how their information will be used, whether participation is voluntary or mandatory, the estimated amount of time it will take to complete the survey, and any risks of participating.⁴ After survey administrators provide this and other relevant information to potential respondents, those potential respondents can provide what is called their "informed consent" to participate.

The NLS program has established a variety of procedures for ensuring respondent confidentiality and obtaining informed consent. As described below, these procedures comply with Federal law and the

⁴ Providing information about risks is particularly relevant for medical studies in which participants could suffer adverse reactions to drugs or treatments. Although the NLSY97 has included health-related topics, information about these topics has been obtained only through questions asked of respondents, rather than through any medical exams or procedures. As such, NLSY97 respondents bear no risk of participating.

policies and guidelines of the U.S. Office of Management and Budget (OMB) and the U.S. Bureau of Labor Statistics.

OMB procedures

OMB is responsible for setting overall statistical policy among Federal agencies. For example, OMB has established standards on collecting information on race and ethnicity, industry, occupation, and geographic location. OMB also has established standards on the manner and timing of data releases for such principal economic indicators as the Gross Domestic Product, the national unemployment rate, and the Consumer Price Index. In addition, OMB sets standards on whether and how much respondents to Federal surveys can be paid for their participation, an issue of particular concern in the NLS program.

Another of OMB's responsibilities is to review the procedures and questionnaires that Federal agencies use in collecting information from 10 or more respondents. Federal data collections reviewed by OMB include administrative data, such as the tax forms that the Internal Revenue Service requires individuals and corporations to complete. OMB also reviews all censuses and surveys that Federal agencies conduct, either directly or through contracts.⁵

OMB examines a variety of issues during these reviews, such as:

- The amount of time (and money, if any) that the agency collecting the information estimates respondents will spend to provide the requested information
- The agency's efforts to reduce the burden to respondents of providing the information
- The purpose and necessity of the data collection, including whether it duplicates the objectives of other Federal data collections
- The ways in which the agency obtains informed consent from potential respondents to participate in the data collection
- The policies and procedures that the agency has established to ensure respondent confidentiality
- The statistical methods used to select representative samples, maximize response rates, and account for nonresponse
- The payment of money or the giving of gifts to respondents

⁵ Surveys that are funded through Federal grants to universities and other organizations generally do not have to undergo this OMB review process unless the grantee in turn contracts with a Federal statistical agency such as the U.S. Census Bureau to collect the data. In place of OMB review, surveys funded through grants typically must undergo a competitive peer-review process established by the agency administering the grant, and that review process examines the procedures for maintaining respondent confidentiality and obtaining the informed consent of the participants. In addition, such surveys also typically are scrutinized by an institutional review board established at the grantee's institution.

- The questionnaire itself, including the quality of its design and whether it includes questions that respondents may regard as sensitive

These OMB reviews are very thorough. From the time an agency prepares an OMB information collection request until the time OMB approves the data collection, the process typically takes 7 months or more and includes multiple layers of review within each agency and at OMB. These reviews are helpful in improving survey quality and ensuring that agencies treat respondents properly, both in terms of providing them with information about the data collection and its uses and in protecting respondent confidentiality.

The review process also provides the general public with two opportunities to submit written comments about the proposed data collection. The agency conducting the data collection publishes a notice in the *Federal Register* describing the data collection and inviting the public to request copies of the information collection request, questionnaires, and other materials that the agency eventually will submit to OMB. The public is invited to submit written comments to the agency sponsoring the data collection within 60 days from the time the *Federal Register* notice is published. In the history of the National Longitudinal Surveys program, the public very rarely has submitted comments to BLS, but when comments are received, they are summarized in the information collection request that ultimately is submitted to OMB.

After the request has been submitted to OMB, the agency sponsoring the data collection then publishes a second notice in the *Federal Register* and invites the public to submit comments directly to OMB within 30 days. Again, in the history of the National Longitudinal Surveys program, the public very rarely, if ever, has submitted comments to OMB.

Once OMB has received the information collection request, they have 60 days to review the package, ask follow-up questions, suggest changes (or occasionally insist upon changes) to the survey questionnaire or procedures, and ultimately grant approval.

After OMB grants approval, the sponsoring agency can begin contacting potential respondents and collecting information from them. The process of contacting potential NLS respondents begins with sending them an advance letter several weeks before interviews are scheduled to begin. The advance letter serves several purposes. The obvious purpose is to inform respondents that an interviewer will be contacting them soon, but BLS and the organizations that conduct the surveys for BLS also use the letter to thank respondents for their previous participation and to encourage them to participate in the upcoming round. Another important objective of the advance letter is to remind respondents that their participation is voluntary and to tell them how much time the interview is expected to take. The letter also explains to respondents how the data will be used and how respondents' confidentiality will be

protected by BLS and the organizations that conduct the surveys for BLS. An example of an advance letter is shown in Figure 1.

2.6 Figure 1. NLSY97 Round 5 Advance Letter

Date

Respondent First and Last Name

Street Address

City, State Zip Code

Dear [Respondent First Name],

Thank you for the time you have spent and the information you have shared through your participation in the National Longitudinal Survey of Youth 1997 (NLSY97). The NLSY97 has been so successful because of conscientious participants like you. You are currently making many important decisions in your life. One of them involves the NLSY97. You will decide if you want to participate again this year. You will decide if you want your voice to be heard. You will decide if you want to have an effect on legislation. You will decide if you want researchers and policy makers to have an accurate view of your generation. Your continued commitment to the NLSY97 is an important decision.

The value of the information we collect increases with each year you choose to participate in the survey. A longitudinal survey such as this provides researchers with a glimpse of how your life has changed over the years. Therefore, it is important that we interview you each year to obtain an accurate picture of your life. For the same reason, we cannot replace you with a new participant.

Soon, the NLSY97 will begin the 6th round of annual interviews. As you may remember, this survey is sponsored by the U.S. Department of Labor and takes about an hour to complete. Your participation is voluntary, and the information you provide is protected under the Privacy Act in accordance with the statement provided on the back of this letter. The survey OMB control number is 1220-0157. Without this number, we would not be able to conduct this survey. (See Disclosure Notice on reverse.)

To help us get in touch with you, please update the information on the card that you received several weeks ago and mail it back in the envelope provided. Please take the time to read the enclosed brochure, which lists some of the organizations that have used the NLSY97, along with a few articles about your generation that have been written using information from the survey.

If you have any questions or feedback regarding the survey, please call our toll-free number at 1-800-789-1346 or e-mail us at nlsy97@norcmail.uchicago.edu. We look forward to seeing you soon.

Sincerely,



Charles R. Pierret
Program Director, Bureau of Labor Statistics

Kymn Kochanek
Project Director, NORC

Institutional review boards

In addition to OMB review, the NLSY97 is reviewed and approved by an institutional review board (IRB) at the institutions that manage and conduct the surveys under contract with BLS. Those institutions are The Ohio State University and the National Opinion Research Center (NORC) at the University of Chicago. BLS and OMB do not require these reviews; rather, they are required under the policies of the universities. Obtaining approval from the IRBs involves completing a form signed by the Principal Investigator, providing a summary of the research project, and submitting a description of the consent procedures and forms used in the survey. Additional documentation includes a copy of any materials used to recruit respondents, a detailed summary of the survey questionnaire, and any other information regarding the risks to humans of participating in the survey.

The NLS project staff at The Ohio State University Center for Human Resource Research (CHRR) and at NORC obtain approval from their respective IRBs prior to the start of each round of data collection. Since each survey includes only an interview and no invasive medical procedures, the IRBs typically focus on respondent compensation, consent procedures, and confidentiality protections for special populations, such as incarcerated or disabled respondents. Prisons, schools, and other institutions where NLS sample members may reside often request the IRB approval statement and application as evidence that appropriate procedures are being followed and to judge whether to permit NLS interviewers to have access to individuals for whom the institutions are responsible.

Federal laws

Two Federal laws govern policies and procedures for protecting respondent confidentiality and obtaining informed consent in the NLSY97: the Privacy Act of 1974 and the Confidential Information Protection and Statistical Efficiency Act (CIPSEA) of 2002.

The Privacy Act of 1974 protects the confidentiality of participants in the NLSY79 and NLSY97. BLS policy requires that advance letters sent to potential respondents in these surveys include on the back of the letter the BLS Privacy Act statement shown in Figure 2. In addition to the Privacy Act statement, BLS also requires that advance letters sent to potential NLSY79 and NLSY97 respondents include on the back of the letters the Disclosure Notice shown in Figure 3.

2.6 Figure 2. Privacy Act Statement Used in the NLSY97

In accordance with the Privacy Act of 1974, as amended (5 U.S.C. 552a), you are hereby notified that this survey is sponsored by the U.S. Department of Labor, Bureau of Labor Statistics, under authority of 29 U.S.C. 2. Your participation is vital to the success of the survey and is voluntary. The information collected in this survey will be used by BLS and other researchers to understand the labor market experience of individuals over time. BLS makes available to the general public and other federal agencies conducting labor force research information that is not individually identifiable. Under written agreements to protect the confidentiality and security of identifying information, BLS provides potentially identifying geographic information to researchers to conduct specific research projects that further the mission and function of BLS. Unless required by law, or necessary for litigation or legal proceedings, and except as provided in the Privacy Act or as indicated in this statement, we will hold personal identifiers (for example, name, address, or Social Security number) in confidence and will not release them to the public. BLS may release identifying data to the National Opinion Research Center (NORC) and The Ohio State University, and contractors, grantees, and their employees or volunteers who are working on this project for BLS and who need access to the material, or to the National Archives and Records Administration or the General Services Administration for records management purposes.

2.6 Figure 3. BLS Disclosure Notice used in the NLSY97

The Bureau of Labor Statistics estimates that it will take an average of 60 minutes per respondent to complete this survey. If you have any comments regarding this estimate or any other aspect of the survey, including suggestions for reducing the time needed to respond, you may send them to the Bureau of Labor Statistics, National Longitudinal Surveys, 2 Massachusetts Ave., N.E., Washington DC 20212.

To be sure, the Privacy Act statement includes language that may be confusing to some respondents, particularly those not familiar with the nuances of legal writing. For that reason, the advance letter itself typically is written in a friendlier, less formal style to inform respondents about how their confidentiality will be protected and how their information will be used. In addition, survey interviewers are trained how to answer questions from respondents about how their privacy will be protected.

Interviewers explain to potential respondents that all the employees who work on the surveys at BLS, NORC, and CHRR are required to sign an affidavit stating that they will not disclose the identities of survey respondents to anyone who does not work on the NLS program and is therefore not legally authorized to have such information. In fact, no one at BLS has access to information about respondents' identities, and only a few staff members at NORC and CHRR who need such information to carry out their job duties have access to information about respondents' identities.

Interviewers also explain that the answers respondents provide will be made available to researchers at BLS and other government agencies, universities, and private research organizations, but only after all personal identifiers—such as names, addresses, Social Security numbers, and places of work—have been removed. In addition, the publicly available data files exclude any information about the states, counties, metropolitan areas, and other more detailed geographic locations where respondents live, making it much more difficult to infer the identities of respondents.

Respondents are told that some researchers are granted special access to data files that include geographic information, but only after those researchers undergo a thorough application process at BLS and sign a written agreement making them official agents of the BLS and requiring them to protect the confidentiality of respondents. In no case are researchers provided with information on the personal identities of respondents.

Finally, the reference in the Privacy Act statement to the National Archives and Records Administration and the General Services Administration may be confusing to some potential respondents, since those Federal agencies are not involved in the administration of the surveys. Interviewers explain to respondents that NLS data and materials will be made available to those agencies because they are responsible for storing the Nation's historical documents, but that the information provided to those agencies does not include respondents' personal identities.

The Confidential Information Protection and Statistical Efficiency Act (CIPSEA), a new Federal law enacted in 2002, also protects the confidentiality of participants in the NLSY97 by ensuring that individuals who provide information to BLS under a pledge of confidentiality for statistical purposes will not have that information disclosed in identifiable form to anyone not authorized to have it. In addition, CIPSEA ensures that the information respondents provide will be used only for statistical purposes. While it always has been the BLS policy to protect respondent data from disclosure through the Privacy Act and by claiming exemptions to the Freedom of Information Act, this new law is important because it specifically protects data collected from respondents for statistical purposes under a pledge of confidentiality. This new law strengthens the ability of BLS to assure respondents that, when they supply information to BLS, their information will be protected. In addition, CIPSEA includes fines and penalties for any knowing and willful disclosure of specific information to unauthorized persons by any officer, employee, or agent of BLS. Since the enactment of the Trade Secrets Act and the Privacy Act, BLS officers, employees, and agents have been subject to criminal penalties for the mishandling of confidential data, and the fines and penalties under CIPSEA are consistent with those prior laws. CIPSEA now makes such fines and penalties uniform across all Federal agencies that collect data for exclusively statistical purposes under a pledge of confidentiality.

The organizations involved in the NLS program have never had an instance in which a respondent's identity was illegally disclosed, but these organizations continuously monitor their security procedures and improve them when necessary. Protecting the privacy of NLS respondents entails considerable responsibilities for BLS, the organizations that conduct the surveys for BLS, and the researchers who use the data. Indeed, researchers in particular may become frustrated that they cannot obtain access to all the data that they want or that they must undergo a long review process at BLS to obtain some types of data. It is important to remember, however, that protecting respondent confidentiality must remain paramount. Any action that might jeopardize respondent confidentiality and erode the confidence of respondents could harm response rates in the NLS program and in other government or academic surveys. Thus, without the safeguards in place to protect respondent confidentiality, researchers would have far less data available to work with than they currently enjoy.

Survey procedures

BLS, NORC, and CHRR are responsible for following the Federal requirements and maintaining their own security procedures. As mentioned earlier, all officers, employees, and agents of BLS are required to sign affidavits stating that they will not disclose the identities of survey respondents to anyone who does not work on the NLS program and is therefore not legally authorized to have such information. Procedures are in place at each contractor to ensure that the data are secure at each point in the survey process. (See the Data Handling section of this chapter for more information.)

Like all contractor staff, field interviewers are agents of BLS and are required to sign the BLS nondisclosure affidavit before working on the NLSY97. All interviewers also must undergo a background check when they are hired. Confidentiality is stressed during training and enforced at all times. Field interviewers receive specific instructions in their reference manuals to remind them of the appropriate procedures when locating or interacting with respondents or contacts.

At the end of each interview, interviewers ask respondents to provide information on family members, friends, or neighbors who can be contacted if the interviewers are unable to locate the sample member in a subsequent round of interviews. The interviewers then use those contacts to help in locating sample members who have moved. When contacting a sample member's relatives, friends, or neighbors about the sample member's whereabouts, interviewers never disclose the name of the survey they are conducting. They are instructed to maintain the confidentiality of any relative, friend, or neighbor who provides information about the sample member's whereabouts.

Answering machines can pose potential problems when contacting sample members because it is difficult to confirm that the interviewer is calling a sample member's correct phone number or that other

household members will not hear the message. For those reasons, interviewers are instructed not to leave messages on answering machines.

When interviewers contact the appropriate household, they ask to speak with the sample member or the parent of a sample member under age 18. Interviewers introduce themselves and state the purpose of the call by saying that they are from the National Opinion Research Center at the University of Chicago and are calling concerning a national survey. The name of the survey is not disclosed to anyone but the sample member.

Special situations. The NLSY97 is a general population survey and includes a variety of sample members with special circumstances, such as incarcerated individuals, other institutionalized persons, disabled persons, those with limited English proficiency, and so forth.

Incarcerated respondents: Incarcerated respondents constitute the largest group requiring special accommodations. Project procedures regarding incarcerated respondents involve extra efforts to protect their confidentiality. These include obtaining a private room to conduct an in-person interview, a guarantee from the institution that a telephone interview is not monitored, parental consent for incarcerated respondents under the age of 18 (in addition to their own consent), and compliance with the regulations of correctional facilities regarding payment for survey participation.

Respondents with limited English proficiency: Some respondents lack fluency in English and are more comfortable using another language. It is not possible to accommodate all the different languages other than English that respondents might speak, but the NLSY97 survey staff historically have made special arrangements for respondents and their parents who speak Spanish, the most commonly spoken language other than English among respondents. NORC staff members translate advance letters and other informational materials into Spanish to enable respondents and the parents of minor respondents to provide their informed consent based on information that is written in the language that they understand the best. Survey questionnaires also have been translated into Spanish to ensure that the questions are administered consistently, rather than having Spanish-speaking interviewers translate the English-language questionnaire during the interview. The first six rounds of the NLSY97 included a Spanish version of the questionnaire, but because the number of respondents who speak only Spanish has continued to decline, it no longer is cost-effective to continue programming a computerized Spanish questionnaire. For that reason, Spanish questionnaires will not be used starting with round 7 (2003-04) of the NLSY97. Advance letters and other informational materials still will be available in Spanish.

Sensitive subjects: The NLSY97 has included questions on income and assets, religion, relationships with parents and other family members, sexual experiences, abortion, drug and alcohol use, criminal activities,

homelessness, runaway episodes, and other topics that are potentially sensitive for respondents to discuss. Respondents are advised at the start of the interview that they can choose not to answer any questions that they prefer not to answer. During training, interviewers undergo exercises to teach them how to allay the concerns of respondents about answering sensitive questions and encourage them to respond. Interviewers are instructed not to coerce respondents into answering questions that they prefer not to answer.

Most questions in the NLSY97 are read to the respondent by an interviewer. The respondent then provides an answer, and the interviewer records that answer on a laptop computer. For especially sensitive questions, some respondents might be reluctant to answer truthfully—or at all—if they have to tell an interviewer their answers, even though interviewers can face criminal and civil penalties if they disclose the respondents' identities or answers to anyone not authorized to receive that information.

Obviously, it is important that respondents answer all questions truthfully, so the NLSY97 includes a self-administered portion of the questionnaire to reduce the potential reluctance to respond to sensitive questions. In this mode of data collection, the interviewer hands the laptop computer to the respondent and asks the respondent to read the questions and enter his or her responses with the keyboard. Sometimes respondents have literacy problems or disabilities that prevent them from reading the questions on the computer screen. For this reason, the interviewer also provides the respondent with a set of headphones that plug into the computer and enable the respondent to listen to a computer-generated recording of someone reading the questions. While the respondent completes this computer-assisted self-interview, the interviewer does not see the respondent's answers and, in fact, does not even see what questions the respondent is answering. Even in the self-administered portion of the survey, however, respondents still have the option not to respond to individual questions.

Guidelines for e-mailing sample members: At the end of each interview, respondents are asked to provide information that will help interviewers contact them during subsequent rounds of the surveys. In addition to the information collected about relatives, friends, or neighbors, interviewers also obtain the e-mail addresses of sample members who have them. During round 5 of the NLSY97 (conducted during 2001-2002), the NLS contractors began using e-mail as a means to contact a small number of sample members who were hard to reach by other means. The following guidelines were enacted to ensure confidentiality:

1. The name of the survey (NLSY97) is not contained in the subject line or text of the e-mail message. Some sample members may share the use of an e-mail address with other household members, so the survey name is omitted from the message to prevent other household members from learning the specific name of the survey.

2. E-mail is sent from one NORC address. Field interviewers are not permitted to use their individual e-mail accounts to contact sample members.

Respondents knowing respondents: One feature of the sample design in the NLSY97 is that there often are multiple respondents within the same original household, usually siblings but occasionally other relatives. It obviously is not possible in these cases to prevent family members from knowing that a relative is in the survey sample, but interviewers take steps to ensure that each respondent's answers remain private and are not revealed to other family members.

Parental consent. Because of the young ages of the NLSY97 cohort in the initial survey years, additional measures were taken to obtain informed consent from minors and their parents. NLSY97 consent procedures in rounds 1–5 included the following protocol:

1. For respondents age 17 and younger, parents were asked to complete a written Parental Permission to Interview Youth form prior to the youth completing the interview. Respondents themselves were asked to complete a written Youth Assent to Participate form prior to completing the interview. This protocol also applied to respondents who were in prison or whose interviews were completed by a proxy. (Proxy interviews are conducted in cases in which a disabled respondent could complete the interview only with the assistance of a parent, guardian, or other caretaker.) Respondents who were age 17 or younger and had attained independence or were considered “self-supportive” were not required to have a parent complete a permission form.
2. Whether or not a parent permission form was required, all youths, regardless of age, signed a consent form prior to completing the interview.

Through the completion of round 5 data collection in May 2002, a significant proportion of NLSY97 sample members required signed parental permission because they had not yet reached age 18 or established legal independence.

Data handling

An important part of maintaining respondent confidentiality is the careful handling and storage of data. Steps taken by BLS, CHRR, and NORC to ensure the confidentiality of all respondents to the National Longitudinal Surveys include maintaining secure networks, restricting access to geographic variables, and topcoding income and asset values.

Network security. The data that are stored and handled at each NLS organization's site are done so with maximum security in place. During data collection, transmission, and storage, password protection and

encryption are used to secure the data. Standard protocols for network security are followed at each organization's site. Detailed information about these arrangements is not provided to the public to prevent anyone from circumventing these safeguards.

Restricting access to geographic information. Geographic information about NLSY97 respondents is available only to researchers who are designated agents of BLS. These researchers must agree in writing to adhere to the BLS confidentiality policy, and their projects must further the mission of BLS and the NLS program to conduct sound, legitimate research in the social sciences. Applicants must provide a clear statement of their research methodology and objectives and explain how the geographic variables are necessary to meet those objectives. For more information about obtaining access to geographic variables in the NLSY97, see the NLS Web site at <<http://www.bls.gov/nls>>.

Topcoding of income and asset variables. Another step taken to ensure the confidentiality of NLSY97 respondents who have unusually high income and asset values is to “topcode” those values in the data set. Values that exceed a certain level are recoded so that they do not exceed the specified level. In each survey round, income and asset variables that include high values are identified for topcoding. For example, the wage and salary income variable is usually topcoded, but variables indicating the amount received from public assistance programs are not. Notes in the codebooks for topcoded income and asset variables provide more information about the exact calculations used to topcode each variable.

Chapter 3: Guide to the NLSY97 Data

This chapter provides some practical information about how NLS variables are collected, created, and arranged in the data set. An explanation of the cohort's hard copy and electronic documentation is also included. The first section describes the different survey instruments used to collect the raw NLSY97 data. This section also explains how question names are assigned. Next, the guide discusses the primary types of NLS variables and the process by which each is assigned a reference number and title that serve to identify it throughout the NLS documentation. The third section reviews the codebook—that is, the information about each variable contained in the data set—and the accompanying paper documentation. This discussion will help users understand how to interpret the various pieces of information presented in the NLS documentation system. Finally, this chapter gives researchers some basic instruction in using the search functions to find variables of interest.

3.1 Survey Instruments & Other Documentation

The primary variables found within the main data set are derived directly from one or more survey instruments. This section explains the conventions used in the NLSY97 documentation to identify questionnaire items from some of the primary survey instruments.

NLSY97 Survey Instruments

The term “survey instrument” is used to refer to the NLSY97 questionnaires that serve as the primary source of information on a given respondent. In round 1, there were separate and distinctly different questionnaires for the household informant (the *Screener, Household Roster, and Nonresident Roster Questionnaire*), the NLSY97 respondent (the *Youth Questionnaire*), and the responding parent (the *Parent Questionnaire*). In rounds 2–5, the *Youth Questionnaire* and the *Household Income Update* were used as survey instruments. Each questionnaire is organized around a set of topical subjects, the titles of which usually appear on either the first page of each section of the questionnaire or as a header. The various survey instruments are described in detail in section 1.4, “Content of the NLSY97.”

User Notes: The questionnaires are critical elements of the NLSY97 documentation system and should be used by researchers to determine the wording of questions, response categories, and the universe of respondents asked a given question.

For each round, NLSY97 questionnaires record (1) interview dates; (2) responses to the topical survey questions; (3) locating information which will assist NORC in finding the respondent for the next interview (not available to users); and (4) interviewer remarks on such topics as the race and gender of the respondent, language in which the interview was conducted, interviewer's impressions, etc. The show card, an interviewing aid used in conjunction with the questionnaire, lists the possible response categories for select questions and helps the respondent keep the more complicated response categories in mind.

Questionnaire Item or Question Name: This generic term identifies the source of data for a given variable. A questionnaire item may be a question, a check item, or an interviewer’s reference item appearing within one of the survey instruments. These items have question names that begin with an abbreviation of the section where each is located. Following the section abbreviation, the question name includes a combination of numbers and letters that identify it within the section. Many questions simply have numbers in numerical order. Some questions, as in the examples in the tables below, have a decimal extension that indicates the question is repeated or looped during the survey. For example, a question about hours worked would be repeated for each employer, with decimal extensions .01 through .09 indicating employers 1–9. Another common extension in question names is *_D*, *_M*, or *_Y* (or *-D*, *-M*, *-Y*), indicating that the variable reports the day, month, or year of a date. If a question is repeated in more than one round, it will have the same question name in each round so that users can easily locate identical questions in the data set across survey years.

3.1 Table 1. Sample Question Names by Youth Questionnaire Section

Section	Rounds 1, 4–5 Question Names	Rounds 2 and 3 Question Names
<i>Youth Questionnaire</i>		
Information	YINF-2560 (round 1 only)	—
Household Information	YHHI-50510.04, YHHI-4100.07~M (rounds 4–5)	YHHI-50510.04, YHHI-4100.07~M
CPS	YCPS-14400 (rounds 1, 4)	—
Schooling	YSCH-22800.01, YSCH-26500	YSCH-2857B, YSCH-33900.01
Peers/Opportunity Sets	YPRS-800 (round 1 only)	—
Time Use	YTIM-2200 (round 1 only)	YTIM-300
Employment	YEMP-1800.02, YEMP-103500	YEMP-200A, YEMP-38313.02
Training	YTRN-800, YTRN-9200.01	YTRN-710, YTRN-7725.02
Health	YHEA-1600	YHEA-2050
Self-Administered	YSAQ-006A, YSAQ-447.03	YSAQ-394, YSAQ-503.02
Marriage	YMAR-2100, YMAR-12700.01	YMAR-729E, YMAR-3050.01
Fertility	YFER-700.04, YFER-14600	YFER-7800, YFER-12100A.01
Child Care	YCCA-2100 (round 5 only)	—
Program Participation	YPRG-1700, YPRG-13500.01~M	YPRG-12700A, YPRG-19400.03~Y
Income / Assets	YINC-2300, YINC-21400.01	YINC-8900, YAST-2696
Expectations	YEXP-900	—
PIAT Math	YPIA-100	YPIA-100
Locator	YLOC-1500	YLOC-350
Interviewer Remarks	YIR-1500	YIR-1740.01
<i>Household Income Update</i>	HIU-5 (rounds 4–5)	HIU-5

3.1 Table 2. Sample Question Names in Screener and Parent Questionnaires

Section	Round 1 Question Names
<i>Screener, Household Roster, and Nonresident Roster Questionnaire</i>	
Screener	SE-9, SE-31B.01
Household Roster	SH-1B, SH-103.05
Nonresident Roster	SN-225.04, SN-337A.02
<i>Parent Questionnaire</i>	
Information	PINF-015_D, PINF-297.01
Family Background	P2-029, P2-108B.01
Calendars	P3-051.01_M, P3-137
Parent Health	P4-027
Income and Assets	P5-073.02, P5-136
Self-Administered	P6-021B, P6-036
Child Calendar	PC8-009_Y, PC8-025, PC8-086.01
Child Health	PC9-014, PC9-039.04
Child Income	PC10-025
Expectations	PC11-013
Family	PC12-010, PC12-012A
Parent Locator	PLOC-018
Parent Interviewer Remarks	PIR-007, PIR-009K

User Notes: Users should be aware that, while the source of the majority of variables in the main data sets is the questionnaire, certain variables are created either from other NLSY97 variables or from information found in an external data source (see “Types of Variables” below).

3.2 Types of Variables

There are five types of variables present in the NLSY97 data. The type of variable affects the title or variable description of each variable and the physical placement of the variable within the codebook.

Types of variables include:

- (1) Direct (or raw) responses from a questionnaire or other survey instrument.
- (2) Symbols and roster items, which are used to guide the interview.
- (3) Constructed variables based on responses to more than one data item. These items are edited for consistency where necessary.
- (4) Constructed variables from data provided on a non-NLS data set.
- (5) Variables provided by NORC or an outside organization.

User Notes: Users should note that survey personnel do not, in general, impute missing values or perform internal consistency checks across waves. Exceptions will be noted.

Variable Descriptions or Variable Titles

Each variable within NLSY97 main file data sets has been assigned an 80-character summary title that serves as the descriptive representation of that variable throughout the hard copy and electronic documentation system. Variable titles are assigned by CHRR archivists who endeavor, within the limitations described below, to capture the core content of the variable and to incorporate universe identifiers that specify the subset of respondents for which each variable is relevant. Some titles indicate the reference periods (e.g., survey year or calendar year) of the variables as well.

Universe Identifiers: If two ostensibly identical variables differ only in their respondent universes, the variable title will include a reference to the applicable universe. The appropriate universe will either be appended in parentheses or identified before the variable title.

Example 1: R00029. “R Do Any Work for Pay Last Week? (R Does Not Own Bus/Farm)”
R00030. “R Do Any Work for Pay or Profit Last Week? (R Owns Bus/Farm)”

Example 2: R01075. “Compensation Received (Start <16) EMP 01”
R01803. “Compensation Received (Start 16+) EMP 01”

User Notes: Users should not presume that two variables with the same or similar titles necessarily have the same (1) universe of respondents or (2) coding categories or (3) time reference period. While the universe identifier conventions discussed above have been utilized, users are urged to consult the questionnaires for skip patterns and exact time periods for a given variable and to factor in the relevant fielding period(s) for the cohort. In addition, variables with similar content may have completely different titles, depending on the type of variable (raw versus created).

Symbols and Roster Items

There are two main types of variables not necessarily represented by a single item in the questionnaire: symbols and roster items. These items are used by the CAPI system during the interview to organize, display, and store information collected during the interview; to determine which question paths the respondent should follow; and to fill in respondent-specific text in various questions. For example, rather than asking about a respondent’s “current employer,” the CAPI software fills in the actual employer name reported earlier in the interview. Many of these symbols and roster items are provided in the data set for user reference; researchers should be aware of the differences between the two types and the uses of each.

Symbols

The NLSY97 CAPI software generates symbols, which are items containing real-time information provided by the respondent during the survey. Symbols can be used to store data derived from one or more questions. For example, if the youth corrects information from the screener about his or her birth date during the administration of the *Youth Questionnaire*, the corrected information replaces the original data in the symbol item. Symbols are used throughout the questionnaire to determine whether certain groups of questions should be asked. For example, the symbol that states whether the youth is independent (Y12!INDEPEN) is later used to determine whether the youth is asked certain income and asset questions.

All symbol variables have “Symbols” as their primary area of interest. In general, question names for round 1 symbol variables begin with “KEY!”; symbols in rounds 2–5 generally have “SYMBOL!” to start their question names.

Rosters

The NLSY97 uses rosters in various sections in which information is collected on a number of persons, schools, or employers. Rosters are an important part of the NLSY97 data set. These grids of information help researchers to analyze data in an efficient and accurate way. However, the structure and use of rosters may be somewhat confusing, so it is vital that researchers understand how they are constructed.

User Notes: In addition to the detailed discussion in the following paragraphs, the introduction to section 4.3, “Employment,” contains an example that illustrates how to use the employer roster in research. Although that example pertains specifically to employers, the basic concepts apply to other NLSY97 rosters. Researchers using any roster data may find the example helpful.

What is a roster? A roster may be thought of as a list—for example, a list of household members, a list of employers, or a list of children. A respondent with two children will have data on the first two lines of the child list, or child roster. A respondent with four employers will have information on the first four lines of the employer roster. In addition to the name of the person or item (which is not released to the public), the roster contains other basic information, such as the age, race, and labor force status of household members or the start date and stop date for each employer.

In the paper-and-pencil interviews (PAPI) of older NLS cohorts, the questionnaires included a chart or grid listing this type of information, like the one shown in Figure 1. For example, in the household roster grid, each household member’s name was entered in a separate row. The interviewer asked the respondent for each member’s date of birth, enrollment status, employment status, etc., filling in the answers in the

appropriate column. This completed household roster contained all the pertinent information about household residents, and researchers could easily use the variables based on this roster to examine characteristics of household members.

3.2 Figure 1. Sample PAPI Roster Grid

What are the names of all family members who are living in your home?

Name	What is __'s relationship to you?	How old is __ today?	(Age 4 and older) Is __ enrolled in school?	(Age 16 and older)		
				How many weeks did __ work in the last 12 months?	How many hours did __ usually work per week?	What kind of work was __ doing in the past 12 months?
Susan	Mother	45	No	50	25	Graphic design
John	Father	49	No	50	40	Banking
Jimmy	Brother	17	Yes	35	15	Food service
Sally	Sister	12	Yes	(n/a)	(n/a)	(n/a)
Robert	Brother	3	(n/a)	(n/a)	(n/a)	(n/a)
Jan	Grandmother	77	No	(n/a)	(n/a)	(n/a)

When the NLS surveys changed to computer-assisted personal interviewing (CAPI), rosters became a very important way of organizing information during the interview. Instead of using an actual grid, however, CAPI questionnaires include a series of questions that gather the same types of information that would have been included in the grid in a paper-and-pencil interview. The computer then moves the answers to these questions into a grid, creating a roster from the information.

After the roster is created, it can be used to guide subsequent portions of the interview. For example, during the interview the NLSY97 questionnaire gathers the names, dates of attendance, and level of school (secondary school or college) for each of the respondent's schools and organizes them into a roster. The rest of the school section then asks questions about the first school on the roster, followed by questions about the second school, then the third, and so on. The information about the level of the school determines whether the respondent is asked questions that apply to high school or college.

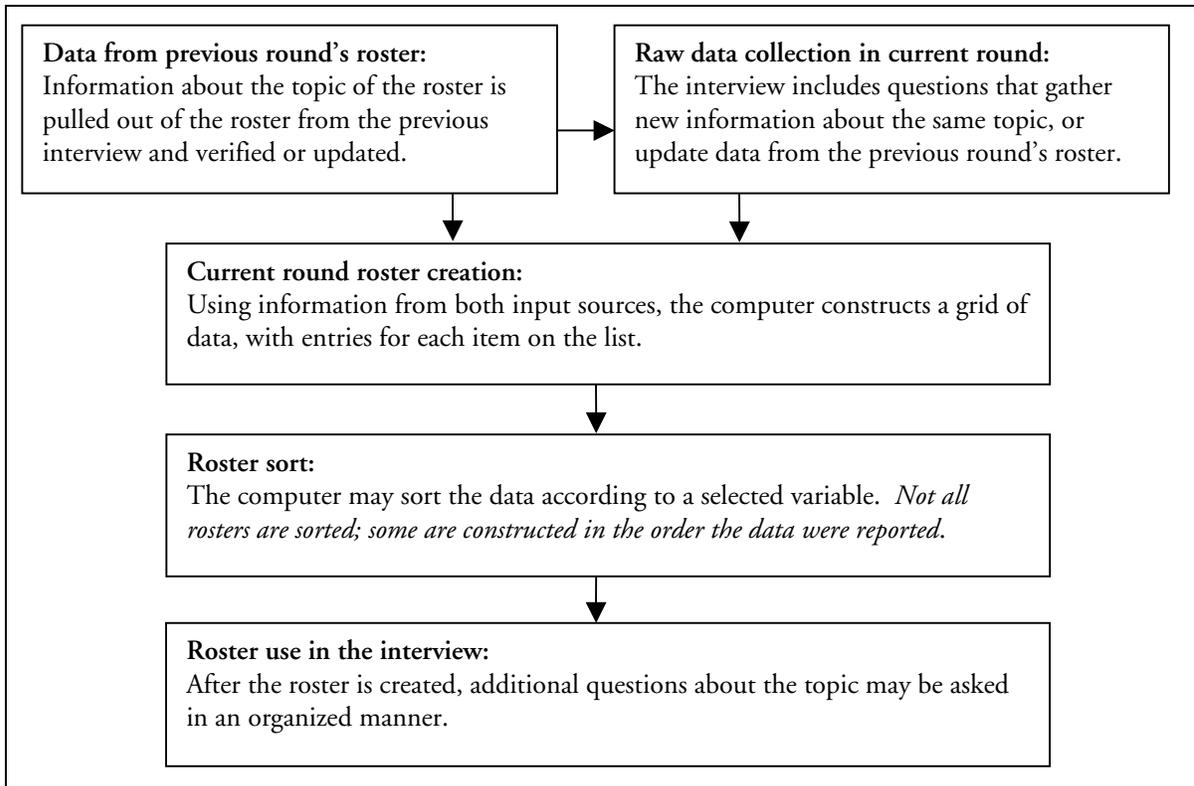
The information from the roster is also presented in the data set as an organized list of data, so that these variables are easy for researchers to access. To the user, the school roster appears as a consolidated block of variables that contains key information such as dates of enrollment, an identification number for the school, and variables indicating the type (private or public) and level (junior high, high school, college) of the school. For example, the variables in the round 2 school roster are listed in Figure 2, along with their reference numbers. Thus, rosters are a way of organizing information both for researchers and for the actual interview so that questions are asked in a logical manner.

3.2 Figure 2. Example: Variables in CAPI-Generated School Roster for Round 2

Question Name	Variable Title	Reference Numbers (one for each school)
NEWSCHOOL_PERIODS.xx	Number of Times R Enrolled in School xx	R24605.–R24610.00
NEWSCHOOL_START1.xx	Month/Year R Start 1st Enrollment in School xx	R24611.00–R24616.01
NEWSCHOOL_START2.xx	Month/Year R Start 2nd Enrollment in School xx	R24617.00–R24620.01
NEWSCHOOL_START3.xx	Month/Year R Start 3rd Enrollment in School xx	R24621.00–R24621.01
NEWSCHOOL_STOP1.xx	Month/Year R End 1st Enrollment in School xx	R24622.00–R24627.01
NEWSCHOOL_STOP2.xx	Month/Year R End 2nd Enrollment in School xx	R24628.00–R24631.01
NEWSCHOOL_STOP3.xx	Month/Year R End 3rd Enrollment in School xx	R24632.00–R24632.01
NEWSCHOOL_SCHCODE.xx	School Code Elementary, Middle, High, College	R24633.–R24638.00
NEWSCHOOL_INTERVIEW.xx	Which Survey Round School xx Reported in	R24639.–R24644.00
NEWSCHOOL_TYPE.xx	Type of School xx R has Attended	R24645.–R24650.00
NEWSCHOOL_PUBID.xx	PUBID of School xx R has Attended	R24651.–R24656.00

How are rosters created during the interview? This section outlines the process used during the interview to create a roster. Rosters may include data from both previous interviews and the current interview. After the roster is created and sorted, it can be used to guide the rest of the interview. Figure 3 provides a pictorial overview of the creation of a roster.

3.2 Figure 3. How Rosters Are Created



Data from previous interviews: As shown in the figure, creation of a roster for the current round often begins with information found in the roster from the previous round. The appropriate respondent-specific data are saved on the interviewer's laptop before he or she administers the survey. When the interview gets to a point where roster information is collected, the data from the previous round's roster are often used as the base for the current roster. The respondent verifies and updates the information. If no changes have occurred since the last interview—for example, if exactly the same people live in the respondent's household—then the current round's roster will be the same as the one from the previous round.

For example, the interviewer reads a list of all of the people on the household roster from the last interview. The respondent first states whether any of those people have moved out of the household and then reports new household members. If any members remain from the previous year, their information—date of birth, gender, race/ethnicity, etc.—is carried over from the previous interview, and any missing data are collected. This method is more efficient than asking the respondent to report all household members every year.

Raw data collection: After the respondent and interviewer review and update the roster from the previous round, the survey collects current information. For example, new people might have moved into the household, so the interviewer asks the respondent about their characteristics. At this point, the respondent is done answering questions that will fill up the data grid on a particular topic.

Roster creation and roster sort: Using the updated roster from the previous round and the new raw data just collected, the computer creates a new roster for the current round. For example, the employer roster contains the following information for each job: a unique identification number for the employer, employment dates, whether the job was current at the interview date, whether the job was in the military, and whether the job was an internship. If the respondent had held the job at the time of the previous interview, the start date and employer identification number are carried over from the old roster, and the other information is taken from the questions at the beginning of the employment section for the current year. Similarly, the household roster contains information from the previous interview about household members reported at that time and data from the current interview about new household members.

In some cases, the computer also sorts the roster and puts the items in order based on a specified variable. For example, in the round 1 household roster, all youths in the age range of the NLSY97 cohort were listed first, and then all other household members were listed from oldest to youngest. The employer roster is sorted by job end date so that the most recent jobs are listed first.

Roster use in the interview: Finally, the roster is used to determine the order in which the other questions about each topic are asked. In most cases, the survey collects far more information than is stored in the

actual roster, and the answers to these questions remain outside the roster as raw data. So that the interview makes sense to the respondent, these additional questions are asked about the people or things on the roster in the order that the people or things are listed.

For example, the respondent first answers questions about industry, occupation, rate of pay, etc., for the first employer listed on the roster. The same questions are then asked about the second job, then the third job, and so on. Similarly, the first set of questions about household members refers to the first person listed on the roster. When all of those questions have been answered, the same questions are asked about the second person, the third person, etc.

How should researchers use the roster data in analysis? The data set is organized so that rosters can easily be found and used in research. Because rosters present key pieces of information in a structured format, they are the best place to obtain that information. All variables found on rosters have “Roster Item” as their main area of interest. Each roster has a unique name that serves as the beginning of the question name for all variables on the roster; the same name appears at the beginning of the variable title for each item on the roster. Different rosters have been used in different rounds, depending on the topics included in the interview and the type of information collected. The roster names and question names are shown in Figure 4.

3.2 Figure 4. Rosters Included Each Round

Roster	Question name	Round 1	Round 2	Round 3	Round 4	Round 5
Household Information	HHI2 (rd. 1), HHI (rds. 2-5)	✓	✓	✓	✓	✓
Nonresident Roster	NONHHI	✓	✓	✓	✓	✓
Youth Information	YOUTH	✓				
School Roster	NEWSCHOOL		✓	✓	✓	✓
Employer Roster	YEMP	✓	✓	✓	✓	✓
Freelance Jobs Roster	FREELANCE		✓	✓	✓	✓
Training Roster	TRAINING			✓	✓	✓
Biological Children Roster	BIOCHILD	✓	✓	✓	✓	
Biological/Adopted Children	BIOADOPT					✓
Parent Household Information	PARHHI	✓				
Parent Youth Information	PARYOUTH	✓				
Partner/Spouse Roster	PARTNERS	✓	✓	✓	✓	✓

Data hint ➔

Researchers can locate rosters in the data by looking at the roster item area of interest, by selecting the appropriate question name, or by searching the “Any Word in Context” index for variables with “ros” or “roster” and the name of the roster of interest in the title.

User Notes: When the NLSY97 data set was initially created, variables could only be assigned to one area of interest. The newer data extraction software permits variables to be linked to multiple areas of interest. However, additional areas have not been assigned to every variable. Because roster variables were initially located in the roster item area of interest, they may not be grouped with the rest of the data on a particular topic. For example, the school roster variables will not appear if the user searches for the “School Experience” area of interest. For this reason, it is very important that researchers become familiar with the rosters used in the data set. If a roster is available on the topic of a particular research project, users should always locate that roster using one of the search techniques mentioned above and examine it before using the other variables that relate to their research.

Using rosters in single-round analyses: When looking at the data set, users will notice that many questions are repeated for each person or thing on the roster, and the titles for these repeated questions include a number. This number indicates the line on the roster that corresponds to the person or item being described in that variable. For example, the question “Self-Employed Business/Industry Job 02” indicates the industry of the second job listed on the respondent’s self-employment roster. The researcher may then want to examine information such as the respondent’s start and stop dates or rate of pay for that job. To find this information, he or she can then look at the data for those items contained in the roster for job #02, or the self-employment job that is on the second line of the roster. For all other questions asked after the roster was created in that same survey year, job #02 will refer to the same self-employment job.

Users should be aware that, in some cases, the information contained in the rosters actually appears in the data set more than once. As Figure 1 suggests, data may first be included at the point in the interview when the information was actually collected. For example, the round 1 screener question SE-28 asked the household informant for the date of birth of each household member. After all the raw data had been gathered, the computer sorted all the answers and created the household roster. At this point the date of birth information is also located in the round 1 roster variables named HHI2_DOB. In the case of the round 1 household roster, both the raw data and roster items are included in the data set.

In other cases, the raw answers may be blanked out of the public use data set. If a reference number is not listed for a given question in the questionnaire, then that raw data item may only be represented in roster form. For example, answers to the raw data questions used to create the employer roster are blanked out and do not appear in the data. In the printed questionnaire, these questions have no reference numbers. However, all of the data collected in these questions (except for confidential information like the name of the employer) appears in the employer roster.

Data hint ➔

Even though the data may appear more than once, **survey staff strongly recommend that researchers use the roster information rather than the raw data whenever possible.** Survey staff are working to eliminate these duplicate sources of information. For example, screener question SE-28 is one of the variables that has already been removed.

For some variables, the roster information may be more accurate because some rosters are updated during the interview if the initial report was inaccurate. When survey staff prepare the data for release, they clean up the rosters if necessary but do not necessarily clean the corresponding raw data. Finally, because many rosters are sorted in a particular order, the number of a person or item on the roster will not match the number in the questions that precede roster creation. For example, in the household screener (the SE questions), person #01 is the first household resident mentioned to the interviewer. In the household roster and all later interview questions, person #01 is the oldest person in the household who was eligible for the NLSY97. Person #01 in the SE questions might be person #05 on the roster. It can be very difficult to determine to which person, school, or job a pre-sort question refers. For all of these reasons, roster data are always preferable to raw data in cases where both are available.

Using rosters from more than one round: Because the NLSY97 is a longitudinal survey, researchers often want to link data across survey rounds. However, household residents, jobs, and so on may move around on the roster in different interviews. That is, a father who was listed third on the roster in round 1 might move to position 2 or 4 in round 2. The unique identification numbers (UIDs) are the key to finding the same person or thing in different rounds. Most of the rosters contain variables assigning a unique number to each person or thing listed. This number never changes and can be used to link roster items across rounds. In some cases, it also makes it possible to link people between two different rosters in the same survey. For example, beginning in round 2 the unique ID listed for a child on the biological children roster is the same one assigned to that child on the household roster. Researchers can therefore examine data on both rosters about the same child.

An additional feature of most unique ID numbers is that they incorporate an indicator of the round in which the person or item was first reported. For example, IDs of roster items reported in round 1 may begin with “1” or “97,” while those first reported in round 2 begin with “2” or “98.” (Beginning with round 3, 4-digit years are used so that IDs begin with “1999” rather than just “99.”) UIDs for people on the household roster are constructed in a slightly different manner; researchers should refer to section 4.6.5, “Household Composition,” for more information.

Created Variables

Created variables generally start with “CV_” in the codebook, as in the ‘Hourly Rate of Pay’ example later in this chapter, with a few exceptions. One major exception is the sampling weight variables, which have question names SAMPLING_WEIGHT and CS_SAMPLING_WEIGHT. Other exceptions to note include the validation variables for rounds 4 and 5, which have question name VALIDR_, and the round 5 timing variables with question name R5_TIM. In addition, the family process variables constructed by Child Trends (see sections 4.5 and 4.6) have question names beginning with “FP_” in the codebook. In the Event History data, all variables are created and can be located in the “Event History” area of interest (see section 4.4 of this guide for more information and question names).

3.3 NLSY97 Documentation

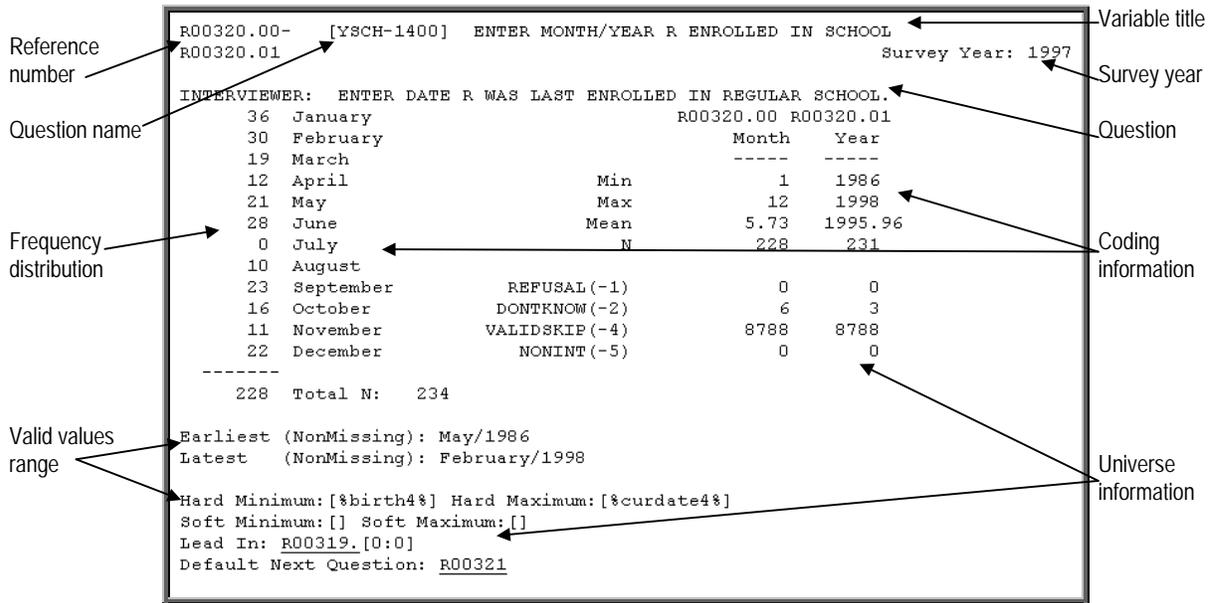
Variables present in the NLSY97 main file are documented via (1) a codebook; (2) accompanying supplemental documentation; and (3) error updates. This section describes the three primary components of the NLSY97 documentation and discusses the important types of information found within each.

Codebook

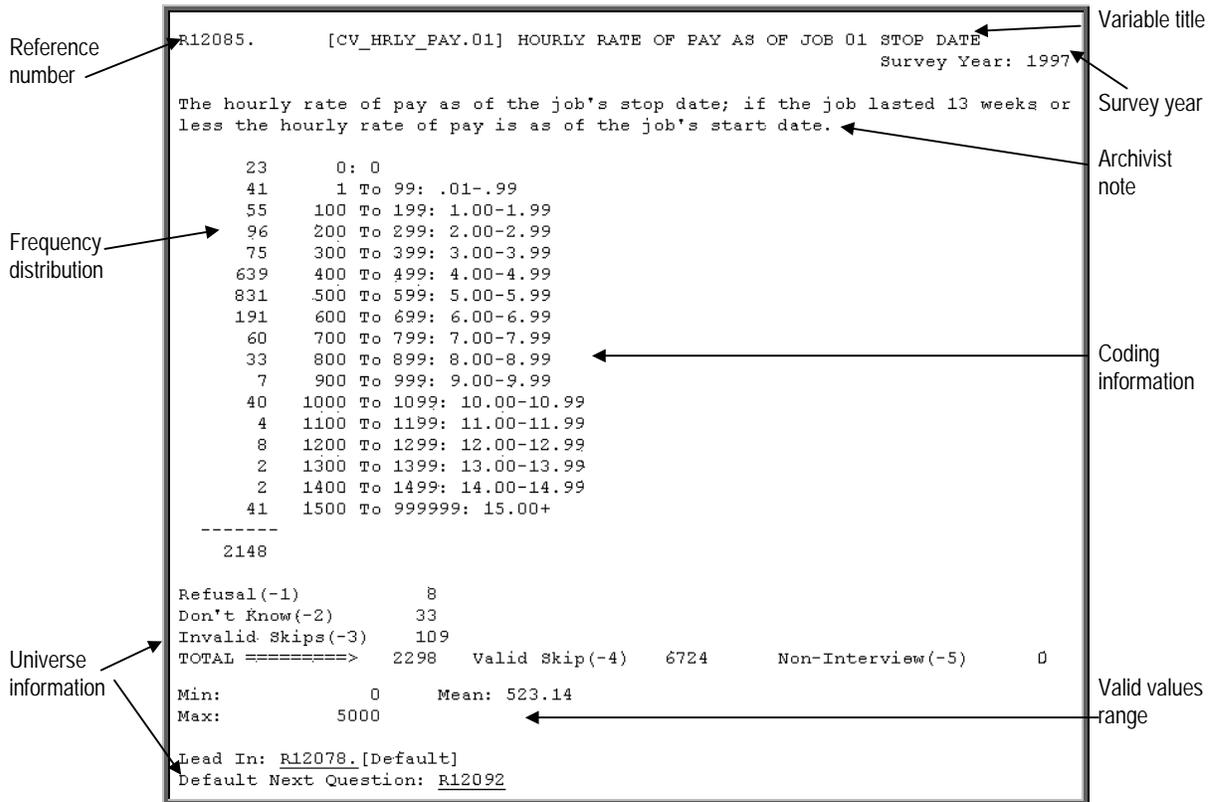
The codebook is the principal element of the NLSY97 documentation system and contains information intended to be complete and self-explanatory for each variable in a data file. The NLS Investigator software allows easy access to each variable’s codebook information and permits the user to print a codebook extract for selected variables.

Every variable is presented as a block of information called a “codeblock.” Sample codeblocks are shown in Figures 1 and 2. Codeblock entries depict the following important information: coding information, frequency distribution, questionnaire items, universe information, valid values range, and question text. Each of the above terms is described more completely in the following pages. Codeblocks for many variables also include special notes designed to assist in the accurate use of data.

