Appendices

Appendix A: Quick Reference Guide

This appendix summarizes some key pieces of information regarding this *User's Guide* and other documentation products for the Young Women, the NLS surveys, and the NLS CD-ROMs. It also provides answers to some common questions about the Young Women and tells users how to get additional help when necessary.

A.1 Guide to NLS Documentation

The *User's Guide* provides in-depth information about the Young Women data. Users requiring general information about the history, administration, or other cohorts of the NLS project should consult the following table for the most appropriate NLS document.

Table A.1 List of NLS Documentation

NLS Handbook	This comprehensive introduction to the NLS gives readers general information about all NLS cohorts and the main topics of investigation for each.
NLS Internet Bibliography http://www.nlsbibliography.org	This on-line searchable database provides citations for research using NLS data.
User's Guides	These cohort-specific guides help researchers understand NLS variables, survey instruments, documentation techniques, and other technical issues.
Questionnaires	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.
Flowcharts	Schematic diagrams depict universe information and skip patterns for many survey instruments.
Codebook Supplements	Supplementary attachments and appendices contain variable creation, description, and coding information not present in the questionnaires.
Compact Disc User's Guides	These guides provide installation, usage, and maintenance instructions for the CD-ROMs.
NLS Web Homepage http://stats.bls.gov/nlshome.htm	This internet site offers an overview of the NLS programs.

With the exception of the bibliography and homepage, which are provided on-line, users may order any of the documents listed above by contacting NLS User Services. Contact information for NLS User Services is provided under "Additional Support" at the end of this appendix.

A.2 About the Young Women CD-ROM

The Young Women survey is a panel data collection—specifically, the same persons are interviewed year after year. Young Women respondents are often asked the same or similar questions in different surveys to gauge the change in behavior over time.

The database is organized by respondent. Information on each respondent is stored in a record. In each respondent-specific record, the variables are arranged in chronological sequence. Users are provided with data extraction software, called CHRRDBA, on each CD-ROM to search this large database and extract the specific variables needed.

Search Strategies

Variables can be selected for extraction either using a Search List (reference number or question number) or Search Index (year or any word in context). The following is a brief description of the search options.

Reference number: Each variable is assigned a 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 number: Each year, a separate survey instrument is used to collect data. A questionnaire item (or question number) refers to the location of a given variable in the printed or electronic questionnaire.

Any word: The database retrieval software allows the user to search for and select those variables whose titles contain any single word or combination of words.

Area of interest: Each variable is assigned to a topical area of interest. For example, questions on a respondent's health and medical insurance are grouped in the "Health" area of interest. Researchers should be aware that an individual question can be linked to only one area of interest, so questions that apply to a common research topic may appear in different areas of interest.

Year: The user can select a specific survey year and choose variables collected in that specific year.

Accessing the Data

This section briefly describes how to access the data from the Young Women CD-ROM. The reader should consult the *NLS Original Cohort Databases Compact Disc User's Guide* for more detailed information about the DOS-based database retrieval system.

Hardware Requirements. Minimum hardware requirements for using the CHRRDBA software to access the Young Women data are (1) an IBM compatible personal computer (PC), running MS-DOS, and (2) a CD-ROM drive connected to the PC.

Extract Instructions. Following is a simplified list of steps to retrieve data:

    (1) Run the DOS Program. If accessing the data from the CD-ROM, double click on CHRRDBA.EXE from the file manager.

    (2) Specify output file path.

    (3) Select "Younger Women 14–24 in 1968."

    (4) Specify a name for the Extract Specification File.

    (5) Select data by either Search Index (word, year) or Search List (reference number, question number).

    (6) Perform an extract. Produce a codebook and/or a SAS or SPSS format extract file.

    (7) Exit the Menu (and software).