3.3 Figure 1. NLSY97 Questionnaire Item Codeblock



3.3 Figure 2. NLSY97 Created Variable Codeblock



Coding Information: Each codeblock entry presents the set of legitimate codes that a variable may assume along with a text entry describing the codes. *Users should note that coding information for a given variable in the NLSY97 codeblock is not necessarily consistent with the codes found within the questionnaire. If the two sources are different, the codebook is current and the questionnaire information should not be used in analysis.* For example, an additional code may be added during data processing if a significant number of respondents gave the same answer to the “other—specify” option in an answer list.

The following types of code entries occur in NLSY97 codeblocks:

Dichotomous (or variables answered yes/no), uniformly coded “Yes” = 1 and “No” = 0. Other dichotomous variables have frequently been reformulated to permit this convention to be followed.

Discrete (Categorical), as in the case of ‘Month Enrolled in School.’

January	May	September
February	June	October
March	July	November
April	August	December

Continuous (Quantitative), as in the case of ‘Hourly Rate of Pay’ in the example above. These variables have continuous data but are presented in the codebook using a convenient frequency distribution. Note that rate of pay variables often have two implied decimal points.

Valid data are generally positive numbers. In a small number of variables, negative responses are possible; users should check the minimum values allowed for each question to clarify whether negative numbers are permissible. The following missing value conventions are used throughout the data:

Noninterview	-5
Valid Skip	-4
Invalid Skip	-3
Don’t Know	-2
Refusal	-1

Frequency Distribution: In the case of discrete (categorical) variables, frequency counts are normally shown in the first column to the left of the code categories. In the case of continuous (quantitative) variables, a distribution of the variable is presented using a convenient class interval. The format of these distributions varies.

Some codeblocks show frequency distributions for more than one variable. For example, the codeblock in Figure 1 above includes responses for both the month and the year that the respondent enrolled in school. Note that this codeblock also lists two reference numbers, R00320.00 and R00320.01. These two variables will be listed separately in the data extraction software, but if either is selected the user will see the codebook page shown above. However, if an extraction is performed, the variable for the month of enrollment will have the data shown in the month distribution, and the variable for the year of enrollment will contain the data shown in the year distribution. These combined codeblocks generally occur for date

variables and for variables that permit multiple responses to a single question (for example, see question YEMP-100300, which asks the respondent to identify all fringe benefits made available by an employer).

Questionnaire Item or Question Name: The question name provides the location of the question in the survey instrument or identifies it as a created variable. In the first example, the question name YSCH-1400 shows that the variable is based on a question in the schooling section of the youth instrument. In the second example, the question name CV_HRLY_PAY.01 indicates that the variable is created. For more information on how question names are assigned, refer to section 3.1, “Survey Instruments & Other Documentation.”

Universe Information: The universe information found in the codebook includes:

- (1) **Universe Totals:** Two totals are presented: (1) The sum of the frequency counts for each coding category is located below the individual codes. (2) The sum of the valid responses plus missing response counts of “refusals,” “don’t knows,” and “invalid skips” can be found in the TOTAL=====> field. The number of respondents who were not asked a question because it did not apply to them—that is, “valid skips (-4)” —is also depicted.
- (2) **Universe Skip Patterns:** The following detailed universe information will enable users to trace the flow of respondents both backward and forward through the CAPI questionnaire:

“Go to # XXXXX,” appended to certain coding categories, indicates that respondents selecting that answer category were routed to the next question specified.

“Lead In(s) # XXXXX” identifies the question or questions immediately preceding the codeblock question through which the universe of respondents was routed. Each lead-in number is followed by the relevant response value indicators, e.g., (Default), (ALL), [1:1], [1:6], etc.

“Default Next Question” specifies the next question that all respondents to the current question will be asked unless some skip condition indicates otherwise.

Valid Values Range: Depicted below the frequency distribution is information relating to the range of valid values for that particular distribution. “MINIMUM” indicates the smallest recorded value exclusive of skips, refusals, and don’t knows. “MAXIMUM” indicates the largest recorded value. As described below, the computer-assisted interview contains internal range checks that limit responses to those between predesignated values, warn interviewers to verify non-normative values, and bolster the information provided by the traditional minimum and maximum fields.

Minimum and Maximum Fields: The MIN and MAX fields define the range of responses, i.e., the minimum and the maximum values, for a data item. The MAX of 5000 (\$50.00) in the ‘Hourly Rate of Pay’ question means that it was the highest value recorded.

Hardmax and Hardmin Fields: Hardmax and Hardmin fields denote the highest and lowest values that were accepted. Dates, e.g., month/day/year of the respondent’s birth (%birth4%) and current interview

(%curdate4%), are often used as Hardmin and Hardmax values in order to restrict responses to certain questions to values within that range, as in the ‘Enter Month/Year R Last Enrolled in School’ example. Responses outside this range must be entered by the interviewer in the comment field; valid numbers are included in the data.

Softmax and Softmin Fields: Softmax and Softmin fields cover ranges where an answer may exceed normal limits yet remain within absolute limits; such answers are accepted after verification. A Softmax set to \$80,000 on an income question will trigger an alert to interviewers that a higher value is unusual.

Income Values: Confidentiality issues restrict release of all income and asset values. To insure respondent confidentiality, the top 2 percent of reported values for many income or asset variables are all converted to one set value. This “topcoded” value is calculated separately for each variable by averaging all the values which exceed the limit for that variable. Calculating topcode values in this way allows statistics such as means to accurately reflect the status of the population under examination without violating respondent privacy.

Verbatim: When an NLSY97 variable is taken directly from the questionnaire, the verbatim of the question or the instructions to interviewers appear beneath the variable title. If a single question is the source for more than one variable, the first variable contains the verbatim, while subsequent variables prompt the user to refer to the variable containing the verbatim.

Archivist information, notes, etc.: Some variables include additional information for users regarding inconsistencies in the data, methods of variable derivation, references to supplemental documentation, and so on. These notes generally appear beneath the variable title or question verbatim.

Supplemental Documentation

Purchasers of the NLSY97 data set must have access to all relevant documentation. Documentation for the NLSY97 includes the following items:

Technical Sampling Report—Youth Survey: This technical manual published by NORC describes the procedures used to select the youth sample. The manual includes weights and standard errors for the initial survey year.

Interviewer Reference Manuals: Accompanying each NLSY97 questionnaire will be an *Interviewer Reference Manual*. In a CAPI survey, interviewers have ready access to general and specific instructions that guide them in the administration of the electronic questionnaire. These “help screens” are physically linked to the appropriate questions throughout the instrument and can be accessed electronically. The

Interviewer Reference Manual reproduces the help screens so that researchers can view the various definitions and other pieces of information used during the interview.

Codebook Supplements: Variable creation procedures and supplemental coding information are provided within the *Codebook Supplement*. This information is **not** available in the hard copy NLSY97 codebooks. The attachments and appendices in the following list can be found in the *NLSY97 Codebook Supplement*.

Attachment:

1. **1990 Census Industrial and Occupational Classification Codes.** This document lists the 3-digit 1990 Census codes used to classify job and training information (Census Bureau, 1990 Census of Population Alphabetical Index of Industries and Occupations, Washington, DC: U.S. Government Printing Office, 1991).

Appendices:

1. **Education Variable Creation.** This document provides the programs for several created variables related to education. These include, among others, enrollment status, type of school, date received diploma, highest grade completed, number of schools attended, and math *PIAT* score.
2. **Employment Variable Creation.** This appendix provides programs for created employment variables, including hourly rate of pay, hourly monetary compensation, number of weeks worked, total tenure at job, and number of jobs held.
3. **Family Background Variable Creation.** This appendix of created variable programs contains those dealing with family background, such as household size, marital status, fertility and child status, marriage and cohabitation history, and citizenship status.
4. **Geographic Variable Creation.** Several variables in the main data set provide information about the respondent's area of residence, permitting researchers to identify key characteristics of the area without needing access to the Geocode CD. Included in this appendix is a summary of the four Census geographic regions, an explanation of the MSA/central city status variable, and the definition for the rural vs. urban variable.
5. **Income and Assets Variable Creation.** This document provides the creation procedures for income and assets created variables. These include household net worth and gross household income, as well as receipt of public assistance.
6. **Event History Creation and Documentation.** This appendix explains the structure of the event history variables and describes the creation process.
7. **Continuous Month Scheme and Crosswalk.** This document explains the structure of the event history month-by-month and week-by-week status arrays and provides crosswalks from continuous month/week numbers to actual month and year dates.
8. **Instrument Rosters.** A number of rosters are used to organize information during various parts of the interview. This appendix identifies these rosters and shows how they were used in different parts of the survey. It also lists the variable names, titles, and reference numbers for the various instrument rosters used in each interview.
9. **Family Process and Adolescent Outcome Measures.** This document, which is provided separately from the *Codebook Supplement*, summarizes the creation procedures for the various scales and indexes created by Child Trends, Inc. The appendix also presents the results of Child Trends' statistical analyses of the scales, indexes, and a number of related attitude and behavior variables.

Geocode Codebook Supplements: Supplemental coding information specific to the Geocode CD is provided within the *Geocode Codebook Supplement*. Information provided within this document is **not** available in the hard copy NLSY97 codebooks.

Attachments:

100. **1990 Census Bureau State and County Codes.** This attachment provides coding information for the state and county variables on the NLSY97 Geocode CD. These variables use the current Federal Information Processing Standards (FIPS) codes.
101. **MSA Codes.** This document lists the Metropolitan Statistical Area (MSA) coding scheme used for NLSY97 geocode variables. It also presents Consolidated Metropolitan Statistical Area (CMSA) codes, New England Consolidated Metropolitan Area (NECMA) codes, and Primary Metropolitan Statistical Area (PMSA) codes.
102. **IPEDS Data and College Identification Codes.** This attachment explains the Integrated Postsecondary Education Data System (IPEDS), and how this and other codes are used to identify the colleges reported by NLSY97 respondents.

Error Updates

Prior to working with an NLSY97 data file, users should make every effort to acquire current information on data or documentation errors. A variety of methods are used to notify users of errors in the data files or documentation and to provide corrected information for those persons who acquired an NLSY97 data set from the Center for Human Resource Research. The most up-to-date list of errors is found on the internet by linking to each cohort's page through the <<http://www.bls.gov/nls>> Web site. Errors discovered after the release of a data file are distributed in hard copy form to current data purchasers along with the data set. Error notices and information on how to acquire the corrected data or documentation also appear in *NLS News*, the quarterly NLS newsletter, available online at <<http://www.bls.gov/nls/nlsnews.htm>>.

3.4 NLS Investigator Search Functions

NLSY97 variables can be accessed via areas of interest or through a search of variable titles for any word. Both search functions provide users with bridging information to the codebook and survey instruments.

Areas of interest. NLSY97 data files are organized so that variables sharing a common factor are stored in unique groupings called “areas of interest.” Users can browse through a given area and examine the variables associated with that topic. NLSY97 areas of interest are listed in Appendix A of this guide.

Any word search. All words, numbers, and symbols found in any variable title form an index in the data set. The “Any Word in Context” search function allows the user to search this index and select NLSY97 variables whose titles contain any single word or combination of words.

User Notes: Any word in context searches for NLSY97 variables are limited by the choice of variable titles. Flexibility in variable title assignment for raw data items is restricted by the wording of the question as it appears in the survey instrument and the maximum allowable length for variable titles.

Chapter 4: Topical Guide to the NLSY97

This chapter summarizes the content of the NLSY97 survey. It presents summary discussions of sets of variables, arranged alphabetically by topic (e.g., “Employment,” then “Event History”). Each topical section begins with a page summarizing the information found within various subtopics in that section. Tables provide general information about age and other universe restrictions that apply to each subtopic.

Within each topic, the descriptions of the subtopics are also arranged alphabetically (e.g., “Fringe Benefits,” then “Gaps in Employment”). Each subtopic details the NLSY97 variables and refers the user to other related supporting documentation. Persons interested in reviewing, for example, variables containing data on a respondent’s rate of pay should turn to the “Employment” section and examine the subtopic “Wages.” A complete list of topics and subtopics can be found on the table of contents page of this guide.

Each subtopic begins with a summary of the content of the data set. This information is presented separately according to the survey instrument with which it was collected; note that all instruments are not included in every subtopic. For more information about the various instruments, see section 1.4, “Content of the NLSY97.” The survey instruments are identified as follows:

Youth Questionnaire. Information collected from the youth respondent is described in this section.

Parent Questionnaire (round 1). Data provided by the responding parent are described in this section. For information on the selection of the responding parent, see section 2.2, “Interview Methods.”

Screener, Household Roster, and Nonresident Roster Questionnaire (round 1). This section describes the information gathered from the household informant during the initial portion of the first interview.

Household Income Update (rounds 2–5). This section describes the information gathered from one of the youth’s parents regarding the income of that parent and his or her spouse or partner.

Created Variables. A number of variables are created based on information provided by the respondents or obtained from other outside sources. Although the data in created variables are not directly collected during the interview, they are described in the summary of the content for each relevant subtopic. The name of each created variable is provided in the text discussion to aid users in finding it in the data set.

Special Data Collections. The School Surveys, Transcript Surveys, and *CAT-ASVAB* data collections are administered at irregular intervals. Variables based on these special data sources are described in sections 4.1.1, 4.1.2, and 4.2.5 of this guide.

In addition to describing the available data, subtopic sections generally provide the following information:

Comparison to Other NLS Surveys: When similar questions have been asked of one of the other NLS survey groups, the subtopic will contain a paragraph briefly describing the comparable information. *Users should note that these descriptions are only intended to aid the user by pointing out areas of similarity; they do not imply that question wording or universes are identical.* Before using any data for comparison purposes, users should consult the *NLS Handbook*, the appropriate *User's Guide*, and the questionnaires for more exact information.

Survey Instruments: A short paragraph summarizes the location of the questions in the various questionnaires.

Related User's Guide Sections: This list directs users to other sections in this *User's Guide* that may contain information relevant to the current subtopic.

Main Areas of Interest: These entries provide the areas of interest in the data set under which users will find the data described in the subtopic section. For more information on areas of interest, including a list of areas, see section A.2, "About the NLSY97 Data Set," in Appendix A of this guide.

Supplemental Areas of Interest: Data related to the main area of interest may be found in these areas. For example, the subtopic "Assets & Debts" refers users to the "College Experience" area of interest, which contains information on student loans.

Finally, subtopic sections may include "user notes," which contain important information about some aspect of the surveys or data file, and "data hints," which direct researchers to a particular variable or variables that should be used for a specific function.

4.1 Aptitude, Achievement & Intelligence Scores

The NLSY97 collects information about a variety of standardized achievement tests commonly taken by young adults in junior high school and high school; these data are described in the “Achievement Tests” subsection. In round 1, most NLSY97 respondents participated in the administration of two national tests, the computerized version of the *Armed Services Vocational Aptitude Battery (ASVAB)* and the *Interest-Finder*, described in the “Administration of the *CAT-ASVAB*” subsection. Finally, certain NLSY97 respondents take the *Peabody Individual Achievement Test (PIAT) Math Assessment* as a part of the interview. The administration of this assessment is discussed in the final subsection.

Table 1 summarizes the *NLSY97 User’s Guide* subtopics available in this section and any global universe restrictions affecting those subtopics. Users should note that certain questions or groups of questions within a subtopic may have additional restrictions. Please consult the NLSY97 questionnaire and codebook for more information on particular questions.

**4.1 Table 1. Aptitude, Achievement & Intelligence Scores
Subtopics and Universe Restrictions**

<i>NLSY97 User’s Guide</i> Subtopic	Round 1 Universe (Age as of 12/31/96)	Rounds 2–5 Universe (Age as of 12/31/96)
4.1.1 Achievement Tests	all ages	11 th grade or higher
4.1.2 Administration of the <i>CAT-ASVAB</i>	all ages	—
4.1.3 <i>PIAT Math Test</i>	all ages; 9 th grade or lower	age 12; 9 th grade or lower in rd. 1

4.1.1 Achievement Tests

Both the youth and the parent questionnaires have collected information about the youth's performance on various standardized achievement tests. In addition, scores are available for some respondents from the transcript survey, which collects transcripts as respondents leave high school. Note that the transcript data are now provided in both the main data set and the event history data set.

Youth Questionnaire. Achievement test information was collected in round 1 from respondents who were enrolled in or who last attended the 9th grade or higher and from respondents with a GED or a high school diploma. The NLSY97 questioned these respondents about three specific achievement tests—the SAT I, American College Test (ACT), and Advanced Placement (AP) tests. For each test reported, the respondent stated the grade level(s) when it was taken and the highest score earned on each test (SAT I Verbal, SAT I Math, and ACT). Those who took AP exams were also asked for the subjects of the tests (e.g., biology, history, economics).

In rounds 2–5, similar information was gathered from respondents who last attended the 11th grade or higher. For the SAT I and ACT, the survey asked for the grade level(s) when each test was taken and the highest score received. In these surveys, respondents who took AP exams were only asked for the subject areas of these tests and did not report their scores.

Parent Questionnaire (round 1). If the youth had been exclusively home-schooled since the seventh grade, the responding parent was asked about achievement tests taken. These questions included whether the youth had ever taken a standardized achievement test (e.g., SAT I, ACT, or California Achievement Test [CAT]). If the NLSY97 respondent had taken one or more of these tests, the parent reported the year when that youth last took a test with a math component and the highest score he or she ever received on that component. The parent then reported the same information for the last test the youth took with a verbal component.

Transcript Survey. In 1999–2000, transcripts were collected for 1,417 respondents. If these transcripts included scores on achievement tests such as the ACT, PSAT, SAT I, SAT II, or AP tests, those scores are reported in the transcript data file. Transcripts will be collected for additional respondents in future rounds, and their scores will be added to the data set. More information is provided in section 4.2.5, “School & Transcript Surveys.”

Comparison to Other NLS Surveys: Achievement test scores were collected during special transcript or school surveys for the NLSY79, the Young Women, and the Young Men. These scores were reported by the schools or coded from the respondent's transcript, rather than being reported by the respondent as in the NLSY97. NLSY79 data include PSAT, SAT I, or ACT scores for about a quarter of civilian

respondents, as well as scores for smaller numbers of respondents on a variety of other achievement and intelligence tests. Available scores for respondents in the Young Men’s and Young Women’s cohorts are primarily from the *California Test of Maturity* and the *Otis/Beta/Gamma*; a few respondents have SAT I or ACT scores recorded, and a wide variety of other tests are also included. For more precise information, see the appropriate cohort’s *User’s Guide*.

Survey Instruments: Questions on achievement tests are found in the schooling section (YSCH) of the *Youth Questionnaire* and in section PC8 of the round 1 *Parent Questionnaire*. The data from the Transcript Survey may also contain achievement tests scores.

<i>Related User’s Guide Sections</i>	<i>Main Area of Interest</i>	<i>Supplemental Areas of Interest</i>
4.1.2 Administration of the CAT-ASVAB	Achievement Tests	College Experience
4.1.3 PIAT Math Test		Ed. Status & Attainment
4.2.1 College Experience		Expectations
4.2.4 School Experience		School Experience
4.2.5 School & Transcript Surveys		Transcript Survey

4.1.2 Administration of the CAT-ASVAB

From the summer of 1997 through the spring of 1998, most NLSY97 round 1 respondents participated in the administration of the computer-adaptive form of the *Armed Services Vocational Aptitude Battery (CAT-ASVAB)*. In this study, also called the Profile of American Youth 1997, the Department of Defense (DOD) used the NLSY97 sample as part of a larger effort to establish new norms for the *CAT-ASVAB*, a military enlistment test. This group also took a separate assessment called the *Interest-Finder (I-F)*, an occupational interest inventory; DOD will use these scores to create norms for this test.

The *Armed Services Vocational Aptitude Battery* measures the respondent’s knowledge and skills in the topical areas listed in Figure 1. The flexibility of the computer-adaptive version allows the test program to route youths through each topical series of questions according to their responses. If a respondent answers a question correctly, the next question will be more difficult; if the respondent answers the question incorrectly, his or her next question will be comparatively easier. The *CAT-ASVAB* also permitted the administration of an “easy” form of the test to the youngest NLSY97 youths, those born in 1983 and 1984. In the easy form, the first question in each section is easier than the standard form.

4.1.2 Figure 1. Contents of the CAT-ASVAB

Arithmetic reasoning	Electronics information	Numerical operations
Assembling objects	General science	Paragraph comprehension
Auto information	Mathematics knowledge	Shop information
Coding speed	Mechanical comprehension	Word knowledge

Several different types of *ASVAB* scores will be created. First, the data includes final ability estimates that summarize the respondent’s performance on each subtest. Because computer-adaptive testing means that respondents do not answer the same questions, their raw scores cannot be directly compared. To provide a score that can be compared across respondents, DOD created the final ability estimates using Item Response Theory (IRT). These scores are computed on a comparable scale and thus, can be compared across respondents—that is, a lower score indicates poorer performance, and a higher score indicates better performance. More information about the statistical methods used to create these scores is provided in appendix 10 in the *NLSY97 Codebook Supplement*.

Two additional sets of scores are scheduled for release in a future round. The Defense Department will create normed scores for each section of the *ASVAB*. These differ from theta scores because they take into account respondent characteristics such as age. Finally, a composite score derived from select sections of the *CAT-ASVAB* will provide an Armed Forces Qualifications Test score (AFQT), a general measure of trainability and a primary criterion of eligibility for Armed Forces enlistment.

In addition to scores on the tests, the *ASVAB* data available in the current data set include the following sets of variables:

- The number of items answered in each section.
- An estimate of the precision of the ability scores (not included for coding speed and numerical operations).
- Answers to the “online questionnaire,” which respondents answered during the *ASVAB* administration. The questions collected information on the respondents’ school experience, family background, reasons for taking the *ASVAB*, and perceptions of the test-taking environment.
- Question timings for each of the items in the online questionnaire.

Interest-Finder

After completing the *CAT-ASVAB*, respondents were given the *Interest-Finder (I-F)*, an inventory designed to measure the relative strengths of a person’s preferences for six vocational areas: realistic, investigative, social, enterprising, artistic, and conventional. For each vocational area, questions ask about the youth’s preferences for related activities, occupations, or training and his or her past training experiences. The 240 items in this interest inventory are evenly spread over the six areas.

Sample Definition and Testing Procedures

During the NLSY97 screening process, two additional nationally representative samples were identified to complete the *CAT-ASVAB*. The first group, the Student Testing Program (STP), consisted of students, screened during the spring and summer of 1997, who expected to be in the 10th through 12th grades in the fall of 1997. Included were many respondents who also participated in the main NLSY97 survey, as well as youths who refused to participate in or were not eligible for the NLSY97. Results from these tests will be standardized to allow educators to counsel students on their career possibilities based on their *I-F* score.

The second sample, the Enlistment Testing Program (ETP), was a nationally representative sample of youths 18 to 23 years old as of June 1, 1997. This group provided the normative information that will be used by the Department of Defense to determine the score distribution of military-eligible youths and will help to assess the impact of these tests on minority and female military eligibility. The sample for this group, none of whom were eligible for the NLSY97, was drawn during the same screening period.

The administration of the *CAT-ASVAB* was conducted by NORC representatives according to standard *ASVAB* procedure guidelines; respondents were paid \$75 for their participation. Groups of five to ten persons were tested under standardized conditions at more than 280 test sites throughout the United States. Many respondents took the *CAT-ASVAB* at Sylvan Learning Centers; others were administered the test at temporary sites including hotels, community centers, and libraries. A total of 7,127 NLSY97

respondents (or 79.3 percent of the NLSY97 sample) completed this test: 5,452, or 80.8 percent, of the cross-sectional sample and 1,675, or 74.9 percent, of the supplemental sample.

User Notes: Respondents who did not take the *ASVAB* are assigned a –4, valid skip, in the data. For most NLSY97 variables, a valid skip indicates that the respondent was not supposed to be asked a particular question. However, all respondents were eligible for the *ASVAB* administration, and a valid skip for these variables means that the respondent chose not to participate.

Comparison to Other NLS Surveys: The *ASVAB*, in paper-and-pencil form, was administered to NLSY79 respondents in 1980. Scores are available for 93.9 percent of that cohort’s members. For more information, consult the *NLS Handbook* or the *NLSY79 User’s Guide*.

Related User’s Guide Sections

- 4.2.1 College Experience
- 4.2.4 School Experience
- 4.2.5 School & Transcript Surveys

Main Area of Interest

Achievement Tests

Supplemental Areas of Interest

- College Experience
- Common Variables
- School Experience

4.1.3 PIAT Math Test

The *Peabody Individual Achievement Test (PIAT)* measures academic achievement of children ages five and over. It is among the most widely used brief assessments of academic achievement, with demonstrably high test-retest reliability and concurrent validity. One of the *PIAT* subtests, the *Mathematics Assessment*, was given to round 1 respondents not yet enrolled in the 10th grade. In rounds 2 through 5, this test was given to respondents who were age 12 as of December 31, 1996, and who were in the 9th grade or lower in round 1. By comparing the test score on the *PIAT Math Assessment* with information on the respondent's math courses from the *Youth Questionnaire* and the transcript survey, researchers may be able to assess correlations between curriculum and math performance.

The version used is the *PIAT-Revised (PIAT-R)*, published in 1989. Administration of this test is relatively straightforward. Respondents enter the assessment at an age-appropriate item (although this is not essential to the scoring) and establish a “basal” by attaining five consecutive correct responses. A “ceiling” is reached when five of seven items are incorrectly answered. Created variables report the youth's performance as a standard score (CV_PIAT_STANDARD_SCORE) and a percentile (CV_PIAT_PERCENTILE_SCORE); creation of these variables is described in the *NLSY97 Codebook Supplement*. Further information about the *PIAT Math Assessment* is found in Markwardt (1989).

Notes on the PIAT-R Math Test: As with any educational or psychological test of this nature, the *PIAT-R* does have certain limitations. To reiterate some of the cautions stated in Markwardt (1989), users of the data should be careful to avoid the following pitfalls:

- This test is not designed for use as a diagnostic test; due to the nature of the test and the wide range of content, it does not include enough items to measure any specific skill area reliably.
- This test is not designed to provide a highly precise assessment of achievement. Although it is a reliable general measure of achievement in the area of mathematics, it cannot make fine distinctions regarding the absolute attainment level of a respondent.
- The items included in the *PIAT-R Math Assessment* were selected to represent a cross-section of various curricula in use across the United States. Although this allows for meaningful comparisons of an individual's scores and the normative national sample, the user must consider curricular and environmental differences of respondents in interpreting the results.

New norms were released in 1998 (Markwardt). The NLSY97 uses the original and the updated norms for rounds 1–3 and the updated norms for rounds 4 and 5. Researchers should use the variables CV_PIAT_STANDARD_SCORE and CV_PIAT_PERCENTILE_SCORE for original norms and the variables CV_PIAT_STANDARD_UPD and CV_PIAT_PERCENTILE_UPD for the updated norms. This math subtest shows a great degree of improvement in student performance. Average performance has risen at grades 5 through 12, and the performance level of above-average students has increased across

almost the entire grade range (grades 2 through 12). Nevertheless, as on most other subtests, the performance of below-average students has declined at grades 1 through 12. Throughout the NLSY97 rounds, the test content and administration are unchanged; only the scoring has been updated.

Comparison to Other NLS Surveys: The Children of the NLSY79 age five and older have been administered an earlier version of the *PIAT Math Assessment*. See the *NLS Handbook* or the *Child and Young Adult Data Users Guide* for more information.

Survey Instruments: This assessment is given during the PIAT section (YPIA) of the youth interview.

<i>Related User's Guide Sections</i>	<i>Main Areas of Interest</i>	<i>Supplemental Area of Interest</i>
4.1.1 Achievement Tests	Created Variables	School Experience
4.2.5 School & Transcript Surveys	PIAT	

Reference

Markwardt, Jr., Frederick C. *Peabody Individual Achievement Test-Revised*. Circle Pines, MN: American Guidance Service, Inc., 1998.

4.2 Education

Data on respondents' educational activities, attitudes, and status are obtained in questions on schooling. These grade-appropriate questions ask about the respondent's educational experience prior to high school, in high school, and in college. Additional questions focus on school-based training programs in which the respondent participates. Finally, the school and transcript surveys gathered data about the schools in the NLSY97 sample areas and the academic records of NLSY97 respondents.

Table 1 summarizes the *NLSY97 User's Guide* subtopics available in this section and any global universe restrictions affecting those subtopics. Users should note that certain questions or groups of questions within a subtopic may have additional restrictions. Please consult the NLSY97 questionnaire and codebook for more information on particular questions.

User Notes: Before reading the topic-specific sections on education, researchers should understand key differences between the round 1 and subsequent surveys, as well as the structure of the school roster in the data from rounds 2–5. The remainder of this introductory section summarizes this important information.

4.2 Table 1. Education Subtopics and Universe Restrictions

NLSY97 User's Guide Subtopic	Round 1 Universe (Age as of 12/31/96)	Round 2 Universe (Age as of 12/31/96)	Round 3 Universe (Age as of 12/31/96)	Rounds 4-5 Universe (Age as of 12/31/96)
4.2.1 College Experience	all ages; enrolled in college	all ages; enrolled in college	all ages; enrolled in college	all ages; enrolled in college
4.2.2 Educational Status & Attainment	all ages	all ages	all ages	all ages
4.2.3 School-Based Learning Programs	all ages; not yet enrolled in college	all ages; currently enrolled in school or college	all ages; currently enrolled in school or college	all ages; currently enrolled in school or college
4.2.4 School Experience	all ages; enrolled in school	all ages; enrolled in school	all ages; enrolled in school	all ages; enrolled in school
4.2.5 School & Transcript Surveys				
School Surveys	96; all high schools in NLSY97 sample areas	—	2000; all high schools in sample areas and any other attended by a respondent in round 2	—
Transcript Survey	—	—	all respondents graduated from high school or age 18 and not enrolled in high school, and signed permission form	—

Note: See section 4.2.2, "Educational Status & Attainment," for the definition of "enrolled."

Round 1 Questionnaire Administration

In round 1, if the youth respondent had not attended college, the schooling section of the youth questionnaire collected information about the school currently or most recently attended by the youth. Thus, the experience questions described in section 4.2.4, “School Experience,” applied only to one middle or secondary school attended by the youth. If the respondent had attended college, the questionnaire permitted the collection of information about all colleges attended, not just the current or most recent college. However, no respondents had attended more than one college as of the round 1 interview date. Therefore, in effect all respondents provided information only about their current or most recent school, regardless of whether that institution was a K-12 school or a college. In addition, if the respondent’s current or most recent school was a college, or if the respondent had completed grade 12 and was not currently enrolled, the survey asked a set of retrospective questions regarding the respondent’s classes and grades in high school, as well as attainment of a high school diploma or GED. Figure 1 shows the information gathered in the round 1 schooling section.

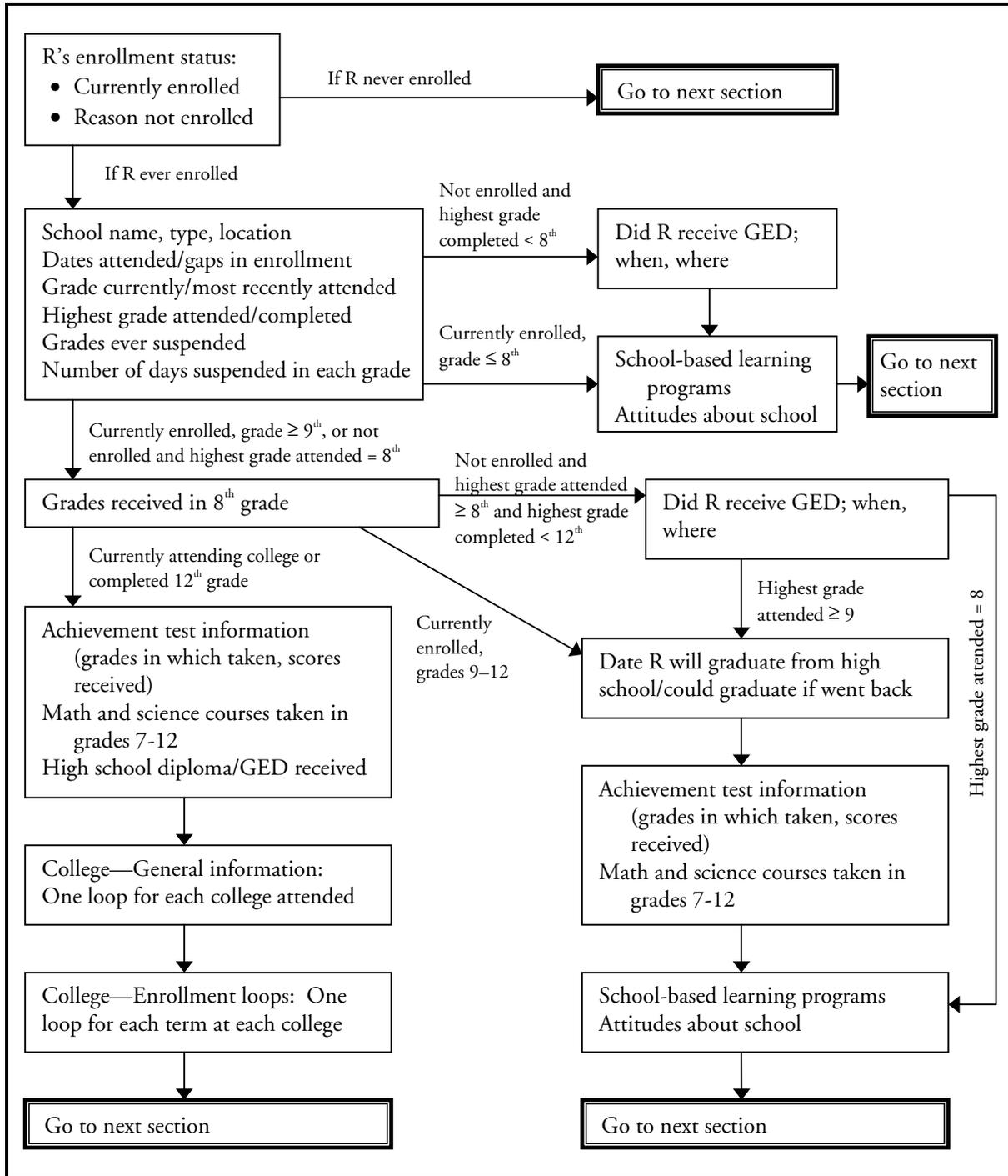
For youths with a parent interview, this school-specific information is supplemented by retrospective data about the youth’s schooling. The parent reports information about all schools attended by the youth since 7th grade and answers questions about gaps in enrollment of one month or more. These data are described in section 4.2.4, “School Experience,” but are not reflected in Figure 1.

Questionnaire Administration in Subsequent Rounds

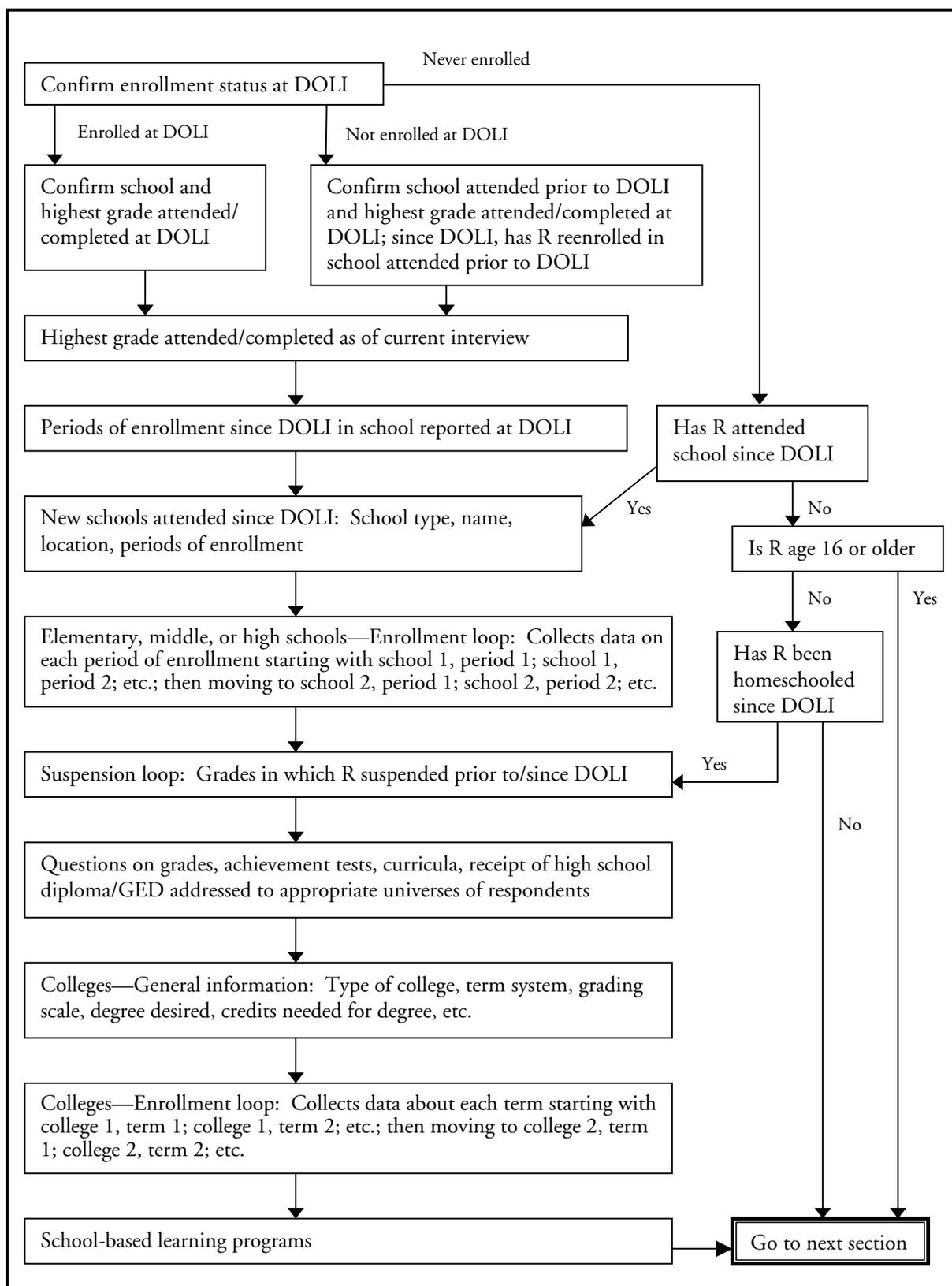
In round 2, the schooling section of the youth questionnaire was redesigned using a retrospective format; the same format was used for rounds 3–5. Rather than asking only about the respondent’s current or most recent school, this section recorded information about all schools the respondent had attended since the last interview. The section first referred to the school (current or most recent) reported by the youth at the time of the last interview and then collected information about all schools attended since that date. The names of these schools and the dates they were attended by the respondent are organized on the NEWSCHOOL roster (school names are not released to the public). This roster also contains variables indicating the reason the respondent left the school; the level of the school (elementary, middle/junior high, high school, college); the type of school (for example, public or private); and the interview during which the school was first reported (that is, in round 1, round 2, etc.). Finally, each school is assigned an identification number (NEWSCHOOL_PUBID) so that users can tell if a respondent attended the same school in more than one round. These variables do not let users identify multiple respondents attending the same school. After the roster was created, the survey proceeded to ask about the respondent’s experiences at each high school or college. More information about the creation and use of rosters in the survey is provided in section 3.2 of this guide.

Figure 2 summarizes the flow of the round 2 schooling section; the section in subsequent rounds is very similar. Note that “DOLI” is an abbreviation for the phrase “date of last interview.”

4.2 Figure 1. Round 1 Schooling Section



4.2 Figure 2. Rounds 2–5 Schooling Section



The NEWSCHOOL roster, a list of the respondent's schools, was a new feature of the round 2 schooling section and was used again in later rounds. For each school attended since the previous interview, this roster contains the dates of up to three periods of enrollment at each school, the level of school attended, and the round in which the school was first reported. More information is provided in chapter 3 about the way rosters are constructed during the interview and how researchers can use rosters in analyses.

Data hint →

The roster is the best source of information for dates of enrollment and school level. Roster items should be used rather than the raw data items whenever possible.

During the interview, at the beginning of the schooling section, the roster is constructed so that the school reported at the last interview date is school #01. If the respondent now reports not attending that school at the date of last interview (that specific school was never attended or the attendance stopped prior to the date of last interview), then the school is deleted from the current round's roster.

The rest of the roster contains data about additional schools attended since the last interview date in the order that the respondent reported them. All of the youth's schools are included on the same list, regardless of whether they are K-12 schools or colleges. Thus, variables for college #05 may actually refer to the first college attended by the youth, not the fifth college attended. During the interview, check items at the beginning of the set of high school questions and the beginning of the set of college questions identify the school level and send the respondent down the correct path for each school attended. For example, the college check item (question YSCH-25144A) will have a value of "1" if the school was a college and a value of "0" otherwise; a value of "1" sends the respondent to the first college question, and a value of "0" skips the respondent past all the college questions for that school. Thus, all respondents with a school in a given roster position will have data for the initial check item in each series of questions, regardless of the school level. Table 2 below demonstrates the implications of the roster numbering system.

Data hint →

Users who wish to examine only K-12 schools or only colleges should use the roster variables called "school code" (question name NEWSCHOOL_SCHCODE.xx) to separate the two levels of institutions.

4.2 Table 2. School Numbers in the NEWSCHOOL Roster

Question topic	School #01	School #02	School #03
Respondent 1: Has attended 1 high school and 2 colleges since the last interview			
School level	High school	College	College
Roster items (Qname NEWS_)	Valid data for "School #01"	Valid data for "School #02"	Valid data for "School #03"
High school questions	Valid data for "School #01"	No data except initial check	No data except initial check
College questions	No data except initial check	Valid data for "College #02" (the youth's 1 st college)	Valid data for "College #03" (the youth's 2 nd college)
Respondent 2: Has attended 1 high school and 1 college since the last interview			
School level	High school	College	—
Roster items (Qname NEWS_)	Valid data for "School #01"	Valid data for "School #02"	—
High school questions	Valid data for "School #01"	No data except initial check	—
College questions	No data except initial check	Valid data for "College #02" (the youth's 1 st college)	—
Respondent 3: Has attended 3 high schools and no colleges since the last interview			
School level	High school	High school	High school
Roster items (Qname NEWS_)	Valid data for "School #01"	Valid data for "School #02"	Valid data for "School #03"
High school questions	Valid data for "School #01"	Valid data for "School #02"	Valid data for "School #03"
College questions	No data except initial check	No data except initial check	No data except initial check
Respondent 4: Has attended 2 colleges and no high schools since the last interview			
School level	College	College	—
Roster items (Qname NEWS_)	Valid data for "School #01"	Valid data for "School #02"	—
High school questions	No data except initial check	No data except initial check	—
College questions	Valid data for "College #01" (the youth's 1 st college)	Valid data for "College #02" (the youth's 2 nd college)	—

4.2.1 College Experience

The NLSY97 contains a number of questions that gather information about the college experiences of respondents. Topics include characteristics of enrollment, such as the degree sought, credits required, and financial aid, and characteristics of the respondent's attendance during each term, including tuition, major, and remedial classes taken.

Youth Questionnaire. For each college attended, the respondent reports the total number of credits he or she earned at that college, as well as the number he or she earned from sources outside that college (e.g., another college, placement test, or college course(s) taken during high school). The respondent then states the type of diploma or degree sought (e.g., associate, vocational, bachelor's) and the total number of credits the college requires to graduate with that degree. The surveys also ask whether each school was public or private; the term system of the school (i.e., semester, quarter, trimester, or other); and the grading scale used to figure students' grade point average (GPA).

After establishing the dates of each term attended, the surveys enter a loop that asks a set of questions about each term. Respondents state the number of course credits taken, the number of credits earned, whether they attended full- or part-time, their GPA, and their primary and secondary majors. The NLSY97 also gathers information on any remedial English, writing, or math classes the respondent may have taken during that term and the total number of hours per week that these classes met.

The survey records information about college costs, although the exact questions have varied. In round 1, respondents first stated the annual tuition and fees for a full-time student. Both in-state and out-of-state rates were reported for publicly supported institutions. Within the term-specific loop, respondents who were attending part-time stated the amount of tuition and fees for that number of credits in that term, and all respondents were asked about the cost of room and board in each term. In rounds 2–5, the annual tuition and fees question was dropped. Questions in the term-specific loops asked respondents to report the amount of tuition and fees for the number of credits taken by the youth. Round 2 also asked about the cost of room and board for each term.

Finally, the respondent is questioned about source(s) of financial aid received during his or her tenure at each college. Non-institutional sources include loans from relatives or friends. Follow-up questions on these sources ask the youth to state the amount he or she was not expected to repay and the amount owed as of the interview date. The youth is also questioned on the amount of financial assistance received from institutional sources such as grants or scholarships, loans, work-study, or employer assistance. Respondents state the amount received from these sources as well as the amount owed to the government from subsidized or other types of loans. Finally, each youth is asked to state the amount that he or she paid toward college expenses using earnings or savings.

In addition, the round 1 survey asked each respondent who reported college enrollment to state the total number of years he or she attended any 2-year colleges and any 4-year colleges and the total number of different colleges attended by the round 1 interview date.

In rounds 2–5, respondents who had earned a degree by the survey date reported the type of degree (e.g., associate, bachelor’s, master’s) and the month and year it was received. Currently enrolled college students also answered the question series on school-to-work programs; see section 4.2.3, “School-Based Learning Programs,” for more information.

User Notes: Researchers should refer to the introduction of the “Education” section for a description of the structure of the schooling section in each round of the survey.

Data hint →

The school roster organizes information about each school in a matrix of data. The roster variables should be used rather than the raw data in analyses. See the introduction to the “Education” section and section 3.2 of this guide for more information on rosters.

Parent Questionnaire (round 1). If any home-schooled youth was age 16 or older at the time the parent was interviewed, the round 1 survey gathered data from the parent on whether the youth had ever enrolled in and received credit for college courses and, if so, the specific years that this occurred.

Created Variables. Two created variables, CV_ASSOC_CREDITS and CV_BA_CREDITS, calculate the fraction of credits toward an associate’s or bachelor’s degree that each respondent attending college has earned as of each round’s interview date. If the respondent has completed an associate’s or bachelor’s degree, the variables CV_AA_DEGREE and CV_BA_DEGREE provide the date that the degree was received in a continuous month format (see section 4.4, “Event History”). Additionally, all colleges are coded with an Integrated Postsecondary Education Data Systems (IPEDS) code to facilitate the extraction of additional data about the institution; these codes are available only on the restricted-use Geocode CD.

Comparison to Other NLS Surveys: NLSY79 respondents have provided information on their major field of study in college, the total number of credits earned at each college attended, and the amounts of educational loans. The NLSY79 Young Adults report their college major, financial aid information, and degree sought, as well as answer a series about their college application process. Young Women and Young Men have stated their major field of study, cost of tuition, and financial aid received. Major field of study and the number of hours and weeks attended since the last interview have been collected from the Mature Women. Older Men simply stated the year they last attended college and the highest degree they received. Consult the *NLS Handbook* or the appropriate cohort’s *User’s Guide* for more information.

Survey Instruments: Questions on college experience are found in the schooling section (YSCH) of the *Youth Questionnaire* and in section PC8 of the round 1 *Parent Questionnaire*.

Related User's Guide Sections

- 4.2.3 School-Based Learning Programs
- 4.8.1 Assets & Debts

Main Areas of Interest

- College Experience
- Created Variables

Supplemental Areas of Interest

- Achievement Tests
- Assets & Debts
- Common Variables
- Ed. Status & Attainment
- School Characteristics
- School Experience

4.2.2 Educational Status & Attainment

Information on NLSY97 respondents' educational experiences is collected in each survey year. This section reviews the primary types of school enrollment, attendance, and attainment data available.

Enrollment

The round 1 *Youth Questionnaire* gathered information on the respondent's enrollment status. Because the fielding period included summer vacation, the survey determined enrollment status through a number of questions. Each respondent was first asked about his or her current enrollment status. Those on summer vacation reported whether (1) they were currently enrolled in summer school, (2) on break from a regular school, or (3) not attending summer school. Respondents who were enrolled at the time of the survey, those on break who reported being enrolled in the spring of 1997, and those who were in summer school and reported being enrolled during the 1996–97 academic school year were classified as “enrolled.” All other respondents were classified as “non-enrolled” (e.g., interviewed during the school year and not enrolled, on break and not enrolled in the spring of 1997, in summer school and not enrolled during the 1996–97 academic year).

In later rounds, the schooling section began by confirming and correcting the enrollment status data collected during the previous round. The survey then proceeded through a series of questions similar to round 1 that determined enrollment status. However, because the field period was during the school year, determination of enrollment status did not need to include consideration of attendance in summer school.

In each survey, those who report a gap in enrollment are questioned about their reason for leaving school and the date at which this separation occurred. All respondents are asked for the name of their current or most recent school and the type of school (e.g., public, parochial, private). The respondent then verifies the location of the primary or secondary school from a list of school names and addresses called the School Finder.¹

Created Variables. A created enrollment status variable summarizes the youth's responses to the questions described above. This single variable, CV_ENROLLSTAT, informs the user about whether the respondent is enrolled; whether enrolled respondents are in grades 1–12 or in a 2- or 4-year college or graduate program; and what level of degree or diploma (or GED certificate) non-enrolled respondents have obtained. A second variable (CV_SCHOOL_TYPE) lists whether the school currently or most recently attended by the respondent (if primary or secondary) is public, private, parochial, or of some other

¹ The School Finder is a record of the names and addresses of primary and secondary schools located in the United States. This information is taken from the “National Education Database,” provided under copyright by Quality Education Data, Inc., and is not released to the public.

type. Additional created variables provide information on the student/teacher ratio, CV_STUDENT_TEACHER_RATIO, and the school size, CV_SCHOOL_SIZE.

Attendance

Youth Questionnaire. If enrolled since date of last interview, respondents are surveyed about gaps within enrollment. These questions ask for gaps since the last interview and the reason respondents left school when each enrollment period ended. Respondents are also asked whether they have ever been suspended from school and the grade level(s) in which this occurred. For each reported incident, the surveys collect information on the duration of the suspension. In rounds 2–5, the survey also asked for retrospective information about any suspensions during the period before the previous interview if the information had not been reported by the respondent in that survey. In round 1, respondents in the 12th grade or lower were also questioned about the number of days they were absent from school during the fall 1996 term.

Parent Questionnaire (round 1). In the round 1 survey, the responding parent answered questions on whether the youth ever missed a month or more of school since the 7th grade (not including summer vacation). For each absence, follow-up questions collected information on the length of the gap, the grade level(s) in which this occurred, and the reason for that absence. Additional information was gathered on whether an NLSY97 youth ever skipped or repeated a grade level. If the responding parent indicated that a youth had either skipped or repeated a grade level, he or she was asked to state the grade level before and after the change. If the parent reported a youth in the 9th grade or higher, he or she was asked whether that youth had ever taken an academic class during a school break. If so, the responding parent was surveyed on the reason for that class.

Created Variables. The total number of grades, through high school, that the youth has ever repeated (CV_GRADES_REPEAT_EVER) and the total number of grades that the youth has ever skipped (CV_GRADES_SKIPPED_EVER) are included in the data file as created variables. Similar information is also available per year using created variables CV_GRADES_REPEAT_YR and CV_GRADES_SKIPPED_YR.

Attainment

The *Youth Questionnaire* asks all respondents to state the highest grade level attended and the highest grade level completed. Beginning in round 2, data from the previous round are confirmed or corrected and information since date of last interview is then collected. If a respondent is currently enrolled in high school, information is collected on the date he or she expects to graduate.

Round 1 respondents who were either enrolled in college, or who were not enrolled but reported their highest grade level attended as 8th grade or higher, were asked whether they had received a high school

diploma. In subsequent rounds, this question was addressed to respondents who listed their highest grade completed as 12th or more or who had attended college. After establishing that a respondent received a high school diploma, both surveys asked for the date the degree was received and the name of the high school where the degree was earned. The respondent then verified the school’s location as provided in the School Finder, which is a list of K-12 schools in the United States and their corresponding addresses. Respondents who reported earning a GED were surveyed on the date and the state where the GED was earned. A follow-up question established the type of program (e.g., job training, adult education, child care program) used to earn the GED. If the respondent reported college enrollment, college-specific information was collected on the type of diploma or degree sought (e.g., associate, bachelor’s). Because some respondents report their GED in the youth training section of the survey, users are encouraged to use the CV_GED variable to obtain more complete information on GED receipt and the date of receipt. Additional questions on training programs that helped the respondent earn a GED are described in section 4.10, “Training.”

Created Variables. Several created variables summarize attainment information. A single variable, CV_HGC_EVER, provides the highest grade completed as of the survey date for each respondent. Similar information is available by year using CV_HGC_YR. CV_HIGHEST_DEGREE_EVER gives the highest degree received as of the survey date. This information is also available by year using CV_HIGHEST_DEGREE_YR. Using a continuous month format, the variables CV_GED and CV_HS_DIPLOMA list the month the GED and high school diploma were received.

User Notes: Researchers should refer to the introduction of the “Education” section for a description of the structure of the schooling section in each round of the survey.

Comparison to Other NLS Surveys: Enrollment status in the past year, current enrollment status, highest grade attended and completed, and degrees and diplomas received have been collected for all cohorts except the Older Men. Respondents in the Older Men cohort provided information on their highest grade completed and highest college degree received. For more details about specific survey years in which these data were gathered, refer to the *NLS Handbook* or the appropriate cohort’s *User’s Guide*.

Survey Instruments: Questions on educational status and attainment are found in the schooling section (YSCH) of the *Youth Questionnaire* and section PC8 of the round 1 *Parent Questionnaire*.

<i>Related User’s Guide Sections</i>	<i>Main Areas of Interest</i>	<i>Supplemental Areas of Interest</i>
4.2.1 College Experience	Created Variables	College Experience
4.2.4 School Experience	Ed. Status & Attainment	Common Variables
4.2.5 School & Transcript Surveys		School Characteristics School Experience

4.4 Event History
4.10 Training

4.2.3 School-Based Learning Programs

Respondents are asked about school programs that help students learn about the world of work. In round 1, all respondents who had attended secondary school reported whether the school had a day when adults came to talk about their jobs. Those who reported a highest grade attended of 9th through 12th grade were then asked about their participation in any school-based learning programs.

These questions were repeated in subsequent rounds, although the universes were slightly different. Respondents currently enrolled in high school first answered the question about adult speakers on a career day. Then, all respondents enrolled at the date of interview at any level were asked about their participation in school-based learning programs. Table 1 lists the types of school-based learning programs included in the NLSY97, along with a short description of each.

4.2.3 Table 1. Definitions of School-Based Learning Programs

Program	Definition
Career Major	A coherent sequence of courses based upon an occupational goal.
Cooperative Education	Students alternate or parallel their academic and vocational studies with a job in a related field.
Internship/ Apprenticeship	Students work for an employer for a short time to learn about a particular industry or occupation.
Job Shadowing	A student follows an employee for one or more days to learn about an occupation or industry.
Mentoring	A student is paired with an employee who assesses his or her performance over a period of time, during which the employee helps the student master certain skills and knowledge.
School-Sponsored Enterprise	The production of goods or services by students for sale or use by others. Enterprises typically involve students in the management of a project.
Tech-Prep	A planned program of study with a defined career focus that links secondary and post-secondary education.

The surveys ask for the characteristics of the most recent program, including the type of program, the number of days or weeks spent at a work site, and the number of hours per day or week spent at the work site. If the respondent was paid for participating in the program, he or she is asked to give the rate of pay. Other questions include whether the respondent took any classes at the work site and whether the employer wrote an evaluation of him or her.

Survey Instruments: Data on school-based learning programs are collected in the schooling section (YSCH) of the *Youth Questionnaire*.

Related User's Guide Sections

4.2.5 School & Transcript
Surveys
4.10 Training

Main Area of Interest

School-Based Learning

Supplemental Areas of Interest

Jobs & Employers
School Experience

4.2.4 School Experience

The NLSY97 surveys gather detailed high school data, including each respondent's course of study, classes, and grades. Created variables, as well as information from the round 1 parent survey, supplement this collection. In addition, round 1 respondents enrolled in grade 12 or lower during the fall of 1996 were asked a series of questions about their experiences at school during the fall of 1996. These questions included the number of times they had something of value stolen from them at school, someone threatened to hurt them, they got into a physical fight at school, or they were late for school without an excuse.

High School Experience

Youth Questionnaire. For respondents who report attending the 9th grade or higher, data are collected on the course of study in high school (e.g., college prep, vocational technical). Respondents are also surveyed on the types of math, science and other (e.g., computer programming, word processing, home economics) courses they have taken, or are currently taking, in the 7th through the 12th grades. In round 1, respondents were also asked whether each math and/or science course they reported was an honors course.

The NLSY97 also questions respondents who attended the 9th grade or higher, or who are no longer enrolled and last attended the 8th grade, on the overall marks they received in the 8th grade (e.g., mostly As, about half As and Bs). Those who are enrolled in college or who are no longer enrolled in a regular school are surveyed on the overall marks they received from the 9th to the 12th grades.

Parent Questionnaire (round 1). During round 1, the responding parent was surveyed on the schools that the youth attended since the 7th grade (including home schooling) and gaps in enrollment of one month or more. Additional educational information included the youth's enrollment in Head Start.

Created Variables. The created variable CV_SCH_ATTEND_YR summarizes the total number of schools the respondent has attended from the 7th through 12th grades as of each round's survey date. An additional created variable (CV_SCHOOL_TYPE) determines whether the respondent's current or most recent school is public, private, parochial, or some other type of school.

User Notes: Researchers should refer to the introduction of the “Education” section for a description of the structure of the schooling section in each round of the survey.

In round 1, the CV_SCH_ATTEND_EVER variable was created using information from the parent questionnaire only. Youths whose parents reported the youth's current grade as sixth or lower are coded as 0; youths with no parent interview are coded as -4. In round 2, the CV_SCH_ATTEND_EVER variable was created using information from the round 2 youth interview (about the period between round 1 and round 2) combined with the round 1 created variable. Youths with no parent interview and youths whose parents incorrectly reported their grade in school as less than 7 may have an undercount of the number of schools attended, because their information is based on the period since the round 1 interview. Subsequent rounds follow the method of creating the round 2 variable.

Data hint →

The school roster organizes information about each school in a matrix of data. The roster variables should be used rather than the raw data in analyses. See the introduction to the "Education" section and section 3.2 of this guide for more information on rosters.

Comparison to Other NLS Surveys: For the NLSY79, information on grades and courses taken in high school is available through the 1980–83 Transcript Surveys. Respondents have also reported their high school curriculum. Data are recorded for the Children of the NLSY79 on participation in Head Start, grades repeated, and characteristics of the school and curriculum. The Original Cohort respondents provided information on the type of their high school curriculum; however, coding categories for curricula were quite different for these cohorts. Mature and Young Women also listed types of mathematics courses taken in high school. For more information, consult the *NLS Handbook* or the appropriate cohort's *User's Guide*.

Survey Instruments: Questions on school experiences are found in the schooling section (YSCH) of the *Youth Questionnaire* and section PC8 of the round 1 *Parent Questionnaire*.

Related User's Guide Sections

- 4.2.1 College Experience
- 4.2.5 School & Transcript Surveys

Main Areas of Interest

- Created Variables
- School Experience

Supplemental Areas of Interest

- College Experience
- Common Variables
- Ed. Status & Attainment
- Household Characteristics
- Time Use

4.2.5 School & Transcript Surveys

To provide a more complete picture of respondents' educational experiences, data collected during the regular survey have been supplemented with two special surveys. School surveys conducted in 1996–97 and 2000–01 provide data about high schools in the areas where NLSY97 respondents live. Transcripts collected in 1999–2000 give information about the classes taken by NLSY97 youths, as well as other related aspects of their high school experiences. Although much of this information cannot be released in the public use data set due to confidentiality concerns, the data that are available will help researchers to examine the school-to-work transition in greater detail. See the detailed descriptions below for more information about access to these data.

School Surveys

In the winter of 1996–97, all high schools in the 147 NLSY97 primary sampling units with a 12th grade, or nearly 7,400 schools, received the first *School Survey* questionnaire. This survey focused on institutional-level attributes such as school policies and management, as well as student-level experience data. This self-administered instrument asked principals (or their proxies) to provide detailed data on the characteristics of the school, the staff, and the student body. Additional information on the school's general practices, graduation policies, and school-to-work programs was also solicited. Seventy-two percent of the schools responded to either the complete school survey or to a shorter critical items questionnaire.

A second school survey was conducted in the winter of 2000–01. All vocational high schools and high schools with a 12th grade in the 147 NLSY97 primary sampling units were included in the survey. Additionally, if an NLSY97 respondent attended a school during round 2 that was not in one of the sampling areas, the school was included in the sample if it met the grade-level requirements. Over 9,600 total schools received *School Survey* questionnaires. Seventy-one percent of the schools completed a *School Survey* questionnaire, either by mail or with an interviewer by telephone.

Due to “births” and “deaths” of schools between 1996 and 2000 and nonresponse in 1996, not all schools in the primary sampling unit areas are present in the data for both years. The retention rate of 1996 schools into the 2000 survey was 74.2 percent (3,900 of 5,253).

The next paragraphs first describe the data collected in 1996–97 and then discuss the questions asked in 2000–01. The content of the 2000 questionnaire overlapped significantly with the 1996 instrument. However, many questions were reworded slightly, and the revised structure of the questionnaire gave less prominence to school-to-work subjects. To reduce the time burden, questionnaire items from the 1996 instrument were modified to encourage respondents to provide approximate values rather than requiring them to consult administrative records for exact figures.

1996 survey. After identifying the type of school (e.g., comprehensive public, Catholic parish, other religious affiliation) and the grade level(s) served by the school, administrators were questioned about school characteristics. Specifics included the length of the school day, the number of days of instruction per year, and the school's facilities (e.g., a library, a health clinic, a computer center, a drug and alcohol prevention program). The survey also collected detailed information about teachers, such as the proportion of full-time to part-time staff; the proportion with a graduate degree; and the staff's racial, ethnic, and gender composition.

School administrators were next asked about the average daily attendance rate, the total enrollment, and the racial, ethnic, and gender composition of the student body. This section of the survey solicited information on the opportunities available to students, such as the type of courses, the course levels, and the programs offered (e.g., remedial math or English, gifted and talented, job placement, summer school). In addition, detailed data were gathered on the 1996 graduating class (e.g., the percentage of students who took the SAT I/ACT, the average score on the SAT I/ACT). Questions were also asked about the requirements for graduation, such as the total number of credits, the number of credits in a subject, or a standardized test score.

Another section of the school survey gathered information on the number of reported incidents of gang activity, vandalism, weapons possession, and student use of alcohol or drugs on school property. In addition, the administrators were asked to rate the amount of influence wielded by certain groups (e.g., state department of education, various school district officials, principal, teachers, parent association) on issues such as establishing the curriculum, hiring full-time teachers, and spending the school's budget.

Finally, the 1996 survey collected information on the general student requirements for enrollment in a number of school-based learning programs (e.g., career major, cooperative education). Information about the provision of school-based learning programs complements questions in the *Youth Questionnaire* that ask the respondents about their enrollment in these programs. See section 4.2.3, "School-Based Learning Programs," for definitions of the various types of programs.

2000 survey. The second *School Survey* instrument began with questions about the school and its overall offerings: grade levels, school type, tuition, program offerings, computer and other facilities, and sources of federal funding. Faculty-related questions asked about the number of teachers with different education and experience levels, any staff development requirements, basic demographics, and the salary schedule. Questions about the student body captured basic demographics, special education and Limited English Proficiency (LEP) enrollments, and remedial and accelerated program offerings, as well as percentages of students involved in such activities as athletics, truancy, and illegal activities. The instrument also asked about the academic requirements of the high school's curriculum and academic outcomes of its graduates.

The career preparation section began with items about how schools help students prepare for careers, how schools support teachers in professional development, and how enrollments have changed in various career preparation programs since the 1994–95 school year. A battery of questions asked about the existence of, enrollment in, duration of, and targeting of several different career preparation programs: apprenticeship, job shadowing, internships, career major/pathway, career academy, school sponsored enterprise, cooperative education, and tech-prep. Most of these programs are also included in section 4.2.3, “School-Based Learning Programs.”

User Notes: Due to the detailed nature of the information collected in the two school surveys, use of this data set is restricted to researchers meeting confidentiality requirements. Interested users should contact BLS at NLS_Info@bls.gov for more information concerning these procedures.

Transcript Survey

Transcript survey data come directly from high school transcripts secured from youths’ high schools after respondents are no longer enrolled as high school students. In 1999–2000, transcripts were obtained and processed for 1,417 youths who had graduated from high school or had reached age 18 and were no longer attending high school. Additional transcript collections are planned for future rounds as more respondents graduate from high school or become age-eligible. School registrars provided a copy of the transcript that the high school maintained, a course catalog (if available), and indicated whether the student was designated for such programs as bilingual education, special education, or gifted/talented programming.

When the transcripts were received from the schools, survey staff coded the information they contained into a standard format. The resulting created variables comprise a history of the respondent’s terms in school, courses taken, and other academic indicators.

There are several different types of variables in the transcript data file. First, at the school level the variable TRANS_SCH_CAT reports whether a course catalog was received from the school to aid in coding. The highest number of schools reported for any respondent is 12, so this variable is repeated 12 times. This course catalog variable also functions as the identification number of the school. During the coding process described below, each school attended by a respondent was assigned an ID number between 1 and 12, with the school that provided the transcript always listed as school #01. These numbers were used in variables that report which school the respondent attended in each term—for example, if a respondent has a value of 4 for term 1, then he or she attended school #04 in the course catalog variables. This ID number does not link to any variables in the main data file.

Second, the transcript file includes information about the respondents that is not associated with a specific term or course. For example, these variables present test scores on a variety of achievement tests (ACT, PSAT, SAT I, SAT II, AP), information on absences and tardies, the student's school completion status, and dates of enrollment. Variables also indicate whether the respondent participated in programs such as gifted, bilingual, or special education.

A number of variables refer to the respondent's terms of enrollment. For up to 18 terms, these items report the beginning and ending dates of the term, the way in which the school year is divided (such as a season, semester, entire year, or another system), the academic year of the term, the respondent's grade level that term, and the number of credits earned. A variable listing the school the respondent attended during that term can be linked to the course catalog variable as described above.

Finally, the transcript file provides details about each course appearing on a student's high school transcript. Course-specific variables include the course code from the Revised Secondary School Taxonomy (SST-R), the grade earned in the course, and the credit value of the course. Because schools use many different grading systems, the course grades were converted into a standard scale that can be compared across respondents. A series of variables called "Recoding Status of Grade" indicates how the grade earned variable for each course was created. This process is described in more detail in appendix 11 of the *NLSY97 Codebook Supplement*.

Comparison to Other NLS Surveys: School surveys have been conducted for the NLSY79, Children of the NLSY79, the Young Women, and the Young Men. Each of these surveys has recorded information on the school's total enrollment, number of books in the school library, qualifications of the staff, and ethnic/racial composition of the faculty and students. The NLSY79 and Children of the NLSY79 surveys also asked about the school's grading system and average daily attendance.

Transcript surveys have been conducted for the NLSY79 and for the Children of the NLSY79. These surveys have included information on course subject matter, enrollment dates, and grades earned. For more precise details about the content of each survey, users should consult the *NLS Handbook* or the appropriate cohort's *User's Guide*.

Survey Instruments: School-level data were collected from schools using the *School Survey* instruments. Transcript variables were created from the respondents' school transcripts as provided by the individual schools.

Related User's Guide Sections

4.1.1 Achievement Tests

4.2.3 School-Based Learning Programs

Main Area of Interest

Transcript Survey

4.2.4 School Experience

4.3 Employment

NLSY97 respondents answer questions about current and previously held jobs; there is no limit to the number or types of jobs a respondent may report. These data are collected about every employer for whom the respondent worked since the last interview so that a complete picture of the respondent's employment can be constructed. Unlike earlier NLS surveys, the employment section distinguishes between four types of jobs: employee-type jobs, freelance jobs, self-employment, and military service.

User Notes: Most age restrictions in the CPS, employment, and training sections of the questionnaire refer to age as of the survey date, rather than age as of December 31, 1996, as in all other sections. However, the check items that route respondents through the rounds 4 and 5 employee and freelance sections are based on age as of December 31, 1996. Users should carefully examine the questionnaire and documentation to ensure that they have correctly identified the age restrictions for a given set of questions. Note that all respondents were at least 14 years old in round 3, so the age 14 restrictions were dropped beginning with that survey.

Users should also keep in mind that the NLSY97 questions based on the Current Population Survey (CPS) were not included in the rounds 2, 3, and 5 surveys.

Employee Jobs

For each employee-type job—defined as a situation in which the respondent has an ongoing relationship with a specific employer—youths age 14 and older are asked about the job's characteristics as of the time they started that job. The survey solicits similar end-date information for each employee-type job lasting more than 13 weeks. For jobs that end after the respondent's 16th birthday (or for on-going jobs held by youths age 16 or older), detailed questions are asked about the workplace. Additional questions similar to the Current Population Survey (CPS), asked of respondents age 15 or older in the rounds 1 and 4 *Youth Questionnaire*, determined respondents' labor force status in the week before the interview.

Freelance Jobs

Questions specific to freelance employment—that is, jobs for which the respondent performed one or a few tasks for several people without a specific boss, or in which the respondent worked for himself or herself—are unique to the NLSY97. This survey captures many typical youth jobs, such as lawn-mowing and baby-sitting, which are often missing from an employment history. In this section of the survey, respondents age 14 and older are asked about their experiences with freelance jobs. For respondents age

12 or 13, the survey asks these questions about all jobs (without explicitly distinguishing between employee and freelance jobs).

Self-Employment

In rounds 1–3, respondents who are age 16 or older and who usually earn \$200 or more per week at a freelance job were considered self-employed. Additional information was collected about those jobs as part of the freelance section of the survey. Starting in round 4, respondents were routed through different paths based on age. Those born in 1980–82 answered questions about self-employment in the regular employee jobs section of the questionnaire (regardless of amount earned at that job). Younger respondents born in 1983–84 who meet the earnings requirement continued to list self-employment in the freelance section. In addition to answering the freelance or employee jobs questions, these respondents were asked to provide more information about their self-employment. See section 4.3.8, “Self-Employment Characteristics,” for more details.

User Notes: Users should note that these data are employer-based, not job-based. As a result, the information collected reflects the time a respondent spent with an employer (or self-employed) and not changes of responsibilities or jobs during that period.

Table 1 summarizes the *NLSY97 User's Guide* subtopics available in this section and any global restrictions affecting those subtopics. Users should note that certain questions or groups of questions within a subtopic may have additional restrictions. Please consult the NLSY97 questionnaire and codebook for more information on particular questions.

4.3 Table 1. Employment Subtopics and Universe Restrictions

NLSY97 User's Guide Subtopic	Round 1 Universe (Age as of interview date)	Round 2 Universe (Age as of interview date)	Round 3 Universe (Age ¹ as of interview date)	Rounds 4-5 Universe (Age as of interview date)
4.3.1 Employers & Jobs	≥14 ≥16 for military service	≥14 ≥16 for military service	all respondents, except ≥16 for military service	all respondents, except ≥16 for military service
4.3.2 Fringe Benefits	≥16, for each employee job lasting ≥ 13 weeks	≥16, for each employee job lasting ≥ 13 weeks	≥16, for each employee job lasting ≥ 13 weeks	≥16, for each employee job lasting ≥ 13 weeks
4.3.3 Gaps in Employment	≥14	≥14	all respondents	all respondents
4.3.4 Industry	≥14 with employee job or ≥16 and self-employed	≥14 with employee job or ≥16 and self-employed	all resp. with employee job or ≥16 and self-employed	all resp. with employee job or ≥16 and self-employed
4.3.5 Job Search CPS Questions Employee Jobs Questions	≥15 ≥16 with employee job	— ≥16 with employee job	— ≥16 with employee job	all respondents in round 4; not asked in round 5 ≥16 with employee job
4.3.6 Labor Force Status	≥15	—	—	all respondents
4.3.7 Occupation	≥14 with employee job or ≥16 and self-employed	≥14 with employee job or ≥16 and self-employed	all resp. with employee job or ≥16 and self-employed	all resp. with employee job or ≥16 and self-employed
4.3.8 Self-Employment Characteristics	≥16 and earn ≥ \$200/week	≥16 and earn ≥ \$200/week	≥16 and earn ≥ \$200/week	≥16 and earn ≥ \$200/week or born in 1980–82
4.3.9 Tenure	all resp. with freelance job or ≥14 with employee job	all resp. with freelance job or ≥14 with employee job	all respondents	all respondents
4.3.10 Time Spent at Work	all resp. with freelance job or ≥14 with employee job	all resp. with freelance job or ≥14 with employee job	all respondents	all respondents
4.3.11 Wages	all resp. with freelance job or ≥14 with employee job	all resp. with freelance job or ≥14 with employee job	all respondents	all respondents
4.3.12 Work Experience	≥14	≥14	all respondents	all respondents

¹ Note that all respondents were at least 14 years old by the round 3 interview date.

The employment sections of the questionnaire are somewhat complex. Before beginning analysis, researchers must understand the structure of each round's questionnaire, particularly the way in which jobs are classified as employee, freelance, or self-employment. It is important to note that this classification depends in part on the survey round and the respondent's age. In rounds 1 and 2, employee jobs were recorded in the first part of the YEMP section, administered only to respondents age 14 or older as of the interview date. The second part of the YEMP section collected information about freelance jobs of respondents age 14 and older and all jobs of respondents age 12 or 13 (the implicit assumption being that respondents younger than 14 are not likely to hold employee jobs). If the respondent was at least 16 years old and made at least \$200/week in a freelance job, the job was classified as self-employment and an extra series of questions was asked during the freelance section.

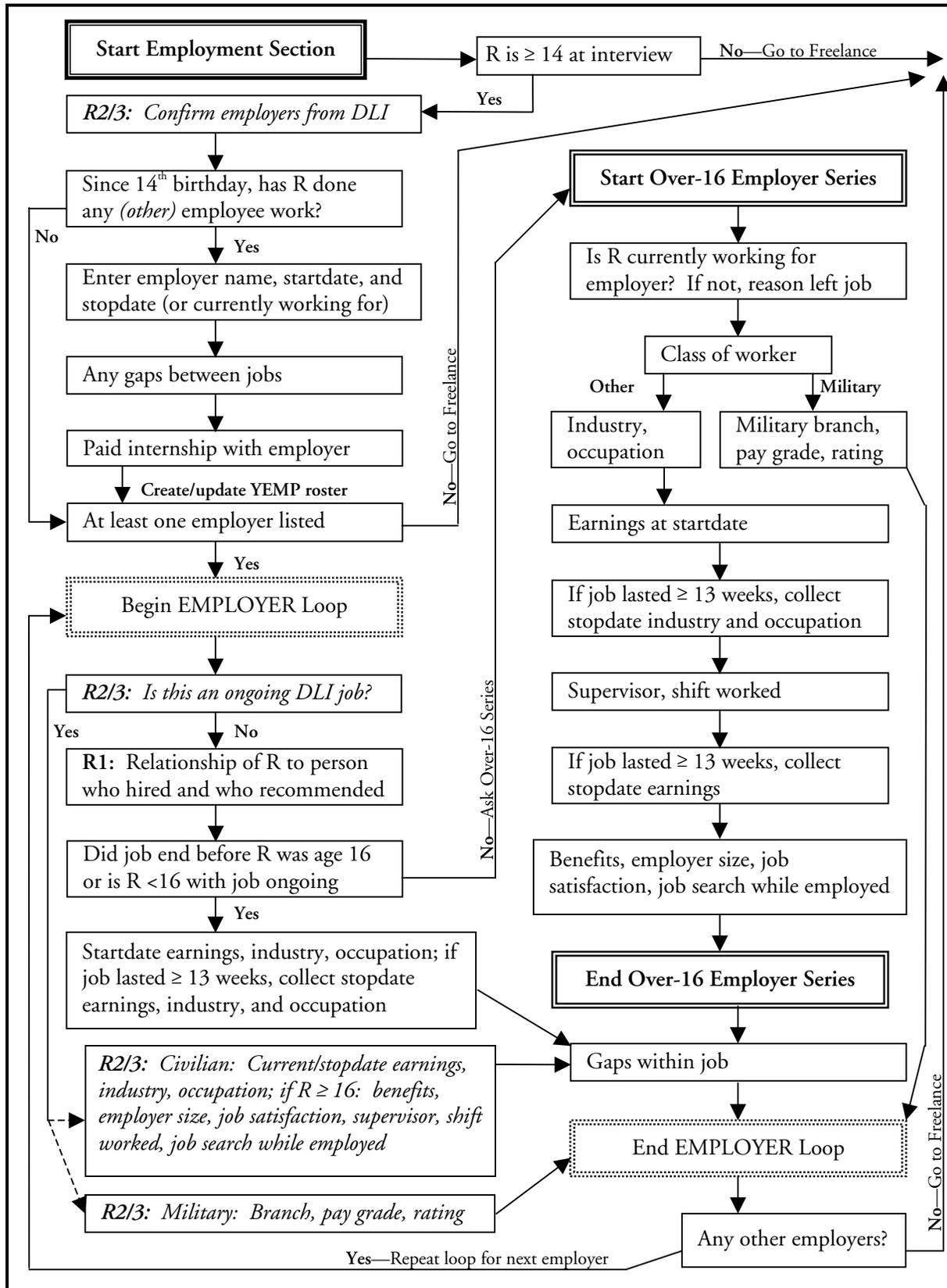
In round 3, all respondents were at least age 14 by the interview date, so the age restriction for employee jobs was no longer necessary. The structure of the section remained largely the same, with a division between employee and freelance jobs. Self-employment was classified in the same way as in the earlier rounds.

In round 4 the section was redesigned. Respondents born in 1980–82 (who were mostly age 18 and older when the round 4 field period began) were asked about employee jobs and self-employment at the same time. In addition, the minimum income requirement from the freelance section no longer applied; jobs could be classified as self-employment regardless of earnings. However, respondents born in 1983–84 (who were mostly age 16 or 17 when the round 4 field period began) continued to describe employee and freelance jobs separately. Data on self-employment jobs were still collected in the freelance section, and freelance jobs still had to meet the income criteria to qualify as self-employment.

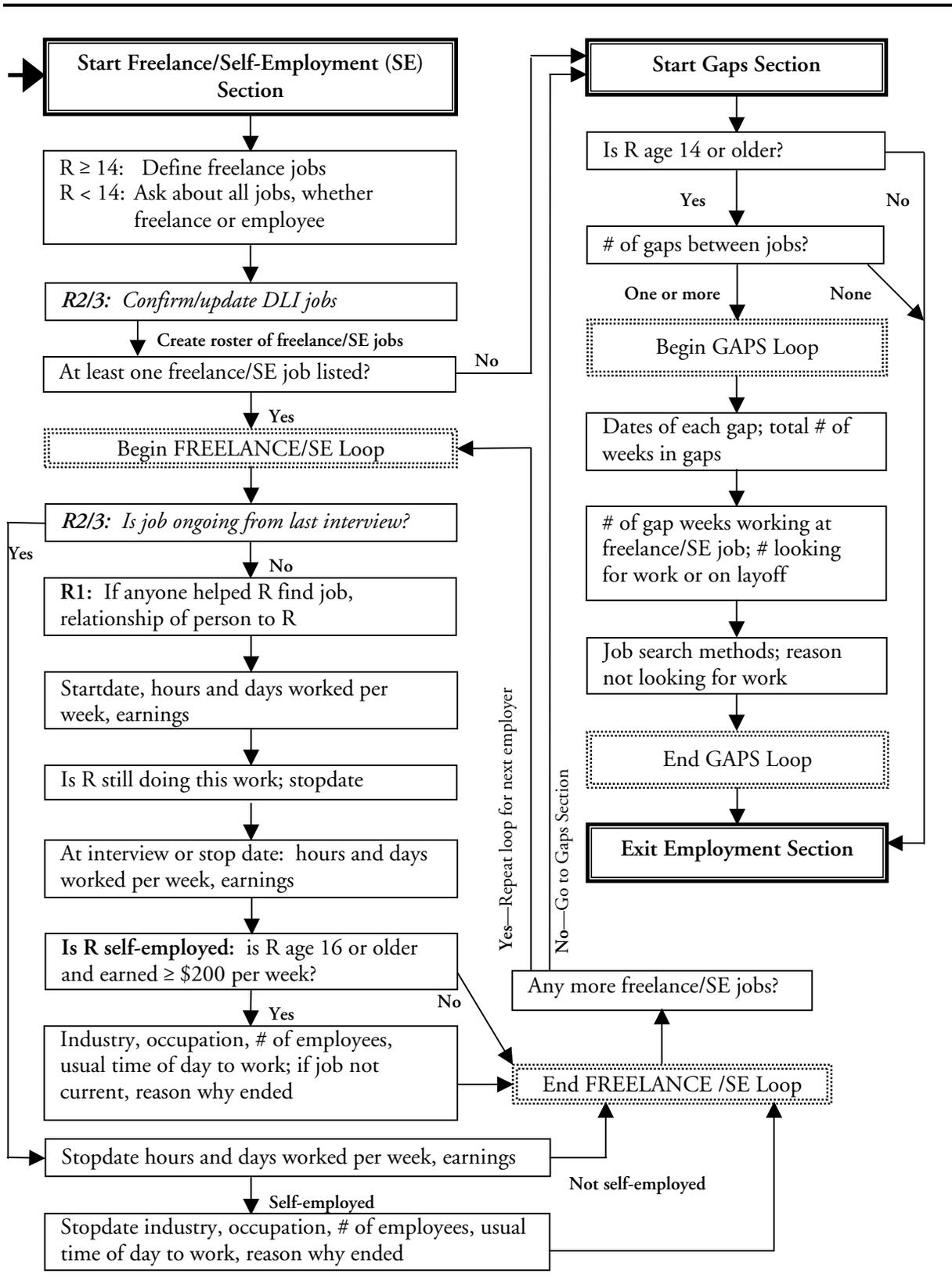
The round 5 employment section followed similar age restrictions and question structure as the round 4 section. For this survey, respondents born in 1980–83 (who were mostly age 18 and older when the round 5 field period began) were asked about employee jobs and self-employment. Respondents born in 1984 (who were younger than age 18 when the field period began) again described employee and freelance jobs separately.

Figures 1 and 2 capture the flow of the employment sections in the various rounds. These figures are intended to picture the major universe restrictions and question topics asked of various groups of respondents about different types of jobs. They are **not** a complete representation of every question in the YEMP section of the questionnaire. Figure 1 focuses on rounds 1–3; users should note that questions asked only in rounds 2 and 3 are indicated in italics. Figure 2 illustrates the change in the structure of the section beginning in round 4, and references to groups of questions that are the same as previous rounds are indicated in italics.

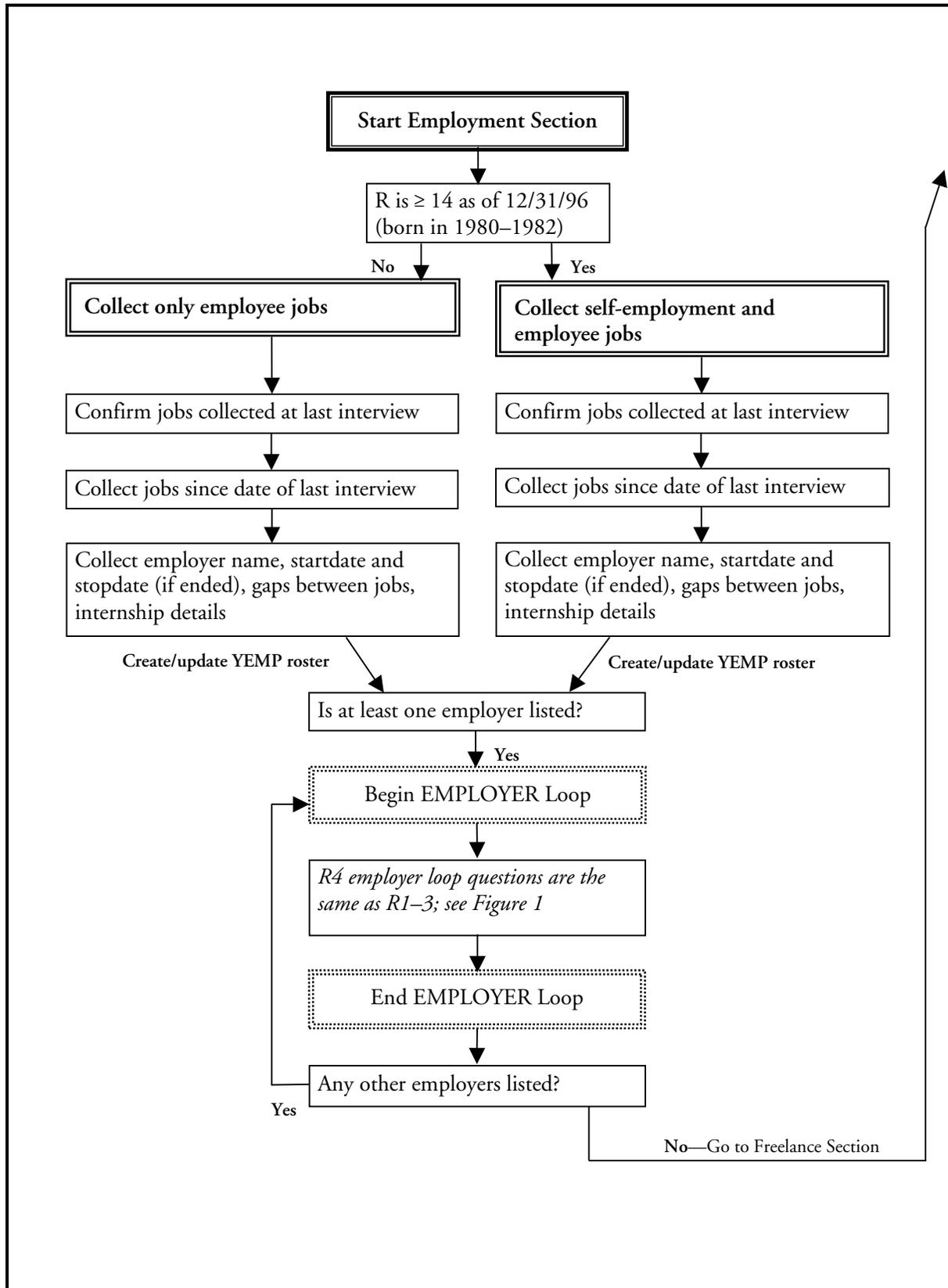
4.3 Figure 1. Rounds 1–3 Employment Flowchart



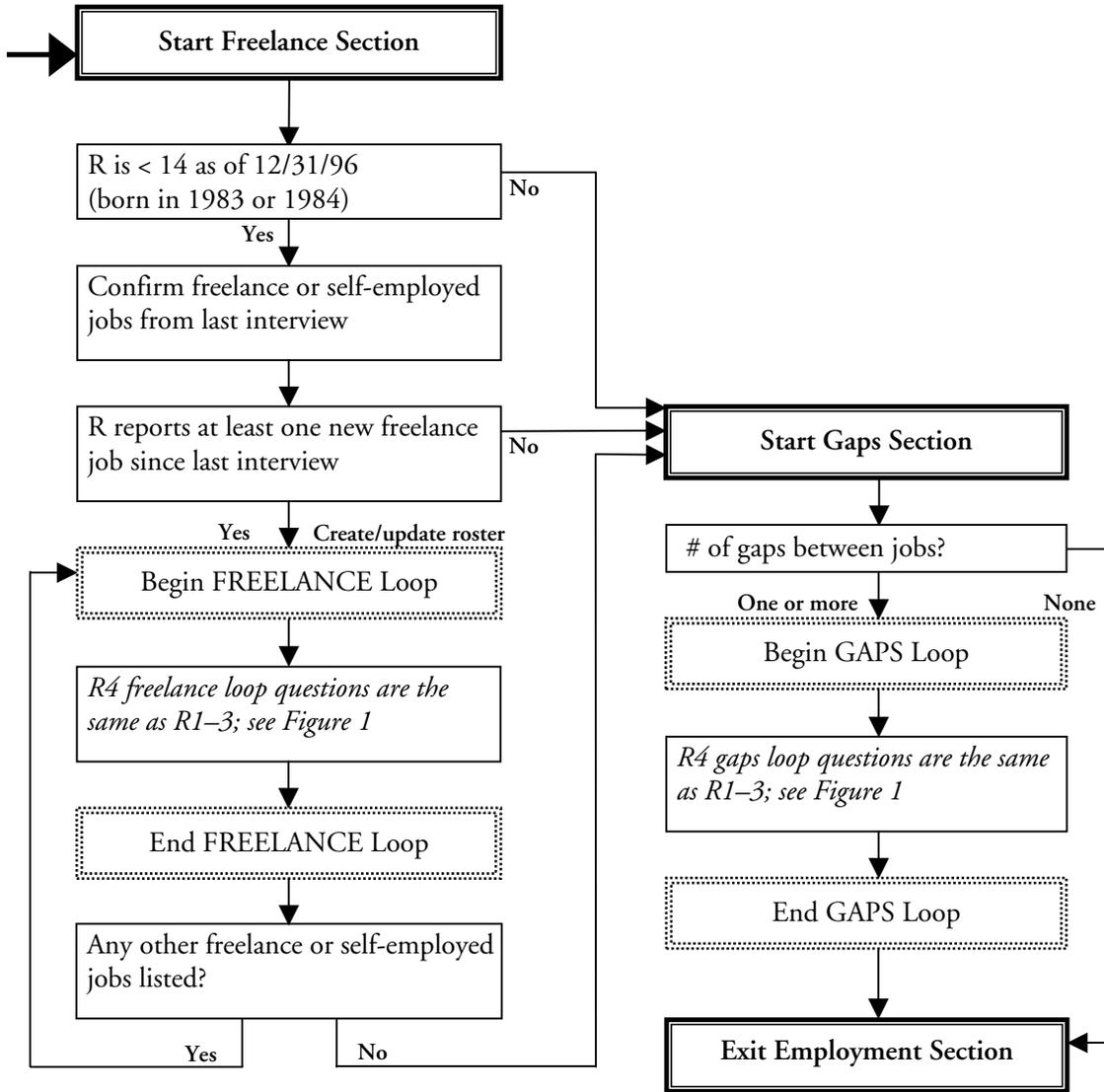
4.3 Figure 1. Rounds 1–3 Employment Flowchart (continued)



4.3 Figure 2. Rounds 4–5 Employment Flowchart



4.3 Figure 2. Rounds 4–5 Employment Flowchart (continued)



User Notes: To associate job information with the correct employer, researchers need to understand how employment information is collected during the interview. The following paragraphs describe how the data are gathered and how employers can be identified in different types of questions and across survey rounds. This discussion of the employer roster provides details specific to this section; general information on rosters is presented in section 3.2 of this guide.

In round 1, any respondent who went through the employee-type jobs section was asked to provide the names of all the employers (including family businesses at which the respondent worked in an unpaid position) for whom he or she had worked since age 14. Then, in the YEMP-1800.xx variables, each employer was assigned a number (e.g., 9701, 9702, and so on through 9707 since the highest number of jobs reported was 7) in the order in which they were reported by the youth. This number is called the unique identification number (UID) for the employer.

After the round 1 employers were assigned a number between 9701 and 9707, the respondent reported the dates he or she started and stopped working for each employer. (These questions are not represented in the data exactly as asked; they are reordered and reported in YEMP_STARTDATE.xx and YEMP_STOPDATE.xx.) At this point, the survey program sorted the jobs by stop date so that the most recent employer was employer #01, the next most recent was employer #02, and so on. Key information about each employer, including the ID number and dates of employment, was organized in the employer roster. Throughout the rest of the employment section, the employer numbers remain constant, so that each variable containing, for example, the phrase “Job #01” or “Employer #01” refers to the same employer for a given respondent. In this case the variables would refer to the first employer on the roster, which is not necessarily the first employer reported by the youth at the beginning of the employment section of the interview.

In rounds 2–5, the employer information was collected in a similar manner. Respondents reported all employers since the last interview date in no particular order. As employers were reported, the CAPI program included a check for whether each employer had been reported in a previous interview. If the respondent reported a new employer, then the YEMP_UID.xx variables contain a new number, as shown in Figure 5. If the employer had been previously reported, the employer kept the same ID number (9701–9707 for round 1 employers, 9801–9809 for round 2 employers, and so on) as it had in previous rounds. This system permits users to link employers across survey rounds, even if there was a break in employment, and to identify the round in which an employer was first reported. After the ID numbers were either continued from a previous round or newly assigned, the roster was sorted according to the stop date of each job. Therefore, employers from different rounds may be mingled on the roster; previous round

employers do **not** necessarily precede current round employers. Note that old employers for whom the respondent has not worked since the last interview do not appear on the current round’s roster.

4.3 Figure 3. NLSY97 Unique ID Numbers

Round	Maximum # Jobs	Unique ID # Range ¹
1	7	9701–9707
2	9	9801–9809
3	9	199901–199909
4	9	200001–200009, 199998, 199999
5	8	200101–200109, 200099

¹ In round 3, the ID number system changed to a 4-digit year.

Data hint →

See the “Self-Employment Characteristics” section of this guide for information on the special case of respondents who had a freelance job classified as self-employment in rounds 1–3 and carried that job over into the regular employee jobs section in round 4. In certain cases, self-employment jobs are assigned a UID of 199999; see section 4.3.8 of this guide for more information.

In addition to retaining the previous ID code to permit linking across rounds, jobs reported at a previous interview retain the start date information from the previous round. For example, if a respondent began a job before the round 1 interview and continued it until the round 2 interview, the round 2 roster will contain the ID code assigned in round 1 and the round 1 start date information. However, all other information in the roster refers only to the time period since the round 1 interview date.

Data hint →

“Employer #01” is **not** necessarily employer number 9701, 9801, 199901, etc. The variables titled YEMP_UID.xx provide a crosswalk between the two systems of identification. For example, if the value of the round 2 variable R24761, ‘YEMP, Employer 01 Unique ID (Ros Item),’ is 9702, then the data regarding employer 9702 from the round 1 interview match with the information reported in the employer #01 variables in round 2.

Treatment of missing values. As mentioned above, the NLSY97 interview collects information from the respondent on the start and stop dates of jobs and the beginning and ending dates of within-job gaps. These dates are transferred onto the individual’s employment roster and additional questions within the survey are asked based on those data. For example, the length of time between jobs is calculated within the CAPI program using the job start and job stop dates, and the respondent is asked follow-up questions about the number of weeks spent actively searching for a job during each gap. If respondents report exact employment dates (e.g., no missing values are reported), the survey program proceeds without any adjustments.

If a respondent does not recall the exact month and day for an employment date, the missing information is imputed and stored in the individual's employment roster. This is done because many questions in the employment section cannot be asked if there is no month and day information, so an imputed month or day is used temporarily so that the section can be completed. For example, if the respondent does not know the start and stop days of the job, "1" is imputed for the start day and "28" for the stop day. Using these temporary days, the survey can ask questions such as those about job search activities during periods of unemployment. As in the case of jobs without missing information, the length of between-job gaps is calculated in the CAPI system using the information in the employer roster. When the respondent's answers include don't know or refuse, the length of between-job gaps is calculated from the imputed dates. Follow-up questions are then asked based on the imputed information.

When the data are being prepared for public release, the original missing values are inserted into the employer roster. At this point the employer roster reflects the actual responses given during the interview and not the temporary imputed values. Therefore, researchers can use the original answers in their analyses. However, they may wish to know what imputed values were substituted so that they can follow the correct question paths and understand the respondent's answers. A complete, detailed explanation of the imputation process is contained in Appendix 6 in the *NLSY97 Codebook Supplement*.

Event history data. The created event history variables (see section 4.4) can be used in conjunction with the main file information about the respondent's employment. Like the main file variables, the event history variables use two systems of identification for a respondent's employers. First, the event history variables contained in the week-by-week status (e.g., EMP_STATUS.01.97, where "01" indicates the first week of the year "97") and dual job (e.g., EMP_DUAL_2.01.97) arrays use the unique ID numbers (UID) for each employer; to associate these employers with characteristic information collected during the interview, researchers must use the YEMP_UID.xx crosswalk variables. A second set of event history variables, those providing start and stop date information (e.g., EMP_START_WEEK.01, EMP_END_WEEK.01, where "01" indicates job #01), use the employer roster line numbers to identify the jobs. The number in the title of these variables refers to the same job as the variables in the main data set with the same number, so users can compare all information about job #02, for example, without any additional ID variables. However, to compare event history start and stop date information about job #02, for example, with information in the event history week-by-week status arrays, researchers must first use the YEMP_UID.xx crosswalk variables to identify the employer ID (9701–9707, 9801–9809, etc.) that matches job #02. See the example below to understand how this process works.

User Notes: The following example illustrates the structure and use of the employer roster. Most aspects of this example apply to other NLSY97 rosters as well. More general information about rosters is available in section 3.2.

Employer Roster Creation in Round 1

Raw data collection: The round 1 survey asked for the names of all employers for whom the respondent had worked since age 14. Assume that a respondent named Emma reported delivering the *Smalltown Press* when she was 14, then switching companies and delivering the *County Register*, and finally working in her parents’ business, Peel’s Corner Store, at the time of the round 1 interview. For this example, the newspaper delivery jobs are assumed to be employee jobs and not freelance-type work. The survey then assigned a unique identification number (UID) in the order the jobs were reported: 9701 for the *Smalltown Press*, 9702 for the *County Register*, and 9703 for Peel’s Store.

Roster creation and roster sort: After the UIDs were assigned, Emma reported the dates she started and stopped working for each employer. At this point, the survey program sorted the jobs according to stop date, so that the most recent employer was employer #01, the next most recent was employer #02, and so on. Therefore, Peel’s Store (UID 9703) became job #01 on the roster, the *County Register* (UID 9702) was listed as job #02, and the *Smalltown Press* (UID 9701) was listed third. Key information about each employer, including the unique ID number and dates of employment, was organized in the employer roster. All of the information about Peel’s Store is located in variables numbered #01 in the title, the *County Register* data are in variables numbered #02, and so on.

Example: Emma’s Round 1 YEMP Roster

Employer	UID	Round 1 Roster Line #
<i>Smalltown Press</i>	9701	03
<i>County Register</i>	9702	02
Peel’s Store	9703	01

Roster use in the interview: Throughout the rest of the employment section, the employer line numbers remain constant, so that each variable containing, for example, the phrase “Job #03” or “Employer #03” refers to Emma’s *Smalltown Press* job. Note that the *Smalltown Press* is **not** the third employer Emma reported at the beginning of the employment section of the interview. It became employer #03 during the roster sort because the other two jobs were more recent.

Roster Creation in Round 2

Data from previous interviews: The employer information was collected in a similar manner in subsequent rounds. Because data were available from the previous interview, they could be used in the construction of the round 2 roster. Before the survey was fielded, survey staff loaded information about each respondent into the interviewers’ laptops. In Emma’s case, part of this information would be the list of employers she reported in round 1.

Raw data collection: During the survey, respondents first provided information about employers who were current at the last interview date. Assume that Emma stated that she worked at Peel’s Store for several months after the round 1 interview. Respondents next reported new employers since the last interview date in no particular order. Emma reported only one additional job, waiting tables at Steed’s Diner after she turned 16. At this point UIDs were given to each employer. Because Peel’s Store was previously reported, it already had a UID—9703—assigned during the last interview. Steed’s Diner was a new employer in round 2, so it was given a UID of 9801.

Roster creation and roster sort: Emma then reported the date she stopped working at each job, and the roster was sorted according to these stop dates. At the round 2 interview, the diner job was more recent, so it was listed as job #01 on the roster, and the store became job #02. At this point, the roster contains information from multiple survey rounds. The UID and start date of the Peel’s Store job are carried over from round 1, while the stop date of the store job and all the information about Steed’s Diner comes from round 2. Because Emma had not worked for the *Smalltown Press* or the *County Register* since the round 1 interview, neither of those employers is listed on the round 2 roster.

Example: Emma’s Round 2 YEMP Roster

Employer	UID	Round 1 Roster Line #	Round 2 Roster Line #
Peel's Store	9703	01	02
Steed's Diner	9801	—	01

Roster use in the interview: Just as in round 1, the employer line numbers remain the same for the rest of the interview. As Emma answered questions about Steed’s Diner, her rate of pay, hours worked, etc., were recorded in the “Employer #01” questions. Peel’s Store data were recorded in the “Employer #2” series.

Roster Creation in Round 3 and Subsequent Surveys

Data from previous interviews: This was collected as it was in round 2. Information reported in previous rounds, including the list of employers previously reported, was loaded into interviewers’ laptops before the survey was fielded.

Raw data collection: During her round 3 interview, Emma reported her ongoing employment at Steed’s Diner, where she had been working in round 2. In addition, she went back to work at the *Smalltown Press* for a 6-month period in between interviews. Both employers retained their original UID numbers, 9801 for the diner and 9701 for the newspaper, despite the break in Emma’s employment at the latter.

Roster creation and roster sort: The roster is again sorted according to the stop date of each job. Since it is a current employer and does not yet have a stop date, Steed’s Diner is listed as job #01 on the round 3 roster. The *Smalltown Press* becomes job #02 since Emma’s employment there had stopped by the date of her round 3 interview. Previous jobs not reported in this round are not listed on this roster.

Example: Emma’s Round 3 YEMP Roster

Employer	UID	Round 2 Roster Line #	Round 3 Roster Line #
Steed’s Diner	9801	01	01
<i>Smalltown Press</i>	9701	—	02

Roster use in the interview: These employer line numbers are in place for the duration of the interview, just as in previous rounds. Information about her employment at the diner is recorded in “Employer #01” questions, while data about the newspaper are recorded in the “Employer #02” series.

Use of the Employer Roster in Analysis

Emma’s information, as organized in the employer rosters, can be used to examine the characteristics of her jobs at the date of each interview or over time. This example focuses primarily on the round 2 employer roster, but subsequent rounds also follow the same sequence for forming the employer roster.

As described above, Emma worked for Peel’s Store and Steed’s Diner during the period between the round 1 and round 2 interviews. Information about these employers was sorted and a roster constructed with the most recent employer appearing first. A researcher using these data would need to be aware of the impact of roster construction.

Because the roster is sorted and employers reported in different rounds may be mixed, variables with “Employer #01” in the title do not necessarily refer to employer number 9701, 9801, etc. The #01 refers solely to the order of the job as listed on the current year’s roster. The unique identification numbers provide a crosswalk between the two systems of identification. The UIDs also allow users to link employers across survey rounds and to identify the round in which an employer was first reported.

For example, Emma’s value for the round 2 variable R24761., “YEMP, Employer 02 Unique ID (Ros Item),” would be 9703—Peel’s Store. The user can identify this as an ID assigned in round 1 because it

starts with “97,” and look at the round 1 UID variables (R05311.–R05317.) to match the employer. In Emma’s case, the comparable variable for employer #01 in round 1 would have UID 9703. Therefore, the researcher knows that information about employer #01 in round 1 refers to the same job as variables about employer #02 in round 2. The variables from the two rounds can then be compared to determine if there were any changes in characteristics such as hours worked, rate of pay, occupation, etc.

The roster line numbers and UID variables in the event history data work in the same way. For example, a researcher might want to know Emma’s employment status in the first and last week of 1998. In the first week of 1998 (variable EMP_STATUS.01.98), Emma was working at her parents’ store, so the status variable would have a value of 9703. Using this UID, researchers can link that job to all of the other information collected during the interview. For example, in the main round 2 data Peel’s Store is job #02, according to variable YEMP_UID.02 (R24761.). Similarly, for the last week of 1998 (variable EMP_STATUS.01.98), when Emma was working at Steed’s Diner, the status variable would have a value of 9801. The job with a UID of 9801 is employer #01 in round 2, so job characteristic data are contained in the employer #01 variables. The second set of event history variables, the start and stop dates of each job, uses the roster line numbers. For these variables, the number in the variable title refers to the same job as in the main data set. For example, the start and stop dates for Peel’s Store in the event history data (variables EMP_START_WEEK.02 and EMP_END_WEEK.02) will also have #02 in the variable title.

4.3.1 Employers & Jobs

In round 1, the NLSY97 collected an employment history for three types of jobs: employee jobs, freelance jobs, and military service (self-employed was collected as part of the freelance section). Subsequent surveys confirm the data gathered in the previous interview and then ask about employers, freelance jobs or self-employment, and military service since the date of last interview. This section highlights key questions asked for each job type; other characteristics of each job are described in the remaining employment subsections (e.g., “Fringe Benefits,” “Industry”).

Data hint →

The employer roster is the only source of information about start and stop dates of employment. The roster also contains flags indicating whether the employer was current at the date of interview, whether the job was in the military, and whether the employer was part of a paid internship experience. The interview questions that collected this information prior to the creation of the roster are not released on the data set but are shown in the questionnaire. Conversely, the roster items appear in the data set but have no questions associated with them in the questionnaire. See the introduction to Employment and section 3.2 for more information on rosters.

Employee Jobs

In rounds 1 and 2, only respondents age 14 and older were eligible to answer the series of questions on employee jobs; in subsequent rounds, all respondents have reached age 14 and are eligible for this section. The surveys collect details on all current and previously held employee jobs at which the respondent has worked since his or her 14th birthday. For each employee job, these respondents are asked about the job’s starting and ending dates and the labor force characteristics (e.g., rate of pay, type of business) of the job at its start date. If the job lasted longer than 13 weeks, the respondent reports this same information as of the job’s stop date (or as of the interview date if the job is on-going). The round 1 survey also asked respondents about their relationship to the person who hired them and to the person who recommended them for the job. There is no limit on the number of employee jobs that the respondent may report.

Additional information is collected from respondents who report a job that ended after their 16th birthday (or, for those age 16 and over, who report an on-going job). The first question determines whether the job is/was in the Armed Forces or a civilian job for a government agency (e.g., local, state, federal), a private or for-profit company, a non-profit organization, or a family business without pay. Respondents in the Armed Forces answer the military service questions described below. The survey then questions civilian respondents about the characteristics of each job and employer. For example, questions are asked about whether the respondent is/was covered by collective bargaining and the gender, race, and age of the respondent’s immediate supervisor. The respondent also reports what type of schedule he or she worked (e.g., regular day shift, split shift, irregular schedule). Other questions in this section gather information about the number of employees working at the same location as the respondent and the number of

employees working for that firm across all locations. One question also gathers the level of job satisfaction for each job (e.g., like it very much, think it is okay). For jobs that have ended, the respondent is asked to state the main reason that he or she left the job.

User Notes: If the job had already been reported during a previous interview, the start date questions were asked at that time. In this case, respondents are asked to report the above information only as it pertains to the stop date or current interview date. However, if the job had been previously reported and the respondent's total tenure was less than 13 weeks, no additional information is collected in the current interview. In this situation the job characteristics data described above are available in the previous round's data.

Some respondents reported new jobs in round 2 that ended before their round 1 interview date. (These jobs should have been reported in round 1 but were overlooked by the respondent.) In these cases no data about job characteristics were collected, but the start and stop dates are still represented in the roster.

Researchers should read the user notes in the introduction to the employment section before using employer characteristic data for analyses. These user notes contain important information about the assignment of employer ID numbers, the creation of the employer roster, and the association of employers across survey years.