Extract Outcomes. Users can extract data for specific subsamples and in various formats. The following is an abbreviated list:

    (1) Delimited or formatted ASCII files for SAS or SPSS

    (2) DBASE formatted file

    (3) Codebook of extracted variables

    (4) Summary of extract

    (5) Subsample data by user-specified equation

A.3 Glossary of NLS Survey Terms

Any word search. This CD-ROM search function allows users to select any word or words and to view all variables which contain those words in their titles.

Area of interest. Variables are grouped by common topical areas. The CD-ROM includes a search function so that users can view variables in an area of interest.

Bureau of Labor Statistics (BLS). This agency of the U.S. Department of Labor sponsors and oversees the National Longitudinal Surveys project.

Center for Human Resource Research (CHRR). A research unit at The Ohio State University, CHRR is responsible for the management of the Original Cohorts, documentation and dissemination of the data, and user services.

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. The codebook contains complete information about all the variables in a data set and is included electronically on the CD-ROM. It comprises a number of codeblocks presenting information about individual variables.

Computer-assisted personal interview (CAPI). These interviews, used since 1995 for both women's cohorts, 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.

Household Record Card. This survey instrument was used during PAPI interviews to collect information about members of the respondent's household. During the interview, demographic information was transferred from the cards to the "Household Roster" section of the main questionnaire, so data from the *Household Record Cards* contain "Household Roster" as part of their variable titles.

Household Screener. This survey instrument was used in 1966 to identify respondents eligible for the Young Women cohort and the other Original Cohorts. It collected demographic information about all members of each surveyed household.

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 19 times through 1999. 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, were selected for inclusion in the newest NLS cohort. The NLSY97 numbers 8,984 respondents, and three interviews have been conducted with the cohort to date.

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 during the 1966 household screening and first surveyed between 1966 and 1968.

Paper-and-pencil interview (PAPI). Traditional paper-and-pencil instruments were used with the Young Women's cohort for each survey through 1993.

Primary Sampling Unit (PSU). A Primary Sampling Unit consists of one or more Standard Metropolitan Statistical Areas (SMSAs), counties (or parishes in some states), parts of counties (parishes), or independent cities. PSUs were the basis for sampling Young Women respondents.

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.

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.

A.4 How to Get Help

Sometimes users have questions about the Young Women data, database retrieval software, or documentation. The following are strategies for finding answers to these questions.

On-line and Paper Documentation Help. CHRR provides both on-line and paper documentation help sources for users. The on-line help is present on the database retrieval software program. Press the F1 function key at any time in the CHRRDBA data retrieval system to receive on-line help. Paper documentation help includes the *NLS of Young Women User's Guide*, the *NLS Handbook*, the *NLS Original Cohort Databases Compact Disc User's Guide,* and assorted supplemental hard copy documentation (see Table A.1). The paper documentation contains the answers to most questions.

Frequently Asked Questions. Table A.2 provides answers to commonly asked questions about the Young Women data and accessing the data.

Additional Support. If questions arise which are not answered in the documentation, contact NLS User Services at:

<div align="center">

NLS User Services
Center for Human Resource Research

</div>

921 Chatham Lane, Suite 100
Columbus, Ohio 43221-2418
614-442-7366

E-mail: usersvc@postoffice.chrr.ohio-state.edu

Table A.2 Frequently Asked Questions

Question	Answer
Is there a distinction between valid and invalid skips? What do the missing value codes mean?	In PAPI interview years (1967–92), there is no distinction between valid and invalid skips in the data. Noninterviews, valid skips, and invalid skips are lumped together as NAs (not applicable) in the codebook. If you wish to separate valid skips from invalid skips, you must design your own program by using the questionnaires and following the skip patterns. In the data, a code of -5 is used to represent NAs and a code of –2 indicates DKs (don't know).
	In CAPI interviews (1995–present), invalid skips are coded as -3, valid skips are coded as -4, and noninterviews are coded as -5. A code of -1 represents a refusal and -2 signifies a response of don't know. More information on the coding of missing responses is presented in section 3.3 of this guide.
Do the sampling weights correct for oversampling?	Yes, the first year weights correct for oversampling. The weights for each subsequent survey year correct for attrition and oversampling.
How can multiple respondent households be identified?	Use the "search any word" feature of the Search Engine to search for all occurrences of the words "identification" and "code." This search will result in a listing of variables which identify other members of the Young Women's cohort who are related to the respondent; it will also provide identification codes for members of other NLS Original Cohorts who are related to the respondent.
In trying to calculate actual work experience for each woman, how can missing years be accounted for?	A complete work history covering all time intervals between surveys for the Young Women cannot be constructed. Details about work experience coverage across survey years and potential gaps in coverage are provided in the "Work Experience" section of this guide.
How can an employer tenure variable be created?	Tenure with an employer can be constructed with information on start and stop dates of the current/last job and intervening jobs between surveys in conjunction with reports on weeks unemployed or out-of-the-labor force. Consult the "Work Experience" section of this guide for further details and possible limitations of the data.
Is there any interview that provides a geographic residence variable smaller than region?	No. This database does not contain a residence variable smaller than region. The regional distinction is South/non-South. However, the surveys through 1988 provide some general information about the respondent's residence (is it in an SMSA, local unemployment rate, etc.) and all surveys include an indicator of whether the respondent moved since the previous interview. For more information, refer to the "Geographic Residence & Environmental Characteristics" section of this guide.
Why is the 'Total Family Income' variable not available in all years?	'Total Family Income' is a KEY variable which is only created for survey years in which a personal interview is conducted. See Table 2.4.1 for a listing of the type of interview by survey year.

Appendix B: Original Cohort Dictionary of Key Words

On the original data tapes, the Young Women's data were searchable by "keyword"—that is, the user could select from a predetermined list of key common words in variable titles. Although this search option has been replaced by the any word search, the keywords remain in the variable titles. This list of keywords used for the Young Women's cohort is provided to help researchers focus their any word searches by employing common words included in variable titles for a given topic.

KEY	Bose
1st Job	Census Division
2 or More	Certificate
35	Child
5 Years	Child Care
6 or More Months	Children
Absent	Class of Worker
ADA	Collective Bargaining
AFDC	College
Age	College Degree
Age 14	College Survey
Age 34	Colleges
Age 35	Commitment to Work
Age 50	Company Training
Alimony	Comparative Job Status
Alternative Job	Contact
Answered	Correctional Institution
Apartment	Counseling
Assets	Counselors
Assistance	Cultural Exposure
Attended	Current Job XX (For each survey 68–87)
Attending	Current or Last Dual Job
Attitude	Current or Last Job XX (For each survey 68–87)
Attitudes	
Automobile	Current School
Benefits	Curriculum
Birth	Days per Week
Birthplace	Debt
Bonds	Dependents
Books	Desire

Desired	Full-Time
Different Area	GED
Diploma	GED Score
Disability	Goal
Discrimination	Grade Attended
Disliked Most	Grade Attending
District Wide	Grade Completed
Dual Job	Graduate
Dual Jobs	Graduated
Duncan	Grandfather
Duncan Index	Grandmother
Education	Guidance
Educational	Head of Household
Employee Association	Health
Employer	Height
Employers	Helping
Employment	Helping Others
Employment Status	High School
Employment Status Recode	High School Subject
Enrolled	Homework
Enrollment	Hourly
Enrollment Status	Hours
Expected	Hours per Day
Expelled	Hours per Week
Expenditure per Pupil	Hours Worked
Extra-Curricular	House
Faculty	Household
Family	Household Chores
Family Member	Household Record
Family Members	Housework
Farm	Husband
Father	Hypothetical Job Offer
Field of Study	Identification Code
Financial Aid	Income
Financial Position	Index of Demand
Food Stamps	Industry
Foreign Language	Interfirm