Freelance Jobs

In the freelance section of the questionnaire, the survey gathers information from respondents age 12 or 13 (at the survey date) about all jobs they held since age 12. Respondents age 14 and older are questioned on freelance jobs, such as snow shoveling or baby-sitting, they have held since the age of 14. For all jobs reported in this section, the respondent is first asked to state the type of job he or she had and the month and year when it began and ended. Details are then collected about the job's characteristics when it began and ended (or as of the date of the interview if the job is current). These questions ask for the usual number of hours worked per week, the amount earned per week, and the number of weekday versus weekend hours. The round 1 survey also recorded whether the respondent had help in finding the job and the relationship of the person who helped. When these responses are combined with the information collected about employee jobs, the respondent's employment history (from age 14) can be constructed by researchers. See section 4.3.12, "Work Experience," for details on employment history data.

Beginning in round 4, older respondents no longer go through the freelance jobs series. See the following paragraphs on self-employment for details.

User Notes: Due to an error in the way freelance jobs were listed on the roster in round 2, about 150 respondents are missing start date information for a freelance job. In some cases the information is available in the round 1 data for jobs that were previously reported. All other data about the freelance job were collected, and this problem was corrected for round 3.

Self-Employment

In rounds 1–3, a respondent who was age 16 or older and reported earning \$200 a week or more at a freelance job was considered to be self-employed. In addition to the data collected for freelance jobs, self-employed respondents also reported the industry and occupation of the job, the number of employees working for them, their usual time of day to work, and the reason the job ended if it was not current.

Beginning in round 4, the structure of the employment section of the questionnaire was different than previous rounds. Older respondents reported both self-employment jobs and employee jobs in the employer loop and skipped past the freelance section. In these cases, details collected for self-employment jobs are the same as those for employee jobs. In round 4 this group was made up of respondents who were born in 1980–82; the round 5 survey included those born in 1980–83. Younger respondents reported only regular employee jobs in the employer loop and continue to list both freelance and self-employment jobs in the freelance section. This follows the same process as previous rounds; information on self-employment jobs is the same as that collected for freelance jobs. This younger group includes respondents born in 1983–84 for round 4 and those born in 1984 for round 5.

Military Service

Respondents first state in which branch of the Armed Forces they serve and whether they serve in the regular forces, the reserves, or the National Guard. The survey then collects occupational and pay information from respondents age 16 or older who report their employer as an active branch of the Armed Forces; these questions are described in section 4.3.7, “Occupation,” and section 4.3.11, “Wages.”

User Notes: Respondents’ answers to the class of worker question (e.g., YEMP-58500) are used to determine whether the employer is the Armed Forces or a regular employer. For each job, this question asks whether the respondent is employed by the government, employed by a private company, employed by a nonprofit organization, working without pay in a family business or farm, or a member of the Armed Forces. The data indicate that some respondents in the military do not correctly answer “member of the Armed Forces.” If another answer is given, the respondent skips the military questions and enters the regular employer series. In addition, a programming error in rounds 2–4 caused some respondents with military employment ongoing since a prior round to be directed to the regular employer questions rather than the military questions. This error was corrected in round 5.

To account for these problems, researchers can use roster variable YEMP_MILFLAG to identify military jobs included on the regular employer roster. This variable allows researchers to include additional respondents in military analyses or to exclude military respondents from analyses of civilian jobs.

Comparison to Other NLS Surveys: For employee jobs, respondents in each cohort have reported the following information in at least some survey years: start and stop dates, labor force characteristics, class of worker, collective bargaining status, and firm size. Young Men, the NLSY79, and Children of the NLSY79 respondents age 15 and older have provided similar information about military service, including pay and occupational data; Older Men reported the dates of any military service. No information on freelance jobs has been collected from the other NLS cohorts, although job information for self-employed respondents has been gathered as a part of the regular employment section. For further details, consult the *NLS Handbook* or the appropriate cohort's *User's Guide*.

Survey Instruments: The employment section (YEMP) of the *Youth Questionnaire* asks these questions.

<i>Related User's Guide Section</i>	<i>Main Areas of Interest</i>	<i>Supplemental Areas of Interest</i>
4.3.8 Self-Employment Characteristics	Jobs & Employers Military	Common Variables Employment Gaps Labor Force Status School-Based Learning Self-Employment Training

4.3.2 Fringe Benefits

Questions on fringe benefits are only asked of respondents who report an employee job lasting at least 13 weeks that ended after the date of their 16th birthday, or who are age 16 and over and report an on-going employee job at which they have worked at least 13 weeks. For each job meeting the above criteria, the interviewer hands the respondent a card listing typical fringe benefits (see Figure 1) and asks him or her to state the benefits available. In addition to these benefits, the survey questions respondents on the number of paid vacation days and paid sick or personal days per year to which they are currently entitled (or were entitled to when the job ended).

4.3.2 Figure 1. Fringe Benefits Available through Employer

Medical, surgical, or hospitalization insurance which covers injuries or major illnesses off the job	A retirement plan other than Social Security
Life insurance that covers your death for reasons not connected to your job	A flexible work schedule
Dental benefits	Tuition reimbursement for certain types of schooling
Paid maternity or paternity leave	Company provided or subsidized child care
Unpaid maternity or paternity leave which allows you to return to the same job, or one similar to it	Employee stock ownership plans

The intent of these questions is to gather information about any benefits to which respondents know they are entitled. If an employer offers a benefit that the respondent chooses not to take, it is still considered available. Benefits that are available to other employees at the company but not to the respondent at the time of the survey are not included.

Comparison to Other NLS Surveys: Information on fringe benefits has been collected for the NLSY79 in each survey except 1981; for the Mature Women in 1977, 1982, 1987, 1989, and 1995–2001; for the Young Women in 1978 and each survey since 1983; and for the Young Men in 1976 and 1981. Users should note, however, that benefits data were collected only for the CPS job (see the appendix for definition) for the NLSY79 through 1993. Additionally, the exact categories of benefits for which information was recorded may vary; generally, less information was collected in earlier years. Consult the *NLS Handbook* or the appropriate cohort’s *User’s Guide* for more information.

Survey Instruments: These questions are found in the employment section (YEMP) of the *Youth Questionnaire*.

Related User’s Guide Section
4.3.3 Gaps in Employment

Main Area of Interest
Fringe Benefits

Supplemental Areas of Interest
Assets and Debts
Employment Gaps
Fertility and Pregnancy

4.3.3 Gaps in Employment

Respondents age 14 and older are asked about gaps within employee-type jobs and gaps between jobs. These periods when a respondent was not working are not counted in the various created variables summarizing total weeks worked and tenure with an employer. For more information on these variables, see section 4.3.9, “Tenure,” and section 4.3.10, “Time Spent at Work.”

Data hint →

The employer roster is the only source of information about start and stop dates of employment. The roster also contains flags indicating whether the employer was current at the date of interview, whether the job was in the military, and whether the employer was part of a paid internship experience. The interview questions that collected this information prior to the creation of the roster are not released on the data set but are shown in the questionnaire. Conversely, the roster items appear in the data set but have no questions associated with them in the questionnaire.

Gaps within Jobs

Regardless of how long the job lasted, the respondent is asked to report any periods of a week or more within an employee job when he or she did not work for the employer, not including paid vacations or sick days. After a within-job gap is established, follow-up questions probe for the main reason for that gap, such as on strike, on layoff, job ended but began again, or unpaid vacation or leave. Next, the respondent is asked to state the number of weeks that he or she spent looking for work or on layoff during the gap. Any respondent who is classified as not looking for work is asked for the reason (e.g., did not want to work, child care problems, vacation). Finally, female respondents who report a job that ended after their 16th birthday are also questioned on employment gaps due to pregnancy or the birth of a child.

Beginning in round 4, older self-employed respondents (those born in 1980–82 for round 4 and those born in 1980–83 for round 5) answer similar questions about gaps within a self-employed job. However, these questions are in a separate series from the questions for employee-type jobs to allow for more appropriate wording.

User Notes: If a respondent reported an employer in round 2 that he or she also reported in round 1, that employer will retain the round 1 ID code to permit linking across rounds (see the user notes in the introduction to the employment section). In addition, the start date information contained in the round 2 employer roster reflects the original (round 1) start date. However, all other information in the roster refers only to the time period since the round 1 interview date. For example, the round 2 roster only contains information about gaps within a job since the last interview; if researchers are interested in all gaps since the beginning of the job, they will need to examine the information collected in round 1 as well. The rosters in rounds 3–5 are similarly constructed; for information prior to the last interview date the researcher must examine the rosters from the previous round.

Gaps between Jobs

In rounds 1–3, the number of weeks that the respondent did not work at any employee job was calculated. For each gap between jobs, the respondent was asked to state the number of weeks he or she spent working at a freelance job or searching for another employee job. Using this information, the total number of weeks spent not working, not looking for work, or not on layoff is computed for each respondent. Those who did not report search activity are questioned on the reason that they did not look for work during that period (e.g., did not want to work, child care problems, vacation). Data are also collected on the type of search activity in which the respondent participated (e.g., contacted employer directly, contacted an employment agency, placed an ad).

In round 4, respondents born in 1983–84 were asked the same questions as all respondents in rounds 1–3. For respondents born in 1980–82, gaps when the respondent was not working at either an employee job or self-employed were identified. For each gap, the respondent was asked if that time was spent searching for another employee job. (In round 4, youths born in 1980–82 were not asked about working at a freelance job during a gap since they did not report this type of job starting in this round.)

Round 5 followed similar age restrictions and question structure as in round 4. For this round, respondents born in 1984 again answered the questions asked in rounds 1–3. Those born in 1980–83 reported gaps when not working at an employee job or when self-employed. Respondents also reported any job search activity during these gaps.

User Notes: In rounds 1–3, the gaps section asked about periods when a respondent was not working at an employee-type job. If the respondents had any freelance jobs or self-employment, part of the gaps series asked how many gap weeks were spent working at those jobs.

In round 4, respondents born in 1980–82 reported employee-type jobs and self-employment at the same time. Both types of jobs are treated equally in the gaps section—that is, the survey program identifies weeks when the respondent was not working at either an employee or self-employed job and asks about the respondent’s activities in those weeks. In theory, these respondents would have been asked about “periods when you weren’t working” rather than “periods when you weren’t working at an employee-type job.” However, due to a programming error, these older respondents did not go through the alternate series but instead were asked the same series as younger respondents. Even though the dates in the questions asked only about weeks when the respondent was not working at any job, the question text referred specifically to “employee-type jobs.” It is not known whether any respondents were confused by this question wording. It appears that most respondents simply reported their reasons for not working and job search activities for the weeks referred to in the question text, as they would have in the alternate question series. This programming error was corrected in round 5.

Comparison to Other NLS Surveys: The NLSY79 records information about gaps both within and between jobs in all survey years. Data include reasons for the gap, the total number of weeks not working, stop and start dates, and job search activity during the gap. Since 1987, female respondents have been asked about periods of paid leave due to pregnancy or the birth of a child. Similar information has been collected from the Children of the NLSY79 age 15 and older since 1994. The Original Cohort respondents have provided data on gaps within jobs since the inception of the surveys; Mature and Young Women also answered questions on gaps between jobs beginning in 1995. For more information, consult the *NLS Handbook* or the appropriate cohort’s *User’s Guide*.

Survey Instruments: These questions are found in the employment section (YEMP) of the *Youth Questionnaire*.

Related User’s Guide Sections

- 4.3.9 Tenure
- 4.3.12 Work Experience

Main Area of Interest

Employment Gaps

Supplemental Areas of Interest

- Fertility
- Fringe Benefits
- Job Search
- Jobs & Employers
- Tenure w/ Employer

4.3.4 Industry

For both employee and self-employed jobs, NORC personnel code respondents' verbatim descriptors of their business or industry according to the 1990 three-digit census industrial classification system. In addition, NORC classifies freelance jobs that do not qualify as self-employment according to the type of work performed.

Employee Jobs

The NLSY97 asks respondents age 14 or older to report the industry of each employer as of the job's start date. Respondent descriptors of the "kind of business or industry at the job's start date"—or, if the respondent is confused by the question, "what did they make or do where you worked"—are the basis of the industrial codes. For employee jobs lasting more than 13 weeks, respondents also describe the business or industry as of the job's stop date (or at the survey date for on-going jobs). In addition, the interviewer codes whether the business was mainly manufacturing, retail trade, wholesale trade, or something else, requesting respondent help if necessary. Survey staff then code each employer's industry using the 1990 Census industrial codes.

User Notes: If the job has already been reported during a previous interview, the start date questions were asked at that time. Respondents are read a description of the industry they reported as of the last interview date and asked if there was any change between the last interview date and the current interview date (or stop date for jobs that ended). If there is no difference, the job is assigned the same industry code; if the respondent reports a change, a new industry code is assigned for the current round. However, if the job had been previously reported and the respondent's total tenure was less than 13 weeks, no additional information was collected in the current interview. In this situation the relevant data are available in the previous round's data.

Additionally, some respondents reported new jobs in round 2 that ended before the round 1 interview date. (These jobs should have been reported in round 1 but were overlooked by the respondent.) In these cases no industry information was collected.

Table 1 below summarizes the industries reported by youths with employee jobs.

4.3.4 Table 1. Industry of NLSY97 Respondents at Current or Most Recent Employee Job

Industry of Job #01	Agriculture, Forestry & Fisheries	Mining	Construction	Manufacturing	Transportation, Public Utilities, Communications	Wholesale Trade	Retail Trade	Finance, Insurance & Real Estate	Business & Repair Svcs	Personal Services	Entertainment & Recreation Services	Professional & Related Services	Public Administration	Military ¹	Missing Data	Total working at an employee job
Round 1																
Male	126	—	96	103	24	28	517	15	109	46	91	106	4	—	27	1292
Female	26	—	27	30	8	17	489	17	42	71	86	147	12	—	22	994
Total	152	—	123	133	32	45	1006	32	151	117	177	253	16	—	49	2286
Round 2																
Male	134	2	197	146	63	37	1092	41	136	57	189	197	10	1	17	2319
Female	41	—	28	58	21	16	1201	40	67	99	160	307	17	—	23	2078
Total	175	2	225	204	84	53	2293	81	203	156	349	504	27	1	40	4397
Round 3																
Male	149	—	222	191	97	41	1350	44	184	54	210	211	12	3	18	2786
Female	33	1	17	93	40	16	1513	61	96	116	189	415	12	1	26	2629
Total	182	1	239	284	137	57	2863	105	280	170	399	626	24	4	44	5415
Round 4																
Male	134	2	276	230	133	55	1470	40	257	68	210	234	17	—	17	3143
Female	34	—	27	119	64	32	1710	77	111	152	182	536	24	1	21	3090
Total	168	2	303	349	197	87	3180	117	368	220	392	770	41	1	38	6233
Round 5																
Male	123	6	336	255	159	55	1401	53	269	78	199	294	31	41	26	3326
Female	32	1	29	129	66	28	1626	117	131	193	192	698	30	11	23	3306
Total	155	7	365	384	225	83	3027	170	400	271	391	992	61	52	49	6632

Note: This table is based on responses about the respondent's current or most recent industry at his or her current or most recent job (e.g., YEMP-37200-COD.01, YEMP-80200-COD.01). The universe is respondents who reported that their current or most recent job was a civilian employee-type job; military, freelance, and self-employment jobs are not included.

¹ Although this table is based on the industry variables for civilian jobs, some respondents reported their industry as military in the civilian question series. These respondents may actually be civilians, or they may be members of the Armed Forces who reported their class of worker incorrectly (see the introduction to the "Employment" section of this guide).

Freelance Jobs

In rounds 1–3, NLSY97 youths not classified as self-employed describe the type of work they perform during the interview. NORC personnel code these verbatim descriptions for release in the data set. Because the Census codes are not appropriate for this freelance employment, the descriptions are classified using a separate coding frame with categories such as babysitting, mowing, pet care, and snow shoveling. The reporting of freelance jobs was limited in round 4 to respondents born in 1983–84 and in round 5 to those born in 1984. The coding of these jobs was changed beginning in round 3. At that time the

freelance codes were redone for rounds 1 and 2 to match the new coding system, which began in round 3. Both the old and new codes are available for these rounds.

Self-Employment

For all freelance jobs reported in rounds 1–3 in which respondents are considered self-employed (age 16 or older and usually earn \$200 or more per week), the survey asked about the business or industry classification of the job. The series of questions determining the industry was similar to that asked for employee jobs. Survey staff then coded the industry of each self-employed job using the 1990 Census industrial codes.

The structure of the questionnaire changed in round 4, based on the age of the respondent. Older respondents (those born in 1980–82 for round 4 and those born in 1980–83 for round 5) reported self-employment jobs in the employee jobs section, answering industry questions like those listed above for employee jobs. These jobs were coded the same way as employee-type jobs. Younger respondents (those born in 1983–84 for round 4 and those born in 1984 for round 5) continued to list self-employment jobs in the freelance section. If a freelance job met the earnings requirement for self-employment, the industry was coded using the 1990 Census codes.

Comparison to Other NLS Surveys: Industry is collected each year from NLSY79 respondents; however, the jobs must meet minimum hours worked and weeks worked requirements before this information is collected. For the NLSY79, industry is coded using the 1970 (through 1993) and/or 1980 (1982-present) industrial classification codes. Industry is also coded using the 1970 and 1990 (for the CPS job only) codes for Children of the NLSY79 age 15 and older. For the Mature and Young Women, industry has been coded using 1960, 1980, and 1990 systems. The industries of Older and Young Men were recorded using 1960 codes for all years; in the final two Older Men surveys, industry was doublecoded using the 1980 system. For more information, consult the *NLS Handbook* or the appropriate cohort's *User's Guide*.

Survey Instruments: These questions are found in the employment section (YEMP) of the *Youth Questionnaire*.

<i>Related User's Guide Section</i>	<i>Main Area of Interest</i>	<i>Supplemental Area of Interest</i>
4.3.7 Occupation	Industry & Occupation	Training

Reference

Census Bureau. *1990 Census of Population Alphabetical Index of Industries and Occupations*. Washington, D.C.: U.S. Government Printing Office, 1991.

4.3.5 Job Search

Job Search in the Weeks before the Survey

In rounds 1 and 4, questions based on the Current Population Survey (CPS) asked NLSY97 respondents age 15 or older about job search (all respondents had reached age 15 in round 4). Respondents who did not report working in the week before the survey stated whether they had been doing anything to find work during the last four weeks. If so, the interviewer recorded the actions that had been taken (see Figure 1). If the respondent looked for a job, follow-up questions then asked if he or she could have started a job in the past week if offered one and the reason why not, if applicable. A key aspect of the CPS questions is the distinction between active and passive search methods. Active methods could result in a job offer without any further action by the job seeker, such as having an interview with an employer or bidding on a contract. Passive search methods, such as picking up a job application or completing a resume, could not result in a job offer without further action on the part of the job seeker. Official statistics like the CPS consider people who have done at least one active method in the last 4 weeks to be unemployed.

Employee Job Search

If the respondent reports gaps between employee-type jobs since the last interview, he or she is asked whether a job search was conducted during each gap. Note that in rounds 1 and 2 only respondents age 14 or older as of the interview date were eligible for gaps questions and therefore for job search questions; all respondents are age-eligible beginning in round 3. These questions ask respondents for the number of weeks during each gap that they spent looking for work. If a job search was conducted, respondents also state what methods they used in their search, choosing from the list shown in Figure 1. If there are any weeks in the gap when respondents were not looking for work and were not self-employed, they are asked why they were not looking for work (respondents born in 1980–82 were skipped past this question in round 4, as were those born in 1980–83 in round 5).

If a job ended after the respondent’s 16th birthday (or, for respondents age 16 and older, if a job is ongoing), the NLSY97 employment questions in each round ask the respondents if they have done anything to look for work while employed. Those who have are asked to choose from a list the specific actions they have taken (see Figure 1).

4.3.5 Figure 1. Possible Job Search Actions

Contacted employer directly	Placed or answered an ad
Contacted employment agency	Checked union or professional registers
Through friends or relatives	Looked at ads
Contacted a school placement center	Employer contacted you directly
Sent out resumes or filled out applications	Other (Specify)

Comparison to Other NLS Surveys: For the NLSY79, information includes active and passive methods of job search used and the total number of weeks spent actively looking for work. Additionally, CPS questions similar to those described above are regularly asked of respondents in each cohort. For further details, refer to the *NLS Handbook* or the appropriate cohort's *User's Guide*.

Survey Instruments: The questions concerning job search in the week before the survey are found in the CPS (YCPS) section of the rounds 1 and 4 *Youth Questionnaires*. The other employee job search questions are in the employment (YEMP) section of the rounds 2–5 questionnaires.

<i>Related User's Guide Sections</i>	<i>Main Area of Interest</i>	<i>Supplemental Areas of Interest</i>
4.3.3 Gaps in Employment	Job Search	Employment Gaps
4.3.6 Labor Force Status		Labor Force Status

4.3.6 Labor Force Status

Borrowing questions from the Current Population Survey (CPS), rounds 1 and 4 of the NLSY97 asked respondents age 15 and older about their labor force status in the week before the interview (note that all respondents had reached age 15 by round 4). These respondents provided information on the characteristics of their current job (e.g., usual number of hours, type of business). Any respondent who did not report working was questioned on job search activities in the previous month. For those who reported looking for work, data on the type of search activity were also collected. Through these questions, the employment status (e.g., working, looking for work, unable to work) of the respondent can be defined according to the Bureau of Labor Statistics (BLS) definition. BLS uses a similar set of CPS questions to determine the monthly labor force participation and unemployment rates for the nation.

User Notes: “CPS employer” is a term used in the NLSY79 to identify the respondent’s current or most recent employer using methods similar to the Current Population Survey. Researchers sometimes limit their analysis to only the CPS employer instead of considering all employers. Unlike the NLSY79, the NLSY97 does not ask specific questions that refer to the CPS employer; all employers are discussed in the employment section of the questionnaire. However, there is information that can be used to determine which job or jobs are current for a given respondent.

1. The current status flag (YEMP_CURFLAG.xx) reports whether the respondent is currently employed at each of the jobs listed in his or her employer roster.
2. If the respondent is not currently employed at any job, the most recent job can be identified through examination of the stop dates.
3. If the respondent is currently employed at more than one job, the usual hours worked variables can be used to determine which job has more hours. (In the NLSY79, the job with more hours is classified as the CPS job if the respondent has more than one job at the same time.)

Created Variables. Using this CPS information, two variables were created for round 1. The first variable, CV_ESR, provides the respondent’s employment status in the week before the survey. CV_ESR_COLLAPSED, the second created variable, collapses this status into four categories: employed, unemployed, not in the labor force, and in the active Armed Forces.

Comparison to Other NLS Surveys: Current labor force status is determined for all respondents in the NLSY79 and the four Original Cohorts for each survey year and since 1994 for Children of the NLSY79 age 15 and older. However, the current labor force status questions were changed for the NLSY79 in 1994 and for the women’s cohorts in 1995 to reflect the redesign of the CPS. Users should employ caution when comparing labor force status from surveys before these years to the NLSY97. For more information, consult the *NLS Handbook* or the appropriate cohort’s *User’s Guide*.

Survey Instruments: These questions are found in the CPS section (YCPS) of the round 1 *Youth Questionnaire*.

Related User's Guide Section
4.3.5 Job Search

Main Areas of Interest
Created Variables
Labor Force Status

Supplemental Areas of Interest
Job Search
Jobs & Employers

4.3.7 Occupation

For both employee and self-employed jobs, NORC personnel code respondents' verbatim descriptors of their occupation according to the 1990 three-digit census occupational classification system. In addition, NORC classifies freelance jobs that do not qualify as self-employment according to the type of work performed.

Employee Jobs

The NLSY97 asks respondents age 14 or older to report their occupation for each employer. The question "what kind of work did you do" elicits information on the occupation when the job started. A follow-up question asks the respondent to describe his or her usual activities or duties on this job. The occupational classification at the job's end date (or at the survey date for on-going jobs) is solicited for all employee jobs lasting more than 13 weeks. Survey staff then code the respondent's occupation at each job using the 1990 Census occupational codes. Table 1 lists the number of youths with employee jobs in each occupation category.

User Notes: If the job has already been reported during a previous interview, the start date questions were asked at that time. Respondents are read a description of the occupation they reported as of the last interview date and asked if there was any change between the last interview date and the current interview date (or stop date for jobs that ended). If there was no difference, the job is assigned the same occupation code; if the respondent reports a change, a new occupation code is assigned for the current round. However, if the job had been previously reported and the respondent's total tenure was less than 13 weeks, no additional information was collected in the current interview. In this situation the relevant data are available in the previous round's data.

Additionally, some respondents reported new jobs in round 2 that ended before the round 1 interview date. (These jobs should have been reported in round 1 but were overlooked by the respondent.) In these cases no occupation information was collected.

4.3.7 Table 1. Occupation of NLSY97 Respondents at Current or Most Recent Employee Job

Occupation of Job #01	Executive, Administrative, and Managerial	Management-Related	Professional Specialty	Technicians and Related Support	Sales	Administrative Support, including Clerical	Private Household	Protective Service	Service, not Private Household or Protective	Farm Operators and Managers	Farm, Not Managerial	Related Agricultural	Precision Production, Craft and Repair	Operators, Fabricators, Laborers	Military ¹	Other	Missing Data	Total working at an employee job
Round 1																		
Male	3	2	42	5	179	64	2	6	426	14	51	76	64	344	—	13	1	1292
Female	1	—	45	3	301	150	25	13	330	2	14	15	9	69	—	10	7	994
Total	4	2	87	8	480	214	27	19	756	16	65	91	73	413	—	23	8	2286
Round 2																		
Male	20	3	52	8	339	124	1	17	839	7	69	98	124	600	1	15	2	2319
Female	14	4	82	3	619	277	13	32	839	2	18	21	15	120	—	17	2	2078
Total	34	7	134	11	958	401	14	49	1678	9	87	119	139	720	1	32	4	4397
Round 3																		
Male	28	9	71	27	377	185	1	31	945	6	69	99	161	769	—	8	—	2786
Female	37	11	103	6	714	401	9	28	1087	1	15	21	14	168	—	12	2	2629
Total	65	20	174	33	1091	586	10	59	2032	7	84	120	175	937	—	20	2	5415
Round 4																		
Male	45	11	82	51	433	223	1	51	948	7	54	107	255	864	—	8	3	3143
Female	58	17	135	29	1008	561	15	36	963	—	13	18	20	208	—	8	1	3090
Total	103	28	217	80	1441	784	16	87	1911	7	67	125	275	1072	—	16	4	6233
Round 5																		
Male	68	13	125	81	485	255	3	70	842	11	49	101	300	878	24	17	5	3327
Female	70	23	166	35	1012	663	44	40	984	—	11	25	32	189	2	10	1	3307
Total	138	36	291	116	1497	918	47	110	1826	11	60	126	332	1067	26	27	6	6634

Note: This table is based on responses about the respondent's current or most recent occupation at his or her current or most recent job (e.g., YEMP-37500-COD.01, YEMP-80500-COD.01). The universe is respondents who reported that their current or most recent job was a civilian employee-type job; military, freelance, and self-employment are not included.

¹ Although this table is based on the occupation variables for civilian jobs, some respondents reported a military occupation in the civilian question series. These respondents may actually be civilians, or they may be members of the Armed Forces who reported their class of worker incorrectly (see the introduction to the "Employment" section of this guide).

Freelance Jobs

In rounds 1–3, NLSY97 youths not classified as self-employed described the type of work they performed. NORC personnel then coded these verbatim descriptions for release in the data set. Because the Census codes are not appropriate for this freelance type of employment, the descriptions were classified using a separate coding frame with categories such as babysitting, mowing, pet care, and snow shoveling. The reporting of freelance jobs was limited in round 4 to respondents born in 1983–84 and in round 5 to those born in 1984. The coding of these jobs followed the same procedure as previous rounds.

Self-Employment

For all freelance jobs reported in rounds 1–3 in which respondents are considered self-employed (age 16 or older, usually earn \$200 or more per week), the survey asked the respondent about his or her occupation or the kind of work performed at that job. The series of questions that determined the occupation is similar to that asked for employee jobs. Survey staff then coded the occupation of each self-employed job using the 1990 Census occupational codes.

The structure of the questionnaire changed in round 4, based on the age of the respondent. Older respondents (those born in 1980–82 for round 4 and those born in 1980–83 for round 5) reported self-employment jobs in the employee jobs section, answering occupation questions like those listed above for employee jobs. These jobs were coded the same way as employee-type jobs. Younger respondents (those born in 1983–84 for round 4 and those born in 1984 for round 5) continued to list self-employment jobs in the freelance section. If a freelance job met the earnings requirement for self-employment, the occupation was coded using the 1990 Census codes.

Military Service

The survey questions respondents in the Armed Forces on the type of job they hold. Members of the Army, Army National Guard, Marine Corps, or their reserves report their Primary Military Occupational Specialty. Coast Guard or Navy (or their reserves) personnel report their Primary Rating. Respondents in the Air Force and Air National Guard (or their reserves) report their Primary Air Force Specialty Code.

User Notes: Some errors were detected in data collection methods for questions on Armed Forces jobs in rounds 1 through 5. These errors were corrected for round 6. Users should also note that military specialty code data may contain errors because these codes contain some alphabetic characters but the coding system used did not permit the input of alphabetic characters.

If the respondent is currently employed at more than one job, the usual hours worked variables can be used to determine which job has more hours. (In the NLSY79, the job with more hours is classified as the CPS job if the re

Comparison to Other NLS Surveys: Occupation is collected each year from NLSY79 respondents; however, the jobs must meet minimum hours worked and weeks worked requirements before this information is collected. For the NLSY79, occupation is coded using the 1970 (through 1993) and/or 1980 (1982–present) occupational classification codes. Occupation is also coded using the 1970 and 1990 (for the CPS job only) codes for Children of the NLSY79 age 15 and older. For the Mature and Young Women, occupation has been coded using 1960, 1980, and 1990 systems in various survey years. The

occupations of Older and Young Men were recorded using the 1960 codes for all years; in the final two Older Men surveys, occupation was doublecoded using the 1980 system. For more information, consult the *NLS Handbook* or the appropriate cohort's *User's Guide*.

Survey Instruments: These questions are found in the employment section (YEMP) of the *Youth Questionnaire*.

Related User's Guide Section

4.3.4 Industry

Main Area of Interest

Industry & Occupation

Supplemental Area of Interest

Training

Reference

Census Bureau. *1990 Census of Population Alphabetical Index of Industries and Occupations*. Washington, D.C.: U.S. Government Printing Office, 1991.

4.3.8 Self-Employment Characteristics

Rounds 1–3: Respondents age 16 or older who usually earn \$200 or more per week at a freelance job are considered to be self-employed. The survey asks these respondents to report the industry of their job, their occupation, and their usual activities. The most common time of day they worked and the number of people who worked for them in that self-employment job are also collected. If the job ended, respondents state the reason (e.g., end of seasonal work, quit to return to school, quit to take another job).

Rounds 4–5: The structure of the questionnaire changed for round 4 and beyond, depending on the respondent’s age. Those who were born in 1980–82 for round 4 (1980–83 for round 5) had mostly reached age 18 by the start of the fielding period. This group reported any self-employment (regardless of earnings) along with regular employee jobs and was skipped past the freelance section. Characteristic data for self-employment is like that collected for employee-type jobs, which is described throughout this section. Younger respondents born in 1983–84 for round 4 (1984 only for round 5) continued to list any self-employment (freelance jobs that met earnings criteria) in the freelance section and reported employee jobs separately, as in previous rounds. Characteristic data for these jobs is the same as that collected in rounds 1–3.

Data hint →

In rounds 1–3, variable YEMP-112800 indicates that a job is classified as self-employment. In rounds 2 and 3, the freelance jobs roster is the best source of information about start and stop dates of self-employment. The roster also contains flags indicating whether the employer was current at the last interview and the current interview, as well as a unique identification code permitting these jobs to be linked across survey rounds. More information about rosters is available in chapter 3 of this guide.

In rounds 4–5, the freelance roster contains information about self-employment jobs (indicated by variable YEMP-112800) for younger respondents. These were born in 1983–84 for round 4 and in 1984 only for round 5. Self-employment jobs for older respondents (those born in 1980–82 for round 4 and in 1980–83 for round 5) are indicated on the regular employment roster by the variable YEMP_SELFEMP.xx. For more details about the structure in rounds 4–5, see the introduction to this Employment chapter.

The transition from the freelance section to the regular employee job section creates a disconnect in unique ID codes. If a respondent has a freelance job at age 16 with enough earnings to qualify as self-employed, and continues that job past his or her 18th birthday, the job will now appear on the regular employer roster rather than the freelance roster. When the job was in the freelance section, it was assigned a freelance UID. However, these UIDs cannot be transferred to the regular employer roster because the numbering system is different. On the other hand, it is misleading to assign a round 4 or round 5 UID, when the self-employed job first appears on the regular employer roster, because the job was preexisting. So, to indicate that a self-employed job existed during a previous interview, it is assigned a UID of 199999 on the round 4 or round 5 roster.

If a self-employed job is new in the round 4 or round 5 interview—it had never been reported in the freelance section—it will be assigned a regular round 4 or round 5 UID.

Comparison to Other NLS Surveys: In other NLS cohorts, information about self-employment is collected using the same series of questions as regular jobs. Respondents who are self-employed can be identified by examining the class of worker variables. For more information, refer to the *NLS Handbook* or the appropriate cohort's *User's Guide*.

Survey Instruments: Freelance and self-employment information is collected in the employment section (YEMP) of the *Youth Questionnaire*.

Related User's Guide Sections

4.3.1 Employers & Jobs

4.3.4 Industry

4.3.7 Occupation

Main Area of Interest

Self-Employment

4.3.9 Tenure

Respondents provide the start and stop dates of each employee and freelance job, as well as military service. The survey also collects information about periods of a week or more when the respondent was not working at a given job (see section 4.3.3, “Gaps in Employment”).

Data hint →

The raw interview questions asking about dates of employment are not released in the data set. Instead, the answers to these questions become part of the employer roster. More information about this process can be found in chapter 3 and in the introduction to the employment section of this guide.

Created Variables. Using the information described above, tenure variables are created. First, the variable CV_WKSWK_JOB_DLI.xx calculates tenure in weeks with an employer for the period between the job’s start date and the job’s stop date, or the current interview date for on-going employee jobs. A second variable (CV_WKSWK_JOB_YR.xx.xx) lists the number of weeks the respondent worked at each employee job in each calendar year. Finally, jobs with a tenure of 13 weeks or less as of the interview date are indicated by CV_JOB<13_WKS.xx. Gaps within jobs are not counted toward total tenure.

If a respondent works for the same employer during more than one interview, the cumulative tenure variable will report the respondent’s total tenure with that employer from the job’s start date through the current interview date. For example, the total tenure reported in the round 2 interview will be added to the tenure previously calculated for the same employer as of the round 1 interview date.

User Notes: If a respondent reported an employer in a later round that he or she also reported in a previous round, that employer will retain the original ID code to permit linking across rounds (see the user notes in the introduction to the employment section). In this type of situation, the start date information contained in the second round’s employer roster reflects the newly reported start date. All other data in the roster also refer to the time period since the last interview date. For example, the round 2 roster only contains information about gaps within a job since the last interview; if researchers are interested in all gaps since the beginning of the job, they should also examine the round 1 information. The roster in round 3 and later rounds is similarly constructed; for information prior to the last interview date the researcher must examine the rosters from the previous round.

Due to an error in the way freelance jobs were listed on the roster in round 2, about 150 respondents are missing start date information for a freelance job. In some cases the information is available in the round 1 data for jobs that were previously reported. All other data about the freelance job were collected, and this problem was corrected for round 3.

Comparison to Other NLS Surveys: Information on tenure is presented in a set of created variables for the NLSY79. As with the NLSY97, if the employer was reported in the previous survey, the total weeks from each interview are added together to create a total tenure variable. For the Mature and Young Women, users may be able to create tenure variables for the later survey years by combining start and stop dates and data on within-job gaps. Consult the *NLS Handbook* or the appropriate cohort's *User's Guide* for details.

Survey Instruments: These questions are found in the employment section (YEMP) of the *Youth Questionnaire*.

Related User's Guide Sections

4.3.3 Gaps in Employment

4.3.10 Time Spent at Work

Main Areas of Interest

Created Variables

Tenure w/ Employer

Supplemental Area of Interest

Time Spent at Work

4.3.10 Time Spent at Work

Respondents age 14 and over with employee jobs and all respondents with freelance jobs provide information about the number of hours worked each week. Those age 16 and over also report the shift usually worked at employee jobs. The number of weeks a respondent has worked for a particular employer and the job's start and stop dates are discussed in section 4.3.9, "Tenure."

Employee Jobs

Respondents age 14 or older report the number of hours per week that they usually worked at the time that a job began. For employee jobs lasting at least 13 weeks, the survey asks about the number of hours per week the respondent usually worked when the job ended (or at the time of the survey for on-going jobs). If the job has already been reported during a previous interview, the start date questions were asked at that time. In this case, respondents are asked only about the stop date or current interview date.

If an employee job ended after the respondent's 16th birthday, data are collected on the respondent's regular shift (e.g., the time of day the shift began and ended, the number of weekdays and/or weekend days usually worked, the number of hours usually worked per weekday and/or weekend day).

See "Work Experience" for information on created variables related to hours worked by the respondent.

User Notes: In rounds 1 and 2, and part of round 3, some respondents reported working overtime hours and then listed zero for the number of overtime hours worked. These cases were skipped out of the questions that collected regular hours worked per week excluding overtime. Therefore, no data were collected on the typical number of hours these respondents worked. This problem was corrected during the round 3 field period.

Information on the usual number of hours worked is also missing in the following cases: (1) if the respondent refused to give an amount in the non-hourly wage questions or (2) if the respondent answered "don't know" to both the amount question and the estimated amount follow-up question in the non-hourly wage section. See section 4.3.11, "Wages," for an explanation of how wage data are collected.

Freelance Jobs

This section gathers information about the usual number of hours the respondent worked per week as of the job's start date. The survey also collects data on the total number of days, the number of weekday/weekend days, and the number of hours per weekday/weekend that the respondent worked when

the job began. Respondents are then asked to provide similar information as of the job's end date (or as of the survey date if the job is on-going).

Self-Employment

In rounds 1–3, respondents who were age 16 or older and reported earning \$200 or more per week at a freelance job were considered self-employed. These jobs were included in the freelance section. Beginning in round 4, the structure of the questionnaire changed and respondents were divided by age. Those born in 1980–82 (or in 1980–83 for round 5) reported self-employment along with regular employee jobs and were skipped past the freelance section (see the Employee Jobs section above for details about time spent at work). Younger respondents born in 1983–84 (or in 1984 only for round 5) continued to list self-employment in the freelance section (see Freelance Jobs above).

User Notes: If a respondent reported an employer in a later round that he or she also reported in a previous round, that employer will retain the original ID code to permit linking across rounds (see the discussion about the creation of the employer roster in the introduction to the employment section). In addition, the start date information contained in the second round's employer roster reflects the original (previous round's) start date. However, all other data in the roster refer only to the time period since the last interview date. For example, the round 2 roster only contains information about gaps within a job since the last interview; if researchers are interested in all gaps since the beginning of the job, they should also examine the round 1 information. The roster in later rounds is similarly constructed; for information prior to the last interview date the researcher must examine the rosters from the previous round. See the "Self-Employment Characteristics" section of this guide for information on the special case of respondents who had a freelance job classified as self-employment and carried that job over into the regular employee jobs section in round 4.

Comparison to Other NLS Surveys: Most interviews of the NLSY79 include information on usual hours worked per week and usual shift worked. The exact time the shift began and ended is available beginning in 1986. The number of hours worked and the usual shift worked have also been collected from Children of the NLSY79 age 15 and older since 1994. The Mature and Young Women regularly report the number of hours usually worked per day and per week; they report exact times shifts begin and end starting in 1995. Older and Young Men reported the usual number of hours per week only. For more information, consult the *NLS Handbook* or the appropriate cohort's *User's Guide*.

Survey Instruments: These questions are found in the employment section (YEMP) of the *Youth Questionnaire*.

Related User's Guide Sections
4.3.3 Gaps in Employment
4.3.9 Tenure

Main Areas of Interest
Created Variables
Time Spent at Work

Supplemental Areas of Interest
Expectations
Tenure w/ Employer
Wages & Compensation

4.3.11 Wages

Employee Jobs

Several questions are used to determine the job's rate of pay as of the start date; the rate may be defined according to different scales (e.g., per month, per week, per day, per hour). Additional information is collected on whether the respondent received any non-wage, non-salary pay (e.g., overtime, tips, commissions, bonuses, incentive pay, other pay) when the job started. For each of these types of pay, the respondent is asked to report the rate that he or she earned. For jobs lasting more than 13 weeks, the respondent is asked to report the same wage information—rate of pay and non-wage, non-salary pay—as of the job's end date (or at the time of the survey for on-going jobs).

User Notes: Respondents are first asked to report a time unit for their rate of pay and then to report the actual dollar amount. However, the dollar amount may not actually match the time unit reported in round 1. This is not clear in the questionnaire because the skip patterns are missing from the relevant question records; the skip patterns are indicated in the codebook. In round 1, respondents reporting a time unit of per day, per week, or other were asked to state their earnings per week in the follow-up dollar amount questions. Respondents who reported a time unit of monthly or semi-monthly were asked to state their earnings per month. The remaining categories (per hour, biweekly, per year) are as reported by the respondent. Affected questions are as follows:

Time unit question	Dollar amount questions
YEMP-19200.x	YEMP-33400.x and YEMP-33500.x
YEMP-38200.x	YEMP-52400.x and YEMP-52500.x
YEMP-62000.x	YEMP-76200.x and YEMP-76300.x
YEMP-83100.x	YEMP-97300.x and YEMP-97400.x

In subsequent rounds, the skip patterns were made clearer and the time units were adjusted to help avoid any confusion. Respondents who reported a time unit of per day were instead asked for their daily earnings, and those who listed a time unit of biweekly reported their earnings as weekly.

Created Variables. Two created variables describe the respondent's hourly compensation at employee-type jobs. The first, CV_HRLY_PAY.xx, contains hourly rate of pay and is the more traditional measure. This variable excludes overtime and performance pay. CV_HRLY_COMPENSATION.xx, a second created variable, calculates the respondent's hourly monetary compensation; this figure includes all extra compensation such as overtime, tips, bonuses, etc. Both variables record the respondent's compensation as of the job's stop date for each job lasting 13 weeks or more. The compensation as of the job's start date is provided for shorter jobs. Jobs lasting less than 13 weeks are identified by a third created variable (CV_JOB<13_WKS.xx).

Freelance Jobs

The set of questions on freelance employment gathers information about the usual number of hours the respondent worked per week and the usual weekly earnings as of the job's start date. Respondents provide similar earnings information as of the job's end date (or at the survey date for on-going jobs). Researchers can use this information to estimate the respondent's hourly rate of pay.

Self-Employment

In rounds 1–3, respondents who were age 16 or older and reported earning \$200 or more per week at a freelance job were considered self-employed. These jobs were included in the freelance section. Beginning in round 4, the structure of the questionnaire changed and respondents were divided by age. Those born in 1980–82 (in 1980–83 for round 5) included self-employment with regular employee jobs and were skipped past the freelance section. For these jobs, respondents reported wage information in the same questions described in Employee Jobs above. Younger respondents born in 1983–84 (in 1984 only for round 5) continued to list self-employment in the freelance section (see Freelance Jobs above) and reported the same wage information as in rounds 1–3.

User Notes: If the employee or freelance job was reported during a previous interview, the rate of pay as of the start date was collected at that time. In this case, respondents are asked only about the rate of pay at the stop date or current interview date. However, if the job had been previously reported and the respondent's total tenure was less than 13 weeks, no additional information was collected in the later round. In this case the relevant information is available in the previous round's data.

Additionally, a few respondents reported new jobs in round 2 that ended before the round 1 interview date. (These jobs should have been reported in round 1 but were overlooked by the respondent.) In these cases no wage information was collected.

Military Service

Respondents in the Armed Forces report their pay grade when they entered and when they left (or at the date of the survey for current members). Pay grades are a pay level classification system associated with ranks of military personnel and used by all branches of the military.

Comparison to Other NLS Surveys: Wage data for the NLSY79 are available but have a few differences from the NLSY97. NLSY79 respondents report receiving non-wage, non-salary pay, but exact amounts are not specified. Further, wage data were collected differently before 1994, when the CPS section of the survey was redesigned. For the Original Cohorts, rate of pay is available for the CPS job and for many dual or intervening jobs. Refer to the *NLS Handbook* or the appropriate cohort's *User's Guide* for details.

Survey Instruments: The employment section (YEMP) of the *Youth Questionnaire* contains these questions.

Related User's Guide Section

4.8.2 Income

Main Areas of Interest

Created Variables
Wages & Compensation

Supplemental Areas of Interest

Income
Self-Employment
Tenure w/ Employer
Time Spent at Work

4.3.12 Work Experience

From the data collected on employee jobs and freelance jobs, researchers can construct a longitudinal record for the respondent’s employment history from age 14 to the current interview date.

Survey staff provide created variables, which summarize different aspects of a respondent’s labor force activity, for employee jobs held since the age of 14. The various sets of variables calculate weeks worked, hours worked, and the number of jobs held across difference time frames. See Table 1 below for the description and title of each variable.

In addition, a smaller set of variables is provided for self-employment since January of the year each respondent turned age 18 and for self-employment and employee jobs combined.

4.3.12 Table 1. Work Experience Created Variables
Prefix of “CV_” (e.g., CV_WKSWK_JOB_DLI.xx)

Variable Description	Employee-Type Job	Self-Employed Job	All Jobs
Weeks worked at each job, calendar year	WKSWK_JOB_YR.xx.xx	—	—
Weeks worked at each job, since last interview	WKSWK_JOB_DLI.xx	—	—
Weeks worked at any job, calendar year	WKSWK_YR_ET.xx	WKSWK_YR_SE.xx	WKSWK_YR_ALL.xx
Weeks worked at any job, since last interview	WKSWK_DLI_ET	WKSWK_DLI_SE	WKSWK_DLI_ALL
Total # of jobs held, calendar year	TTL_JOB_YR_ET.xx	TTL_JOB_YR_SE.xx	TTL_JOB_YR_ALL.xx
Total # of jobs held, since age 14	TTL_JOB_EVER (age 14–19) TTL_JOB_EVER (age 20+)	—	—
Total # of weeks worked, since age 14	WKSWK_EVER (age 14–19) WKSWK_EVER (age 20+)	—	—
Total # of hours worked, calendar year	HOURS_WK_YR_ET.xx	HOURS_WK_YR_SE.xx	HOURS_WK_YR_ALL.x x
Total # of hours worked, since age 14	HOURS_WK_EVER (age 14–19) HOURS_WK_EVER (age 20+)	—	—

Comparison to Other NLS Surveys: Total number of weeks worked and total weeks of tenure variables have been constructed for each cohort; consult the appropriate cohort's *User's Guide* for more information on the precise survey years in which these variables are available.

Survey Instruments: The questions used to construct the created variables are found in the employment section (YEMP) of the *Youth Questionnaire*.

Related User's Guide Sections

4.3.9 Tenure

4.4 Event History

Main Area of Interest

Created Variables

4.4 Event History

This section describes the NLSY97 event history data, which contains created variables summarizing the month and year in which major life events occurred for each respondent, along with all main file data. Variables cover topics such as marital status, enrollment, employment status, and program participation. The user can create an array for an individual respondent showing his or her status (e.g., single, married, receiving government assistance) at a point in time or over time.

Employment

The first set of event history arrays provides information on the respondent's employment on a weekly basis. These arrays include information about employer jobs and self-employed jobs held at age 18 and above; freelance jobs are not included in the arrays. All employment arrays provide information starting when the respondent turned 14 and ending in the week that he or she was interviewed in the most recent survey round. The arrays are presented using a continuous week and year naming scheme. In this format, the first week of January 1980 is numbered week 1, the second week of January 1980 is numbered week 2, and so on through the end of the year; the week numbers then start over for the first week of January 1981. Weeks are listed by exact date as well.

User Notes: Some respondents do not provide complete information about start and stop dates of employment during the interview. When the event history variables are created, survey staff must account for these missing data. For example, if a respondent reported the month and year that a job began and ended but did not know the exact days, the 1st is imputed for the start date and the 28th for the end date. Substituting in this way permits the creation of employment variables that closely approximate the true conditions. Due to a coding error, if a respondent could not provide the exact days that a job began and ended but reported that the start date was in the month of February, the 28th is imputed as the start date. Detailed information about the imputation rules is provided in Appendix 6 in the *NLSY97 Codebook Supplement*.

EMP_STATUS. This main array presents the employment status of a respondent in a particular week. Respondents may be classified as:

- not associated with an employer, not actively searching for an employee job
- not working (unemployment vs. out of labor force cannot be determined)
- associated with an employer, periods not working for employer are missing
- unemployed
- out of the labor force

- in active military service
- working for an employer, indicated by the employer's ID number

Deny Variables. Deny variables in the employment status section flag employment status data that a respondent denies reporting in a previous survey round.

EMP_DUAL_JOB#. If a respondent holds more than one employee job during a week, the ID number of the second employee job is presented in the dual jobs arrays. These arrays contain only the job number of the overlapping job; labor force status information is only included in the main array. For example, if a respondent held two employee jobs (e.g., the first and third jobs listed on the employer roster), during the 52nd week of 1997, the employer number for the first job is recorded in the EMP_STATUS array and the employer number for the third job is recorded in the EMP_DUAL_2 array.

EMP_HOURS. This final array calculates the total number of hours worked by a respondent at any employer job during a given week. Hours per week worked at each job are assumed constant except during a reported gap, when the hours for that job are assumed to be zero. Each week is assigned a code of -3 (invalid skip) when any of the jobs has an indeterminate month or year.

Start/Stop Date Variables. In addition to the three arrays, the employment event history includes a set of variables that provides the start and stop date of each job and each gap within a job in a continuous week and year format. For example, if the respondent started job #01 on May 4, 1997 (the 19th week of the year), the variable for the start week would have a value of 19 and the variable for the start year would have a value of 1997. These continuous week variables will aid researchers in making comparisons to the status arrays, which are reported in the same format. A crosswalk between the continuous week numbers and the actual dates is provided in Appendix 7 of the *NLSY97 Codebook Supplement*.

EMP_BK_WKS_XXXX. This variable provides, for interview years 2000 and 2001, the total number of weeks prior to the previous interview date that a back-reported job started. The weeks prior to the previous interview date are not updated in the employment event history arrays, however information for the weeks that occur after the previous interview date are included in the arrays.

EMP_BK_STATUS_XXXX. This variable gives the number of weeks from a round 4 back-reported job's start date to the date of last interview that a nonworking status (e.g., 1, 2, 3, 4, or 5) would have changed in the EMP_STATUS arrays to the back-reported job's employer ID had this job been included in the last interview. The weeks prior to and including the previous interview date are not updated in the employment event history arrays, but information for this job's weeks that occur after the previous interview date are included in the arrays. This variable is available for interview years 2000 and 2001.

EMP_BK_HOURS_XXXX. This variable gives the number of hours per week from a round 4 back-reported job's start date to the date of last interview that would have been included in the EMP_HOURS arrays had this job been included in the last interview. The hours per week prior to and including the previous interview date are not updated in the employment event history arrays, but information for this job's hours per week that occur after the previous interview date are included in the arrays. This variable is available for interview years 2000 and 2001.

User Notes: The created event history variables can be used in conjunction with the main file information about the respondent's employment. Like the main file variables (see the introduction to section 4.3, "Employment"), the event history variables use two systems of identification for a respondent's employers. First, the event history variables contained in the week-by-week status and dual job arrays use the unique ID numbers (UID) for each employer. To associate these employers with job characteristic information collected during the interview, which numbers jobs as job #01, job #02, etc., researchers must use the YEMP_UID.xx crosswalk variables from the employer roster. A second set of event history variables, those providing start and stop date information, use the job #01 numbering convention for a specific round. The number in the title of these variables refers to the same job as the variables in the main data set with the same number, so users can compare all information about job #02, for example, without any additional ID variables. However, to compare event history start and stop date information about job #01, for example, with information in the event history week-by-week status arrays, researchers must first use the YEMP_UID.xx crosswalk variables to identify the employer ID (9701–9707, 9801–9809, 199901–199909, etc.) that matches job #01. See the example in the introduction to section 4.3 for more details.

Marital Status

Three NLSY97 marital and cohabitation arrays record changes in the respondent's marital and cohabitation status. These arrays are presented using a continuous month timeline. This timeline labels January 1980 as month 1, February 1980 as month 2, and so on. Thus, a respondent born in month 4 (April 1980) might have a cohabitation that began in month 193 (January 1997) and ended in month 198 (June 1997). All marital/cohabitation arrays provide information beginning in the month that the respondent turned 14 and ending in the month that he or she was last interviewed. Additionally, the beginning dates of the youth's first marriage and first cohabitation are provided in two created continuous month variables: CV_FIRST_MARRY_MONTH and CV_FIRST_COHAB_MONTH. A crosswalk between the continuous month numbers and the actual dates is provided in Appendix 7 of the *Codebook Supplement*.

MAR_STATUS. The main array presents the status (e.g., never married/not cohabiting, cohabiting, married, divorced) of a respondent during a particular month. Marital status takes precedence over cohabiting; for example, if a respondent is divorced and living with another partner, the status listed in this array will be “divorced.” Respondents who are married but not living with their spouse are coded as married. There is no separate code for annulments; if a respondent reports this event, the marriage is maintained and the marital status code after the annulment is “divorced.”

MAR_COHABITATION. This second array details the partner that the respondent is living with in a particular month. For example, if the respondent is cohabiting, the variable for each month identifies whether the respondent lives with partner 1, partner 2, spouse 1, spouse 2, etc. In these variables, 1 and 2 refer to the respondent’s partners/spouses in chronological order. The numbers do not necessarily refer to the same person as the spouse/partner questions asked directly of the respondent during the survey. Users can distinguish between partners and spouses because partner IDs begin with “1” (e.g., 101, 102) and spouse IDs begin with “2” (e.g., 201, 202).

MAR_PARTNER_LINK. The third array links the cohabiting partner or spouse to the partner using the ID found on the partners roster. This array allows the researcher to identify characteristics of the respondent’s partner and to link them with spells of marriage or cohabitation. For example, a researcher might look at the MAR_COHABITATION variable for the 10th month of 1998 and determine that a respondent was living with his second partner in that month because the variable’s value is 102. If the researcher checks the value of MAR_PARTNER_LINK for the same month and year, the respondent might have a value of 9801, indicating that the partner in the event history arrays that month is the first partner reported in the round 2 survey. The researcher can then examine the round 2 variables with “Spouse/Partner 01” in the title to determine that person’s characteristics, such as race, ethnicity, age, religion, and so on. However, if there is a significant gap between relationship spells—for example, if the respondent was married and then divorced a spouse before round 1 and then began cohabiting with the same person in round 3—the survey would not necessarily identify this as the same person.

Deny Variables. Deny variables in the marital status section flag respondents who deny a relationship reported in a previous survey round.

Program Participation

Program participation arrays are constructed individually for three need-based programs—AFDC, Food Stamps, and WIC. The AFDC array includes all federal and state programs created under Temporary Assistance to Needy Families (TANF) or any government program for needy families that replaces AFDC. All other need-based programs (e.g., SSI, other) are combined into a fourth program participation array entitled Other. Three arrays are created for each program type. All program participation arrays use the

same continuous month format as the marital status arrays and provide information starting in the month that the respondent turned 14 and ending in the month that he or she was interviewed.

In addition, arrays are available for two employment-based programs. Unemployment Insurance is included in all rounds, and Worker's Compensation is only included in rounds 1–3 (see section 4.8.3 for information on Worker's Compensation questions).

STATUS. The main array (e.g., AFDC_STATUS) presents the status—receiving or not—of the respondent during each month. A value of '1' in the status array indicates months of receipt; a value of '0' indicates months that a respondent did not receive that benefit but was above the age of 14 (other eligibility requirements such as income level or presence of children are **not** considered). Respondents not age-eligible for the program have a value of –4. An edit variable (e.g., AFDC_EDIT_DATE) flags respondent-reported and imputed dates; the edit flags are described in the *NLSY97 Codebook Supplement*.

AMOUNT RECEIVED. If a respondent reports receiving benefits in a particular month, a second array presents the amount received that month (e.g., AFDC_AMT). A second set of edit variables (e.g., AFDC_EDIT_AMT) flags problematic values and explains any editing performed on these variables. More information about this editing process is available in the *NLSY97 Codebook Supplement*.

HOUSEHOLD MEMBERS RECEIVING. If a respondent reports receiving benefits in a particular month, except for Unemployment Insurance, the household members who benefited that month (e.g., respondent only, child only, respondent and child) are recorded in a third array (e.g., AFDC_HH). This array condenses the set of answers from the survey questions that collect this information; for example, see YPRG-18300.01_001 to YPRG-18300.01_005 for AFDC.

Deny Variables. Deny variables in the program participation section flag respondents who deny previously reported receipt of assistance.

Schooling

A set of variables provides information on the respondent's educational experiences beginning in 1980, when the first information is available in the survey, through the current interview. From 1980 through the round 1 interview date, the variables report schooling information on a yearly basis. Data from rounds 2–5 have both monthly and yearly schooling event histories. This approach permits the combination of information from the youth questionnaire, which collects more detailed data, and from the parent questionnaire, which presents information only for each year. In general, these variables refer to the school year rather than the calendar year.

Yearly Schooling Variables

SCH_YEAR_TO_GRADE, SCH_GRADE_TO_YEAR. These arrays present the grade the respondent attended during the school year and the school year during which the respondent attended a certain grade. For example, `SCH_YEAR_TO_GRADE.1990` refers to the grade attended by the respondent during the school year that starts in fall 1990 and ends in spring 1991. Similarly, if the respondent attended fourth grade in 1992–93, then `SCH_GRADE_TO_YEAR.4` would have the value 1992.

SCH_CHANGES. This set of variables counts the number of times that the school the respondent attended changed during the school year. For example, `SCH_CHANGES.1990` shows how many different schools the respondent attended during the school year that started in fall 1990 and ended in spring 1991.

SCH_MNTHS_MISSED. Not including summer vacation, this array presents the number of months during the school year that the respondent did not attend school.

SCH_SUMMER_SCHOOL. These variables show whether the respondent attended extra school classes, such as summer school, during an educational break in a given school year.

SCH_GRADE_PROGRESS, SCH_YEAR_PROGRESS. These arrays report whether the respondent was skipped ahead or demoted during a given grade in school or during a given school year.

SCH_SUSPENSIONS. This array counts the number of days during the school year the respondent was suspended from school. For example, if `SCH_SUSPENSIONS.1990` has a value of 3 then the respondent was suspended from school for three days during the 1990–91 school year.

Monthly Schooling Variables

SCH_STATUS. This array reports the respondent's enrollment status during each month from the round 1 interview date through the current interview date. Coding categories include unknown, not enrolled, in grades K to 12, in college, on vacation, expelled, and other.

SCH_TERM. These variables report the respondent's school type (public, private, or religious) and grade for each month in the time period. Researchers should consult the *NLSY97 Codebook Supplement* for exact information on the coding structure used in this array.

SCH_ID. These variables permit users to link array information to the school roster in the main data file and access other information about the school. For each month that the respondent was enrolled in the `SCH_STATUS` array, the corresponding monthly variable in this array contains an identification code.

Users should refer to the *NLSY97 Codebook Supplement* for exact information on using the code to match event history data with main file data.

SCH_DUAL_XXXX. A small number of NLSY97 respondents went to two different schools in the same month. Because only the first school can be reported in the other arrays, this variable flags these special cases. There is only one variable for each school for the period between each interview; the exact month when the overlap occurred is not indicated, and overlap may have occurred in more than one month.

Deny Variables. Deny variables in the schooling section identify respondents who deny ever attending a school reported in a previous interview round.

Comparison to Other NLS Surveys: The main NLSY79 data set includes information on each respondent's program participation history, presented using a continuous month timeline. These variables indicate the types of assistance received, the months each type was received, and the average monthly benefits. The NLSY79 Work History data presents employment status information in a format similar to the NLSY97 employment information, using a continuous week timeline. For more information, refer to the *NLS Handbook* or the *NLSY79 User's Guide*.

Related User's Guide Sections

4.2.2 Educational Status & Attainment

4.3.12 Work Experience

4.8.3 Program Participation

4.9.3 Marital & Marriage-Like Relationships

Main Area of Interest

Event History

4.5 Expectations, Attitudes, Behaviors & Time Use

Information in this section of the guide focuses on behavioral factors, such as the respondent's current outlook, perceptions of the future, and interactions with his or her parents. In the self-administered section of the *Youth Questionnaire*, a computer-assisted self-interview collects sensitive data that reflects anti-social behavior (e.g., substance use, delinquency, criminal activity), as well as details about attitudes and dating or sexual activity. Rounding out this section are data on respondents' typical weekly activities.

Table 1 summarizes the *NLSY97 User's Guide* subtopics available in this section and any global universe restrictions affecting those subtopics. Users should note that certain questions or groups of questions within a subtopic may have additional restrictions. Please consult the NLSY97 questionnaire and codebook for more information on particular questions.

4.5 Table 1. Expectations, Attitudes, Behaviors & Time Use Subtopics and Universe Restrictions

<i>NLSY97 User's Guide</i> Subtopic	Round 1 Universe	Round 2 Universe	Round 3 Universe	Round 4 Universe	Round 5 Universe
4.5.1 Alcohol Use	all ages	all ages	all ages	all ages	all ages
4.5.2 Attitudes					
Justice System	all ages	all ages	all ages	all ages	all ages
Parents	12-14	12-14	12-14	12-14	12-14
Peers	all ages	—	—	—	—
School	all ages; enrolled in school	—	—	—	—
Self	12-14	—	—	—	—
4.5.3 Autonomy & Parental Control	12, 13	12, 13	12, 13	12, 13	12, 13
4.5.4 Cigarette Use	all ages	all ages	all ages	all ages	all ages
4.5.5 Crime, Delinquency & Arrest	all ages	all ages	all ages	all ages	all ages
4.5.6 Drug Use	all ages	all ages	all ages	all ages	all ages
4.5.7 Expectations	15, 16	—	—	all ages	all ages
4.5.8 Sexual Activity & Dating					
Dating	all ages	all ages	all ages	all ages	all ages
Intercourse and Pregnancy	14-16	≥14 as of 12/31/97	all ages	all ages	all ages
4.5.9 Time Use	12-14 16—not enrolled, not employed	≥15—not enrolled not employed	≥15—not enrolled not employed	—	—

Note: Age restrictions as of 12/31/1996 unless stated otherwise.

Created Variables. In addition to the variables created by CHRR, Child Trends, Inc., an organization involved in the NLSY97 questionnaire design process, has created a number of scales and indexes from several groups of variables described in this section. These scales and indexes are intended to aid researchers in using the various data items relating to attitudes and behaviors. These variables are described in the following sections:

- Alcohol Use
- Attitudes
- Autonomy & Parental Control
- Cigarette Use
- Crime, Delinquency & Arrest
- Drug Use
- Sexual Activity & Dating

Although these Child Trends created variables are described only briefly in this guide, interested researchers may obtain a detailed discussion of the creation procedures from NLS User Services. This document, Appendix 9 of the *NLSY97 Codebook Supplement*, also summarizes statistical analyses of the scales and indexes, as well as related data items, performed by Child Trends researchers. (Although considered part of the *Codebook Supplement*, this appendix is a separate physical document.)

4.5.1 Alcohol Use

The focus of the alcohol use section is the frequency of beverage consumption in a recent period. An alcoholic beverage is defined as a can or bottle of beer, a glass of wine, a mixed drink, or a shot of liquor. The respondent is instructed not to include any “childhood sips that you might have had from an older person’s drink.”

The round 1 survey first established whether the respondent had ever consumed an alcoholic beverage and asked for the respondent’s age the first time. Respondents who answer don’t know, or who refuse to answer this retrospective question, continue to receive the question in later rounds until a valid answer is given. If a respondent answers affirmatively to having ever consumed alcohol, but cannot or will not provide the age at which they first drank, they continue to receive the age question in subsequent rounds. In rounds 2–5, respondents who have previously provided a valid answer to the question about having ever consumed an alcoholic beverage were asked whether they had a drink of an alcoholic beverage since the date of last interview. Respondents who had consumed alcohol were asked a series of follow-up questions in these rounds about the quantity and frequency of alcohol use in the 30 days prior to the interview. Specific questions included the number of days the respondent had one or more alcoholic beverage(s), how many drinks he or she usually had per day, and the number of days where the respondent had five or more drinks on the same occasion. In all rounds, the survey also elicits information on the number of days (in the 30 days prior to the interview) that the respondent had something alcoholic to drink right before or during school or work.

4.5.1 Table 1. Age Respondent First Drank Alcohol

Age	Male	Female	Total
≤ 10	380	192	572
11	184	125	309
12	311	304	615
13	394	374	768
14	540	586	1126
15	557	633	1190
16	435	420	855
17	388	360	748
18	284	274	558
19	150	138	288
20	64	56	120
21	29	26	55
22	3	1	4
Total	3719	3489	7208

Note: Universe is limited to respondents who reported ever drinking or drinking since the last interview. In rounds 2–5, respondents who had not drunk alcohol as of the previous round but did report drinking after that date were not directly asked their age when they first drank. These respondents were assigned their age at the round 2–5 interview date (CV_AGE_INT_DATE) for this table. This table is based on KEY!SEX, YSAQ-364, and YSAQ-364D.

Additional round 1 questions gathered information on the respondent’s perception of the effects of alcohol on various health conditions. Users can find a description of these questions in section 4.7, “Health.”

Created Variables. In rounds 1–4, the response to the “ever drank alcohol” question was used as part of the Substance Use Index (FP_YYSUBSI) created by Child Trends, Inc.; see the *Codebook Supplement Appendix 9*, available from NLS User Services, for more information and statistical analysis.

Comparison to Other NLS Surveys: NLSY79 respondents have provided information about the quantity and frequency of their alcohol consumption in the 1982–85, 1988–89, 1992, 1994, and 2002 surveys; the Children of the NLSY79 age 10 and older (including the Young Adults) responded to similar questions in 1988–2002. The 1989 and 1995–2001 interviews of the Mature Women, the 1991–2001 surveys of the Young Women, and the 1990 survey of Older Men also gathered data on the frequency and quantity of alcohol consumption. For more information, refer to the *NLS Handbook* or the appropriate cohort’s *User’s Guide*.

Survey Instruments: Questions on alcohol use are found in the self-administered section (YSAQ) of the *Youth Questionnaire*.

Related User’s Guide Sections
4.5.4 Cigarette Use
4.5.6 Drug Use
4.7 Health

Main Areas of Interest
Family Process Measures
Substance Use

Supplemental Areas of Interest
Expectations
Health

4.5.2 Attitudes

The NLSY97 *Youth Questionnaire* collects information about the respondents' attitudes toward the justice system and toward their parents in each round. The round 1 *Youth Questionnaire* also asked about respondents' attitudes toward school and perceptions of their peers and themselves. In some cases, the round 1 *Parent Questionnaire* asked similar questions of the responding parent.

Justice System

In the self-administered section of each survey, a series of questions focuses on the respondent's beliefs about the criminal justice system. The survey first queries respondents about their perceived odds of being arrested if they stole a car. Supposing that he or she had in fact been arrested for stealing a car, the respondent was then asked to report his or her perceived odds of being released without charges, being released with only a fine, and serving time in jail.

Parents

The NLSY97 obtains a great amount of detail about the respondent's relationship with his or her parents in each interview. Complementary information was gathered from the responding parent in round 1. This section also collects information about the relationship the respondent's parents have with each other. Users should note that "parent" in these descriptions generally refers to the respondent's parent or parent-figure and not only to biological parents; exceptions are clearly stated.

The universes of respondents who answered these questions and the parents to whom the questions referred changed after round 1. Table 1 presents the universes and reference parents for each of the groups of parent questions; the content of the questions is then described after the table.

A related question not listed in Table 1 collects information about the person (e.g., biological mother, brother or sister, boyfriend/girlfriend, teacher or school counselor, no one) the respondent would choose to go to if he or she had an emotional or personal relationship problem. This question is asked of all respondents in rounds 1–4.

4.5.2 Table 1. Universe and Reference Parents for Parent Attitude Questions

Question topic	Round 1		Rounds 2-5	
	Respondent universe	Reference parents	Respondent universe	Reference parents
1. Contact with absent parents	All respondents with absent parents	Absent bio or adoptive mother and father	- All respondents with absent parents - Respondents born 1983–84 in R5	Absent bio or adoptive mother and father
2. Parent's supportiveness/ permissiveness (not round 5)	All respondents	Resident parents/ parent-figures; absent parents in contact in past 12 months	All respondents	Resident parents/ parent-figures; absent parents in contact since last interview
3. R's opinion of parent	Respondents born in 1982–84	Resident parents/ parent-figures; absent parents in contact in past 12 months	- Respondents born in 1982–84 - All respondents in R5	Resident parents/ parent-figures
4. Parent's knowledge and behavior toward R	Respondents born in 1982–84	Resident parents/ parent-figures; absent parents in contact in past 12 months	- Respondents born in 1982–84 - Respondents born in 1983–84 in R5	Resident parents/ parent-figures
5. Parent relationship behaviors (not round 4)	Respondents born in 1982–84	Both bio parents, if reside with R, or bio parent R lives with and spouse/partner	Respondents born in 1983 or 1984	Both bio parents, if reside with R, or bio parent R lives with and spouse/partner
6. Contact between R's biological parents (not rounds 4–5)	Respondents born in 1982–84 living with only bio dad or not living with either bio parent	Bio parents	Respondents born in 1983 or 1984 living with only one or neither bio parent	Bio parents

Youth Questionnaire. The following list of question topics includes numbers that refer to the question groups in Table 1.

(1) *Contact with absent parent.* In round 1, the survey first determined the specific month and year that the NLSY97 respondent last lived with any absent biological mother or father or adoptive mother or father. Each round established whether the most recent contact with that parent was within a month, a week, or a day. A number of questions then focus on the respondent's contact with the absent parent during the past 12 months (round 1) or since the last interview (rounds 2–5). For each of these absent parents, the respondent is asked the following questions:

- Number of times youth contacted or tried to contact absent parent by mail or phone
- Number of times youth received a card, letter, or phone call from absent parent
- Number of times youth visited absent parent without spending the night
- Number of times youth stayed overnight at the absent parent's home

(2) *Parent's supportiveness and permissiveness.* This series first asks whether the respondent feels that each parent is supportive of him or her. A second question gathers information on whether the parent is

permissive or strict about making sure the respondent did what he or she was supposed to do. This series was not included in round 5.

- (3) *Respondent's opinion of parent.* For a group of younger respondents, additional data are collected on the respondent's opinion of each parent. Respondents state whether they think highly of the parent, want to be like him or her, and enjoy spending time with him or her.
- (4) *Parent's knowledge and behavior toward respondent.* Information about the parent's behavior toward the respondent, collected from a group of younger respondents, includes how often the parent does each of the following:
- Praises the respondent for doing well
 - Criticizes the respondent or the respondent's ideas
 - Helps the respondent to do things important to the respondent
 - Blames the respondent for his or her problems
 - Makes plans with the respondent and cancels for no good reason

In addition, the survey gathers data on the respondent's opinion of how well his or her parent knows the respondent's close friends, the parents of the respondent's close friends, the people the respondent is with when not at home, and the respondent's teachers/school activities. Specific questions in this series vary among surveys.

- (5) *Parent relationship behaviors.* These questions ask about the frequency with which each parent displays the following behaviors:
- Blames his or her partner for problems
 - Criticizes his or her partner or the partner's ideas
 - Encourages his or her partner to do things that the partner considers important
 - Expresses affection or love for his or her partner
 - Is fair and willing to compromise when they disagree
 - Screams and yells at his or her partner when angry

This series was not included in round 4.

- (6) *Contact between respondent's biological parents.* These questions determine the frequency of contact between the respondent's biological parents and the level of friendliness or hostility in their relationship. This series was not included in rounds 4–5.

Several question series about the respondent's parents were added in round 5 for older respondents born in 1980–82. The first questioned how often the respondent asked his or her parent for advice or help on education/training/job decisions and on friendships or close personal relationships. Another series asked about the parent's knowledge concerning the respondent's goals and aspirations in life, values and spiritual

beliefs, friends, and what the respondent is really like as person. In addition, these respondents were asked how often their family gathers to celebrate family events like birthdays, holidays, weddings, etc. A follow-up question asked the respondent to rate the importance of attending these events to him or her.

Parent Questionnaire (round 1). The NLSY97 also asked two of these series of questions in the parent survey. First, if the youth was born in 1982, 1983, or 1984, the responding parent described the frequency with which his or her spouse or partner displayed the behaviors listed in #5 above. If the responding parent was one of the youth's biological parents and the other biological parent did not reside in the household, the responding parent answered the second set of questions (in #6 above) about the frequency of contact with the other parent and the friendliness or hostility of their relationship.

Created Variables. Child Trends, Inc., has created a number of scales based on the variables described in this section. The following are available:

- (1) Parents' Marital Relationship–Youth Report (FP_YMFRELAT, FP_YFMRELAT); all rounds. This pair of scales uses the variables described in question topic 5 in Table 1 above. One scale summarizes the residential mother's relationship with the residential father; the other describes the father's relationship with the mother.
- (2) Parents' Marital Relationship–Parent Report (FP_PPRELAT); round 1 only. This scale uses the set of comparable parent questions on relationship quality to measure the responding parent's perception of the relationship with his or her spouse.
- (3) Parent-Youth Relationship–Youth Report (FP_YMSUPP and FP_YFSUPP in all rounds; FP_YNRMSUPP and FP_YNRFSUPP in round 1 only). Using the data collected in topics 3 and 4 in Table 1 above, this set of scales indicates the quality of the youth's relationship with each of the parents referred to in the questionnaire.
- (4) Parenting Styles–Youth Report (FP_YMPSTYL, FP_YFPSTYL, FP_YNRMPSTYL, FP_YNRFPSTYL); all rounds. Based on the supportiveness/permisiveness questions (see topic 2 in Table 1 above), this set of scales describes the parenting styles of the parents/parent-figures referred to in the questionnaire.
- (5) Parental Monitoring–Youth Report (FP_YMMONIT and FP_YFMONIT in all rounds; FP_YNRMMONIT and FP_YNRFMONIT in round 1 only). This set of scales describes the awareness of the youth's activities for up to four parents/parent-figures (see topic 4 in Table 1 above).

Interested researchers should consult *Codebook Supplement Appendix 9*, available from NLS User Services, for more information on the above scales. In addition, this appendix describes statistical analyses performed by Child Trends on the scales and on four related groups of round 1 variables: (1) the contact between the youth respondent's separated biological parents as reported by the youth (see topic 6 in Table 1 above); (2) the similar contact questions answered by the responding parent; (3) the sources of personal/emotional support question addressed to youth respondents; and (4) the supportive/permisive questions asked of all youth respondents.

Peers

Youth Questionnaire. A series of questions in the round 1 survey asked respondents about their perception of their peers' activities and behavior. Respondents who were enrolled at the time of the survey were asked to estimate the percentage of peers in their grade involved in the activities listed below; respondents who were not enrolled were asked the same questions about the percentage of peers in their grade when last enrolled.

- Belong to a gang that participates in illegal activities
- Cut classes or skip school
- Do volunteer work
- Ever use marijuana, inhalants, or other drugs
- Get drunk at least once per month
- Go to church or religious services on a regular basis
- Participate in organized sports, clubs, or school activities
- Plan to go to college
- Smoke cigarettes

Respondents who were age 15 or older were also asked to state the percentage of their peers they believed to have ever had sexual intercourse.

School

Youth Questionnaire. The round 1 NLSY97 survey attempted to ascertain the impact that school has had on the feelings of well-being experienced by various youths. To this end, respondents who were enrolled at the time of the survey were asked to agree or disagree with the following statements regarding their school's environment and their teachers:

- Discipline is fair
- Disruptions by other students get in the way of my learning
- I feel safe at this school
- Students are fairly graded
- Teachers are good
- Teachers are interested in the students
- There is a lot of cheating on tests and assignments

Self

Youth Questionnaire. In round 1, a series of questions designed to elicit information on the respondent's attitude about him- or herself was asked of respondents born in 1982, 1983, and 1984. These respondents were asked whether they strongly disagreed, disagreed, agreed, or strongly agreed with the following statements:

- I hardly ever expect things to go my way
- I rarely count on good things happening to me
- I'm always optimistic about my future

- In uncertain times, I usually expect the best

The round 1 NLSY97 survey asked the same youths questions similar to those found in Achenbach’s *Youth Self Report*. For each behavior, respondents were asked how well the statement described them over the past six months (not true, somewhat or sometimes true, or often true). The separate questions asked of male and female respondents are listed in Figure 1.

4.5.2 Figure 1. Youth Descriptors of Attitude toward Self

Female respondents	Male respondents
You lie or cheat	You lie or cheat
Your school work is poor	You don’t get along with other kids
You have trouble sleeping	You have trouble concentrating or paying attention
You are unhappy, sad, or depressed	You are unhappy, sad, or depressed

The round 4 self-administered section of the survey included a series that addressed respondents’ mental health. The questions asked how often they felt the following ways during the last month before the interview: very nervous, calm and peaceful, downhearted and blue, happy, or so down in the dumps that nothing could cheer them up. Respondents answered all, most, some, or none of the time for each feeling.

Parent Questionnaire (round 1). Parents were also asked a set of questions similar to Achenbach’s *Youth Self Report*. For each NLSY97 youth born in 1983 or 1984, the responding parent was asked how well each statement described the youth’s behavior in the past six months.

Created Variables. Child Trends, Inc., created the following scales based on the Achenbach questions in the round 1 youth and parent questionnaires:

- (1) Behavioral and Emotional Problems Scales–Youth Report (FP_YYFBEHS, FP_YYMBEHS). This pair of scales, one for male and one for female respondents, summarizes the data collected by the Achenbach *Youth Self-Report* variables described above.
- (2) Behavioral and Emotional Problems Scales–Parent Report (FP_PYFBEHS, FP_PYMBEHS). Based on the comparable Achenbach questions in the parent survey, this pair of scales measures behavioral and emotional problems separately for male and female youths.

Child Trends, Inc., also created a mental health scale (FP_YMNTHLTH) based on the round 4 series described above. This scale summarizes the responses into a score indicating whether the respondent has positive mental health or has emotional problems.

Codebook Supplement Appendix 9, available from NLS User Services, provides interested researchers with more information. In addition to describing the creation procedures for these scales, the appendix summarizes statistical analyses performed by Child Trends on these variables.

Comparison to Other NLS Surveys: The Children of the NLSY79 age 10 and older (including the Young Adults) have answered questions about their attitudes toward their school and themselves, about their relationship with their parents, and about their parents' relationship with each other. Users should note, however, that the wording of these questions is not identical to the NLSY97. For more information, consult the *NLS Handbook*.

Survey Instruments: Questions from the *Youth Questionnaire* are found in the self-administered (YSAQ), peers (YPRS), and schooling (YSCH) sections in round 1 and the self-administered section (YSAQ) in rounds 2, 3, and 4. In the round 1 *Parent Questionnaire*, these questions are found in sections P6 and PC12.

<i>Related User's Guide Sections</i>	<i>Main Areas of Interest</i>	<i>Supplemental Areas of Interest</i>
4.2.4 School Experience	Attitudes	Marriage & Cohabitation
4.2.5 School & Transcript Surveys	Family Process Measures	Parent Current Status
4.5.3 Autonomy & Parental Control	Illegal Activity & Arrest	Parent Retrospective
4.5.5 Crime, Delinquency & Arrest		

Reference

Achenbach, T. M. *Manual for the Youth Self-Report and 1991 Profile*. Burlington, VT: University of Vermont Department of Psychiatry, 1991.

4.5.3 Autonomy & Parental Control

Both the NLSY97 respondents and their responding parents have been asked about the level of autonomy or parental control that determines the rules, if any, regarding the youth's activities. These questions allow for an opportunity to compare the respondents' and their parent's perceptions about rule setting and discipline.

Youth Questionnaire. In each round, respondents born in 1983 or 1984 are questioned about the person or persons who make decisions concerning their activities, including how late they may stay out at night, the kinds of TV shows or movies they may watch, and who they are allowed to “hang out” with. For each rule that is reportedly made only by the respondent's parents or jointly by the parents and the respondent, a follow-up question asks about the number of times the respondent broke the rules in the last 30 days. Information is also collected on what action the parent(s) would take if they found out that the respondent had violated the rules, as well as which person would be most likely to handle a rule violation.

Parent Questionnaire (round 1). Responding parents of youths born in 1983 or 1984 were asked a set of autonomy and control questions similar to those found in the *Youth Questionnaire*, allowing comparison of youth and parent perceptions. For each rule detailed above, the responding parent was first surveyed on who makes that particular rule. Follow-up questions asked about the number of times the youth broke the rules in the last 30 days if the rule was made only by the parent or jointly by the parent and the youth. The responding parent was also asked to state the person who would most likely handle a rule violation.

Created Variables. Child Trends, Inc., has created several indexes based on the variables described in this section. The following are available:

- (1) Control/Autonomy Limit-Setting–Youth Report (FP_YHLIMITS); all rounds. This index is based on the questions asking who sets the limits on three different kinds of activities. The index indicates the youth's perception of the amount of control the parent exerts in setting limits.
- (2) Control/Autonomy Limit-Setting–Parent Report (FP_PHLIMITS); round 1 only. This comparable index uses the same questions asked of the responding parent to indicate the parent's perception of the amount of control exerted.
- (3) Control/Autonomy Limit-Breaking–Youth Report (FP_YHBROKED); all rounds. This index indicates whether the youth reported breaking any of the limits in the past 30 days.
- (4) Control/Autonomy Limit-Breaking–Parent Report (FP_PHBROKED); round 1 only. This comparable index reports whether the responding parent stated that the youth broke any limits in the past 30 days.

Interested users should consult *Codebook Supplement Appendix 9*, available from NLS User Services, for details on the above indexes. In addition, this appendix describes statistical analyses performed by Child Trends on these indexes and on several related variables: (1) who handles a limit violation, as reported by

the youth; (2) who handles a violation, according to the parent; and (3) the youth's report of the consequences expected if a limit is broken.

Comparison to Other NLS Surveys: The Children of the NLSY79 age 10–14 have provided information about who makes rules regarding their behavior, although not about the consequences of violating the rules. For more information, refer to the *NLS Handbook*.

Survey Instruments: Questions on autonomy and control are found in the self-administered section (YSAQ) of the *Youth Questionnaire*. In the round 1 *Parent Questionnaire*, this information was collected in section PC12.

Related User's Guide Section
4.5.2 Attitudes

Main Areas of Interest
Autonomy & Control
Family Process Measures

4.5.4 Cigarette Use

In round 1, all respondents were first asked whether they had ever smoked an entire cigarette; if so, they reported their age at that time. Respondents who answer don't know, or who refuse to answer this retrospective question, continue to receive the question in later rounds until a valid answer is given. If a respondent answers affirmatively to having ever smoked a cigarette, but cannot or will not provide the age at which they first smoked, they continue to receive the age question in subsequent rounds.

In rounds 2–5, respondents who have previously provided a valid answer to the question about having ever smoked a cigarette were asked whether they had smoked a cigarette since the date of last interview. Additional data were then collected in each round on the number of days they smoked during the 30 days prior to the interview. Finally, the respondents were questioned about the number of cigarettes they usually had on the days they smoked during those 30 days.

4.5.4 Table 1. Age Respondent First Smoked a Cigarette

Age	Male	Female	Total
≤ 10	397	234	631
11	223	150	373
12	372	331	703
13	352	383	735
14	378	422	800
15	322	351	673
16	257	265	522
17	273	195	468
18	201	179	380
19	130	94	224
20	53	33	86
21	18	25	43
22	1	2	3
Total	2977	2664	5641

Note: Universe is limited to respondents who reported ever smoking or smoking since the last interview. In rounds 2–5, respondents who had not smoked as of the previous round but did report having their first cigarette after that date were not directly asked their age at the time they first smoked. These respondents were assigned their age at the round 2–5 interview date (CV_AGE_INT_DATE) for this table. This table is based on KEY!SEX, YSAQ-360, and YSAQ-360C.

In addition, information was collected in round 1 regarding the respondent's perception of the effects of smoking on various health conditions; details can be found in section 4.7, "Health."

Created Variables. In each round, the response to the “ever smoked a cigarette” question is used as part of the Substance Use Index (FP_YYSUBSI) created by Child Trends, Inc.; see *Codebook Supplement Appendix 9*, available from NLS User Services, for more information and statistical analysis.

Comparison to Other NLS Surveys: The 1984, 1992, 1994, and 1998 surveys of the NLSY79 included questions on the quantity and frequency of cigarette use. Children of the NLSY79 age 10 and older (including the Young Adults) have provided this information in the 1988–2002 surveys. Detailed information on cigarette use was gathered from the Mature Women in 1989 and from the Young Women in 1991 and 1993; these respondents also stated whether they currently smoked in 1995–2001. Older Men provided retrospective smoking data in 1990. Consult the *NLS Handbook* or the appropriate cohort’s *User’s Guide* for further details.

Survey Instruments: Questions on cigarette use are found in the self-administered section (YSAQ) of the *Youth Questionnaire*.

<i>Related User’s Guide Sections</i>	<i>Main Areas of Interest</i>	<i>Supplemental Area of Interest</i>
4.5.1 Alcohol Use	Family Process Measures	Health
4.5.6 Drug Use	Substance Use	
4.7 Health		

4.5.5 Crime, Delinquency & Arrest

The NLSY97 survey questions on crime, delinquency, and arrest can be divided into two areas: questions about active participation and questions about perceptions of personal safety and other antecedents to delinquent behavior. This section focuses on the youth's participation in delinquent and criminal behaviors; other details can be found in section 4.6.4, "Household & Neighborhood Environment." The following information on delinquency, crime, and arrest is collected in the self-administered section of the youth instrument.

Within the criminal justice system, "delinquent" behavior is usually differentiated from "criminal" behavior on the basis of an offender's age. For example, a 12-year-old who has damaged property would likely be charged with a delinquent offense; an older youth might be charged with either a delinquent or criminal offense at the discretion of the judge. Since an offense is only classified as criminal or delinquent if the offender actually appears in court, a different distinction is made for the purposes of this discussion. The descriptions of delinquency in this section include behaviors, such as running away from home, which are unlawful for minors but not necessarily for those ages 18 and older. The discussions of crime and arrest in this section describe the questions on behaviors that are considered criminal activity for adults, even though a respondent who has committed one of these offenses may have actually been charged as a juvenile rather than an adult.

Delinquency

The first set of delinquency questions addresses the respondent's history of running away from home and staying away at least overnight. Data include the age when this first occurred, the total number of times the respondent has ever run away, and the total number of times he or she has run away since the date of last interview. In rounds 2–5, youths who were age 17 or older at the date of last interview or who were not living with any parents were skipped past the runaway questions.

Next, the respondents answer a series of questions on whether they have ever carried a handgun, defined as any firearm other than a rifle or a shotgun, and if they have carried one since the last interview (in rounds 2–5). The series continues by asking their age the first time they did so, the number of days out of the past 30 that they carried the gun, and whether they carried the gun to school during this period.

In a separate set of questions, respondents are surveyed on their involvement in gang activity. Those who belonged to a gang are asked to state the age when they joined the gang and their membership status in the previous 12 months (round 1) or since the date of the last interview (subsequent rounds). Those who did not belong to a gang during that time period state their age when they last belonged to a gang. Additional information on gang activity in the respondent's extended family, school, and neighborhood is found in sections 4.6.4, "Household & Neighborhood Environment," and 4.5.2, "Attitudes."

Crime

The survey asks about participation in and the intensity of various criminal activities in the previous year (round 1) or since the date of last interview (subsequent rounds). The types of criminal activities included are:

- Purposely damaged or destroyed property not belonging to the respondent
- Stole something worth less than \$50
- Stole something worth \$50 or more (including a car)
- Other property crimes, including fencing stolen property, possessing or receiving stolen property, or selling something for more than it was worth
- Attacked or assaulted someone
- Sold or helped to sell marijuana, hashish, or other hard drugs

Youths who stole something state whether they stole from a store, snatched someone's purse/wallet or picked someone's pocket, went into a locked building to steal something, or used a weapon during the act. Additionally, respondents who report stealing something worth \$50 or more are asked if they stole a car.

In round 1, follow-up questions about a particular activity (except stealing something worth less than \$50) determined the respondent's age the first time and the number of times he or she participated in the past 12 months. Respondents who reported no involvement during the previous 12 months were asked to state their age when they last participated in that criminal activity. If the respondent committed other property crimes, sold or helped to sell drugs, or stole something worth \$50 or more, he or she was asked about any monetary rewards in 1996—the total cash received or the total cash he or she would have received—from selling these items. Respondents who reported having sold or helped to sell drugs stated whether they were involved in selling marijuana or hashish, other hard drugs, or both, as well as the amount of income received from these sales.

Reference periods and question details shifted in round 2. In this survey, respondents were questioned about the number of times they participated in each activity since the date of last interview. However, respondents who answered "don't know" or "refused" in round 1 were again asked about criminal activities in which they had ever participated. Those who stole something (worth more or less than \$50) or participated in the other property crimes listed above were asked for the amount of any monetary rewards in the previous calendar year—the total cash received or the total cash he or she would have received—from selling these items. Respondents who reported selling or helping to sell drugs distinguished whether they were involved in selling marijuana or hashish, other hard drugs, or both, and listed the amount of income received from these activities. Rounds 3–5 followed the same pattern as round 2.

Arrests

NLSY97 youth respondents are asked whether they have ever been arrested by the police or taken into custody for an illegal or delinquent offense (not including arrests for minor traffic violations) and the total number of times this has happened. Round 1 then collected information on the number of arrests before the respondent's 12th birthday and the respondent's age at the first arrest. Further data were gathered on the respondent's most recent and first arrests. (If the respondent reported only one arrest, this information was collected as the respondent's most recent arrest; the first arrest data fields contain missing values.) Rounds 2–5 gathered the number of arrests since the last interview, including the month and year of the first arrest if it occurred during that time.

For the first and most recent arrests (round 1) or a continuous loop of all arrests (subsequent rounds), the respondent provides the month and year of the arrest and states whether the police charged him or her with an offense. A series of questions then determines the type of offenses with which the respondent was charged; he or she could enumerate multiple charges. The list of possible arrest charges includes assault, burglary, destruction of property, possession or use of illicit drugs, sale or trafficking of illicit drugs, a major traffic offense, and a public order offense.

For each offense with which they have been charged, respondents report whether they were convicted for or pled guilty to that offense. A follow-up question distinguishes whether the respondent went to juvenile court, adult court, both, or neither as a result of the charges. Respondents then describe the sentence. For those sentenced to jail, an adult corrections institution, a juvenile corrections institution, or a reform/training school, the surveys record the month and year they began their sentence and the month and year their sentence ended (if appropriate). Similar questions ask about the month and year community service or probation began and ended. If the respondent did not go to court, was not convicted of, and did not plead guilty to any charges, he or she states the arrest outcome: no further action, pre-court diversion program/counseling, appearance before a judge with no further action, or other.

Finally, a description of the questions on the attitudes of the respondent toward justice is found in section 4.5.2, "Attitudes"; the use of drugs by the respondent is described in section 4.5.6, "Drug Use"; and descriptions of survey questions on the safety of the respondent's household environment are found in section 4.6.4, "Household & Neighborhood Environment."

Created Variables. A Delinquency Index (FP_YYCRIMI), created for each round by Child Trends, Inc., sums the responses to the questions asking if the youth ever participated in various criminal/delinquent activities. More information and a statistical analysis of the index can be found in *Codebook Supplement* Appendix 9, available from NLS User Services.

Comparison to Other NLS Surveys: Information on crime, delinquency, and arrest records was collected from the NLSY79 in a special self-report supplement during the 1980 interview. This supplement detailed respondents' participation in and income from criminal activities and their contacts with the criminal justice system. The 1988–2002 surveys of the Children of the NLSY79 asked those age 10 and older (including the Young Adults) to report on participation in various illegal activities. For more information, consult the *NLS Handbook* or the *NLSY79 User's Guide*.

Survey Instruments: Questions on crime, delinquency, and arrest are found in the self-administered section (YSAQ) of the *Youth Questionnaire*.

Related User's Guide Sections

4.5.2 Attitudes
4.5.6 Drug Use
4.5.7 Expectations
4.6.4 Household & Neighborhood Environment

Main Areas of Interest

Family Process Measures
Illegal Activity & Arrest

Supplemental Areas of Interest

Expectations
Household Characteristics
Income
Substance Use

4.5.6 Drug Use

In this section, all respondents are surveyed on their experience with marijuana, which is also referred to in the survey as “grass” or “pot.” In round 1, respondents were first asked whether they had ever used marijuana; if so, they reported the age at which they first used the drug. Respondents who answer don’t know, or who refuse to answer this retrospective question, continue to receive the question in later rounds until a valid answer is given. If a respondent answers affirmatively to having ever used marijuana, but cannot or will not provide the age at which they first used it, they continue to receive the age question in subsequent rounds. In rounds 2–5, respondents who have previously provided a valid answer to the question about having ever used marijuana were asked whether they had used marijuana since the last interview, even if only once. Each survey then collected additional information on the number of days the respondent smoked marijuana in the 30 days prior to the interview. Finally, questions determined the number of times the respondent used marijuana right before or during school or work in those 30 days.

4.5.6 Table 1. Age Respondent First Used Marijuana

Age	Male	Female	Total
≤ 10	71	22	93
11	57	23	80
12	151	121	272
13	238	171	409
14	331	337	668
15	363	307	670
16	319	279	598
17	336	246	582
18	265	221	486
19	135	133	268
20	60	61	121
21	18	24	42
22	1	—	1
Total	2345	1945	4290

Note: Universe is limited to respondents who reported ever using marijuana or using marijuana since the previous interview. In rounds 2–5, respondents who had not used marijuana as of the previous round but did report their first use after that date were not directly asked their age at first use. These respondents were assigned their age at the round 2–5 interview date (CV_AGE_INT_DATE) for this. This table is based on KEY!SEX, YSAQ-370, and YSAQ-370C.

The drug use questions in rounds 2–5 were more extensive than those in round 1. In addition to the series on marijuana use, the questionnaire asked about the use of any drugs such as cocaine, crack, heroin, or any other substance not prescribed by a doctor in order to get high or to achieve an altered state. Details collected include the respondent’s age when he or she first used any of these drugs and the number

of times since the last interview that the respondent took some drug or other substance. The number of times in the last 30 days that the respondent used a drug or other substance right before school or during school or work hours was also gathered. Additional questions on whether the respondent ever sold drugs are described in section 4.5.5, “Crime, Delinquency & Arrest.”

4.5.6 Table 2. Age Respondent First Used Cocaine/Other Drugs

Age	Male	Female	Total
≤ 10	28	16	44
11	15	9	24
12	36	33	69
13	50	63	113
14	53	51	104
15	75	75	150
16	102	89	191
17	127	99	226
18	126	108	234
19	78	70	148
20	52	49	101
21	20	9	29
22	1	—	1
Total	763	671	1434

Note: Universe is limited to respondents who reported ever using cocaine/other drugs or using cocaine/other drugs since the previous interview. First asked in round 2. In rounds 3–5, respondents who had not used cocaine/other drugs as of the previous round but did report their first use after that date were not directly asked their age at first use. These respondents were assigned their age at the round 3–5 interview date (CV_AGE_INT_DATE) for this. This table is based on KEY!SEX, YSAQ-372C, and YSAQ-372CC.

Created Variables. In each round, the response to the “ever smoked marijuana” question is used as part of the Substance Use Index (FP YYSUBSI) created by Child Trends, Inc. See *Codebook Supplement Appendix 9*, available from NLS User Services, for more details and statistical analysis.

Comparison to Other NLS Surveys: The NLSY79 first included drug use questions in 1980, when respondents stated the number of times they had ever smoked marijuana or used other drugs. In 1984, 1988, 1992, 1994, and 1998, NLSY79 respondents reported their age at first use, lifetime use, and frequency of use in the past 30 days. Children of the NLSY79 age 10 and older (including the Young Adults) have been asked about their use of various drugs in the past 3 months since the 1988 survey. For more information, refer to the *NLS Handbook* or the *NLSY79 User’s Guide*.

Survey Instruments: All the questions on drug use are found in the self-administered section (YSAQ) of the *Youth Questionnaire*.

Related User's Guide Sections

4.5.1 Alcohol Use

4.5.4 Cigarette Use

4.5.5 Crime, Delinquency &
Arrest

Main Areas of Interest

Family Process Measures

Substance Use

Supplemental Areas of Interest

Illegal Activity & Arrest

Income

4.5.7 Expectations

Both the NLSY97 respondents and their responding parents have been asked to state their beliefs about events that may occur in the youth's future. Since these questions were asked of both groups, researchers may be able to compare differences in their perceptions.

Youth Questionnaire. In round 1, respondents who were born in 1980 or 1981 were surveyed on their beliefs about the future. Asked to assess the probability that certain events would occur in a specified time period, the respondents were instructed to use a scale from 0 (impossible) to 100 (certain).

Expectations questions were included again in the round 4 survey. In this round, all respondents were asked to answer a similar series of questions. Table 1 lists the events and the time periods included in each round.

In round 5, respondents were randomly divided into four experimental groups of roughly equal size that were used to study response rate for this type of question. Instead of answering expectations questions in a single section (as in previous rounds), they encountered them separately in each appropriate topical section. Expectations about education were asked in the schooling section, expectations about pregnancy were asked in the fertility part of the self-administered section, etc. The variable SYMBOL!TXGROUP indicates the experimental group to which each respondent was assigned. A non-experimental group of respondents skipped these questions in the topical sections and instead were guided through the expectations section as in previous rounds.

Parent Questionnaire (round 1). Responding parents of youths born in 1980 or 1981 were asked to assess the probability that certain events, described in Table 1, would occur in the life of the youth within a specified time period.

Comparison to Other NLS Surveys: Respondents in the NLSY79, Children of the NLSY79, Young Women, and Young Men have all answered questions about their educational and employment expectations for the future; however, the specific questions and reference periods have varied widely. Users should consult the *NLS Handbook* or the appropriate cohort's *User's Guide* for more precise information.

Survey Instruments: These questions are found in the expectations section (YEXP) of the round 1 *Youth Questionnaire* and section PC11 of the round 1 *Parent Questionnaire*.

Related User's Guide Section
4.5.5 Crime, Delinquency & Arrest

Main Area of Interest
Expectations

Supplemental Areas of Interest
Ed. Status & Attainment
Fertility and Pregnancy
Illegal Activity & Arrest
School Experience
Substance Use
Time Spent at Work

4.5.7 Table 1. Topics of Expectations Questions

Event and Time Period	Youth, Round 1	Responding Parent, Rd 1	Youth, Round 4	Youth, Round 5
One year from the interview date				
Arrested (whether rightly or wrongly)	✓		✓	
Dead from any cause (e.g., crime, illness, accident)	✓			
Enrolled in a regular school	✓	✓	✓	
Pregnant or have gotten someone pregnant	✓		✓	✓
Seriously drunk at least once	✓		✓	
Victim of a violent crime	✓		✓	
Working for pay for > 20 hours per week and enrolled	✓	✓	✓	
Working for pay for > 20 hours per week and not enrolled	✓		✓	
Get married			✓	✓
Moved from current residence				✓
By the youth's 20th birthday				
Mothered or fathered a baby	✓	✓		
Dead from any cause (e.g., crime, illness, accident)	✓			
Received a high school diploma	✓	✓		
Served time in jail or prison	✓	✓		
By the youth's 30th birthday				
Earned a four-year college degree	✓	✓		✓
Working for pay for > 20 hours per week	✓	✓		
Five years from the interview date				
Arrested (whether rightly or wrongly)			✓	
Enrolled in a regular school			✓	✓
Pregnant or have gotten someone pregnant			✓	✓
Seriously drunk at least once			✓	
Victim of a violent crime			✓	
Working for pay for > 20 hours per week and enrolled			✓	

Chapter 4: Expectations, Attitudes, Behaviors & Time Use—Expectations

Working for pay for > 20 hours per week and not enrolled			✓	
Get married			✓	✓
Across entire life				
Have any/have more children, how many				✓

Note: Round 1 youth who were born in 1980–81 were asked these questions; by round 4, all respondents were age-eligible.

4.5.8 Sexual Activity & Dating

Dating

In the self-administered section of the youth instrument, all respondents are asked whether they have ever been on a date or unsupervised social outing with a member of the opposite sex. Those who have are questioned on their age the first time they went on a date. Other questions include how often the respondent dated and the number of different people he or she went out with on a date in the past year (round 1) or since the last interview (subsequent rounds).

Intercourse and Pregnancy

In rounds 1 and 2, data on sexual activity were obtained from respondents who were at least 14 years old; all respondents were eligible in later rounds. These respondents are first asked if they have ever had sexual intercourse and if they have had intercourse since the last interview (in rounds 2–5). If a respondent reports ever having sexual intercourse, questions are asked about the first experience (e.g., the respondent's age, whether the respondent and the partner used birth control). Those who report not using any method of birth control are questioned on whether they wanted a pregnancy at that time. Rounds 2–5 also included a question that specifically asked whether respondents used a condom that first time. Additional information is collected about the respondent's total number of partners and the number of partners in the previous year (round 1) or since the last interview (subsequent rounds). Complementing information on the number of times that the respondent had sexual intercourse in the past 12 months (round 1) or since the last interview (subsequent rounds), the surveys gather the frequency of his or her birth control use during that same period. Rounds 2–5 also asked specifically about the frequency of condom use. Finally, these respondents state the method of birth control used most often.

Additional questions were included in the rounds 4–5 surveys concerning the respondent's first sexual experience. These questions asked for the month and year this experience occurred; the age of the respondent's partner at that time; the location where it occurred (e.g., family's home, partner's home, hotel); the time of day; the status of the relationship at that time (e.g., dating, living together, engaged); the ethnicity and race of the partner; whether they discussed birth control before that point; and, if they used birth control, the methods they used that first time.

The same respondents eligible for questions on sexual activity also answer questions on pregnancies in the self-administered section. This information on the respondent's pregnancies, including those that ended in miscarriage or abortion, is described in section 4.9.2. Additionally, information on the respondent's knowledge about pregnancy and sexually transmitted diseases is found in section 4.7, "Health."

4.5.8 Table 1. Age Respondent First Had Sex

Age	Male	Female	Total
≤ 10	160	30	190
11	62	13	75
12	158	68	226
13	341	202	543
14	552	463	1015
15	619	659	1278
16	617	700	1317
17	400	479	879
18	255	322	577
19	86	125	211
20	34	38	72
21	10	11	21
22	1	1	2
Total	3295	3111	6406

Note: This table is limited to respondents who reported ever having sex. The universe for these questions was youths age 14 or older as of 12/31/96 in round 1 or 12/31/97 in round 2; 3,578 respondents were skipped out of these questions in round 1 and 1,693 respondents were skipped out in round 2 because of their age. All respondents were eligible starting in round 3. Table is based on CV_AGE_INT_DATE, KEYISEX, and YSAQ-300.

Created Variables. Child Trends, Inc., created these three variables for the round 4 data: if respondents had ever had sex by the round 4 interview (FP_SEXEVER), their age at first sex (FP_SEXAGE), and the round they first reported having had sex (FP_SEXYEAR). All are included in the main file data set; see the *Codebook Supplement Appendix 9*, available from NLS User Services, for more information and statistical analysis.

Comparison to Other NLS Surveys: Information on sexual activity was gathered from NLSY79 respondents in the 1983–85 surveys; data on contraception have been regularly collected since 1982. NLSY79 Children age 13 and older have reported on their sexual activity beginning in 1988. Children age 15 and older have also provided information on contraceptive use since 1994. For further details, consult the *NLS Handbook* or the *NLSY79 User's Guide*.

Survey Instruments: Questions on dating and sexual activity are found in the self-administered section (YSAQ) of the *Youth Questionnaire*.

Related User's Guide Sections
 4.7 Health
 4.9.2 Fertility, Pregnancy &
 Children

Main Area of Interest
 Sexual Activity

Supplemental Areas of Interest
 Attitudes
 Fertility and Pregnancy
 Health

4.5.9 Time Use

In the round 1 series of questions on time use, respondents born in 1982, 1983, or 1984 assessed the amount of time in a typical week that they spent doing homework (if enrolled), taking extra classes or lessons, watching TV, and reading for pleasure. Youths born in 1980 who were not enrolled or employed were asked about the time they usually woke up, whether they went to any place on a regular basis, and the times that they left for and returned from that place.

In rounds 2 and 3, the time use series was addressed to respondents born in 1980 or 1981 who were not enrolled or employed. They were asked for the usual times that they got up, left their home, arrived home, and went to sleep on a typical weekday. Then the survey collected the main activities these respondents participated in and/or the places they went between waking and sleeping, as well as the number of hours spent at these places during a typical weekday. In the final question in this section, respondents reported the main activities or places where they spent time during the week before the survey.

Respondents were not asked time use questions in rounds 4 or 5.

Comparison to Other NLS Surveys: In 1981, NLSY79 respondents provided detailed data on their use of time in the past seven days. This section recorded time spent working, attending school or training programs, sleeping, watching TV, caring for children, and completing household chores. For more information, refer to the *NLS Handbook* or the *NLSY79 User's Guide*.

Survey Instruments: These questions are found in the time use section (YTIM) of the *Youth Questionnaire*.

Main Area of Interest

Time Use

Supplemental Areas of Interest

Health

School Experience

4.6 Family Background & Demographic Characteristics

This section discusses a variety of demographic, family, and neighborhood factors that potentially affect a young person's labor market attachment and educational attainment. Data supplied by the household informant, the youth, and the parent provide researchers with a wealth of background information about the youth's family members and household environment.

Table 1 summarizes the *NLSY97 User's Guide* subtopics available in this section and any global universe restrictions affecting those subtopics. Users should note that certain questions or groups of questions within a subtopic may have additional restrictions. Please consult the NLSY97 questionnaire and codebook for more information on particular questions.

4.6 Table 1. Family Background & Demographic Characteristics Subtopics and Universe Restrictions

<i>NLSY97 User's Guide</i> Subtopic	Round 1 Universe (Age as of 12/31/96)	Rounds 2–5 Universe (Age as of 12/31/96)
4.6.1 Age of Respondent	all ages	all ages
4.6.2 Gender	all ages	all ages
4.6.3 Geographic Indicators	all ages	all ages
4.6.4 Household & Neighborhood Environment	all ages	≤14
4.6.5 Household Composition	all ages	all ages
4.6.6 Nonresident Characteristics	all ages	—
4.6.7 Parent Characteristics		
Background	all ages; parent int. conducted	—
Status	all ages; parent int. conducted	<i>Household Income Update completed</i>
History	all ages; parent int. conducted	—
4.6.8 Race & Ethnicity	all ages	—
4.6.9 Youth History	all ages; parent int. conducted	all ages

Created Variables. In addition to the variables created by CHRR, Child Trends, Inc., an organization involved in the NLSY97 questionnaire design process, has created a number of scales and indexes from several groups of variables described in this section. These scales and indexes are intended to aid researchers in using the various data items relating to environmental and parent characteristics. These variables are described in the following sections:

- Household & Neighborhood Environment
- Parent Characteristics

Although these Child Trends created variables are described only briefly in this guide, interested researchers may obtain a detailed discussion of the creation procedures from NLS User Services. This document, *NLSY97 Codebook Supplement Appendix 9*, also summarizes statistical analyses of the scales and indexes, as well as related data items, performed by Child Trends researchers for the round 1 data. (Although considered part of the *Codebook Supplement*, this appendix is physically a separate document.)