Interruption	Marriages
Interview Date	Medical
Interview Length	Medical Insurance
Interview Method	Method of Finding
Interview Status	Method of Seeking
Intrafirm	Mortgages
IQ Category	Most Recent College
IQ Score	Most Recent Job
IQ Test Name	Mother
Job	Motivation
Job Characteristics	Move
Job Satisfaction	Moved Backward
Job Supervision	Mutual Funds
Jobs	Nationality
Knowledge	Newspapers
Labor Force	Noninterview
Labor Force Status	Not Empld XX (For each survey 68–87)
Labor Market	Occupation
Last Interview	Occupation Desired
Last Job	Occupational
Last Year in High School	OLF
Layoff	OLF XX (For each survey 68–87)
Leisure	On the Job
Liability	On the Job Training
Library	Out of School
Library Card	Overtime
Liked Most	Overtime Pay
Limit	Parents
Limitations	Part-Time
Locate	Partner
Location	Pension
Longest Job	Plan to Seek
Loss of Current Job	Plans
Magazines	Probation
Marital Status	Progressed
Market Value	Promotion
Marriage	Property

Public Assistance	SSI
Race	Stocks
Rate of Pay	Study
Reading Material	Summer
Real Estate	Survey Week
Religion	Survey Week XX (For each survey 68–87)
Rent	Suspended
Residence	Teacher
Responsibility	Teachers
Restrictions	Tenure
Retirement	Training
Return to College	Transfer
Return to High School	Transportation
Return to School	Travel
Returned to School	Tuition
Rotter Scale	Typing or Shorthand
Round Trip	Unemployed
Same Area	Unemployed XX (For each survey 67–87)
Sampling Weight	Unemployment
Savings	Unemployment Compensation
Savings Bonds	Union
School Quality	Vocational
School Survey	Volunteer
Seek	Wages
Seeking	Weeks in Labor Force
Self-Employed	Weeks Not Working
Separation	Weeks OLF
Services	Weeks Unemployed
Sex	Weeks Worked
Sibling	Weight
Siblings	Welfare
SMSA	Wife
Social Security	Women Working
Spell Not Working	Work Schedule
Spells of Unemployment	

Appendix C: Unpacking Multiple Entries

Responses to multiple entry questions found in early years of the surveys of the four Original Cohorts were coded in a geometric progression format to conserve space on the tape. Variables such as 'Method of Seeking Employment,' 'Method of Finding Current or Last Job,' 'Type of Financial Aid Received,' 'Type of Child Care Arrangement,' and numerous health-related questions have been formatted in this way since the surveys began. Multiple entry items are identified by an asterisk under the source code box in the questionnaire and by a special detailed codeblock in the documentation. These responses need to be "unpacked" before they can be used in analysis. Although the example below pertains specifically to the Mature Women's cohort, it is applicable to the Young Women as well.

Example: Codes for the variable R03380., 'Fringe Benefits at Current Job 77,' range from 1 (the respondent reported only one such benefit, "medical insurance") to 259 (the respondent reported "medical insurance," "life insurance," and "paid sick leave") to 1023 (the respondent reported that she had access to all of the benefits listed). Although there are several different ways to sort out which respondent has positive answers on which components, this appendix provides one example in SAS and one example in SPSS.