4.6.1 Age of Respondent

The NLSY97 collected information on the respondent’s age and date of birth in the round 1 *Screener, Household Roster, and Nonresident Roster Questionnaire*. This information was then verified or corrected by the respondent in the initial *Youth Questionnaire* and by the responding parent in the round 1 *Parent Questionnaire*. Date of birth and age were not recollected from respondents in subsequent rounds.

Created Variables. Created variables summarize the information described above, presenting the ages of the respondents as of the interview date in both a regular year format (CV_AGE_INT_DATE) and in a continuous month scheme [CV_AGE(MONTHS)_INT_DATE]. See section 4.4, “Event History,” for a more detailed description of the continuous month format. In round 1, the ages of respondents were also calculated as of December 31, 1996 (CV_AGE_12/31/96).

4.6.1 Table 1. Age of NLSY97 Respondents as of 12/31/96

Age (birth year)	Male	Female	Total
12 (1984)	911	860	1771
13 (1983)	934	873	1807
14 (1982)	953	888	1841
15 (1981)	947	927	1874
16 (1980)	854	837	1691
Total	4599	4385	8984

Note: Table based on KEY!SEX and KEY!AGEDOL.

4.6.1 Table 2. Age of NLSY97 Respondents as of Interview Date

Age	Round 1	Round 2	Round 3	Round 4	Round 5
12	1169	—	—	—	—
13	1726	105	—	—	—
14	1858	1597	109	—	—
15	1877	1676	1664	58	—
16	1719	1736	1632	1555	67
17	614	1747	1728	1667	1548
18	21	1416	1622	1657	1620
19	—	109	1387	1606	1610
20	—	—	67	1402	1588
21	—	—	—	136	1339
22	—	—	—	—	111
Total	8984	8386	8209	8081	7883
Noninterview	—	598	775	903	1101

Note: Table based on CV_AGE_INT_DATE.

Data hint →

Researchers may notice that there are two round 1 variables that present the youth's age: KEY!AGE (R05366.) and CV_AGE_INT_DATE (R11941.). The KEY!AGE variable is a symbol generated during the administration of the *Screener, Household Roster, and Nonresident Roster Questionnaire* (round 1). Thus, it initially represented the respondent's age at the screener date. During the youth interview, respondents had an opportunity to correct their age if it had changed since the administration of the screener or was reported incorrectly by the household informant. If a new age was provided, the KEY!AGE variable was updated. However, approximately 400 respondents either had a birthday between the screener date and the youth interview date or had incorrect information collected during the screener but did **not** correct their age information during the interview.

The created age at interview date variable, CV_AGE_INT_DATE, presents the youth's age at the time the round 1 youth questionnaire was administered. This variable was constructed using birth date and interview date information collected during the interview. Because KEY!AGE may not always accurately reflect the respondent's age at the **youth** interview date, researchers may wish to use the created variable for most research purposes. Table 3 below shows the difference in ages between KEY!AGE and the created age variable.

There are also two age variables in the round 2 data, SYMBOL!KEY!AGE (R25162.) and CV_AGE_INT_DATE (R25535.). For almost all respondents, these variables present identical ages, so either can be used. However, two respondents (identification numbers 471 and 476) have inaccurate ages in the symbol variable due to a minor data error. For these two respondents, the created variable contains accurate age information.

In the round 3 data, SYMBOL!KEY!AGE and CV_AGE_INT_DATE present identical information.

The round 4 data has one different age between these two variables: ID number 9022. The CV_AGE_INT_DATE variable has the correct information.

No changes were made between round 4 and round 5.

4.6.1 Table 3. Comparison of Round 1 Variables KEY!AGE and CV_AGE_INT_DATE

KEY!AGE	CV_AGE_INT_DATE							Total
	12	13	14	15	16	17	18	
12	1168	63						1231
13	1	1663	79	1				1744
14			1779	79	1			1859
15				1796	93			1889
16				1	1625	91	1	1718
17						523	8	531
18							12	12
Total	1169	1726	1858	1877	1719	614	21	8984

Note: Table based on KEY!AGE (R05366.) and CV_AGE_INT_DATE (R11941.).

User Notes: Because some respondents were first interviewed during the refielding period in March through May 1998, there are respondents in the sample who had already turned 18 by their round 1 interview date. These respondents were still eligible for the sample because they were age 16 or younger as of December 31, 1996. See section 2.2, “Interview Methods,” for more information.

The youth respondent and the responding parent both had an opportunity to verify or correct the respondent’s age as reported in the *Screener, Household Roster, and Nonresident Roster Questionnaire*. However, if the youth and parent did not provide the same age and/or date of birth, no attempt was made to correct this discrepancy. Users should make decisions about these cases on an individual basis. The created variables for age are based on the youth data.

Comparison to Other NLS Surveys: Age data are available for all NLS cohorts. These variables include both the age of the respondents as of a fixed date during the initial survey year and as of the interview date in various years. Date of birth is also available for all cohorts. Consult the *NLS Handbook* or the appropriate cohort’s *User’s Guide* for information about accessing these data.

Survey Instruments: This information was collected in the household roster section of the *Screener, Household Roster, and Nonresident Roster Questionnaire* (round 1). Verification or correction of the information is found in the youth information section (YINF) of the round 1 *Youth Questionnaire* and the PINF section of the *Parent Questionnaire* (round 1).

Related User’s Guide Sections

- 4.6.5 Household Composition
- 4.6.6 Nonresident Characteristics
- 4.6.7 Parent Characteristics

Main Areas of Interest

- Created Variables
- Demographic Indicators

Supplemental Area of Interest

- Common Variables

4.6.2 Gender

Information on the respondent’s gender was first collected in the *Screener, Household Roster, and Nonresident Roster Questionnaire* (round 1). This information was then verified or corrected by the respondent in the round 1 *Youth Questionnaire* and by the responding parent in the *Parent Questionnaire* (round 1). The interviewer recorded the respondent’s gender in the rounds 2–5 interviewer remarks sections (YIR). Table 1 summarizes the composition of the cohort by gender and sample type.

4.6.2 Table 1. NLSY97 Composition by Gender and Sample Type

Sample type	Male	Female	Total
Round 1			
Cross-sectional	3459	3289	6748
Supplemental	1140	1096	2236
Total	4599	4385	8984
Round 2			
Cross-sectional	3213	3066	6279
Supplemental	1070	1037	2107
Total	4283	4103	8386
Round 3			
Cross-sectional	3144	3029	6173
Supplemental	1026	1010	2036
Total	4170	4039	8209
Round 4			
Cross-sectional	3098	2957	6055
Supplemental	1019	1007	2026
Total	4117	3964	8081
Round 5			
Cross-sectional	3012	2907	5919
Supplemental	977	987	1964
Total	3989	3894	7883

Note: Table is based on CV_SAMPLE_TYPE, KEY!SEX, and RNI.

Information on the gender of household members is found in section 4.6.5, “Household Composition.”

Comparison to Other NLS Surveys: Information on gender is available for both the NLSY79 and the Children of the NLSY79. Gender is implicit by membership in the Original Cohorts.

Survey Instruments: This information was collected in the household roster section of the round 1 *Screener, Household Roster, and Nonresident Roster Questionnaire*. Verification or correction of the information is found in the youth information section (YINF) of the round 1 *Youth Questionnaire*, the

PINF section of the round 1 *Parent Questionnaire*, and the interviewer remarks section (YIR) of the rounds 2, 3, and 4 *Youth Questionnaires*.

Related User's Guide Sections

- 4.6.5 Household Composition
- 4.6.6 Nonresident Characteristics

Main Area of Interest

Demographic Indicators

Supplemental Area of Interest

Common Variables

4.6.3 Geographic Indicators

The variables in this section are created to describe the respondent's area of residence. The software package Maptitude (V4.5) is used in the creation of the NLSY97 geographic variables. This program assigns latitude and longitude information to respondent addresses. Latitude and longitude are then used to link respondent addresses to standard geographic information such as state, county, and metropolitan statistical area. See the *NLSY97 Geocode Codebook Supplement* for more information about the use of Maptitude.

Main File Created Variables. The NLSY97 main data set includes several created variables that provide general information about the respondent's permanent residence. The main variables indicate whether the respondent lives in an urban or rural area (CV_URBAN-RURAL), whether the respondent lives in a Metropolitan Statistical Area (CV_MSA), and in which Census region the respondent resides (CV_CENSUS_REGION). The states comprising each Census region are listed in Figure 1.

In addition, the same three variables were created in round 1 to describe the respondent's residence as of his or her 12th birthday; another counts the number of residences in which the respondent has lived from his or her 12th birthday until the survey date (all rounds). See section 4.6.9, "Youth History," for more details on these variables.

The variable CV_MIGRATE.XX is available for rounds 2 through 5 and describes all moves made by the respondent, including moves within a county, within a state to a different county, between states, and to and from a foreign country.

4.6.3 Figure 1. Number of NLSY97 Respondents and States by Census Region

Census Division	# of Respondents					States
	Rd 1	Rd 2 ¹	Rd 3 ¹	Rd 4 ¹	Rd 5 ¹	
Region 1: Northeast	1585	1451	1403	1390	1332	Connecticut Maine Massachusetts New Hampshire New Jersey New York Pennsylvania Rhode Island Vermont
Region 2: North Central	2050	1903	1836	1774	1735	Illinois Indiana Iowa Kansas Michigan Minnesota Missouri Nebraska North Dakota Ohio South Dakota Wisconsin
Region 3: South	3359	3160	3116	3074	3003	Alabama Arkansas Delaware District of Columbia Florida Georgia Kentucky Louisiana Maryland Mississippi North Carolina Oklahoma South Carolina Tennessee Texas Virginia West Virginia
Region 4: West	1990	1846	1834	1808	1768	Alaska Arizona California Colorado Hawaii Idaho Montana Nevada New Mexico Oregon Utah Washington Wyoming

Note: Table based on CV_CENSUS_REGION.

¹ Round 2 numbers do not total to 8386, the number of respondents interviewed. Data on geographic location were missing for 4 respondents, and 22 respondents resided outside of the United States. Similarly, the round 3 numbers do not total to 8209 because data were missing for 2 respondents and 18 respondents resided abroad. In round 4, data were missing for 21 respondents and 14 respondents resided abroad, so the numbers do not total to 8081. In round 5, data were missing for 32 respondents and 13 respondents resided abroad, so the numbers do not total to 7883.

Geocode CD. In addition to all main file data, the NLSY97 Geocode CD provides a variety of statistics for the counties where respondents lived at each interview date. Most of these data are based on the 1994 edition of the U.S. Census Bureau’s *County and City Data Book* (the most recent edition available). The first group of geocode variables lists basic demographic information for respondents’ counties. These data include land area in square miles; population by race, age, and gender; and birth and death rates. Another variable reports the percent of persons in that county who lived in a different house and/or state in 1990 compared to their residence in 1985, providing information about migration rates for the respondent’s area.

Factors that might influence the respondent’s education and employment outcomes are the focus of several other geocode variables. These provide the number of serious crimes, households with children, female householders with no spouse present, persons with high school or college degrees, and families below the poverty level. A pair of variables summarizes medical availability for each county, reporting the number of active nonfederal physicians and community hospital beds.

Geocode variables associated with economic and labor force issues include the size of the county’s civilian labor force, the percent employed in various industries, and the percent of workers age 16 and over with

jobs outside their county of residence. Income variables include per capita money income for the respondent's county, per capita personal income, and median family money income. The unemployment rates for the respondent's metropolitan areas or for the portions of the state not in an MSA (for non-MSA respondents) are also reported.

The final geocode variables involve colleges attended by respondents. Survey staff use information from the Integrated Postsecondary Education Data Systems (IPEDS) to provide users with a code identifying each college attended by the respondent and its location. These codes can be used to associate the NLSY97 respondent's college with various characteristics of the institution contained in the IPEDS database. More information is located in Attachment 102 of the *NLSY97 Geocode Codebook Supplement*.

User Notes: Researchers should be aware that the geocoding procedure changed between rounds 2 and 3 of the NLSY97. All geocode variables from all rounds have been recreated using the new, more accurate method. Regardless of the round of interest, researchers should use data from a round 3 or later Geocode CD. For more information on geocoding procedures, refer to Appendix 5 in the main file *Codebook Supplement* or the introduction to the *Geocode Codebook Supplement* (the latter is distributed only with a Geocode CD).

Due to the detailed nature of the information found on the Geocode CD, use of this data set is restricted to those meeting confidentiality requirements. The Geocode CD is released only to those who satisfactorily complete the Bureau of Labor Statistics geocode agreement procedure. The application for this geocode agreement is now available online at <<http://www.bls.gov/nls/geocodeapp.htm>>. The round 5 Geocode CD is scheduled for release in September 2003.

Comparison to Other NLS Surveys: Data on the respondent's area of residence are available for all cohorts. For the NLSY79, information on the respondent's region of residence and geographic mobility is available to all users on the main public data file; more detailed information (e.g., state and county of residence) is available on the restricted-use Geocode CD. Data on NLSY79 Children are available through the mother's record. Region of residence and geographic mobility of Original Cohort respondents are provided for most survey years. For more complete information, refer to the *NLS Handbook* or the appropriate cohort's *User's Guide*.

Main Area of Interest: Created Variables

4.6.4 Household & Neighborhood Environment

In the self-administered section of the *Youth Questionnaire* in rounds 1–4, respondents born in 1982–84 report on their household environment in the past month. These questions inquire as to whether the respondent usually had a computer available and a quiet place to study at home in the past month. Round 5 respondents born in 1983–84 were only asked if they had a quiet place to study at home in the past month. Round 1 also asked if the respondent usually had electricity and heat at home when needed and if a dictionary was usually available at home. Respondents in this age group (12–14 years old) living with a parent-figure or guardian at the survey date are also asked about their household environment in a typical week. In this set of questions, respondents report the number of days in a typical week the housework gets done (round 1 only), their family eats dinner together, their family does something religious together, and their family does something fun together. In round 1, youths also stated the number of days in a typical week that they heard gunshots in their neighborhood.

A separate set of questions in round 1 asked all youths, regardless of age, about events they may have experienced before turning 12 years old. These events include: the respondent’s house or apartment was broken into, he or she was the victim of repeated bullying, and he or she saw someone get shot or shot at with a gun. A follow-up asks the respondent to state his or her relationship (e.g., relative, friend, stranger) to the person who was shot or shot at. The same questions, referring to the youth’s experiences from age 12 to age 18, were repeated in subsequent rounds for older respondents. The questions are also asked in the youth’s first interview after turning 18.

In addition, the survey questions all youths on whether there were gangs in their neighborhood or school at the time of the survey. A gang is defined as a group that hangs out together, wears gang colors or clothes, has set clear boundaries of its territory or turf, and protects its members and turf against other rival gangs through fighting or threats. A separate set of questions asks whether any of the respondent’s brothers, sisters, cousins, or friends belongs to a gang. For information on the respondent’s participation in a gang, see section 4.5.5, “Crime, Delinquency & Arrest.”

Created Variables. Child Trends, Inc., has created several indexes based on the variables described in this section. The following are available:

- (1) Index of Family Routines–Youth Report (FP_YHROUTIN); rounds 1 through 4. Based on the questions about the number of days per week the youth’s family eats together, does housework, does something fun together, and does something religious together, this index indicates the number of days per week spent in routine activities with the family.
- (2) Family/Home Risk Index (FP_ADHRISKI); round 1 only. This index combines responses to a number of questions in the youth and parent interviews, as well as the interviewer remarks section of both the youth and parent questionnaires, to provide an overall assessment of the youth’s environment. These variables include the quality of the youth’s physical home environment, risk

factors in the youth's neighborhood, the availability of study materials in the youth's home, the religious behavior of the responding parent and the youth's family, the parents' involvement with the youth's schooling, family routines in an average week, time spent watching TV in an average week, the responding parent's physical or mental disabilities, and the parental monitoring and youth-parent relationship scales described in section 4.5.2, "Attitudes."

- Physical Environment Risk Index (FP_ADPENVRI); round 1 only. This index uses a subset of the items from the Family/Home Risk Index to assess risk factors in the youth's physical environment.
- Enriching Environment Index (FP_ADENRCHI); round 1 only. This index uses a second subset of the items from the Family/Home Risk Index to identify opportunities for enriching educational activities in the youth's environment.

Interested researchers should consult *Codebook Supplement Appendix 9*, available from NLS User Services, for more information on the above indexes. In addition to describing the creation procedures for these indexes, the appendix summarizes statistical analyses performed by Child Trends on these variables.

Comparison to Other NLS Surveys: For the NLSY79, information on the respondent's perception of the neighborhood environment is available only for NLSY79 women. Data on both household and neighborhood environments are collected from Children of the NLSY79 age 10 and older (including the Young Adults). For more information, refer to the *NLS Handbook* or the *NLSY79 User's Guide*.

Survey Instruments: These questions are found in the self-administered section (YSAQ) of the *Youth Questionnaire*.

Related User's Guide Section
4.5.5 Crime, Delinquency &
Arrest

Main Areas of Interest
Family Process Measures
Household Characteristics

Supplemental Areas of Interest
Children
Demographic Indicators
Ed. Status & Attainment
Marriage & Cohabitation
Parent Current Status

4.6.5 Household Composition

The NLSY97 collects basic demographic information about each member of the respondent's household. In addition, the survey establishes the relationships among household members. Most data were gathered from the household informant in the round 1 household roster and updated by the youth respondent in subsequent rounds.

Data hint →

During the interview, the survey program organizes the respondent's answers to questions about household members into a "household roster." This grid of data is the best source of information about household members and should be used instead of the raw data in analyses. Details about the creation and use of the household roster are provided at the end of this section; general information about rosters is available in section 3.2 of this guide.

User Notes: The respondent's household is based on what the respondent considers to be his or her permanent household as reported in the household information section. This is not necessarily the same as where he or she is living at the time of the survey. So respondents who are in the military (or away at college or incarcerated) may report their spouse or children as being in the household, even though the answers in the fertility and marriage sections have the respondent separated from them.

Screener, Household Roster, and Nonresident Roster Questionnaire (round 1). If a potentially eligible youth lived in the household based on age and date of birth information, the extended screener solicited basic demographic data for each household occupant. In this section, the household informant reported the gender, ethnicity (e.g., Hispanic, Latino, of Spanish origin), and race (e.g., white; black or African American; American Indian, Eskimo, or Aleut; Asian or Pacific Islander; or other) of each occupant.

The household roster, also administered to the household informant, collected further demographic information on household occupants. The survey recorded the highest grade level completed by each household occupant. A follow-up question established the highest degree received by those age 17 and above. This section also collected the marital status of each occupant. Employment information included the number of weeks that each resident above the age of 15 was self-employed or worked for pay at an employee job in 1996. Additional information was collected on the resident's usual number of hours worked per week during that period and his or her current employment status. Finally, the household informant reported the current or most recent occupation of each household resident who either worked at least one week during 1996 or had worked for pay since age 18.

The round 1 household roster further established the relationship of each person in the household to the youth and to each other. Follow-up questions verified the exact relationship; see Figure 1 for definitions of

relationships. For example, if the household informant identified an occupant only as a “mother,” an additional question asked the household informant if this person was a biological, adoptive, step-, or foster mother. The survey collected the same type of information for a person only identified as a “father.” If occupants were listed as half siblings, the interviewer questioned the household informant on whether they shared a biological mother or a biological father. Another set of questions determined if full siblings whose reported birth dates differed by a month or less were multiple births; if they were the same gender, the household informant was asked if they were identical or fraternal twins. In addition, a person listed only as a “grandmother” was further identified as a maternal, paternal, or social grandmother. The survey solicited similar information for a household occupant listed only as a grandfather, a great-grandmother, or a great-grandfather.

4.6.5 Figure 1. Definitions of Relationships in Household Roster

Biological Relationship: Two people are related by blood or conception and birth. For example, one’s biological father is the same as one’s natural father or the man who made one’s biological mother pregnant.

Step Relationship: Two people are related through a marriage where the husband and/or wife had children with another partner. For example, a stepchild is the biological offspring of one’s spouse (with some other partner). A stepmother is the wife of one’s biological father (if he is not married to one’s biological mother). A stepbrother is the biological son of one’s stepmother who is not the biological son of one’s biological father.

Adoptive Relationship: The permanent legal rights and duties with respect to a child have been transferred from one person or institution to another. The parental rights to an adopted child have been permanently and legally transferred from the birth parents to the adoptive parents. Any other children of the adoptive parents become adopted siblings of the adopted child.

Foster Relationship: Someone assumes a legal and financial obligation for a child but there is no permanent adoptive relationship.

In-law Relationship: Two people are related through marriage. A mother-in-law is the mother of one’s spouse. A son-in-law is the spouse of one’s daughter. A sister-in-law is the sister of one’s spouse or the spouse of one’s brother.

Full Relationship: The youth and his or her siblings share the same biological mother and biological father.

Half Relationship: Siblings share only one common biological parent. Half-siblings have the same biological mother but different biological fathers, or vice versa.

Social Relationship: A person functions in a particular family role but is not biologically related. For example, one’s social grandparent would be someone who functions as a grandparent but is not biologically related.

Source: Interviewer Reference Manual for *Screener, Household Roster, and Nonresident Roster Questionnaire*, 1997.

Parent Questionnaire (round 1). For the round 1 survey, the responding parent provided additional information on the 1996 earnings (self-employment earnings and earnings from an employee job) of each household member older than 14 at the time of the survey. Another question determined the income each household member received from any other sources such as Social Security, pensions, welfare, interest, gifts, etc.

Youth Questionnaire. In rounds 2 through 5, the household information section of this instrument verified characteristics of the respondent’s parents listed in the previous round and asked the youth to describe any

times since the last interview when he or she did not live with each parent. Parents who no longer lived with the youth were moved to the nonresident roster (see section 4.6.6, “Nonresident Characteristics”).

The household information section next asked the respondent to review the list of household members from the last interview. If any members moved out of the household or were deceased, this information was recorded and the person was moved to the nonresident roster. The nonresident roster is maintained across rounds so that if anyone moves back into the household, that person can be identified and matched to their old information by ID number, which remains the same. If anyone from a previous round nonresident roster moved into the household, they were moved onto the household roster but kept the same ID number so that they can be tracked across rounds. Finally, the respondent reported any other new household members, listing their age or birth date and relationship to the respondent. All of the information about current members from the previous round’s roster and about new members was then used to create the household roster for the current round.

After the roster was created, the respondent provided additional information about the household members. The questions in this section gathered the gender and race/ethnicity for all new household members, all members transferred from the nonresident roster, and any previous members for whom the information was missing. Marital status, employment status, and highest degree received were recorded for all household members age 16 or older. Current enrollment status was collected for all members age 4 or older; highest grade attended was gathered for all new members and anyone currently enrolled in school. Respondents reported their relationship to any members from the previous round who were not blood relatives or whose relationship was missing. Finally, if the new household member was a stepparent, an adoptive parent, or the partner of the respondent’s parent, the date he or she joined the household was recorded.

Created Variables. Using this information, a single created variable (CV_YTH_REL_HH_CURRENT) reports the relationship of the youth to the primary adults in the household at the time of the survey (e.g., both biological parents, biological mother, adoptive parent[s]). In round 1, this variable was also created for various points in the youth’s childhood (CV_YTH_REL_HH_AGE_x); see section 4.6.9, “Youth History,” for information. Users should be aware that, on the round 1 release, these variables were named CV_HH_REL_BIRTH (for current household composition) and CV_HH_REL_AGE_x. They were subsequently renamed to more accurately reflect the information presented.

Additional created variables identify the household size (CV_HH_SIZE), the number of household members under the age of 6 (CV_HH_UNDER_6), and the number of household members under the age of 18 as of the survey date (CV_HH_UNDER_18). Created variables describing household net worth

and total household income are discussed in section 4.6.7, “Parent Characteristics,” section 4.8.1, “Assets & Debts,” and section 4.8.2, “Income.”

User Notes: The classification of a stepparent who has legally adopted a child may differ between the *Screener, Household Roster, and Nonresident Roster Questionnaire* and the other survey instruments. In the former, help screens instructed interviewers to categorize such parent-figures as stepparents rather than adoptive parents. Since the respondent defines this relationship in the self-administered section of the *Youth Questionnaire*, he or she may choose to list this parent-figure as an adoptive parent.

Users should exercise caution when drawing conclusions based on household characteristics using NLSY97 data. The large number of multiple respondent households in the sample may skew the data on certain characteristics if the analysis is performed at the respondent level rather than at the household level. See section 1.3, “NLSY97 Sample,” for more information on multiple respondent households.

Most household information was collected during the administration of the round 1 *Screener, Household Roster, and Nonresident Roster Questionnaire*. This part of the survey also asked questions about any nonresident relatives of the NLSY97-eligible youth, including biological, step-, and adoptive parents and biological children; these data are discussed in detail in section 4.6.6, “Nonresident Characteristics.” The following paragraphs contain a description of the collection and organization of the household roster information. Researchers interested in using these data should first read this discussion. Section 3.2 in this guide contains a general description of NLSY97 rosters.

To organize the information about household residents and nonresident relatives, two rosters were created during the administration of the *Screener, Household Roster, and Nonresident Roster Questionnaire*. The first, the household roster, includes information for all current residents of the respondent’s permanent household. For easy identification, all variables on the round 1 household roster were assigned question names that begin with “HHI2_.” The second key roster is the nonresident roster, which presents information about the youth’s nonresident relatives. All question names of nonresident roster variables begin with “NONHHI_.” Note that all household roster variables from subsequent rounds have names that begin with “HHI_” (the “2” is dropped).

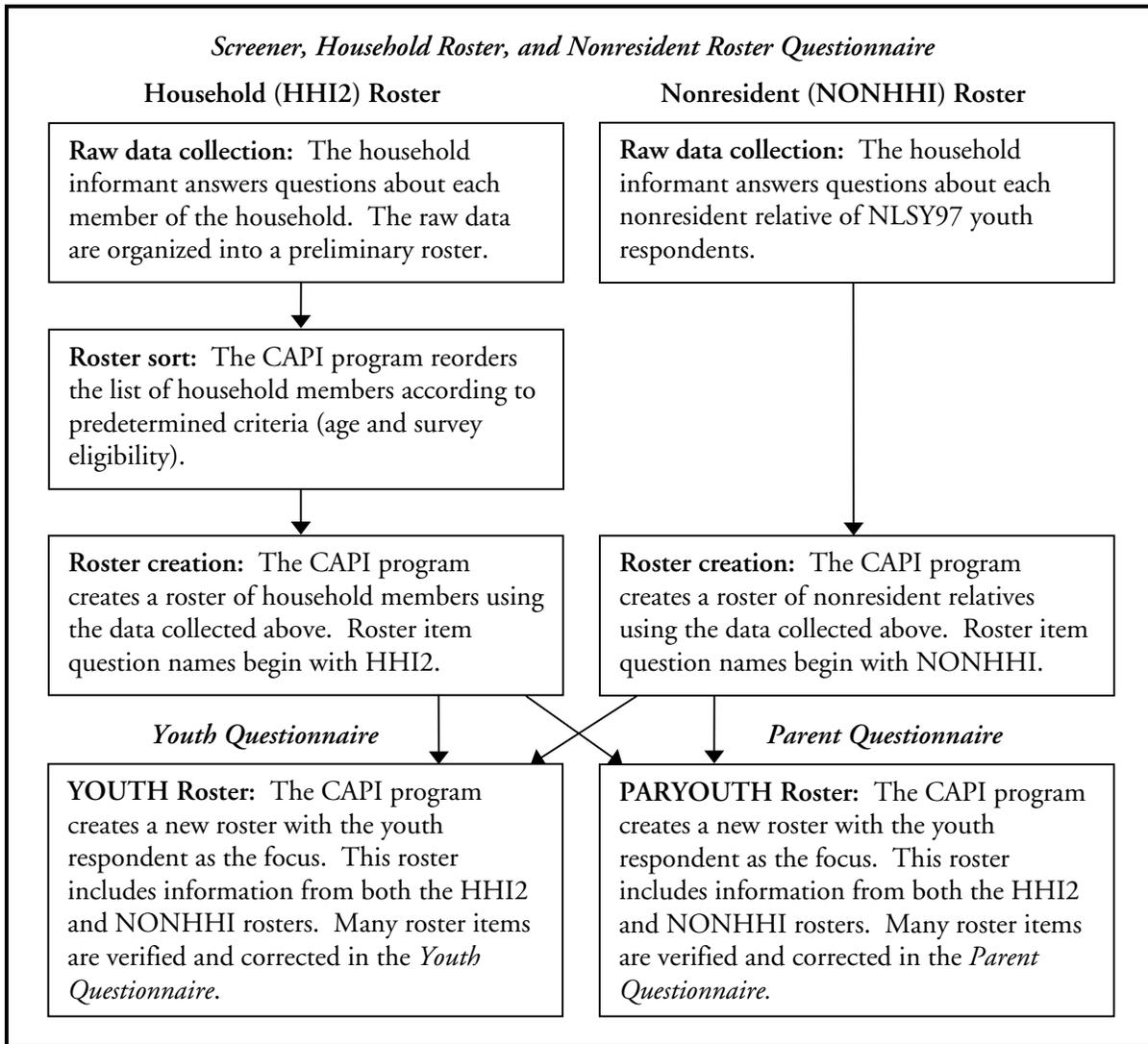
Two additional rosters were then created using household and nonresident data provided by the household informant. These youth and parent rosters essentially modify the household and nonresident rosters with the NLSY97 youth respondent as the focus. For example, an item on both new rosters identifies the line where each NLSY97 youth’s biological mother is located. The HHI2 data are not reused because there

may be more than one youth and more than one responding parent in a given household, each requiring their own roster (see section 1.3 of this guide for details about multiple respondent households).

The youth roster includes information specific to the NLSY97-eligible youth, as well as some data collected regarding the youth's parents. Many of the roster items are later verified and corrected if necessary during the youth interview. For example, youths are asked if their age as reported by the household informant is correct, and they provide the correct age if the information is inaccurate. All items on the youth roster have question names beginning with "YOUTH_." Similarly, the parent roster is created using *Screener, Household Roster, and Nonresident Roster Questionnaire* data about the youth respondent and the responding parent; it is updated during the parent interview. The question names of parent roster items begin with "PARYOUTH_." The parent roster was created only if a parent interview was conducted.

Figure 2 provides a pictorial representation of how the rosters described above are created and used.

4.6.5 Figure 2. Creation of Round 1 Rosters Based on Screener Data



Users should be aware that much of the information contained in the rosters may appear in the data set more than once. As Figure 2 suggests, data will first be included at the point in the interview when the information was actually collected. For example, screener question SE-28 asked the household informant for the date of birth of each household member. After all the raw data had been gathered, the computer sorted all the answers and created the rosters described above. If there were errors in the original answers and the youth respondent or responding parent provided corrected information, the roster items were often changed to reflect the most up-to-date information. Additionally, because the data were sorted before the creation of the roster, the ID number listed for a person in the screener questions does not necessarily identify the same person as the ID number in the household roster and elsewhere during the interview. To associate screener information with household roster information, researchers must use

variables R10978.–R10993., which provide a crosswalk between the two sets of ID numbers. This process can be avoided by using the roster items rather than the raw interview data.

Data hint →

Survey staff strongly recommend that researchers use the roster information whenever possible, as it is more accurate and easier to use than the raw data. Roster items are presented as consolidated blocks of data in the data set and can be identified through their unique question names as described above.

In previous rounds, some questions such as screener question SE-28 were included in the data, but they have now been removed. Survey staff continue to eliminate duplicate sources of information in the data.

Using the household roster. The household roster contains the data described above for each household member and organizes it in a matrix form for use by researchers. A key variable in the household roster is the ID number of the household member (HHI2_ID.xx—the “.xx” indicates that this variable is repeated for each household member, beginning with HHI2_ID.01, HHI2_ID.02, and so on). This variable identifies the line number of the household member on the roster. For example, if the NLSY97-eligible youth is listed first on the roster (that is, has a value of 1 for the variable HHI2_ID.01), then all other HHI2 variables that refer to household member 1 contain information about the youth. If the youth’s father is second on the roster, or has a value of 2 for the variable HHI2_ID.02, then his information is presented in the HHI2 variables referring to household member 2.

Users should be aware that the ID numbers, or line numbers, were assigned by the computer in a specific order. The household informant reported information about household members in no particular order. After the raw data were collected from the household informant, the computer first identified youths eligible for the NLSY97 and put them at the top of the list of residents. If there was more than one eligible youth, the respondents are listed from oldest to youngest. No household has more than five youth respondents, so no youth respondent has a household ID number higher than 5. After listing the NLSY97-eligible youths, the computer sorted everyone else in the household from oldest to youngest. Therefore, if an older relative such as a grandparent lived in the household, he or she will be listed next, followed in many cases by the youth’s parents and then any siblings not eligible for the survey.

The relationship variables on the round 1 household roster provide information about the relationship of every household member to every other household member. For example, consider a household with three members: the respondent (ID number 1), his father (ID number 3), and his grandmother (ID number 2). The resulting relationship variables are depicted in Table 1. In later rounds, relationship variables indicate the relationship of each household member to the respondent but not relationships between household members.

4.6.5 Table 1. Example Structure of the Round 1 HHI2 Relationship Data

Line (ID) number	1	2	3
1	HHI2_REL1.01 (relationship of 1 to 1): identity ¹	HHI2_REL1.02 (relationship of 1 to 2): grandson	HHI2_REL1.03 (relationship of 1 to 3): son
2	HHI2_REL2.01 (relationship of 2 to 1): paternal grandmother	HHI2_REL2.02 (relationship of 2 to 2): identity ¹	HHI2_REL2.03 (relationship of 2 to 3): mother
3	HHI2_REL3.01 (relationship of 3 to 1): father	HHI2_REL3.02 (relationship of 3 to 2): son	HHI2_REL3.03 (relationship of 3 to 3): identity ¹

¹ A code of "identity" for these variables indicates a relationship of "self."

By sorting through the relationship variables, researchers can identify all people in the household with a particular relationship to each other. For example, a user might want to count the number of children of the oldest woman in the household. After identifying which household member is the oldest woman, the researcher can look at each of the relationship variables for that member and see which have a code of son or daughter. If the woman has an ID number of 3, the researcher would write a program that checked the variables for the relationship of member 3 to member 1, member 3 to member 2, and so on. Each member with a code of 49 (daughter) or 50 (son) is a biological child of the woman.

The household roster also includes variables that identify specific types of relationships among household members. For each household member, these variables provide the ID number of that person's biological mother (HHI2_MOMID), biological father (HHI2_DADID), and spouse or partner (HHI2_SPOUSEID or HHI2_PARTNERID) if they also live in the household.

In addition to the set of variables indicating the relationship of each household member to every other person in the household, the round 1 household roster includes a set of variables called HHI2_RELY. These variables provide the relationship of each person in the household to the youth respondent, eliminating the need for the detailed programming described above if the youth respondent is the person of interest. For example, to determine how the youth respondent is related to person 2, researchers can look at the variable HHI2_RELY.02. In the case presented in table 1 above, this variable would have a value of 4, indicating that household member 2 is the youth respondent's father.

User Notes: Users should note that a number of inconsistencies were discovered during a review of the round 1 relationship data, and the relationship codes in the household roster were substantially revised for the release of the round 2 data. Relationships involving NLSY97 youth respondents were given top priority; some relationships between other household members were updated in the process. Survey staff place greater confidence in the accuracy of codes for relationships involving youth respondents.

Additional corrections were made to relationship and other variables on the household roster for the round 3 Event History release. Researchers should use the most recent data before drawing conclusions in their research. Information about the specific variables affected is provided in the NLSY97 errata notices; contact NLS User Services for details.

Linking the rosters to other data from the same round. Most research requires linking variables from the household and nonresident rosters to other data collected during the parent and youth portions of the survey. This section describes how to identify the youth respondent, responding parent, and household informant, as well as the steps necessary to identify other key relatives of the youth and responding parent.

The youth respondent can be identified by using R05334. (YOUTH_ID.01). This variable provides the line number of the youth respondent on the round 1 household roster. For example, if the value of R05334. is 1, the youth's ID number for the household roster is 1. All information about household member 1 on the roster pertains to the youth. If the value of R05334. is 2, then information about household member 2 pertains to the youth, and so on. As noted above, no NLSY97 youth respondent has an ID number higher than 5, so researchers who just want youth information will only need to examine data for the first five members on the household roster.

The parent roster (PARYOUTH) also contains a variable with the youth ID number on the household roster. However, researchers are advised to use the youth roster variable because it was created during the youth interview.

Identification of the responding parent requires a similar process. Researchers should use variable R07350. (PARYOUTH_PARENTID), which gives the ID number of the parent selected to be the responding parent at the end of the screener interview. Users should note that the youth roster also contains an ID variable called "ID of R 01 Resp Parent." However, because this variable is based solely on the screener and does not contain any updated information from the parent questionnaire, researchers are advised not to use this variable to identify the responding parent. (For information about identifying the responding parent's spouse or partner, refer to section 4.6.7, "Parent Characteristics.")

Finally, the household informant is fairly easy to identify. Variable R05381. (INFORMANT!ID) provides the ID number of the informant. As with the parent and youth, this number is the position of the informant on the household roster.

Researchers also may want to identify key relatives of the NLSY97 youth, particularly the youth's parents, even if they were not respondents to any part of the survey. To facilitate this process, the round 1 data

include identification variables that indicate the ID number of a given person on the household and nonresident rosters. Table 2 lists key youth roster ID variables available in the round 1 data set.

4.6.5 Table 2. Round 1 ID Variables for Key Relatives of NLSY97 Respondents

Reference number	Question name (all begin with YOUTH)	Description
R05318.	_ADOPDADID.01	ID # of youth's resident adoptive father
R05319.	_ADOPMOMID.01	ID # of youth's resident adoptive mother
R05323.	_DADID.01	ID # of youth's resident biological father
R05327.	_FOSTDADID.01	ID # of youth's resident foster father
R05328.	_FOSTMOMID.01	ID # of youth's resident foster mother
R05336.	_MOMID.01	ID # of youth's resident biological mother
R05339.	_NONR1ID.01	ID # of youth's 1 st non-responding resident parent
R05344.	_NONR2ID.01	ID # of youth's 2 nd non-responding resident parent
R05350.	_NRDADID.01	ID # of youth's nonresident biological father
R05351.	_NRMOMID.01	ID # of youth's nonresident biological mother
R05358.	_SPOPARID.01	ID # of youth's resident spouse or partner
R05359.	_STEPDADID.01	ID # of youth's resident stepfather
R05360.	_STPEMOMID.01	ID # of youth's resident stepmother

For instance, the youth respondent answers questions about his or her biological mother. If researchers want to examine the characteristics of the biological mother contained in the household roster or nonresident roster (depending on her residence), they would first look at the MOMID variable. If the mother lives in the household, this variable will have a valid value. For example, a value of 5 means that all the roster variables for household member number 5 contain information about the mother. If there is no positive value for this variable, the next step is to look at the NRMOMID variable to obtain the position of the biological mother on the nonresident roster. As with the household roster, the value of this variable indicates the ID number of the biological mother on the nonresident roster.

User Notes: Researchers should be aware that the ID variables described above were substantially revised for the release of the round 2 data. The comparable variables in the round 1 data should **not** be used for analyses. Additionally, researchers may notice that similarly titled variables are contained in the parent roster (PARYOUTH). However, the youth variables (YOUTH) are more accurate because they were adjusted to reflect corrected relationship data. Therefore, researchers are advised to use the youth roster variables to identify the youth's parents.

Linking individuals on the rosters across survey rounds. Many researchers want to examine changes in the youth’s household over time. Each round includes the collection of information about members of the youth’s household; as in round 1, these data are organized in a roster format. Users should be aware that individuals will not necessarily remain at the same place on the roster in different rounds. For example, a father who had a line ID number of “3” in round 1 might move to position 2 or 4 in round 2. Therefore, each household member is assigned a second, separate identification code, called a unique ID (UID). This unique ID will remain constant across survey rounds, even if household members move to a different place on the roster, so that researchers can identify a given household member in more than one round. For household members in round 1, UIDs are contained in questions HHI2_UID.01–HHI2_UID.16 on the household roster.

As mentioned above, individuals may move between the household roster and the nonresident roster. If such a move occurs, the person keeps the same UID number. Researchers can use the UID number to track individuals as they move in and out of the household. However, users should be aware that no updated information is collected about nonresidents other than biological, step-, or adoptive parents after round 1 (see section 4.6.6, “Nonresident Characteristics,” for more information).

The UID number can be used to identify the round in which the household resident was first reported. UIDs that begin with “1” indicate that the person was first reported in round 1 on the household roster. An initial digit of “2” means that the person was first reported on the round 1 nonresident relative roster. If the UID begins with “1998,” then the member was first listed on the household roster in round 2, and “1999” indicates a first appearance on the round 3 household roster and so on. For example, a respondent’s round 5 household roster might contain the people in Table 3.

4.6.5 Table 3. Examples of UID Codes

Line number	Unique ID	Round first reported
1	102	Round 1 Household Roster
2	202	Round 1 Nonresident Roster
3	199801	Round 2
4	199901	Round 3
5	200001	Round 4
6	200101	Round 5

Comparison to Other NLS Surveys: Information on the respondent’s household is available for all cohorts for most survey years. Data generally include the age, gender, relationship to the respondent, and educational attainment of all occupants; the enrollment status of those of school age; and the occupation and weeks worked of residents age 14 and older. In the pre-1980 surveys of the Original Cohorts, data

were generally collected only for family members living in the respondent's household and not for unrelated household members. For more complete information, refer to the *NLS Handbook* or the appropriate cohort's *User's Guide*.

Survey Instruments: These questions are found in the screener and household roster sections of the round 1 *Screener, Household Roster, and Nonresident Roster Questionnaire* and section P5 of the round 1 *Parent Questionnaire*. In rounds 2–5, they are found in the household information section (YHHI) of the *Youth Questionnaire*.

<i>Related User's Guide Section</i>	<i>Main Areas of Interest</i>	<i>Supplemental Areas of Interest</i>
4.6.4 Household & Neighborhood Environment	Created Variables Household Characteristics	Screener Extended Screener Household

4.6.6 Nonresident Characteristics

In round 1, the household informant provided information about key relatives of the youth who lived outside of the household. In subsequent rounds, the household information section of the *Youth Questionnaire* collected limited information about biological, step-, or adoptive parents who lived outside of the youth’s household. This section describes only *Screener* and *Youth Questionnaire* information. The round 1 *Parent Questionnaire* gathered additional background information about nonresident parents; these data are described in section 4.6.7, “Parent Characteristics.”

Screener, Household Roster, and Nonresident Roster Questionnaire (round 1). This section describes the data available from the nonresident roster portion of this instrument. It gathered the relationship and status of the youth’s nonresident immediate family, including biological, adopted, or stepparents; siblings; spouses; partners; or children. The demographic data collected for these family members are summarized in Table 1 below. Please note that some characteristics, such as gender and marital status, may be inferred from the nonresident’s relationship to the youth.

4.6.6 Table 1. Data Collected for Nonresident Relatives in Round 1

Relationship	Gender	Race/ Ethnicity	Age	Marital Status	Highest Grade Level Attained	Highest Degree Attained ¹	Employment Status in 1996
Biological parent		✓	✓	✓	✓	✓	✓
Adoptive parent ²		✓	✓		✓	✓	✓
Stepparent ²					✓	✓	
Full or half sibling ³	✓		✓		✓	✓	
Spouse of youth					✓	✓	✓
Children of youth ⁴	✓				✓	✓	
Parent of youth’s children				✓			

¹ This information was only collected if that relative was age 17 or older.

² Information on step- or adoptive parents was collected only if the parent lived with one of the youth’s biological parents.

³ Data on full- or half-siblings were collected only if they do not live with either of the youth’s biological parents.

⁴ This information was only collected if the youth is age 14 or older.

Data hint →

During the interview, the survey program organized the informant’s answers to questions about nonresident relatives into a “nonresident roster.” This grid of data is the best source of information about these relatives and should be used instead of the raw data in analyses. Details about the creation and use of the nonresident roster are provided at the end of this section; general information about rosters is available in section 3.2.

Additional information about a nonresident biological parent includes the number of miles the youth lived from that parent, the members of that parent's household, and the relationship of each member of that parent's household to the youth. For each nonresident biological child of the youth, the survey collected information on the person the child lived with (e.g., other biological parent, foster parent, adoptive parent, other) and the distance between the child's current residence and the youth's current residence. If nonresident children lived with the other biological parent, the respondent provided additional data on the individuals the other parent lived with (e.g., no other adults, other biological parent's spouse, other biological parent's partner, other biological parent's parents, other relatives).

If the youth's biological parent, full sibling, or child was reported as being deceased, additional information on his or her gender, the year he or she died, and his or her age at death was also gathered.

User Notes: Significant corrections were made to variables on the non-resident roster for the round 3 Event History release. Researchers should use the most recent data before drawing conclusions in their research. Information about the specific variables affected is provided in the NLSY97 errata notices; contact NLS User Services for details.

Youth Questionnaire (rounds 2–5). Although the full set of nonresident questions was not repeated after round 1, subsequent questionnaires collected limited information on biological, step-, and adoptive parents who resided elsewhere. This section focused mainly on whether the youth lived with the nonresident parent at any time between the youth interviews. A series of questions also established whether any nonresident parents shared custody of the youth. Marital status of nonresident parents is collected in round 5 in the Youth Household Information section. The variable indicating the relationship of the parent is PREVNONHHIPARENT_RELY.

User Notes: After data on nonresident relatives were collected during the administration of the round 1 *Screener, Household Roster, and Nonresident Roster Questionnaire*, the information was organized into a nonresident relative roster. This roster was maintained in subsequent rounds even though no additional questions were asked about most nonresident relatives. The following paragraphs contain a description of the collection and organization of this information and explain how the roster functions in rounds after the initial interview. Before using these data, researchers should first read this discussion. Researchers may also want to refer to the user notes in section 4.6.5, "Household Composition," which contain information helpful in using the nonresident roster, and to section 3.2, which contains general information about NLSY97 rosters.

To organize the information about household residents and nonresident relatives, two rosters were created during the administration of the *Screener, Household Roster, and Nonresident Roster Questionnaire*. The first, the household roster, included information for all residents of the respondent’s permanent household at the time of the round 1 interview. For easy identification, all variables on the household roster were assigned question names that begin with “HHI2_.” The second key roster is the nonresident roster, which presents information about the youth’s nonresident relatives. All question names of nonresident roster variables begin with “NONHHI_.”

Two additional rosters, described in detail in section 4.6.5, were then created using screener data provided by the household informant, modified with the youth respondent as the focus. The YOUTH roster was used during the youth interview, and the PARYOUTH roster was used during the parent interview (if a parent interview was conducted). Figure 2 in section 4.6.5 provides a pictorial representation of how the rosters described above are created and used.

Users should be aware that much of the information contained in the rosters actually appears in the data set more than once, first at the point in the interview when the raw data were actually collected and then in the roster. Survey staff strongly recommend that researchers use the roster information whenever possible, as it is more accurate and easier to use than the raw data. Roster items can be identified through their unique question names as described above.

Using the nonresident roster. The nonresident roster is similar to—but somewhat simpler than—the household roster. Although the roster is not sorted in any particular manner, each person about whom information is collected appears at the same point in the roster for each item. Thus, researchers can easily identify the key characteristics of a given nonresident relative. Different information was collected about nonresident relatives depending on their relationship to the youth, as depicted in Table 1 above.

Researchers often want to identify the respondent’s biological mother and father so that they can use the information collected in the nonresident roster and connect it to data collected in other sections of the survey. To facilitate this process, the round 1 data include identification variables that indicate the ID numbers of the nonresident mother (R05351.) and father (R05350.) on the roster. For example, if the youth’s biological mother does not reside in the youth’s household, the nonresident mother ID variable provides the position of the biological mother on the nonresident roster. If the value of this variable is 3, then all nonresident roster variables referring to person 3 contain information about the youth’s mother.

Data hint →

Users should note that these variables were substantially revised for the release of the round 2 data and again for the round 3 data; they should **not** use the variables from the round 1 data set in analyses.

User Notes: Although questions about nonresident relatives were not asked in rounds 2–5, the nonresident roster remains part of the data for these rounds. This roster does contain updated information about the respondent’s nonresident biological, step-, or adoptive parents. In addition, anyone who was previously a member of the youth’s household but who left that household is moved to the nonresident roster. This permits the unique ID code (UID) of that person to be maintained across rounds. If the person later returns to the respondent’s household, he or she will be moved back onto the household roster. The UID code can also be used to track nonresident parents across rounds if their information is updated.

Comparison to Other NLS Surveys: For the Children of the NLSY79, data are available about the education, occupation, and geographic residence of the child’s nonresident biological father. Two surveys of the NLSY79, two surveys of the Mature Women, and six surveys of the Young Women have collected demographic information on the respondent’s nonresident siblings. For more complete information, refer to the *NLS Handbook* or the appropriate cohort’s *User’s Guide*.

Survey Instruments: Questions are found in the nonresident roster section of the round 1 *Screener, Household Roster, and Nonresident Roster Questionnaire*.

<i>Related User’s Guide Sections</i>	<i>Main Area of Interest</i>	<i>Supplemental Areas of Interest</i>
4.6.5 Household Composition	Non-Res. Characteristics	Parent Current Status
4.6.7 Parent Characteristics		Parent Retrospective
4.9 Marital History & Fertility		Screener Non-Resident

4.6.7 Parent Characteristics

This section details information collected from the responding parent in the round 1 *Parent Questionnaire*, as well as data gathered in the *Youth Questionnaire* and *Household Income Update* in subsequent rounds. First, several types of background information provided by the responding parent are described. Data on the status of NLSY97 respondents' parent-figures at the interview date are then detailed. The third subsection describes the retrospective information collected about the responding parent.

Background

This subsection of the *Parent Questionnaire* (round 1) first established the responding parent's origin or descent, birth date, and place of birth. If the parent was born in the United States or a U.S. territory, a follow-up question asked for the state or territory of birth. Parents not born in the United States were questioned on their country of birth and the year they first came to live in the United States.

User Notes: Researchers should note that the commonwealth of Puerto Rico is listed as both a country (P2-002) and as a U.S. territory (P2-007). In addition, Quebec is listed as a country separate from Canada. The United Kingdom of Great Britain and Northern Ireland is not listed; respondents specify England, Northern Ireland, Scotland, or Wales.

The survey also collected information about the responding parent's number of siblings, the religion in which he or she was raised, and the year of birth and highest grade level completed by his or her parents. Further questions focused on whether this parent lived with both biological parents while growing up, and if not, the parent he or she did live with, the reason he or she did not live with both parents, and the ages that he or she lived apart from one or both parents. The type of community in which the responding parent lived when he or she was 14 was also gathered. The responding parent reported this same information for his or her current spouse or partner. An additional question determined the number of times the responding parent's current spouse or partner had been legally married. If the NLSY97 respondent had had contact with a nonresident biological parent since the age of 10, the same set of questions was asked about that nonresident parent.

Status

The responding parent was questioned about his or her present employment status, marital status, and highest grade completed, as well as whether any language other than English was spoken at home. These data were also provided for the responding parent's spouse or partner. Additional questions about religion were addressed to parents of youths born in 1982, 1983, and 1984. In the self-administered section of the questionnaire, these parents responded to the following statements about their religiosity:

- I don't need religion to have good values

- The Bible/Koran/Torah/religious teachings should be obeyed in every situation
- I often ask God to help me make decisions
- God has nothing to do with what happens to me personally
- I pray more than once per day

Created Variables. A scale created by Child Trends, Inc., is based on the religious beliefs and practices questions in round 1 and indicates the religiosity of the responding parent (FP_PPRELIG). See *Codebook Supplement Appendix 9*, available from NLS User Services, for more information.

Attitudes: In the self-administered section of the *Parent Questionnaire* (round 1), responding parents of youths born in 1982, 1983, and 1984 expressed their level of agreement with a series of statements about their attitude toward themselves. The same questions were asked of the youth respondent, as described in section 4.5.2, “Attitudes.” Data on the responding parent’s attitude towards his or her partner or spouse and the youth’s view of their relationship are also described in section 4.5.2.

Health: The NLSY97 collected information on the responding parent’s and his or her current spouse’s/partner’s general health. Data were gathered on the presence of any long-term health problem that limited the type or amount of employment possible. The parent then reported any serious long-standing medical condition (e.g., severe heart problem, cancer, potentially life-threatening disease) since the birth of his or her oldest child. The parent’s height and weight were also collected. If the NLSY97 respondent was not currently living with his or her biological parents but had been in contact with them since age 10, the responding parent was asked about the height and weight of those parents. In multiple respondent households, these questions were asked if the oldest respondent who was a child of the biological parents had had contact since age 10.

Income and Assets: In round 1, information was gathered on the responding parent’s earnings from a job—including military service—and from a farm, business, or professional practice during the 1996 calendar year. Similar data were collected for his or her current spouse or partner. Additional questions gathered information about the combined income received from other sources by the responding parent and the spouse or partner during 1996. Sources listed were interest or dividend income, including any amount that has been reinvested or credited; child support payments; and other income. Finally, data were collected on amounts received from various government programs in 1996.

For a more comprehensive picture of the family’s financial situation, the responding parent was also asked about assets and debts. The parent first stated whether the family owned or rented their dwelling (e.g., ranch or farm, mobile home, house or apartment). After determining the percentage of the dwelling and the land owned by the family, a follow-up question asked about the present value of these holdings. The

survey then established whether there was a mortgage, land contract, or any other type of loan that used the property as collateral, such as a second mortgage or a home equity loan, and the amount still owed. If the family neither rented nor owned the dwelling, the living situation was determined (e.g., housing part of job compensation, temporary living arrangements while house is under repair).

Additional asset information included the amount that would be received (less any outstanding debts) if any business partnership/professional practice or any real estate, such as a second home or land, was sold. The responding parent also stated whether he or she owned each of the assets listed in Figure 1; a follow-up question ascertained the total dollar value of each asset reported. Finally, to balance information on assets, the responding parent reported the amount still owed on any vehicles, any educational loans for a child, and other debt including credit cards or bank loans.

4.6.7 Figure 1. Assets Held by the Responding Parent

Bills and bonds including CDs, government savings bonds, or treasury bills
Checking, savings, or money market accounts
Educational IRA accounts or other prepaid tuition savings accounts
Other assets including money owed by others, life insurance policies, and precious metals
Pension or retirement savings
Shares in publicly-held corporations or mutual funds
Vehicles

User Notes: Researchers using income and asset data should be aware of survey practices with respect to item nonresponse and topcoding high asset values. To reduce the impact of item nonresponse and capture at least some information, responding parents who don't know or refuse to provide the exact value for an income or asset question are asked to select the appropriate range from a predefined list. Topcoding of parent and household income information uses the same system as youth income variables, as discussed in the introduction to section 4.8 of this guide.

Created Variables. Several created variables summarize the household's financial situation. First, the net worth of the household is calculated by subtracting total debts from total assets. If the youth was independent, this calculation is based on information provided by the youth and is reported in the variable CV_HH_NET_WORTH_Y; if the youth was not independent and a parent was interviewed in round 1, the calculation is based on parent data and the results reported in CV_HH_NET_WORTH_P. The second variable, CV_INCOME_GROSS_YR, provides gross household income in the previous year. Although only one variable is created, the data for the variable, like net worth, could be taken from the youth questionnaire if the respondent was independent and from the parent questionnaire if the youth was

not independent. The variable CV_HH_INCOME_SOURCE identifies whether the income variable was created from youth or parent information. Finally, a poverty status variable (CV_HH_POV_RATIO) creates a ratio comparing the gross household income variable to the federal poverty level for the previous year, taking household size into account. After round 1, all household income and net worth information are based on data provided by the youth respondent (if he or she is independent). See section 4.8.1, “Assets & Debts,” and 4.8.2, “Income,” for details.

Household Income Update. The surveys following round 1 include this form, administered to one of the respondent’s parents, which is designed to gather basic income information concerning the respondent’s parent and his or her spouse/partner in the absence of a detailed parent questionnaire. All respondents who live with a parent at the interview date are eligible to have this questionnaire administered. It collects the parent’s total pre-tax income from wages, salaries, commissions, and tips during the past calendar year; the same data for the parent’s spouse or partner; and the total pre-tax amount of any other income received (i.e., farm or business income, inheritances, child support, government programs). In round 2, parents of 7,601 respondents answered at least one question from the *Household Income Update*; parents of 5,488 respondents answered at least one question in round 3; and 5,225 parents of respondents answered at least one question in round 4. In round 5, parents of 4,090 respondents answered at least one question. The administration of this one-page paper instrument is described in section 2.2 of this guide.

User Notes: The *Household Income Update* was administered only once in households with more than one NLSY97 respondent. In round 2, the information provided by the parent was attached to the records of all the NLSY97 respondents who lived in the round 1 household. This means, for example, that if one respondent had left the household to live with a different parent between rounds 1 and 2 but the siblings remained in the original household, that respondent would have the same *Household Income Update* information that was provided by the parent in the original household. In round 3 and subsequent rounds, the instrument reflects the youth’s actual household at the time of that survey.

Youth Questionnaire. In rounds 2–5, this instrument included basic questions about the income of the respondent’s parents. These separately referred to the youth’s biological mother, biological father, nonbiological mother, and nonbiological father as applicable. The youth was asked for that parent’s total income from a job, farm, business, or professional partnership. This information, along with the parent’s responses to the *Household Income Update*, may allow researchers to compare a youth’s perception of parental income with the amount the parent reports.

History

In the round 1 parent calendar section, retrospective data were collected on the parent's employment and marriage histories and history of participation in government programs for low-income households (e.g., AFDC or ADC, SSI, food stamps).

Employment: The round 1 NLSY97 collected a history of the responding parent's spells of employment since the parent's 18th birthday or the youth's date of birth, whichever was earlier. In this context, a spell was defined as three or more months of employment without a break of six or more months, regardless of employer. After establishing the start and stop dates for each spell, the responding parent reported the usual number of hours worked per week during each spell. The responding parent also provided work history information for each spouse who lived with the oldest NLSY97 youth during his or her marriage to the responding parent. The survey collected information on whether the spouse was working at the time of the marriage or for a period of three consecutive months or more during the marriage. If these conditions were met, the responding parent provided information on the spouse's spells of employment and the average number of hours per week worked during each spell.

Marriage: Questions on the responding parent's marital history collected information on the length of each of the responding parent's marriages. In addition, the survey gathered data on any changes in the marital status (e.g., legal separation) during each marriage.

Program Participation: A series of questions recorded the number of years during the previous five years that a responding parent participated in various government programs targeting low income households (e.g., AFDC or ADC, SSI, WIC).

Created Variables. Although no questions in the round 1 survey directly collected a fertility history from the responding parent, two created variables based on information from the household and nonresident relative rosters relate to this part of the parent's personal history. These created variables provide the age of the youth's biological mother when she gave birth to her first child (CV_BIO_MOM_AGE_CHILD1) and the mother's age when she gave birth to the youth respondent (CV_BIO_MOM_AGE_YOUTH).

User Notes: Users should exercise caution when drawing conclusions based on parent characteristics in the NLSY97 data. The large number of multiple respondent households in the sample may skew the data on certain characteristics if the analysis is performed at the respondent level rather than at the parent level. See section 1.3, "NLSY97 Sample," for more information on multiple respondent households.

Basic demographic information about the responding parent and his or her spouse or partner is also available in the household roster. The content and structure of the household roster are discussed in detail in section 4.6.5, “Household Composition.” The current user note describes how to identify the responding parent’s spouse or partner so that household roster information may be linked to other data about the spouse or partner.

To find the responding parent’s spouse or partner on the household roster, researchers must use the HHI2 relationship variables. Researchers should begin this process by determining the ID number of the responding parent (R07350.). The value of this variable indicates the position of the responding parent on the roster. The next step is to look at the household relationship variables for the responding parent. Codes of 1 (wife), 2 (husband), and 69 (partner) indicate that the household member is the responding parent’s spouse or partner. For example, if the responding parent is number 3, the user would examine the variables for the relationship of member 3 to member 1, member 3 to member 2, member 3 to member 4, and so on. If the variable for member 4 had a code of 2, then member 4 would be the responding parent’s husband.

Comparison to Other NLS Surveys: Data provided by the respondent about his or her parents are available for all cohorts. For the NLSY79, information includes the parents’ birthplace, educational attainment, occupation, age, and life status. Information about the mothers of NLSY79 Children is quite detailed, as the mothers are main respondents for the NLSY79; basic demographic data on the fathers are available as well. Original Cohort respondents have provided information about the country of birth and life status of their parents, as well as the educational attainment and occupation of their parents during their teenage years. More recently, the Mature and Young Women also provided information about their parents’ health and income and about transfers of time and money to and from their parents. For more complete information, refer to the *NLS Handbook* or the appropriate cohort’s *User’s Guide*.

Survey Instruments: These questions are found in sections P2 through P6 of the round 1 *Parent Questionnaire*. The rounds 2–5 questions about parent income are located in the income (YINC) section of the *Youth Questionnaire* and in the *Household Income Update*, a brief questionnaire administered to one of the respondent’s parents.

Related User’s Guide Sections

- 4.5.2 Attitudes
- 4.6.5 Household Composition
- 4.6.6 Nonresident Characteristics

Main Areas of Interest

- Created Variables
- Family Process Measures
- Parent Background
- Parent Current Status
- Parent Family Background
- Parent Retrospective

Supplemental Areas of Interest

- Attitudes
- Household Characteristics
- Income
- Non-Res. Characteristics

4.6.8 Race & Ethnicity

Data on the respondent's race and ethnicity were collected in the *Screener, Household Roster, and Nonresident Roster Questionnaire* (round 1) and were based on the household informant's identification. Using the household roster variables, the survey program created KEY!RACE, which describes the respondent's race, and KEY!ETHNICITY, which identifies respondents of Hispanic origin. These variables can be combined to create a single race/ethnicity variable; however, there are a number of missing observations. Researchers may prefer to use a created variable, KEY!RACE_ETHNICITY, described in the user notes below. Table 1 in section 1.3 summarizes the racial/ethnic composition of the sample.

Created Variables. A round 1 created variable (CV_CITIZENSHIP) provides the respondent's U.S. citizenship status based on his or her parents' residence at the time of the respondent's birth. The variable KEY!RACE_ETHNICITY indicates respondents' race and Hispanic ethnicity.

User Notes: In the supplemental PSUs, the information from the *Screener, Household Roster, and Nonresident Roster Questionnaire* was used to determine whether a youth was eligible for inclusion in the black and Hispanic oversample. No non-black/non-Hispanic youths from the supplemental PSUs were included in the survey. For more information on sampling procedures, refer to section 2.1, "Sample Design & Screening Process."

To simplify the race/ethnicity identification process, survey staff created a single combined variable, KEY!RACE_ETHNICITY (R14826.). This variable is based on KEY!RACE, KEY!ETHNICITY, household roster information from the HHI2 variables, and biological parent race/ethnicity information. All respondents are classified as Hispanic, black, non-black/non-Hispanic, or mixed race/non-Hispanic; there are no missing values. While a respondent can be of Hispanic ethnicity and still be of any race, Hispanic ethnicity was given priority in the creation of this variable. Users who wish to identify, for example, blacks of Hispanic origin must create their own variable from the screener information.

Although the questionnaire contains response categories for white; black or African-American; American Indian, Eskimo, or Aleut; and Asian or Pacific Islander races, some household informants refused to select one race and identified household members only as "mixed race." This category was added in the codebook.

Comparison to Other NLS Surveys: Race is available for all cohorts; ethnicity is available for all cohorts except the Older Men and Young Men. Users should be aware that coding categories for race and

ethnicity have varied among cohorts and over time. For more complete information, refer to the *NLS Handbook* or the appropriate cohort's *User's Guide*.

Survey Instruments: This information was collected in the extended screener section of the round 1 *Screener, Household Roster, and Nonresident Roster Questionnaire*.

Related User's Guide Sections

4.6.5 Household Composition

4.6.6 Nonresident Characteristics

4.6.7 Parent Characteristics

Main Areas of Interest

Created Variables

Demographic Indicators

Supplemental Areas of Interest

Household Characteristics

Non-Res. Characteristics

Parent Background

4.6.9 Youth History

Parent Questionnaire (round 1). In the youth history section, the responding parent first reported the NLSY97 respondent's housing situation since his or her birth. If the youth lived with anyone else before the responding parent, the survey collected a history of living situations. This history recorded the relationship of each adult the youth lived with and in what month and year the situation changed. All round 1 responding parents were asked whether the youth had ever lived apart from them for three months or more (since the youth first lived with them). For each spell apart, the parent stated the place where or the person with whom the youth lived, whether the arrangement was still in effect, and the year the arrangement changed.

The survey also considered residential moves (not necessarily related to a change in living situation). These questions recorded the total number of different places the youth lived prior to his or her 12th birthday and since his or her 12th birthday.

For each NLSY97 youth not living with both biological parents, the survey asked whether the responding parent holds legal responsibility or legal custody for the youth. If only the youth's biological mother was listed on the birth certificate, follow-up questions asked whether the biological father had ever been legally identified (e.g., blood test, court ruling, signed legal document). For adopted youths, data were collected on whether the youth ever lived with his or her biological mother or father.

Screener, Household Roster, and Nonresident Roster Questionnaire (round 1). Information included the distance the youth lived from his or her biological parents (if they were alive) or the date that the youth's biological parents died.

Youth Questionnaire. In rounds 2–5, the respondent answered questions about residential moves to a different city, county, or state since the date of last interview. These data were used to update the total number of residences variable described below. These surveys also collected the periods the youth lived with each parent between interviews.

Created Variables. A single created variable (CV_TTL_RESIDENCES) calculates the total number of different residences in which the youth has lived since the age of 12. A series of round 1 variables, CV_YTH_REL_HH_AGE_x, reports the relationship of the youth to the primary adults in the household (e.g., both biological parents, biological mother, adoptive parent[s]) at ages 2, 6, and 12. Users should be aware that, on the round 1 release, these variables were named CV_HH_REL_AGE_x; they were renamed for subsequent releases to more accurately reflect the content of the variable. A similar variable, CV_YTH_REL_HH_CURRENT (formerly CV_HH_REL_BIRTH), provides the same information as of the interview date for each survey round. Finally, a series of round 1 created variables describing the

youth's geographic residence at age 12 indicates in which Census region the youth resided (CV_CENSUS_REGION_AGE_12), whether the youth lived in an MSA (CV_MSA_AGE_12), and whether the youth lived in an urban or rural area (CV_URBAN-RURAL_AGE_12). For more information about geographic variables, see section 4.6.3, "Geographic Indicators."

User Notes: Researchers should be aware that the geocoding procedure changed between rounds 3 and 4 of the NLSY97. All geocode variables from all rounds have been recreated using the new, more accurate method. Regardless of the round of interest, researchers should use data from a round 4 Geocode CD. For more information on geocoding procedures, refer to Appendix 5 in the main file *Codebook Supplement* or the introduction to the *Geocode Codebook Supplement* (the latter is distributed only with a Geocode CD).

Comparison to Other NLS Surveys: The *Childhood Residence Calendar*, a supplement to the 1988 NLSY79 survey, provided detailed retrospective data on the residential history of each respondent from birth to age 18. For more information, refer to the *NLS Handbook* or the *NLSY79 User's Guide*.

Survey Instruments: Questions are found in section PC8 of the round 1 *Parent Questionnaire* and in the nonresident roster section of the round 1 *Screener, Household Roster, and Nonresident Roster Questionnaire*. Residential move updates are collected in the youth information section (YINF) of subsequent youth questionnaires.

Related User's Guide Sections

4.2.4 School Experience

4.5.2 Attitudes

4.6.3 Geographic Indicators

Main Areas of Interest

Created Variables

Geographic Indicators

4.7 Health

The NLSY97 collects information on the general health of respondents; in round 1, selected respondents also provided data on their health practices and knowledge.

4.7 Table 1. Health Subtopics and Universe Restrictions

<i>NLSY97 User's Guide</i> Subtopics	Round 1 Universe (Age as of 12/31/96)	Rounds 2–5 Universe (Age as of 12/31/96)
Status	all ages	all ages
Practices and Knowledge	13	—

Status

Youth Questionnaire. All youths are asked to report their height and their weight and to state the level of their general health in every round. In the self-administered section, youths further describe their weight (very underweight, slightly underweight, about the right weight, slightly overweight, very overweight) and their current weight strategy (lose weight, gain weight, stay the same weight, not doing anything about weight). Youths also state whether they have entered puberty and their age at the time of onset.

In round 1, if the respondent did not live with a parent/guardian at the time of the survey, he or she was questioned on the source of any health insurance coverage that included physician or hospital care.

The round 4 survey included a question series concerning how often the respondent felt certain ways during the month before the interview date. Respondents reported the frequency of being nervous, feeling calm and peaceful, feeling downhearted and blue, being happy, and feeling so down in the dumps that nothing could cheer them up. Answer choices were all, most, some, or none of the time during that month. This series is located in the self-administered section of the questionnaire in questions YSAQ282B–YSAQ282G.

Parent Questionnaire (round 1). The round 1 parent interview provided additional information about the youth's general health. The responding parent reported any past or present medical condition(s) that limited the youth's ability to attend school regularly, to do regular school work, or to work at a job for pay. The survey specifically asked whether the youth suffered from any of the following conditions:

- Chronic health condition or life-threatening disease
- Learning disability
- Part of body missing or deformed

- Physical, emotional, or mental condition
- Trouble seeing, hearing, or speaking

For each health condition, the responding parent was asked how old the youth was when the condition was first noticed and whether the youth was currently limited by the condition. Data on the youth's health insurance coverage were also collected from the responding parent.

Additional information about the general health of the parent, his or her partner or spouse, and, in some cases, the youth's absent biological parents is described in section 4.6.7, "Parent Characteristics."

Practices and Knowledge

The round 1 NLSY97 interview included a series of questions, addressed to youths born in 1983, on respondents' health practices and knowledge. The health practices questions asked about the number of days in a typical school week that the youth ate breakfast, the number of days in a typical week the youth ate green vegetables or fruits, the number of days in a typical week that the youth engaged in exercise lasting 30 minutes or more, and the percent of the time that the youth wore a seatbelt.

Further round 1 questions assessed these youths' knowledge of pregnancy and sexually transmitted diseases. Two questions asked the youth to choose the most effective method of preventing pregnancy and sexually transmitted diseases—withdrawal, condom, or birth control pill. The youth also stated at which point in the female menstrual cycle he or she believed pregnancy was most likely to occur. Respondent answers to these questions are depicted in Table 2.

4.7 Table 2. Knowledge about Pregnancy and STDs among Youths Born in 1983 (Round 1)

	Male	Female	Total
Which of these methods is the most effective for preventing pregnancy?			
Withdrawal	283	256	539
Condom	379	271	650
Birth control pill	204	279	483
Which of these methods is the most effective for preventing sexually transmitted diseases like AIDS or gonorrhea?			
Withdrawal	258	218	476
Condom	588	559	1147
Birth control pill	34	32	66
When during the female monthly cycle of menstrual periods is pregnancy most likely to occur?			
Right before the period begins	104	165	269
During the period	107	88	195
About a week after the period begins	84	59	143
About two weeks after the period begins	57	84	141
Anytime during the month, makes no difference	85	148	233
Don't know	481	317	798

In addition, these youths were asked to state their opinions on whether smoking cigarettes contributed to getting heart disease or getting AIDS. The youths were also surveyed on whether having five or more drinks of alcohol once or twice per week increased the risk of damaging the liver, getting heart disease, getting arthritis, becoming addicted to alcohol, or harming an unborn child.

Parent Questionnaire (round 1). If a youth was born in 1983, the responding parent was questioned on his or her perception of the effects of drinking alcohol on the same health conditions listed in the youth portion of the survey (see above).

Comparison to Other NLS Surveys: Respondents in each cohort have answered questions about their health; however, the specific questions have varied widely as the health sections were modified to reflect the respondents' varying life cycle stages. Health insurance information has been collected from respondents in all cohorts except the Young Men. In 1984, NLSY79 respondents answered health knowledge questions about when pregnancy occurs; the NLSY79 children age 10 and over (10–14 in 1994 and 1996) have responded to this question each year since 1988. Users should refer to the *NLS Handbook* or the appropriate cohort's *User's Guide* for more precise information.

Survey Instruments: Questions in the *Youth Questionnaire* are found in the health (YHEA) and self-administered (YSAQ) sections. In the round 1 *Parent Questionnaire*, these questions are found in sections P6 and PC9.

<i>Related User's Guide Sections</i>	<i>Main Area of Interest</i>	<i>Supplemental Areas of Interest</i>
4.5.1 Alcohol Use	Health	Fertility and Pregnancy
4.5.4 Cigarette Use		Sexual Activity
4.5.8 Sexual Activity & Dating		Substance Use
4.6.7 Parent Characteristics		

4.8 Income, Program Participation & Assets

The NLSY97 survey collects a large amount of detailed information on the income of youth respondents, which complements the labor market data described in section 4.3, “Employment.” Some youths also answer questions on their assets, debts, and participation in government programs for low-income individuals and families. In addition, information is available for other members of the household, as described in section 4.6.5, “Household Composition,” and section 4.6.7, “Parent Characteristics.”

Table 1 summarizes the *NLSY97 User’s Guide* subtopics available in this section and any global universe restrictions affecting those subtopics. Users should note that certain questions or groups of questions within a subtopic may have additional restrictions. Please consult the NLSY97 questionnaire and codebook for more information on particular questions.

**4.8 Table 1. Income, Program Participation & Assets
Subtopics and Universe Restrictions**

<i>NLSY97 User’s Guide</i> Subtopics	Rounds 1–3 Universe	Rounds 4–5 Universe
4.8.1 Assets & Debts	all ages; independent	independent or turned 18 or 20 years old since last interview
4.8.2 Income	all ages	all ages
4.8.3 Program Participation	all ages; independent	all ages; independent

Note: See 4.8 Figure 1 for NLSY97 independence criteria.

User Notes: Many of the questions on these topics can be cognitively difficult or feel intrusive to respondents. To reduce the proportion of missing (“don’t know” or “refused”) data, respondents who do not provide exact dollar answers to questions are asked follow-up questions designed to elicit approximate information. For many income categories, these respondents are asked to select the applicable category from a predefined list of ranges. The approach for asset questions is slightly different: The initial question asks the respondent to provide an exact value, but if he or she is unable or unwilling to do so, interviewers are instructed to ask the respondent to define a range for the value using whatever values he or she feels are appropriate. If the respondent doesn’t know or refuses to provide either an exact value or a range, a follow-up question asks him or her to select the appropriate range from a predefined list. This will provide researchers with some information on income, asset, and debt amounts when the respondent is reluctant or unable to furnish an exact figure.

Because the NLSY97 is a public use data set distributed widely throughout the research and public policy communities, the survey takes extensive measures to protect the confidentiality of respondents. One aspect of assuring confidentiality is the “topcoding” of the highest income and asset values. In each survey round, income and asset variables that include high values are identified for topcoding. (For example, the wage and salary income variable is usually topcoded, but variables indicating the amount received from public assistance programs are not.) For each of these variables, the top 2 percent of reported values are topcoded and replaced with the mean of the high values. Calculating topcode values in this way allows statistics such as means to accurately reflect the status of the population under examination without violating respondent privacy. Notes in the codeblocks for topcoded income and asset variables provide more information about the exact calculations used to topcode each variable.

Independence criteria. In all rounds, certain criteria apply to all NLSY97 respondents. Youths are considered independent if they have had a child, are enrolled in a 4-year college, are no longer enrolled in school, are not living with any parents or parent-figures, or have ever been married or are in a marriage-like relationship at the time of the survey. Reaching the age of 18 is another criteria for independence, but the reference date for that age has varied between surveys and questionnaire sections. See Figure 1 below for more details.

4.8 Figure 1. NLSY97 Independence Criteria

Age Criteria by Survey Round		
Survey	Income and Program Participation	Assets
Round 1	No youths age 18	No youths age 18
Round 2	No youths age 18	No youths age 18
Round 3	1980 birth cohort eligible (age 18 as of 12/31/1998)	1980 birth cohort eligible (age 18 as of 12/31/1998)
Round 4	1981 birth cohort eligible (age 18 as of 12/31/1999)	Eligible if age 18 as of interview date
Round 5	1982 birth cohort eligible (age 18 as of 12/31/2000)	Eligible if age 18 as of interview date
Other Criteria		
<p>All rounds, regardless of age: The youth is independent if he or she has had a child, is enrolled in a 4-year college, is no longer enrolled in school, is not living with any parents or parent-figures, or has ever been married or is in a marriage-like relationship at the time of the survey.</p>		

4.8.1 Assets & Debts

Independent youths¹ are surveyed on their current asset holdings and the amount of debt they currently owe. The assets section of the questionnaire first determines if respondents rent or own their dwelling place. These questions are tailored to respondents living on a farm or ranch, in a mobile home, or in another living situation (e.g., house, apartment). Renters are then asked to state the amount they pay in rent and utilities (such as heat, electricity, and water) per month. In rounds 2–5, respondents who own their own dwelling were asked for more details, such as the percentage of the residence they owned, its value, and whether they have made improvements. Independent youths who were eligible to go through the assets section in round 1 numbered 429; this group increased to 1,835 in round 2 and grew to 3,105 in round 3. With the new age restrictions described below, 4,164 respondents were eligible in round 4 and 3,966 respondents were eligible in round 5.

As described in the introduction to this section, the eligibility requirements for answering the assets questions changed beginning in round 4:

Rounds 1–3. Respondents are asked assets questions if they are age 18 as of the previous calendar year reference date (12/31/1998 for round 3), or if they meet one of the other independence criteria (e.g., got married, had a child).

Data hint →

Rounds 4–5. Respondents are asked assets questions if they are newly independent or if they are age 18 as of the interview date. In addition, respondents are asked assets questions again in the first interview after they turn age 20.

Therefore, not all independent respondents will reach the assets section in each round. For example, in round 3 John Doe is age 16 and gets married; because he is considered newly independent, he answers the assets section of the questionnaire. In round 4, he is age 17 but does not yet meet the age requirement; he skips the assets section. In round 5, he is age 18 and now eligible to again answer the assets questions.

Respondents who report owning part of their residential property are asked to estimate the percentage of the property held by them and the current value of their holdings. The respondent is asked if the property is mortgaged and, if so, the amount still owed on that mortgage or land contract. Additional questions determine the amount of property debt. The survey first asks respondents if a friend or relative loaned money to buy (or remodel) the residence. Follow-up questions include the relationship of the lender to the youth, the amount of money each lender provided, and the amount the respondent still owes. Next, these respondents are asked if the property is secured by other loans such as second mortgages or home equity loans. If the respondent reports either type of loan, he or she is asked to state the amount of the

¹ To be considered independent, respondents must meet one characteristic from a list of criteria; see the introduction to section 4.8 for more information.

loan currently outstanding. Those who own property are surveyed on the amount they paid in property taxes in the past calendar year. A follow-up question asks for the average amount each respondent pays for utilities in a month.

After determining the respondent’s residential ownership status, the second part of this section surveys other asset holdings, such as a business or vehicle, as well as current debt. For each asset held or debt owed by the respondent, he or she reports the amount the asset is worth or the amount currently owed. If a respondent reports a loan of \$200 or more, follow-up questions determine the relationship of the lender to the respondent, the amount loaned, and the amount owed. Figure 1 lists NLSY97 asset and debt categories used in each round.

4.8.1 Figure 1. NLSY97 Asset and Debt Categories

Asset Categories	
Business, partnership, or professional practice	Stock
Other real estate (e.g., second home, rental real estate, or land)	Vehicles
Pension or retirement accounts (e.g., 401k, IRA, stock ownership)	Household furnishings
Savings accounts, checking accounts, money market accounts	Other assets
Certificates of deposit, savings bonds, treasury bills	
Debt Categories	
Vehicle debt	
Other debts	
Non-educational loans of at least \$200 from friends or relatives (up to 5 loans recorded)	

In rounds 2–5, additional questions referred to the spouse or partner of the youth. Respondents who were married or began living with a partner since the date of last interview were asked for the amount that would be left over if the spouse/partner sold their assets and paid off any debts. A second question, directed toward all respondents with a spouse or partner, asked whether that person had assets, savings, or investments that the respondent did not share. A follow-up question recorded the value of these assets.

Parent Questionnaire (round 1). The responding parent provided information on his or her assets and debts in round 1. These data are described in section 4.6.7, “Parent Characteristics.”

Created Variables. Based on the information provided by independent youths, a created variable calculates the net worth of the respondent’s household (CV_HH_NET_WORTH_Y), defined as the amount of total assets minus the amount of total debts. Asset values used in this calculation are property or money from estates; ownership of a ranch, mobile home, house, business, or second piece of real estate; a retirement pension; savings accounts; stocks; the value of owned vehicles and furniture, and other assets. Debt values used are a mortgage, a second mortgage, loans owed, amount owed on a car, and other loans.

A similar net worth variable is calculated based on parent data; see section 4.6.7, “Parent Characteristics,” for details.

User Notes: Researchers using asset data should be aware of survey practices with respect to item nonresponse and topcoding high asset values. These issues are discussed in the introduction to this section of the guide.

Comparison to Other NLS Surveys: Information on assets is regularly collected from each cohort (but only from NLSY79 Children age 15 and older). Users should note, however, that the assets included have varied widely over time and among cohorts. Data on the respondent’s debts have been collected from each cohort on a less regular basis. For more details, including specific years, consult the *NLS Handbook* or the appropriate cohort’s *User’s Guide*.

Survey Instruments: These questions are found in the income and assets section (YINC) of the round 1 *Youth Questionnaire* and in the assets section (YAST) of the *Youth Questionnaire* beginning in round 2.

<i>Related User’s Guide Sections</i>	<i>Main Areas of Interest</i>	<i>Supplemental Areas of Interest</i>
4.2.1 College Experience	Assets & Debts	College Experience
4.8.2 Income	Created Variables	Income Training

4.8.2 Income

Youth Questionnaire. The NLSY97 collects gross wage and salary data for the past calendar year from all respondents. In addition to income from a regular job, the surveys instruct the respondent to include income from odd jobs, temporary or seasonal work, and military positions. Information is also gathered on allowances during the past calendar year, including the amount received and periodicity of the allowance.

Independent youths¹ answer more extensive questions about income in the past calendar year in addition to wages and salary. First, they report any income received from self-employment, including a farm, business, or professional practice. Respondents who are married or living with a partner report the amount of wage and salary income the spouse/partner received, as well as the amount earned from any self-employment. A series of questions then records income received by the respondent and spouse or partner together from the following sources: child support; interest payments; dividend payments; rental properties; inheritance, trust, or annuity income; and other income received in the last calendar year. Some sources considered “other income” include Social Security, alimony, or survivor’s benefits. If respondents or their spouses/partners have any sources of income, they state whether they claimed an Earned Income Tax Credit (EITC) on their federal income tax return for the past calendar year.

Independent youths also report parental transfers in the past year. Aside from allowances and parental loans, respondents state the amount of money they received from each parent or guardian. Additionally, round 1 respondents were asked if they made regular payments to their parents during the previous year. For respondents living at home, this survey collected information on money paid for room and board.

Older youths answer questions about the income of their parents or guardians. In round 1, this group was composed of respondents age 14 and older as of 12/31/96; round 2 collected this data from youths who were age 14 or older as of 12/31/97. In rounds 3 and 4, all respondents were age-eligible for these questions. For each parent or guardian listed, the respondent is asked if that person received any earnings or income during the previous calendar year. Follow-up questions gather information on the amount the parent or guardian earned. Finally, independent respondents are asked about the income of all other individuals age 14 or older who currently reside in the household.

Parent Questionnaire (round 1). Data on the earnings and income of the youth in 1996 were collected from the responding parent. The survey also established the amount of financial support given to an independent youth by the parent.

¹ To be considered independent, respondents must meet one characteristic from a list of criteria; see the introduction to section 4.8 for more information.

The responding parent provided information about his or her own income, the income of his or her spouse, and the income of each household occupant age 14 or older. These data are described in section 4.6.5, “Household Composition,” and section 4.6.7, “Parent Characteristics.”

Household Income Update. In rounds 2–5, a parent of the respondent answered this short paper instrument concerning his or her income and that of his or her spouse/partner. See section 4.6.7, “Parent Characteristics,” for more information.

User Notes: Researchers using income data should be aware of survey practices with respect to item nonresponse and topcoding high income values. These issues are discussed in the introduction to this section of the guide.

Created Variables. If the respondent is independent, a variable based on the data he or she provided in the income section of the youth questionnaire (YINC) calculates total household income (CV_INCOME_GROSS_YR) in the previous calendar year. Several questions are combined to create this income variable: non-farm and farm wages, the wages of the respondent’s spouse/partner, child support, interest and dividends from stocks or mutual funds, rental income, retirement pension/alimony/Social Security payments, parents’ income if the respondent resides with them, monetary gifts (other than allowance) from parents, public support sources, and other income. A second variable creates a poverty status ratio comparing total household income (from the youth CV_INCOME_GROSS_YR variable) to the federal poverty level for the previous year (CV_HH_POV_RATIO). In round 1, these two variables could be based on either youth or parent data as described in section 4.6.7, “Parent Characteristics.” Therefore, a third variable, CV_HH_INCOME_SOURCE, indicates whether the information in round 1 was taken from the youth or parent interview.

Comparison to Other NLS Surveys: Information on income is regularly collected from each cohort. Users should note, however, that the income sources included have varied widely over time and among cohorts. For more information, refer to the *NLS Handbook* or the appropriate cohort’s *User’s Guide*.

Survey Instruments: These questions are found in the income and assets section (YINC) of the round 1 *Youth Questionnaire*, the income section (YINC) of the round 2–5 *Youth Questionnaires*, and section PC10 of the round 1 *Parent Questionnaire*.

Related User's Guide Sections
4.6.5 Household Composition
4.6.7 Parent Characteristics
4.8.1 Assets & Debts
4.8.3 Program Participation

Main Areas of Interest
Created Variables
Income

Supplemental Areas of Interest
Children
Household Characteristics
Illegal Activity & Arrest
Parent Current Status
Program Participation
Wages & Compensation

4.8.3 Program Participation

Independent youths¹ provide information about participation in government programs. The respondents are told these questions refer only to payments made directly to themselves and their spouses; answers should not include payments made to other members of the household. For each program listed in Figure 1, the surveys ask whether the respondent or his or her spouse had participated in the program since the last interview (for respondents who had answered these questions in a previous interview) or had ever participated (for respondents going through this section for the first time). If so, respondents report whether they are currently receiving payments from that program. Questions concerning unemployment compensation or Worker’s Compensation receipt ask about the respondent and spouse separately.

4.8.3 Figure 1. Government Assistance Programs

Round 1	Rounds 2–5
Aid to Families with Dependent Children (AFDC) or Aid to Dependent Children (ADC)	Aid to Families with Dependent Children (AFDC) or Aid to Dependent Children (ADC)
Cuban/Haitian or Indian assistance	Food stamp benefits
Emergency assistance	Unemployment Compensation
Food stamp benefits	Women, Infants, and Children (WIC)
Foster child support	Worker’s Compensation
General assistance	Other welfare programs, including:
Low Income Home Energy Assistance Program (LIHEAP)	<ul style="list-style-type: none"> • Cuban/Haitian or Indian assistance • Emergency assistance • General assistance • Supplemental Security Income (SSI)
Supplemental Security Income (SSI)	
Unemployment Compensation	
Women, Infants, and Children (WIC)	
Worker’s Compensation	
Other welfare programs (none specified)	

For each program reported, the survey collects information on the duration of each spell of receipt (e.g., stop and start dates) and the amount of aid the respondent or his or her spouse/partner received for each spell. For certain programs (e.g., AFDC or ADC, WIC, food stamps), the survey asks the respondent to state the members of the household covered by the program.

User Notes: Users who encounter problems with the program participation data should be aware that inconsistencies exist for some respondents with more than one receiving spell. Contact NLS User Services for more details.

These respondents are then questioned on whether they lived in public housing, received public housing assistance, or received rental vouchers. The wording in these questions directs respondents to report only

¹ To be considered independent, respondents must meet one characteristic from a list of criteria; see the introduction to section 4.8 for more information.

the amount they directly received from public housing benefits, as opposed to indirect benefits (e.g., living with a parent). Similar to the information gathered for other programs, respondents are asked to state dates they participated in these programs and the dollar amount of the benefits they received.

Created Variables. To summarize the above information for the user, three series of created variables provide data about total benefit amounts received and time periods of receipt. Table 1 shows the variables available for unemployment insurance, Worker’s Compensation, and any other government program.

4.8.3 Table 1. Created Program Participation Summary Variables

Type of Information	Any government program except UI and WC	Unemployment Insurance (UI)	Worker’s Compensation (WC)
Total amount received in calendar year	CV_AMT_GOVNT_PGM_PCY.xx	CV_AMT_UI_YR.xx	CV_AMT_WC_YR.xx
Number of months/weeks received in calendar year	CV_GOVNT_PRG_YR.xx (months)	CV_UI_YR.xx	CV_WC_WKS_YR.xx (weeks)
Number of months/weeks ever received	CV_GOVNT_PRG_EVER (months)	CV_UI_EVER	CV_WC_WKS (weeks)
Number of spells of receipt each year	—	CV_UI_SPELLS_YR.xx	—

In addition, the Event History data set contains a number of created variables tracing the respondent’s participation in various government programs over time. Interested researchers should refer to section 4.4, “Event History.”

User Notes: In the round 1 survey, youths with a current partner who did not meet any of the other independence criteria were incorrectly skipped past the program participation section. These youths are coded as -3, invalid skip, in the data.

In round 2, respondents who were receiving benefits from a given program at the time of the round 1 interview and who were currently receiving benefits at the round 2 interview date reported whether they had stopped receiving benefits at any point between the interviews. If the respondent reported a stop date, he or she was **not** subsequently asked to report the start date of the current spell (that is, the period of receipt after the gap through the current interview date). Information was gathered on the current spell regarding amount and household member(s) receiving. This was corrected for all types of programs except food stamps in rounds 3 and 4.

For the Worker’s Compensation variables, the data suggest that some respondents are reporting the dates the payment was actually received as opposed to the period of time to which the payment applied. For example, if a respondent was out of work for six months but received a lump sum payment a year later, he or she might report either the date the lump sum was paid or the dates he or she was unable to work.

Comparison to Other NLS Surveys: Information on public assistance has been regularly collected for each cohort (but only for NLSY79 Children age 15 and older) as part of the income section. Additionally, the NLSY79 data include a series of created variables detailing spells of program participation and amounts of assistance received. Consult the *NLS Handbook* or the appropriate cohort's *User's Guide* for details.

Survey Instruments: These questions are found in the program participation section (YPRG) of the *Youth Questionnaire*.

Related User's Guide Sections

4.4 Event History
4.6.5 Household Composition
4.6.7 Parent Characteristics
4.8.2 Income

Main Areas of Interest

Created Variables
Program Participation

Supplemental Area of Interest

Income

4.9 Marital History & Fertility

This section details the NLSY97 data on respondents' experiences with pregnancy, children, and marriage or marriage-like relationships. Throughout the survey, a marriage-like relationship is defined as a sexual relationship in which partners of the opposite sex live together.

Table 1 summarizes the *NLSY97 User's Guide* subtopics available in this section and any global universe restrictions affecting those subtopics. Users should note that certain questions or groups of questions within a subtopic may have additional restrictions. Please consult the NLSY97 questionnaire and codebook for more information on particular questions.

4.9 Table 1. Marital History & Fertility Subtopics and Universe Restrictions

<i>NLSY97 User's Guide</i> Subtopic	Round 1 Universe	Round 2 Universe	Round 3 Universe	Round 4 Universe	Round 5 Universe
4.9.1 Child Care	—	—	—	—	female with resident child under age 13; female born in 1983 or 1984 whether she has children or not
4.9.2 Fertility, Pregnancy & Children					
Pregnancies not ending in live birth	≥14 as of 12/31/1996 and female	≥14 as of 12/31/1997	all ages	all ages	all ages
Fertility and children	all ages	all ages	all ages	all ages	all ages
Adopted children	child reported on household roster	17 as of 12/31/1997 or child on HH roster	≥17 as of 12/31/1998 or child on HH roster	≥17 as of 12/31/1999 or child on HH roster	all ages
4.9.3 Marital/Marriage-Like Relationships	16 as of 12/31/1996	≥16 as of 12/31/1997	≥16 as of 12/31/1998	≥16 as of 12/31/1999	all ages

4.9.1 Child Care

This section discusses information collected about respondents' use of child care.

In the round 5 *Youth Questionnaire*, female respondents who have children (biological, adopted, and step) under the age of 13 years of age at the time of the interview are asked a series of questions on child care. They report on the number of hours the mother is in school, work, or training or traveling to one of these. For each child, the mother is asked about child care arrangements used, mode of care, hours in each child care arrangement, location of care (for those in the care of a relative), travel time to care, and the amount paid for care. These female respondents also answer questions about who was responsible for taking the child to care and picking the child up from care. In addition, these respondents reported on whether an employer, outside agency, or anyone other than the respondent pays for all or part of the care. If another entity does pay, the respondent tells which employer, outside agency, or person pays and how much they contribute toward the care.

Female respondents born in 1983 or 1984 who have no children, and *all* women with children who do not use relatives as a source of child care, are asked a series of questions related to relatives and child care. These respondents report on the number of relatives within 15 and 45 minutes of the respondent's home and on the possibility of these relatives caring for the respondent's children. Respondents then answer questions on the amount of pay a relative would require to provide child care, if any.

The child care section also asks these women how much they would expect to pay for someone to come to their home to care for their child and how much they would be willing to pay. Similar questions are asked about family day care and child care centers. For both these types of care, respondents report on the travel time to the closest center, the amount they would expect to pay, and the amount they would be willing to pay to use the services.

For the questions about relatives in the child care section, the definition of "relative" includes in-laws and relatives of the child's father.

Comparison to Other NLS Surveys: Information on child care has been collected from the NLSY79, Children of the NLSY79 age 15 and older, and the Original Cohorts. Data include types and locations of child care arrangements and extent of child care responsibility. Refer to the *NLS Handbook* or each cohort's *User's Guide* for exact survey years and the types of information available.

Survey Instruments: These questions are found in the child care section (YCCA) of the *Youth Questionnaire*.

Related User's Guide Section

4.9.2 Fertility, Pregnancy &
Children

Main Areas of Interest

Children
Created Variables
Fertility and Pregnancy

Supplemental Area of Interest

Youth History

4.9.2 Fertility, Pregnancy & Children

This section discusses information about fertility, pregnancy, and children collected in the self-administered and fertility sections of the youth instrument. For information about sexual activity, refer to section 4.5.8, “Sexual Activity & Dating.”

In the self-administered section of the *Youth Questionnaire*, female respondents who were age 14 by the end of the year before the interview are asked a series of questions on pregnancy. Those who report having ever been pregnant are asked whether the pregnancy resulted in a child born alive. If the pregnancy did not end in a live birth, follow-up questions ask for the month and year the pregnancy ended, the number of months into the pregnancy it ended, the youth’s age at that time, and the outcome (e.g., stillbirth, miscarriage, abortion). Pregnant respondents are asked how far along they are in their pregnancy; if not sure, they are asked for the date of their last period. The rounds 2–5 interviews also included questions for male respondents. This self-administered series asked if the youth had ever made a woman pregnant, his age at that time, if any woman was currently carrying his child, the total number of women he had ever impregnated, the total number of pregnancies, the number ending in a live birth, and the number ending in an abortion.

The fertility section of the questionnaire asks all respondents if they have given birth to or fathered any children. For each live birth, the survey records the birth date, gender, mortality status, and primary residence of the child. Female respondents are asked if they are legally responsible for their child; male respondents are asked the same question about the birth mother’s legal responsibility.

User Notes: The respondent’s household is based on what the respondent considers to be his or her permanent household as reported in the household information section. This is not necessarily the same as where he or she is living at the time of the survey. So respondents who are in the military (or away at college or incarcerated) may report their spouse or children as being in the household, even though the answers in the fertility and marriage sections have the respondent separated from them.

In addition, the round 1 NLSY97 survey gathered detailed information concerning legal custody and legal responsibility for each child listed by the respondent. The survey defined “custody” as the place the child was living and “responsibility” as the person who made major decisions about the child’s life. After these questions, male respondents were asked if they had ever legally acknowledged paternity by signing the birth certificate, applying for insurance, or using other methods that would establish their relationship to the child. Female respondents were asked the same type of paternity questions about the child’s father.

For each child, the survey instruments then ask for the characteristics of the child's other biological parent at the time of conception. Data recorded include race, age, school enrollment, labor market status, welfare reciprocity, and highest grade completed. These details were expanded in round 2 with the addition of highest degree received and religious preference. In each round, the respondent is then asked about the type of relationship he or she had with the other person. The first question asks how often the couple had sex (on-going or only once/twice), while the second question asks about the parents' relationship at the time of the birth (i.e., not seen each other for a month or more, still in contact but not sexual, still sexual relationship).

Another set of questions in rounds 4–5 addressed only male respondents who had fathered a child. First, the respondent provided current information about the mother of each child, such as her enrollment and employment status, program participation status, and income. He also stated whether he currently had a close relationship with the mother. The survey then asked about the respondent's relationship with each of his children. Respondents reported whether they participated in prenatal activities such as going with the mother to the doctor, buying things before the baby was born, and being present at the delivery. If the respondent had seen the child in the past month, he stated how often he performed activities such as bathing or dressing the child, preparing a meal for the child, or reading books to the child. Finally, male respondents provided information about child support. Respondents first reported the amount of support awarded in a child support agreement and then stated whether they had provided additional informal support, such as performing household repairs, buying clothes for the child, or buying household items or gifts for anyone in the household, in the past 12 months.

User Notes: Basic information about the respondent's children is collected and organized into the BIOCHILD roster. In general, this roster functions in the same way as other NLSY97 rosters, as described in section 3.2 of this guide. However, users should be aware that the loop numbers for the series of questions about new biological children do **not** correspond to the numbers on the roster. The procedure for matching questions about new children to the correct child on the BIOCHILD roster is described in the *NLSY97 Codebook Supplement*.

Questions about adopted children have been present in each survey. Information collected includes the child's birth date, gender, residence, date the child died or left the respondent's household, and whether the respondent is still the legal adoptive parent. Note that no respondents reached this series in round 1 because of an age restriction and no respondents reported having any adopted children in round 2. A few adopted children were reported by respondents in rounds 3 and 4. In round 5, only two adopted children appear on the BIOADOPTCHILD roster.

Created Variables. The status of the respondent’s biological children is identified in a created variable (CV_CHILD_STATUS.xx). Coding categories include put up for adoption, deceased, nonresident in foster care, nonresident not adopted or in foster care, and resident in the respondent’s household. Other variables include each biological child’s date of birth and, if applicable, date of death. These variables record both the actual dates (CV_CHILD_BIRTH/DEATH_DATE.xx_M/Y,) and the dates in the continuous month scheme (CV_CHILD_BIRTH/DEATH_MONTH.xx) described in section 4.4, “Event History.” Finally, two variables report the number of biological children ever born to the respondent who reside in the household (CV_BIO_CHILD_HH) and the number who do not reside in the household (CV_BIO_CHILD_NR).

User Notes: Users should be aware that, in some cases, problems arose in the marriage and fertility sections due to inaccuracies in relationship codes in the household roster. If a household member was identified on the roster as the respondent’s child or spouse, the marriage and fertility sections asked questions about that person. If the relationship code in the roster was incorrect and the person was not a spouse or child, the youth respondent would correct the relationship at that point.

CHRR staff subsequently corrected a number of problems in the roster relationship codes. Thus, it may appear that respondents were asked marriage and fertility questions even though there was no one in the household to ask these questions about. Users may not be able to trace a child or spouse back to the roster due to these corrections. Researchers should contact NLS User Services for more details.

Comparison to Other NLS Surveys: Information on fertility and on the status of children has been regularly collected from the NLSY79, Children of the NLSY79 age 15 and older, and the Original Cohorts. Data include number of children, dates of birth, gender, and life status. Refer to the *NLS Handbook* or each cohort’s *User’s Guide* for exact survey years and the types of information available.

Survey Instruments: Questions on pregnancy are found in the self-administered section (YSAQ) of the *Youth Questionnaire*; questions on children are found in the fertility (YFER) section.

<i>Related User’s Guide Section</i>	<i>Main Areas of Interest</i>	<i>Supplemental Area of Interest</i>
4.5.8 Sexual Activity & Dating	Children Created Variables Fertility and Pregnancy	Sexual Activity

4.9.3 Marital & Marriage-Like Relationships

The surveys ask respondents age 16 or older at the end of the calendar year before the survey about their marital status at the time of the survey. In addition, information is gathered about all changes in the respondent’s marital and cohabitation status that occurred between the interviews. In the NLSY97, a marriage-like relationship is defined as a sexual relationship in which partners of the opposite sex live together. In rounds 2–5, the partners must have lived together for at least one month.

4.9.3 Table 1. Number of Respondents Ever Married or Never Married and Cohabiting

	Never Married, Cohabiting at Interview Date			Ever Married		
	Male	Female	Total	Male	Female	Total
Round 1	4	10	14	—	4	4
Round 2	34	78	112	6	35	41
Round 3	76	179	255	24	83	107
Round 4	170	310	480	61	164	225
Round 5	248	429	677	121	282	403

Notes: This table is based on CV_MARSTAT. The two categories in the table are exclusive. For example, a respondent might be divorced or separated but currently cohabiting with a new partner; this respondent would be represented in only the “ever married” category.

Each survey asks the respondent to report the name (or initials, if the respondent is reluctant to provide a full name) of any spouse or partner meeting the above definition. The surveys then collect specific information on each partner. For all spouses and partners, the respondent reports when he or she started and stopped living with that individual. If the spouse or partner is not residing in the respondent’s household, the youth is asked for additional information. This includes the partner’s age, race, and highest grade of schooling at the time the relationship began; whether the partner was enrolled in school, working, or receiving government assistance when they began living together; and the partner’s religious preference and the frequency with which they attended religious services (rounds 1 and 2).

User Notes: The respondent’s household is based on what the respondent considers to be his or her permanent household, not where he or she is living at the time of the survey. So respondents who are in the military (or away at college or incarcerated) may report their spouse or children as being in the household, even though the answers in the fertility and marriage sections have the respondent separated from them.

After the list of spouses and partners is compiled, the questionnaires go back through the list and determine the dates at which changes in marital status occurred during each relationship. Respondents

state whether they were married, divorced, reunited, separated, or had their marriage annulled at that date. Additional questions ensure that the time periods are captured if respondents break up and then live again with the same partner.

Round 1 also included a series that provided a detailed picture of a respondent's life with his or her current spouse/partner. Respondents were asked to rate how often they and their partner engaged in behavior such as screaming or yelling, compromising to resolve a disagreement, or expressing affection for each other. Respondents answered once for themselves and once for each partner so that answers are provided for both sides of the relationship, although only from the respondent's point of view.

Created Variables. Based on questions in this section, two created variables identify the current marital status of age-eligible respondents. The first, CV_MARSTAT_COLLAPSED, classifies respondents in five basic marital categories—never married, married, separated, divorced, and widowed. The other, CV_MARSTAT, uses more detailed categories to provide cohabitation information as well (e.g., never married, cohabiting; married, spouse absent; separated, cohabiting). The respondent's total number of marriages (CV_MARRIAGES_TTL) and total number of cohabitation experiences (CV_COHAB_TTL) are also created. The beginning date of the respondent's first marriage is provided in actual month and year variables (CV_FIRST_MARRY_DATE_M/Y) and as a continuous month scheme variable (CV_FIRST_MARRY_MONTH). The beginning date of the respondent's first cohabitation has the same created variables available, with "COHAB" in place of "MARRY" in the title. Users should note that the created marriage date variables refer exclusively to spouses; the created cohabitation date variables refer only to partners, including partners who lived together and later married.

In addition, the Event History data set contains a number of created variables tracing the respondent's marriage and cohabitation history over time. Unlike the created variables described above, these event history variables include both spouses and partners in the same variable. Interested researchers should refer to section 4.4, "Event History," for descriptions of these variables and of the continuous month scheme mentioned above.

User Notes: Users should be aware that, in some cases, problems arose in the marriage and fertility sections due to inaccuracies in relationship codes in the household roster. If a household member was identified on the roster as the respondent's child or spouse, the marriage and fertility sections asked questions about that person. If the relationship code in the roster was incorrect and the person was not a spouse or child, the youth respondent would correct the relationship at that point in the interview.

CHRR staff subsequently corrected a number of problems in the roster relationship codes. Thus, it may appear that respondents were asked marriage and fertility questions even though there was no one in the household to ask these questions about. Users may not be able to trace a child or spouse back to the roster due to these corrections. Researchers should contact NLS User Services for more details.

Comparison to Other NLS Surveys: Information on marital status, history, and transitions has been collected for each cohort (but only for NLSY79 Children age 15 and older). For more information, refer to the *NLS Handbook* or the appropriate cohort's *User's Guide*.

Survey Instruments: These questions are found in the marriage section (YMAR) of the *Youth Questionnaire*.

Related User's Guide Sections

4.4 Event History
4.6.5 Household Composition
4.9.2 Fertility, Pregnancy &
Children

Main Areas of Interest

Created Variables
Demographic Indicators
Marriage & Cohabitation

Supplemental Areas of Interest

Children
Household Characteristics

4.10 Training

This section describes the types of information collected within the training section of the NLSY97 questionnaires. Through these questions, the surveys examine formal training experiences of respondents outside of their regular schooling. Alone, these data do not provide a complete picture of the entire training investment of an individual because they do not cover regular schooling or informal training; however, the data will help researchers to examine which types of formal training programs assist youths in acquiring skills and securing employment. The training section, addressed only to respondents 16 years of age and older as of the survey date, explores what kind of training youths obtain, where and when they are trained, how the training is paid for, and what skills are acquired.

4.10 Table 1. Training Subtopics and Universe Restrictions

<i>NLSY97 User's Guide</i> Subtopic	Universe for Rounds 1–5 (Age as of each round's interview date) ¹
Training	≥16

¹ Note that age restrictions in the employment and training sections of the questionnaire refer to age as of the survey date, rather than age as of December 31, 1996, as in other sections.

In general, the training section collects information on each training program with which a respondent is involved. The *Youth Questionnaire* asks all respondents who were at least 16 years old whether they had ever participated in any occupational training programs outside of their regular schooling. For each program, the survey then collects basic information, including the type of program, start and stop dates, time devoted to the training, periods of nonattendance lasting a week or longer during training sessions of at least two weeks, and whether the program was completed (and if not, the reason). In rounds 2–5, the survey confirmed the data from the previous round and then collected these details about programs taken since the last interview.