Program 1: Unpacking Fringe Benefits Data in SAS

This SAS program unpacks fringe benefits from the variable "fringe." It creates 10 (dichotomous) dummy variables indicating the presence or absence of each of the 10 benefits. Each dummy is set to missing if fringe is missing (coded at -998 or -999). Note that the variables are created in reverse order from the codeblock, i.e., MEDICAL is code 1 on the tape and FRINGE10 in the program. The program statements listed below can be modified by the user to include the expanded set of fringe benefits available in later survey years as well as to unpack other multiple entry variables by extending the dummy, the counter, and the number of variables to agree with the total number of responses listed in the codeblock in the documentation.

```
data benefits;
infile 'D:\documents\requests\unpack.dat' lrecl=4;
input
 R0338000  4.;
  if R0338000 = -998 then R0338000 = .;
  if R0338000 = -999 then R0338000 = .;
label R0338000 = "FRINGE BNFTS CUR_JOB_77";
array fringe fringe01-fringe10;
do over fringe; if R0338000 ne . then fringe=0; end;
all=R0338000;
```

```
if all ge 512 then do; fringe10=1; all=all-512; end;
if all ge 256 then do; fringe09=1; all=all-256; end;
if all ge 128 then do; fringe08=1; all=all-128; end;
if all ge  64 then do; fringe07=1; all=all- 64; end;
if all ge  32 then do; fringe06=1; all=all- 32; end;
if all ge  16 then do; fringe05=1; all=all- 16; end;
if all ge   8 then do; fringe04=1; all=all-  8; end;
if all ge   4 then do; fringe03=1; all=all-  4; end;
if all ge   2 then do; fringe02=1; all=all-  2; end;
if all ge   1 then do; fringe01=1; all=all-  1; end;

label fringe01='medical,surgi';
label fringe02='life insuranc';
label fringe03='a retirement  ';
label fringe04='training/educ';
label fringe05='profit sharin';
label fringe06='stock options';
label fringe07='free….meals';
label fringe08='free…..mdse';
label fringe09='paid sick lea';
label fringe10='paid vacation';
run;
```

Program 2: Unpacking Fringe Benefits Data in SPSS

The SPSS program works in the same way as the SAS program. Users of this alternative package can follow this template.

```
/* UNPACKING 1981 YOUNG MEN FRINGE BENEFITS: SPSS/

compute FB1=0                              compute FB7=0
variable labels FB1 '81 NONE'              variable labels FB7 '81 STOCK'
compute FB2=0                              compute FB8=0
variable labels FB2 '81 FLEX HRS'          variable labels FB8 '81 PROFT'
compute FB3=0                              compute FB9=0
variable labels FB3 '81 PAID VACATION'     variable labels FB9 '81 TRED'
compute FB4=0                              compute FB10=0
variable labels FB4 '81 PD SICK'           variable labels FB10 '81 RETR'
compute FB5=0                              compute FB11=0
variable labels FB5 '81 FR MERCH'          variable labels FB11= '81 LIFE'
compute FB6=0                              compute FB12=0
variable labels FB6 '81 FR MEALS'          variable labels FB12 '81 HLTH'
```

compute FB81a=FB81
variable labels FB81a 'variable for none'

do if (2048 le FB81)
compute FB1=1
compute FB81=FB81-2048
else
compute FB1=-4
end if

do if (1024 le FB81)
compute FB2=1
compute FB81=FB81-1024
else
compute FB2=-4
end if

do if (512 le FB81)
compute FB3=1
compute FB81=FB81-512
else
compute FB=-4
end if

do if (256 le FB81)
compute FB4=1
compute FB81=FB81-256
else
compute FB4=-4
endif

do if (128 le FB81)
compute FB5=1
compute FB81=FB81-128
else
compute FB5=-4
end if

do if (64 le FB81)
compute FB6=1
compute FB81=FB81-64
else

compute FB6=-4
end if

do if (32 le FB81)
compute FB7=1
compute FB81=FB81-32
else compute FB7=-4
end if

do if (16 le FB81)
compute FB8=1
compute FB81=FB81-16
else
compute FB8=-4
end if

do if (8 le FB81)
compute FB9=1
compute FB81=FB81-8
else
compute FB9=-4
end if

do if (4 le FB81)
compute FB10=1
compute FB81=FB81-4
else
compute FB10=-4
end if

do if (2 le FB81)
compute FB11=1
compute FB81=FB81-2
else
compute FB11=-4
end if

do if (1 le FB81)
compute FB12=1
compute FB81=FB81-1
else
compute FB12=-4
end if