An extended series of questions explores the financial aspects of the training program. Specifically, respondents are first asked if any aspect of the training was provided by a government program. They also state whether they received any allowance or additional money for participating in the training. Respondents are then asked whether they or their families paid any fees for the training program and whether they used any student financial aid or loans—and the amount of that aid or loan—to help pay for the training. Examples of financial aid or loans include Pell Grants, Stafford Loans, Supplemental Educational Opportunity Grants (SEOGs), or scholarships. Finally, respondents are asked whether an

employer helped to pay for the training. If an employer did pay, the respondent is questioned on whether the training was required by the employer and the main reason that he or she undertook the training (e.g., promotion, obtain a license/certificate, look for a new job).

The next series of questions explores the content of the training program. Information is collected on whether any skill tests (i.e., reading, writing, math) were given upon enrollment in the program and what types of assessments (e.g., exams, skill demonstrations) were used to measure progress in the program. The surveys also explore the types of certificates, licenses, or degrees the respondent received from the program and the types of skills learned through the program. Respondents are questioned about any classes taken while enrolled in the training program to improve their reading/writing/math skills, to get their GED or high school diploma, or to learn English.

The surveys also collect data on additional activities the respondent was participating in while involved in a training program; these include vocational training for a specific job, job search assistance classes, on-the-job training, work experience, or other classroom training for a certain job. Specific questions about these activities gather which one the respondent did the most, the job trained for if involved in on-the-job training, and the types of services offered in the job search assistance classes. The skills learned in job training sessions are also recorded; respondents report learning or improving their computer skills, team work skills, health or safety procedures, equipment operation or repair, sales or managerial skills, sensitivity training, etc.

Types of Training Providers. Information is collected on the type of organization providing the training in which NLSY97 respondents participated. Types of training providers include the following:

- Adult Basic Education (pre-GED program)
- Apprenticeship program
- Business or secretarial school
- Community or junior college
- Correspondence course
- Formal company training run by employer
- GED program
- Government training
- Nursing school (LPN or RN)
- Seminar or training program outside of work
- Seminar or training program at work run by someone other than employer
- Vocational rehabilitation center
- Vocational, technical, or trade school
- School-based: K-12, including ROTC (round 1 only)

In addition, information on school coursework and school-based learning programs is recorded in the schooling section of the instrument and is collected for all respondents. See section 4.2.3, “School-Based Learning Programs,” and section 4.2.4, “School Experience,” for more information.

User Notes: Beginning in round 3, some of the information about the respondent’s training programs is organized in the TRAINING roster. Information about the structure and use of rosters is provided in section 3.2 of this guide.

Created Variables. To summarize information on training programs, several variables are created for each round. The first variable (CV_TRN_CERT) states whether the respondent received a vocational license or certificate other than a GED as a result of any training program. For the most recent certificate or license, a pair of variables (CV_TRN_CERT_DATE) lists the month and year that it was received.

Comparison to Other NLS Surveys: Information on training programs has been collected from the NLSY79 respondents in each survey. Although the earlier surveys focused on government training programs, the questions in later surveys are very similar to the NLSY97. Children of the NLSY79 age 15 and older were asked comparable questions in 1994–2002. Questions about training both on and off the job have also been asked of Original Cohort respondents in most surveys. Users should consult the *NLS Handbook* or the appropriate cohort’s *User’s Guide* for more information about exact questions and universes.

Survey Instruments: These questions are found in the training section (YTRN) of the *Youth Questionnaire*.

<i>Related User’s Guide Sections</i>	<i>Main Area of Interest</i>	<i>Supplemental Areas of Interest</i>
4.2.3 School-Based Learning Programs	Training	Achievement Tests
4.2.4 School Experience		Assets & Debts
4.8.1 Assets & Debts		Income
		Industry & Occupation
		Jobs & Employers

Appendices

Appendix A: Quick Reference Guide

This appendix summarizes some key pieces of information regarding documentation products for the NLSY97, the NLS surveys, and the NLS data sets. It also provides an index to the created variables discussed in this guide and concludes with information about how to find additional help when necessary.

A.1 Guide to NLS Documentation

The *User's Guide* provides in-depth information about the NLSY97 data. Users requiring general information about the history, administration, or other cohorts of the NLS project, or technical information about topics such as sampling, should consult Table 1 for the most appropriate document.

A.1 Table 1. List of NLS Documentation

<p>NLS Handbook http://www.bls.gov/nls/handbook/nlshndbk.htm</p>	<p>This comprehensive introduction to the NLS gives readers general information about all NLS cohorts and their main topics of investigation.</p>
<p>NLS Web Homepage http://www.bls.gov/nls</p>	<p>This internet site offers an overview of the NLS programs.</p>
<p>NLS Internet Bibliography http://www.nlsbibliography.org</p>	<p>This on-line searchable database provides citations for research using NLS data.</p>
<p>NLS News http://www.bls.gov/nls/nlsnews.htm</p>	<p><i>NLS News</i>, a quarterly newsletter, is mailed at no charge to all data users and other interested persons. The newsletter contains informative articles about the NLS data, announces the release of new data, and reports any errors discovered in the data or documentation. Past issues of the <i>NLS News</i> are available on the NLS internet site.</p>
<p>NLS User's Guides http://www.bls.gov/nls, follow link for cohort of interest</p>	<p>These cohort-specific guides help researchers understand NLS variables, survey instruments, documentation techniques, and other technical issues.</p>
<p>Questionnaires available for download online at http://www.bls.gov/nls/nlsorder.htm</p>	<p>The complete set of survey instruments used with the cohort in each survey year allows researchers to view questions, supplemental information, and household interview forms.</p>
<p>Codebook Supplements available for download online at http://www.bls.gov/nls/nlsorder.htm</p>	<p>Supplementary attachments and appendices contain variable creation, description, and coding information not present in the questionnaire.</p>
<p>Technical Sampling Reports</p>	<p>These manuals and addenda describe the sampling procedure used to select the initial survey participants, including information about weights and standard errors.</p>
<p>Error updates via the internet http://www.bls.gov/nls, follow link for cohort of interest</p>	<p>Information about data errors and ways to obtain corrected data is available for each cohort.</p>

With the exception of the bibliography and the NLS Web site, which are provided on-line, users may obtain any of the documents listed in Table 1 by contacting NLS User Services or by going to

<<http://www.bls.gov/nls>> and choosing the “Ordering Data” link. Contact information for NLS User Services is provided under “Additional Support” at the end of this appendix.

A.2 About the NLSY97 Data Set

The NLSY97 is a panel data collection—specifically, the same respondents are interviewed at regular intervals, and many of the same questions are asked in each interview. This longitudinal approach permits researchers to examine changes in a respondent’s behavior over time. Each data set contains data extraction software to search this large database and extract the specific variables needed.

Search Strategies

Variables can be selected for extraction using either a search list (reference number or question name) or a search index (any word in context, area of interest, or survey year). Search indexes may also be combined to produce more precise results. Below is a brief description of each search option; the *NLS Compact Disc User’s Guide* contains more detailed instructions.

Reference Number: Each variable is assigned a unique reference number that determines its relative chronological position in the database. That reference number never changes—even when waves of data are added and the database is revised. The reference number is the equivalent of a unique variable name.

Question Name: Each year, a separate survey instrument is used to collect data. A questionnaire item (or question name) identifies the variable in the questionnaire. The question names for the same questions are identical from year to year, so users can easily identify a question of interest in various survey rounds.

Any Word in Context: The database retrieval software allows the user to search for and select all variables with a title containing any single word or combination of words.

Area of Interest: Variables related to a common topic are stored in organizational groupings called “areas of interest.” Figure 1 lists the current areas of interest in the NLSY97 data set.

Survey Year: The user can choose a survey year and select variables collected in that year.

Hardware Requirements

To access the NLSY97 data, users need an IBM-compatible personal computer (PC) running Microsoft Windows and sufficient hard drive space to accommodate the data set. The amount of space required for each data set is noted both on the download page and in the ReadMe file on the CD.

Users can extract data for specific subsamples and in various formats. Step-by-step instructions for extracting data and detailed information about examining the codebook can be found in the *NLS Compact Disc User's Guide*.

A.2 Figure 1. NLSY97 Areas of Interest

Achievement Tests	Income	School Characteristics
Assets & Debts	Industry & Occupation	School Experience
Attitudes	Interaction btwn Parents	Screener
Autonomy & Control	Interviewer Remarks	Screener Extended
Child Calendars	Job Search	Screener Household
Child Care	Jobs & Employers	Screener Non-Resident
Child Family Background	Labor Force Status	Screener Parent
Children	Locator	Self-Employment
College Experience	Machine Check	Sexual Activity
Common Variables	Marriage & Cohabitation	Substance Use
Contact w/Non-Res. Parent	Military	Symbols
Created Variables	Non-Res. Characteristics	Tenure w/ Employer
Demographic Indicators	Parent Background	Time Spent at Work
Ed. Status & Attainment	Parent Current Status	Time Use
Employment Gaps	Parent Family Background	Training
Expectations	Parent Information	Transcript Survey
Family Process Measures	Parent Locator	Validation
Fertility and Pregnancy	Parent Retrospective	Wages & Compensation
Fringe Benefits	PIAT	Work Experience
Geographic Indicators	Program Participation	Youth History
Health	Roster Item	Youth Locator
Household Characteristics	Sample Design & Screening	Youth Schooling
Illegal Activity & Arrest	School-Based Learning	Youth Self-Administered

A.3 How to Get Help

Users may have questions about the NLSY97 data, database retrieval software, or documentation. The following are strategies for finding answers to these questions.

On-line and Paper Documentation Help: NLS User Services provides both on-line and paper documentation help sources for users. Choosing “Help” in the menu bar of the Investigator program leads the user to an on-line manual or a download feature for on-line help. Paper documentation includes this *NLSY97 User's Guide*, the *NLS Handbook*, the *NLS Compact Disc User's Guide*, and assorted supplemental hard copy documentation (see A.1 Table 1).

Additional Support: For questions not answered in the documentation, contact NLS User Services:

NLS User Services
 Center for Human Resource Research
 921 Chatham Lane, Suite 100
 Columbus, Ohio 43221-2418
 614-442-7366
 E-mail: usersvc@postoffice.chrr.ohio-state.edu

A.4 Index to Created Variables, Roster Items, and Symbols

This index provides users with an easy reference to discussions in this guide of the various roster, symbol, and created variables in the NLSY97 data set. Researchers working with any created, roster, or symbol variables are encouraged to consult the pages listed below and review all available information before proceeding with analysis. Note that a number of question names are truncated; for example, “CV_ESR” represents both CV_ESR and CV_ESR_COLLAPSED. Additionally, the variables created by Child Trends, Inc., are not included here; interested researchers should look up “Child Trends, Inc.” in the main index to this guide or consult Appendix 9 in the *NLSY97 Codebook Supplement*.

AFDC_ (AFDC event history)	143	CV_MARSTAT	228
BIOCHILD roster	51, 225	CV_MSA	179, 208
CS_PANEL_WEIGHT	38	CV_PIAT	73
CS_SAMPLING_WEIGHT	38, 39	CV_REFIELD_YOUTH	24
CV_AA_DEGREE	83	CV_SAMPLE_TYPE	7, 22
CV_AGE	175, 176	CV_SCH_ATTEND_YR	89
CV_AMT_GOVNT_PGM_PCY	221	CV_SCHOOL_TYPE	85, 89
CV_AMT_UI_YR	221	CV_TRN_CERT	232
CV_AMT_WC_YR	221	CV_TTL_RESIDENCES	207
CV_ASSOC_CREDITS	83	CV_UI	221
CV_BA_CREDITS	83	CV_URBAN-RURAL	179, 208
CV_BA_DEGREE	83	CV_WC_WKS	221
CV_BIO_CHILD	225	CV_WKSWK	130
CV_BIO_MOM_AGE	203	CV_YTH_REL_HH	186, 207
CV_CENSUS_REGION	179, 208	FDSTMP_ (food stamps event history)	143
CV_CHILD	225	FREELANCE roster	51, 112, 128
CV_CITIZENSHIP	205	HHI or HHI2 (household) roster	28, 51, 187, 189, 190, 192, 193, 197, 204, 205, 226, 228
CV_COHAB_TTL	228	INFORMANT!ID	192
CV_ENROLLSTAT	85	KEY!AGE	176
CV_ESR	123	KEY!ETHNICITY	205
CV_FIRST_COHAB	142, 228	KEY!RACE	205
CV_FIRST_MARRY	142, 228	KEY!RACE_ETHNICITY	205
CV_GED	87	MAR_ (marriage event history)	142, 143
CV_GOVNT_PRG	221	NEWSCHOOL roster	51, 75, 77, 80, 81, 83, 90
CV_GRADES	86	NONHHI (nonresident) roster	51, 187, 193, 197, 198
CV_HGC_EVER	87	OTHER_ (other public assistance event history)	143
CV_HH_INCOME_SOURCE	202, 219	PARHHI roster	51
CV_HH_NET_WORTH	201, 216	PARYOUTH roster	28, 51, 188, 191, 197
CV_HH_POV_RATIO	202, 219	RNI (reason for noninterview)	35
CV_HH_REL	186, 207	SAMPLING_PANEL_WEIGHT	38
CV_HH_SIZE	186	SAMPLING_WEIGHT	38, 39
CV_HH_UNDER	186	SAMPLING_WEIGHT_CC	38
CV_HIGHEST_DEGREE_EVER	87	SCH_ (schooling event history)	144, 145
CV_HRLY_COMPENSATION	135	SYMBOL!KEY!AGE	176
CV_HRLY_PAY	135	TRAINING roster	51, 232
CV_HS_DIPLOMA	87	WIC_ (WIC event history)	143
CV_INCOME_GROSS_YR	201, 219	YEMP (employer) roster	51, 103, 110, 123, 130, 133, 142
CV_INTERVIEW_DATE	25	YOUTH roster	28, 51, 188, 192, 197
CV_JOB<13_WKS	130, 135		
CV_MARRIAGES_TTL	228		

Appendix B: NLSY97 Glossary of Terms

This glossary is provided to aid researchers in understanding the terms used in the NLSY97 questionnaires and *User's Guide*. The glossary is divided into two sections. The first defines terms used by survey staff in processing the data and providing information to researchers—the “jargon” of the NLS program. Section B2 includes terms used during the interview that relate to specific questions. The definitions in this section are available to interviewers and respondents during survey administration.

B.1 NLS Survey Terms

This section defines terms related to survey management and data dissemination that researchers should understand when working with the NLS cohorts.

Any word search

This search function allows users to select any word or words and to view all variables that contain those words in their titles.

Area of interest

Variables are grouped by common topical areas. The data set includes a search function so that users can view variables in an area of interest.

Armed Services Vocational Aptitude Battery (ASVAB)

This test measures the applicant's knowledge and skills in a number of topical areas and is used by the military to produce a general standard of the trainability and enlistment eligibility for the Armed Forces. The computer adaptive form of the test, the *CAT-ASVAB*, was administered to NLSY97 respondents during round 1. This test was sponsored by the Department of Defense.

Audio computer-assisted self-interview (ACASI)

Self-administered portions of NLSY97 instruments use an ACASI system that allows the respondent to listen to the questions and response categories being read over headphones. ACASI sections can be heard in either English or Spanish.

Bureau of Labor Statistics (BLS)

This agency of the U.S. Department of Labor sponsors and oversees the National Longitudinal Surveys project.

CAT-ASVAB

See *Armed Services Vocational Aptitude Battery (ASVAB)*.

Center for Human Resource Research (CHRR)

A research unit at The Ohio State University, CHRR is responsible for NLSY97 variable creation, documentation and dissemination of the data, and user services.

Check item

Check questions are used during the course of the interview to identify the sample universe and denote the skip pattern for a particular set of questions. These questions are answered based on previous responses or direct observations, without asking the respondent.

Children of the NLSY79

This survey group comprises all children born to female NLSY79 respondents. The group was first surveyed in 1986 and has been reinterviewed biennially. Since 1994, a separate survey has been administered to the children age 15 and older, referred to as the “Young Adults.”

Codeblock

Information about each variable is presented in a consistent form called a codeblock. Most codeblocks provide users with the variable title, reference number, question number, survey year, coding information, and a frequency distribution.

Codebook

In the NLS Investigator software, the codebook contains complete information about all the variables in a data set. It comprises a number of codeblocks presenting information about each variable.

Cohort

A cohort is a group of people who share similar characteristics at a particular point in time.

Computer-assisted personal interview (CAPI)

These interviews are administered using a survey instrument on a laptop computer. CAPI allows for more complex questionnaire programming, bounded interviewing, and faster data dissemination than with PAPI interviews.

Created variables

Created by survey staff after the interview, this type of variable may be based on information provided by the respondents or obtained from other outside sources. These variables can be found under the “Created Variables” area of interest in the data set.

Current Population Survey (CPS)

Conducted by the Census Bureau, these monthly interviews collect up-to-date information about the number of persons in the country who are employed, unemployed, or not in the labor force in a specific survey week. Labor force status questions included in rounds 1 and 4 of the NLSY97 were the same as the CPS questions used to determine employment status and were administered as closely as possible to the method used by the Census Bureau so that findings are comparable.

Employment Status Recode (ESR)

Available in round 1 only, ESR is a variable created from information collected on employment status in the week before the survey. It provides a standardized employment status code for each member of the sample.

Geocode

Geographic codes, published by the National Bureau of Standards as the *Federal Information Processing Standards (FIPS)*, are assigned to NLSY97 respondents' residences. They can be used to match NLS data with other data (e.g., *City and County Data Books*). To preserve respondent confidentiality, geocode data are available only to researchers who satisfactorily complete a BLS accessing agreement.

Household Income Update

This one-page paper instrument, administered to a parent of respondents still living at home, collects information about the income of the parent and his or her spouse or partner.

Household unit

A single room, or group of rooms, that is intended as separate living quarters for a family, a group of unrelated persons living together, or for a person living alone.

Independent youth

A respondent with at least one of the following characteristics: is age 18 or older, has had a child, is enrolled in a 4-year college, has ever been married or is in a marriage-like relationship at the time of the survey, is no longer enrolled in school, or is living outside his or her parent-figure's home.

Industry Census code

A 3-digit numeric code from the 1990 Industrial Classification System used by the Census Bureau to assign codes to industries (Census Bureau. *1990 Census of Population Alphabetical Index of Industries and Occupations*. Washington, DC: U.S. Government Printing Office, 1991).

Int.

This is an abbreviation for interview.

Longitudinal study

A longitudinal study follows the same group of individuals over an extended period of time.

Mature Women cohort

This group of 5,083 respondents, ages 30–44 on March 31, 1967, was first interviewed in 1967 and has been surveyed 20 times through 2001. The Mature Women cohort is one of the four NLS Original Cohorts.

National Longitudinal Survey of Youth 1979 (NLSY79)

This group of 12,686 male and female respondents was first interviewed in 1979 and has been reinterviewed 18 times through 2000. Respondents in this cohort were ages 14–21 as of December 31, 1978.

National Longitudinal Survey of Youth 1997 (NLSY97)

Respondents ages 12–16 as of December 31, 1996 (born in 1980–1984), were selected for inclusion in the newest NLS cohort. The NLSY97 numbers 8,984 respondents, and four interviews have been conducted with the cohort to date.

National Opinion Research Center (NORC)

Associated with the University of Chicago, NORC is responsible for the management of the NLSY97 cohort and collection of the data.

Occupation Census code

A 3-digit numeric code from the 1990 Occupational Classification System used by the Census Bureau to assign codes to occupations (Census Bureau. *1990 Census of Population Alphabetical Index of Industries and Occupations*. Washington, DC: U.S. Government Printing Office, 1991).

Older Men cohort

This group of 5,020 respondents, ages 45–59 on March 31, 1966, was first interviewed in 1966 and subsequently surveyed 12 additional times before its discontinuation in 1990. The Older Men cohort is one of the four NLS Original Cohorts.

Original Cohorts

The four cohorts (Older Men, Mature Women, Young Men, and Young Women) selected in 1966 and first surveyed between 1966 and 1968.

Paper-and-pencil interview (PAPI)

Although the NLSY97 surveys have been almost entirely administered using computerized instruments, traditional paper-and-pencil instruments were used for some screening interviews and for the administration of the *Household Income Update*.

Peabody Individual Achievement Test Math Assessment (PIAT Math)

The *Peabody Individual Achievement Test Math Assessment*, which is given to a subset of the NLSY97 cohort, measures the respondent's ability in mathematics. The math assessment is one subtest in the full *PIAT* battery.

Primary Sampling Unit (PSU)

A primary sampling unit (PSU), as defined by NORC, is composed of either a single county (or the equivalent), a group of counties, or an SMSA and is based on population and area constraints. PSUs are the basic unit of sample selection for the NLSY97.

Profiles of American Youth (PAY97)

The name given to describe the *ASVAB* and *Interest-Finder* tests taken by the NLSY97 respondents from summer 1997 to spring 1998. The *ASVAB*, a military enlistment test, was administered to NLSY97 respondents and to other youths selected during the same screening process by the Department of Defense in an effort to establish new norms for the recruits entering the Armed Forces. The *Interest-Finder*, an occupational interest inventory, aids career counselors in helping youths with occupational decisions.

R

This is an abbreviation for Respondent.

Reference number

A reference number is a unique identifying number beginning with "R," which is assigned to each variable in the data set. Reference numbers never change after they are assigned to the variables from an interview even as additional information is added to the data set from later surveys.

Sampling weights

A numeric value calculated on the basis of each respondent's characteristics (gender, ethnicity, year of birth, sample type, and location). The value, which is adjusted for differential probabilities of selection into the sample and for attrition, allows for population estimates.

School surveys

Supplemental surveys of the high schools in the NLSY97 primary sampling units, conducted in 1996 and 2000–01. The school surveys collected information about school characteristics (enrollment, library size, vocational programs, ethnic diversity of students and teachers, etc.) and school-based learning programs.

Screener, Household Roster, and Nonresident Roster Questionnaire

The simple and extended screener sections were administered in round 1 to the initial sample of households to locate respondents eligible for the NLSY97. NORC interviewers screened 75,291 households to identify eligible respondents based on age, race, and ethnicity. If an eligible youth resided in the household, the remainder of the questionnaire (household roster and nonresident roster) was administered to collect basic demographic information about other household residents and key nonresident relatives.

Standard Metropolitan Statistical Area (SMSA)

An SMSA, defined by the Census Bureau, consists of one or more entire counties containing at least one city (or twin cities) having a population of 50,000 or more plus adjacent metropolitan counties that are economically and socially integrated with the central city. In New England, towns and cities rather than counties are the basis for defining an SMSA. SMSA boundaries may cross state lines.

Transcript survey

Survey staff collect high school transcripts from respondents' high schools after respondents are no longer enrolled as high school students and code information from these transcripts for use by researchers. In 1999–2000, transcripts were obtained and processed for 1,417 youths who had graduated from high school or had reached age 18 and were no longer attending high school. Additional transcript collections are planned for future rounds as more respondents graduate from high school or become age-eligible.

Young Men cohort

This group of 5,225 respondents, ages 14–24 on March 31, 1966, was first interviewed in 1966 and subsequently surveyed 11 additional times before its discontinuation in 1981. The Young Men cohort is one of the four NLS Original Cohorts.

Young Women cohort

This group of 5,159 respondents, ages 14–24 on December 31, 1967, was first interviewed in 1968 and has been surveyed 20 times through 1999. The Young Women cohort is one of the four NLS Original Cohorts.

B.2 Terms Used in NLSY97 Interviews

Most of the terms in this section are defined for interviewers and respondents in “help screens”—screens that may be referred to during an interview if a respondent or an interviewer needs further clarification of a term or concept. A few terms are defined within the text of the applicable question. These terms may be found in the *Youth Questionnaire*, the round 1 *Parent Questionnaire*, or the round 1 *Screener, Household Roster, and Nonresident Roster Questionnaire*.

Active forces

An umbrella term used to refer to the regular branches of the Armed Forces (Army, Navy, Marines, and Air Force) as a group. The Reserve and Guard components of these branches are not included.

Active job search

An active job search is one that could have resulted in a job offer without further action on the part of the job seeker. Methods include:

- Contacting public employment agency
- Contacting private employment agency
- Contacting employer directly
- Sending out resumes or filling out applications
- Contacting friends or relatives
- Contacting school/college or university employment center
- Placing or answering ads
- Checking union/professional registers
- Other active method (for example, bid on a contract)

See also *Passive Job Search*.

Adult basic education

This is remedial education, such as basic literacy or math skills, which does not include work directed toward passing the GED test.

Aid to Families with Dependent Children (AFDC); Aid to Dependent Children (ADC)

Commonly referred to as welfare, these two federal programs help families with dependent children by providing financial assistance each month. Eligible children are from low-income households and lack support of one or both parents. Since the implementation of welfare reform, these programs have been replaced by the TANF program on the state or county level.

Annuity

A type of asset that provides a fixed payment each month or year until a person dies. Most pensions are annuities; annuities can also be purchased directly from insurance companies. Types of annuities include the following: deferred annuities, fixed annuities, variable annuities, immediate annuities, and annuity certain.

Apprenticeship program

A formal program in which a person agrees to work in a skilled trade or art in return for wages and training for a prescribed period of time.

Armed Forces

This includes the active forces, the reserves, and the guard. Please note that civilian employees of the Armed Forces are considered federal employees.

Bonus

A bonus is a sum of money or an equivalent (stocks, company shares) given in addition to usual compensation, normally for outstanding performance and service. See also *Commission* and *Tips*.

Business

A business exists when one or more of the following conditions is met: (1) machinery or equipment of substantial value is used in conducting the business; (2) an office, store, or other place of business is maintained; or (3) the business is advertised by listing in the classified section of the phone book, displaying a sign, distributing cards or leaflets, or otherwise publicizing that the work or service is offered to the general public.

Business school

A training program that provides the trainee with business skills, not to be confused with business classes in college or graduate school. This type of program does not contribute to an undergraduate or professional degree.

Certificate of deposit (CD)

A certificate of deposit or a savings certificate is issued for a specific deposited amount at a fixed rate of interest for a given time period. Certificates can be issued by a bank, credit union, or savings & loan association. There is usually a required minimum deposit; the maturity period can vary from a few weeks to several years; there is a penalty for early withdrawal; and deposits are usually insured.

Collateral

Collateral includes property, goods, or other assets which must be pledged as part of a loan. If the borrower defaults on the loan, the lender has the right to “foreclose on” (take) the collateral.

College or university

A college or university is an educational institution that gives credit toward an academic degree at the post-secondary and/or post-graduate levels. A college or university offers candidates an associate’s, bachelor’s, master’s, or doctoral degree in a specific academic concentration.

Commission

A commission is an amount or percentage of money that is given in addition to one's regular salary. In some jobs, usually sales, the base pay may be quite low, with most of a person's earnings in the form of commissions. See also *Bonus* and *Tips*.

Community or junior college

These colleges provide courses that may result in a 2-year or associate's degree; these courses are considered regular school. If the courses are taken to improve job skills and do not count toward an academic degree, these courses are considered training rather than regular schooling.

Dividends

Dividend payments, usually in the form of money or stock, are paid to shareholders of a corporation or some other cooperative society or membership club. The dividend amount is decided by the board of directors and is usually paid on a quarterly basis. Dividends must be declared as income in the year they are received.

Employed

See *Labor Force and Employment Status*.

Employee-type job

An employee-type job is one in which an individual is working as an employee; that is, he or she has an on-going relationship with a particular employer. See also *Freelance Job*.

Estate

The assets, debts, and property left by someone at the time of his or her death. Life insurance proceeds are generally added to the estate.

Even Start

Federally funded program that provides families with early childhood education for their young children, adult education for parents, and parenting education and support.

Expelled

Expulsion occurs when a student is permanently prohibited from attending a learning institution, usually due to misconduct or breaking the school's rules. See also *Suspended*.

Freelance job

A type of employment where the individual does not have long-term commitments to any one employer. The individual can act independently without regard or deference to authority. The freelancer may work on one or more tasks for several people and does not have a "boss" (for example, baby-sitting or mowing

lawns), or the freelancer may work for himself or herself (for example, running a business). See also *Employee-type job*.

Futures contract*

A purchase of the right to buy or sell commodities, such as wheat, corn, and cocoa, at a specific point in the future. These contracts are traded on futures exchanges like the Chicago Board of Trade.

General Educational Development Diploma (GED)

A certificate, considered by some to be an equivalent to a high school diploma, which is obtained as a result of taking the General Educational Development Test or GED. The test provides a valid means of measuring the educational proficiency of individuals taking the test in comparison with high school graduates. (Also known as a general equivalency diploma.)

Government employee

A government employee can be one of three types: federal, state, or local. Respondents are classified according to the following definitions:

Federal government employees are respondents working for any branch of the federal government, including those elected to paid federal offices, civilian employees of the Armed Forces, and some members of the National Guard. Employees of international organizations (e.g., United Nations) and employees of foreign governments (e.g., the French Embassy) are also included.

State government employees include paid state officials, state police, and employees of state universities and colleges.

Local government employees are respondents employed by cities, towns, counties, and other local areas, including employees of city-owned businesses, water and sewer services, etc. Employees of public elementary and secondary schools also work for the local government.

Government training

Training sponsored by the government for the general public (e.g., JTPA, Job Corps). Note that training sponsored by the government for its employees (e.g., Police Academy, computer courses) is considered company training.

Head Start

Head Start is a federally funded program that provides comprehensive developmental services for disadvantaged, low-income children from birth until they reach school age.

Home equity loan

Any loan which uses the collateral of a person's home and which has a fixed number of payments and a fixed amount borrowed.

* Not included in survey help screens; listed here for user reference.

Incentive pay

In some jobs, employees receive extra money or other compensation for reaching or exceeding certain levels of performance, such as meeting established sales quotas in a given month or year.

Individual Retirement Account (IRA)

An Individual Retirement Account is a personal, tax-deferred savings account to which any individual under 70½ years of age can transfer funds to save and invest for retirement. The earnings are tax deferred until drawn upon. One can begin withdrawing from his or her IRAs after age 59½. Withdrawals prior to this age are generally subject to a tax penalty.

Inheritance

Money or property that is legally transmissible to an heir is considered inheritance.

Job

A job exists when there is a definite arrangement for regular work every week, or every month, for pay or other compensation (e.g., profits; anticipated profits; or pay “in kind” such as meals, living quarters, or supplies received in lieu of cash from someone other than a family member). A formal, definite arrangement with one or more employers to work on a continuing basis for a specified number of hours per week or days per month, but on an irregular schedule during the week or month, is also a job. Generally speaking, any job that is usually 35 hours or more per week is considered full-time. In some occupations (for example, air traffic controllers), usual weekly schedules of less than 35 hours per week are considered to be full-time.

Job Corps

A program designed to assist economically disadvantaged young men and women who need and can benefit from an intensive educational and vocational training program. The intent is to help them become more responsible, employable, and productive citizens. The program is operated in a residential group setting.

Job search

To be considered as looking for work a person must have conducted an active search for a job within the four weeks prior to the interview week. See *Active Job Search* and *Passive Job Search*.

Job Training Partnership Act (JTPA)

Legislation enacted in the fall of 1983 authorizing federal funds for employment and training. The emphasis is on private sector participation in training, without subsidies for the training participants.

Labor force and employment status

All civilian respondents who were either employed or unemployed during the survey week are defined as being in the labor force. Variables using these definitions are only available for the round 1 survey.

Employed, unemployed, and out of the labor force are defined as follows:

Employed: All respondents who during the survey week were either (1) at work—those who did any work for pay or profit or worked without pay for 15 hours or more on a family farm or business—or (2) with a job but not at work—those who did not work and were not looking for work, but had a job or business from which they were temporarily absent because they were taking time off for various reasons—are classified as employed.

Unemployed: All respondents who did not work at all during the survey week and either (1) were looking or had looked for a job in the four-week period prior to the survey; (2) were waiting to be recalled to a job from which they had been laid off; or (3) were waiting to report to a new job within 30 days are considered to be unemployed.

Out of the Labor Force (OLF): All respondents who were neither employed nor unemployed during the survey week are considered to be out of the labor force.

Land contract

Also known as a conditional sale agreement. A contract between a property owner and a potential purchaser, where they agree on the price of the home and the purchaser moves into the property. However, the seller retains legal rights to the property. The buyer makes regular monthly payments, which in some cases complete the sale of the property. In other cases, there is a large balloon payment due which completes the sale. These are often used when the buyer does not have enough money for a down payment.

Layoff

Respondents are classified as on layoff if they are waiting to be recalled to a job from which they were temporarily separated for business-related reasons, such as temporary drops in demand, business downturns, plant remodeling, material shortages, and inventory taking. They must either have been given a date to report back to work or expect to be recalled to their job within six months.

Low Income Home Energy Assistance Program (LIHEAP)

This federal program assists low-income households in meeting their costs for home energy. The funds may be used for the cost of running energy utilities, low cost weatherization (preparing the home for harsh weather), or other energy-related home repairs.

Managed investment account

A portfolio of stocks and bonds which is managed by a professional agent, usually for a fee based on the value of the assets in the account.

Margin loans through a stock broker

These are stock purchases where the purchaser borrows part of the purchase price from a bank or brokerage firm. “Margin loan” refers to the amount borrowed for the purchase. Technically, any time stock is used as collateral for a loan, even if it is already owned, it is called a margin loan.

Medicaid

A federal health program which pays for health costs of eligible low-income families. Most families receiving welfare payments are eligible, as are certain institutionalized persons.

Medicare

A federal health insurance program which pays for most hospital expenses. Most persons age 65 and older are eligible for Medicare. Persons under age 65 with major kidney failure and those receiving Social Security Disability payments are also eligible for Medicare.

Money market account

An account that invests in commercial paper, banker’s acceptances, repurchase agreements, government securities, certificates of deposit, and other liquid assets that pay money market rates of interest. The net asset value of a money market fund remains at \$1 per share, with only the interest rate going up and down.

Mortgage

This is any loan that is collateralized by a home or other real estate. The lender files a lien with the court system, which gives them the right to foreclose on (take over) the property should the borrower default on the loan.

Non-profit

This type of organization is not maintained for the making of profit.

Out of the labor force (OLF)

See Labor Force and Employment Status.

Passive job search

A passive job search is one that utilizes only methods which could not result in a job offer without any further action on the part of the job seeker. Methods include:

Looking at ads, but not responding

Attending job training programs (for example, how to write a resume, how to dress for an interview)

Other passive methods (for example, studying for a Real Estate license or picking up a job application)

See also *Active job search*.

Present value

The amount of money that sale of an asset would bring if it were sold at the time of the interview.

Primary ratings^{*}

An alphanumeric jobs/skills classification system for the Navy, Naval Reserves, Coast Guard, and Coast Guard Reserves that describes the specific job for which a person in those branches of the military is trained or to which the person is assigned.

Private employee

A private employee works for wages, salary, commission, tips, piece-rates, or pay in kind; this applies regardless of the occupation. Respondents who work for settlement houses, churches, unions, and other private nonprofit organizations are also classified as private employees.

Public housing assistance

Includes rental certificates or vouchers provided by the government for low-income families.

Regular school

A school that provides credit toward an academic diploma or degree, such as an elementary school, junior high school, high school, college, or university, as contrasted with special schools that offer certificates rather than diplomas or degrees. Schools or programs (such as technical schools, nursing programs, etc.) that provide credits that can be applied toward a regular academic diploma or degree are also considered regular school. Courses taken to prepare for the GED are not included as regular school.

Summer vacation

Respondents on summer vacation are considered enrolled. In ambiguous cases, the respondent is considered to be enrolled if he or she intends to return to school (has not dropped out or formally left school).

^{*} Not included in survey help screens; listed here for user reference.

Supplemental Security Income (SSI)

Supplemental Social Security is a form of public assistance paid to blind and disabled persons who are not covered under the Social Security program and to some people who are over age 65 whose income falls below a certain level. The program is administered through the Social Security Administration. This program should not be confused with state supplemental assistance or with regular Social Security benefits; checks are labeled “Supplemental Security Income Payment.”

Suspended

Suspension is a temporary prohibition of a student from access to school grounds. It is used as a way of punishing the student for engaging in misconduct or breaking the school’s rules. See also *Expelled*.

Talent search

A talent search is a program that focuses on educational advancement.

Temporary Assistance for Needy Families (TANF)

Commonly referred to as welfare, this program helps families with dependent children by providing financial assistance each month. States and territories operate the programs. There is a time limit on the cash assistance, and work is a major component of TANF. Work requirements under the program help recipients gain the experience need to find a job and become self-sufficient.

Tips

Payments given in addition to base wage or salary for a service performed or anticipated. See also *Bonus* and *Commission*.

Treasury notes, bills, or bonds

These long-term debt agents are issued by the U.S. government and can be owned directly or as part of a mutual fund. Bills mature in less than a year; notes mature in 1–5 years; and bonds generally have a maturity period of more than 5 years.

Trust

A trust is an arrangement whereby the right to financial assets or property is held by one party, the “trustee” (or manager), for the benefit of another, the “beneficiary.” Trusts are often used as a way in which a parent or grandparent can distribute his or her assets to a child or grandchild at a particular point in that person’s life. Trusts require the drafting of legal documents by an attorney.

Unemployed

See *Labor Force and Employment Status*.

Unemployment Compensation (or Unemployment Insurance) benefits

Payments made by state or local governments to help replace wages lost by a worker who was laid off or fired from her/his job. Benefits continue only for a set period of time and are taxable by the federal government.

Upward Bound

This program is designed to prepare youth for successful postsecondary studies.

Vocational or technical institute

This training program provides the trainee with vocational or technical skills and is not to be confused with technical classes in college. This type of program does not contribute to a college degree. Examples of vocational or technical institutes include beauty school, auto mechanics training, etc.

Vocational rehabilitation center

A facility offering specialized training to prepare disabled persons to enter or re-enter the work force.

Women, Infants, and Children Supplemental Nutrition Program (WIC)

WIC is a federal assistance program designed to provide such supplemental dietary products as milk, butter, peanut butter, and orange juice to pregnant women, nursing mothers, infants, and children.

Work

Work includes any activity for wages or salary; for profit or fees; or for payment “in kind” such as meals, living quarters, or supplies received in lieu of cash from someone other than a family member. One hour or more of such activity constitutes work. Work also includes unpaid activity of at least 15 hours a week on a family farm or business. Respondents working as civilian employees of the Armed Forces or the National Guard are classified as working.

Worker’s Compensation or Worker’s Comp

Money paid to workers who have been hurt or injured on the job. These benefits are not taxable. (Note: Worker’s Compensation includes payments from insurance companies only when the employer, not the employee, paid the premiums.)

Youth Build

Program to help disadvantaged young adults who have dropped out of high school obtain the education and employment skills necessary to achieve economic self-sufficiency.

Index

A

Abortion, 246
 Achievement scores, 80–81, 85–86, 107
 Adoption, 203, 205, 215, 226, 247
 Age, 27, 32, 33, 108, 176, 191–94, 202, 214, 222, 247, 249
 Aid to Families with Dependent Children (AFDC), 240, 270. *See also* Public assistance
 Alcohol, 163–64, 169, 184, 230
 Allowance, 237
 Any word search, 73, 260, 264
 Aptitude scores, 79–86
 Archivist note, 68, 71
 Area of interest, 64, 73, 78, 260, 261, 264
 Armed Forces. *See* Military
 Armed Forces Qualifications Test score (AFQT), 82
Armed Services Vocational Aptitude Battery (ASVAB), 3, 8, 9, 12, 29, 82–84, 264, 268
 Arrests, 178, 184
 Assets, 11, 70, 72, 219, 233, 234–36
 Attitudes, 11, 124, 165–71, 219
 Audio computer-assisted self-interview (ACASI), 27, 264

B

Behavior, rules, 172–73
 Benefits. *See* Fringe benefits
 Bibliography, i, 259
 Birth control. *See* Contraception
 Birth date. *See* Date of birth
 Birthplace, 224
 Blacks, 6, 19, 224–25
 Bonus. *See* Non-wage, non-salary pay
 Bureau of Labor Statistics (BLS), 3, 5, 38, 264

C

CAT-ASVAB. *See* *Armed Services Vocational Aptitude Battery (ASVAB)*
 CD-ROMs, 13, 73, 260

Center for Human Resource Research (CHRR), 6, 265
 Certifications, 253, 254
 Child care, 12, 127
 Child support, 219, 237, 247
 Child Trends, Inc., 66, 72, 162, 164, 168, 170, 172, 175, 178, 181, 190, 200, 219
 Children, 64, 215, 244–48
 Children of the NLSY79, 4, 5, 78, 265
 Cigarette use, 169, 174–75, 230
 Citizenship, 72, 224
 Class of worker, 123, 126, 273
 Codeblock, 67, 68, 66–71, 265
 Codebook, 66–71, 265
Codebook Supplement, 71–72, 154, 162, 190, 259
 Cohabitation, 243, 249–51
 College, 94–96, 99, 169, 199, 271, 272
 Commission. *See* Non-wage, non-salary pay
 Computer, 105, 200, 253
 Computer-assisted personal interviews (CAPI), 23, 25, 60, 265
 Continuous month/week array, 72, 154, 155, 156, 158, 191, 248, 250
 Contraception, 186, 229, 230
 County, 198
 Created variables, 57, 66, 68, 71, 72, 77, 263, 265
 Crime, 8, 11, 102, 105, 165, 177, 178, 184, 200
 Cross-sectional sample, 6, 19
 Current Population Survey (CPS), 108, 134, 136, 266
 Curriculum, 94, 102, 105, 106

D

Data hints, 14, 78
 Data sets, 13
 Data, extracting, 260
 Date of birth, 191–94, 218
 Date of interview, 22, 23, 191
 Dating, 186
 Death, 35, 184, 215, 226, 248
 Debts, 219, 234–36
 Degrees, 94, 95, 97, 99, 184, 202, 214, 247, 253

Delinquency, *11*, 102, 105, 176, 178
 Department of Defense, 8, 19, 82
 Department of Education, 8
 Department of Justice, 8
 Department of Labor, 8
 Design effects, 39, 71
 Diploma, 97, 99, *184*, 253
 Disability, 28, 228
 Discipline, 169, 172–73
 Divorce. *See* Marriage—Status
 Dropped respondents, 33, 37
 Drug use, 169, 178, 180–82

E

Education, *11*, 71, 87–107, 158–60, 202
 Employee jobs, *11*, *64*, 108, 123, 127, 131, 134, 138, 144, 146, 149, 155, 272, 277
 Employers, 109, 123–26, 154–56. *See also* Jobs
 Firm size, 123
 Number, 71, 152
 Enlistment Testing Program (ETP), 18, 19, 83
 Enrollment, 71, 92, 97, 107, 159, *184*, 247, 249, 277
 Error updates, 73, 259
 Ethnicity, 6, 9, 18, 33, 202, *214*, 224–25
 Event history, 66, 72, 118, 122, 154–60, 241, 248, 250
 Event History CD-ROM, 154
 Expectations, *11*, *12*, 183–85
 Extended screener. *See* Screener, Household Roster, and Nonresident Roster Questionnaire

F

Family, 200
 Background, 189
 Father, 226. *See also* Parents
 Fertility, *11*, *72*, *184*, 222, 244–48
Field Interviewer Reference Manual, 71
 Fielding period, 18, 22, 29, 58. *See also* Refielding period
 Financial aid, 94, 252
 Food stamps, *240*
 Freelance jobs, *11*, *64*, 108, 124, 132, 139, 144, 146, 150, 272. *See also* Self-employment

Fringe benefits, 127
 Funding agencies, i, 4, 8

G

Gangs, 169, 176, 200
 Gaps in employment, 128–30, 144, 155, 222
 Gatekeepers. *See* Screener, Household Roster, and Nonresident Roster Questionnaire
 Gender, 195, 202, *214*
 General Equivalency Diploma (GED), 97, 99, 253, 273
 Geocode, 72, 198, 266
 Geocode CD-ROM, *13*, 198, 199, 227. *See also* CD-ROMs
Geocode Codebook Supplement, 72
 Geographic mobility, *12*, 226
 Geographic regions, 72, *198*, 227
 Government training, 252, 273, 274
 Grade, 98, 159, 202, *214*, 218, 247, 249
 Grandparents, 218
 Guns, 176, 200

H

Hardmax. *See* Maximum
 Hardmin. *See* Minimum
 Hardware requirements, 260
 Health, *11*, *12*, 219, 228–31
 Health insurance, 127, 228, 276
 Height, 219, 228
 Help, i, 262
 High school, 12, 30, 99, 102, 106
 Highest grade attended/completed. *See* Grade
 Hispanics, 6, 19, 23, 224–25
 Home schooling, 80, 95
 Homework, 188, 200
 Hours worked, 101, 146–48, 152, 155, 202, 222
 Household
 Composition, *11*, *72*, 202–13, 248
 Environment, 28, 200–201
 Members, 10, 60, *64*, 158, 202–5, 237
 Residence, 20, 234, 266
Household Income Update, 13, 30, 55, 77, 221, 238, 266
 Household informant, 9, 23

Household roster. *See* Roster; *Screener, Household Roster, and Nonresident Roster Questionnaire*

I

Identification code (roster line number), 208, 210, 211, 216, 223
 Identification number. *See* Unique identification code (UID)
 Incentive experiment, 21
 Income, 11, 12, 13, 30, 70, 72, 177, 199, 203, 219, 221, 233, 237–39
 Independent youth, 233, 234, 237, 240, 266
 Industry, 71, 125, 131–33, 142, 198, 266
 Integrated Postsecondary Education Data Systems (IPEDS), 95, 199
 Intelligence scores, 79–86
Interest-Finder (I-F), 8, 12, 29, 83
 Internet bibliography. *See* Bibliography
 Internet sites, i, 3, 6, 73, 259
Interviewer Reference Manual. See Field Interviewer Reference Manual
 Interviewer remarks, 28, 55, 195
 Invalid skips, 69
 Item nonresponse, 38, 69, 118, 154, 220, 232

J

Jail, 178, 184
 Jobs, 11, 108–53, 154–56, 184, 222, 274. *See also*
 Employers
 Placement, 123
 Satisfaction, 124
 Search, 129, 134–35, 136, 253, 270, 274, 276
 Justice system, 165

L

Labor force status, 11, 108, 134, 136–37, 154, 202, 214, 218, 247, 249, 266, 275
 Labor unions, 123
 Language, 23, 25, 28, 218
 Layoff, 128, 275
 Licenses, 253, 254
 Loans, 94, 234, 252, 273, 276

M

Main file CD-ROM, 13. *See also* CD-ROMs
 Major field of study, 94
 Maptitude, 197
 Marriage, 11, 184, 249–51
 History, 222
 Status, 72, 202, 214, 218, 243
 Mature Women, 4, 5, 78, 267
 Maximum, 70
 Mental health, 170, 171, 229. *See also* Self-esteem
 Metropolitan Statistical Area (MSA), 72, 227, 269
 Military, 82, 123, 125, 140, 144, 150, 270, 271, 277
 Minimum, 70
 Mother. *See* Parents
 Multiple respondent households, 7, 8, 21, 27, 205, 221, 222

N

National Institute of Child Health and Human Development (NICHD), 4, 8
 National Longitudinal Survey of Youth 1979 (NLSY79), 4, 5, 78, 267
 National Longitudinal Survey of Youth 1997 (NLSY97), 3, 4, 5, 267
 National Longitudinal Surveys (NLS), 3–6
 National Opinion Research Center (NORC), 6, 267
 Nationality, 218
 Neighborhood environment, 28, 200–201
NLS News, 3, 13, 73, 259
 NLS User Services, i, 262
 Noninterview, 26, 33–36, 33–36, 37, 69
 Nonresident roster. *See* Roster; *Screener, Household Roster, and Nonresident Roster Questionnaire*
 Nonresponse. *See* Item nonresponse, Noninterview
 Non-wage, non-salary pay, 149, 271, 272, 274, 278

O

Occupation, 71, 125, 138–41, 142, 202, 267, 277
 Older Men, 4, 5, 78, 267
 On-the-job training, 253
 Original Cohorts, 4, 267
 Overtime, 146, 149. *See also* Non-wage, non-salary pay

P

Paper screener. *See* Screener, Household Roster, and Nonresident Roster Questionnaire

Paper-and-pencil interviews (PAPI), 268

Parent Questionnaire, 11, 12, 25, 55, 77

Parents, 11, 13, 30, 165–69, 172–73, 183, 203, 210, 215, 216, 218–23, 226, 237, 238

Partner, 186, 215, 235, 243, 247, 249–51

Paternity, 226, 246

Peabody Individual Achievement Test (PIAT)—
Mathematics Assessment, 11, 71, 85–86, 268

Peers, 11, 169

Poverty status, 221, 238

Pregnancy, 11, 128, 184, 186, 229, 230, 244–48

Primary Sampling Unit (PSU), 17, 20, 28, 40, 268

Program participation. *See* Public assistance

Proxy screener. *See* Screener, Household Roster, and Nonresident Roster Questionnaire

Puberty, 11, 228

Public assistance, 11, 72, 219, 222, 240–42, 247, 249, 275

Public housing, 240, 277

Q

Quality Education Data (QED) file, 29, 31, 97

Question name, 9, 10, 11, 13, 56–57, 59, 64, 66, 67, 69, 260

Questionnaire. *See* Survey-Instruments

R

Race, 6, 9, 18, 33, 202, 214, 224–25, 247, 249

Rate of pay, 71, 101, 149–51

Reference number, 67, 68, 260, 268

Reference periods, 22, 58

Refueling period, 22, 193

Relationships, 165–69, 186–87, 202, 203, 208, 209, 211, 216, 219, 223, 226, 247, 250

Religion, 169, 200, 201, 218, 219, 247, 249

Residence, 72, 197–99, 204, 226, 234

Responding parent, 26, 27, 64, 218–23

Response rate, 20, 22, 29, 31, 33–36, 33–36. *See also* Noninterview

Rosters, 59–66, 72, 89–92, 93, 95, 103, 117, 116–22, 123, 125, 128, 142, 144, 147, 156, 159, 205–12, 214, 216, 215–17, 223, 247, 263

S

Salary. *See* Rate of pay, Wages

Sample

- Design, 6, 71
- Size, 5, 7, 18, 32

Sampling weights, 38, 37–39, 66, 71, 259, 268

School, 11, 12, 61, 64, 169, 277

- Attendance, 98, 105, 159
- Grades, 102, 107
- Type, 71, 92, 94, 97, 102, 105, 159

School Finder, 97, 99

School surveys, 8, 12, 28, 30, 104–6, 269

School-based learning programs, 95, 101, 105, 106

Screen and go, 23

Screener, Household Roster, and Nonresident Roster Questionnaire, 9–10, 23–25, 55, 77, 202, 269

- Extended screener, 9
- Gatekeepers, 25
- Household roster, 10, 205–12, 216, 223, 248, 250
- Nonresident roster, 10, 205, 214–17
- Paper screener, 24
- Proxy screener, 24
- Simple screener, 9, 18
- Telephone screener, 25

Screening process, 20. *See also* Screener, Household Roster, and Nonresident Roster Questionnaire

Search strategies, 73, 260

Self-employment, 109, 125, 133, 140, 142–43, 147, 150, 237, 271. *See also* Freelance jobs

Self-esteem, 169, 219

Sex. *See* Gender

Sexual activity, 11, 169, 186–87, 229, 247

Shift worked, 123, 142, 146

Show card, 55

Siblings, 7, 8, 21, 203, 215, 218

Sick leave, 127

Simple screener. *See* Screener, Household Roster, and Nonresident Roster Questionnaire

Skip patterns, 69

Smoking. *See* Cigarette use

Softmax. *See* Maximum
 Softmin. *See* Minimum
 Spouse, 215, 235, 247, 249–51
 Standard errors, 259
 Standardized tests, 80–81, 105, 107
 State, 198
 Student Testing Program (STP), 18, 19, 83
 Substance use, 28, 163–64, 169, 174–75, 230
 Supplemental sample, 4, 6, 9, 19, 224
 Supplemental Security Income (SSI), 240, 277
 Survey
 Instruments, 8–13, 23–28, 29, 55–57, 78
 Suspension, 98, 159, 278
 Symbol variables, 59, 263

T

Technical Review Committee, 6
Technical Sampling Report, 71, 259
 Telephone screener. *See* *Screener, Household Roster, and Nonresident Roster Questionnaire*
 Television, 172, 188, 201
 Temporary Assistance to Needy Families (TANF), 157. *See also* Aid to Families with Dependent Children (AFDC)
 Tenure, 71, 128, 144–45, 152
 Time use, 11, 188
 Tips. *See* Non-wage, non-salary pay
 Topcoding, 70, 220, 233
 Training, 11, 64, 99, 101, 252–54, 271, 279
 Transcript survey, 3, 12, 30, 80, 106–7, 269
 Transfers, 237
 Tuition, 94, 127

U

Unemployment compensation, 240, 241, 278

Unemployment insurance, 158
 Unemployment rate, 199
 Unique identification code (UID), 66, 116, 117, 121, 128, 156, 212, 217
 Universe, 55, 58, 67, 68, 69
 User notes, 14, 78
 User Services. *See* NLS User Services
 Utilities, 200, 234

V

Vacation, 127
 Valid skips, 69, 84
 Valid values, 67, 68, 69, 70
 Validation reinterviews, 31
 Variables, 57–66
 Titles, 58, 64, 67, 68, 73

W

Wages, 149–51, 203, 237
 Weeks worked, 71, 101, 128, 144–45, 152, 202
 Weight, 219, 228
 Welfare. *See* Public assistance
 Women, Infants, and Children (WIC), 240, 279
 Work experience, 152
 Work history, 152, 154–56, 222
 Worker's compensation, 158, 240, 241, 242, 279
 Work-study, 94

Y

Young Adult Children of the NLSY79, 4, 5, 78
 Young Men, 4, 5, 78, 269
 Young Women, 4, 5, 78, 269
Youth Questionnaire, 10, 11, 25, 29, 55, 77