
**CHAPTER 5: DATA FILES &
DOCUMENTATION**

The NLSY79 Child and Young Adult data release contains child-specific information from the Child and Young Adult surveys as well as the mother's main Youth interviews. Certain variables are derived from the longitudinal record of each NLSY79 mother while others represent the questions administered during the Child and Young Adult interviews and the responses from each assessment. Detailed information on the types of data available for the NLSY79 Children and Young Adults can be found by examining print or on-line field instruments and by searching the codebook indices. Instructions on how to search the database can be found in the NLS Investigator's User's Manual that accompanies the data file. Section 3.4 of the *NLSY79 User's Guide* also describes the search functions. This guide is available electronically at <<http://www.bls.gov/nls/79guide/nls79usg.htm>>. Users who are interested in items based on data from the mother's record are encouraged to access copies of the main Youth questionnaires and to review the NLSY79 main Youth documentation.

The NLSY79 Child data file includes demographic and family background, pre-and postnatal health history, home environment reports, childcare and schooling experiences, items and scores from the biennial child assessments, and all items from the child "10 and older" self-report questionnaire. The Young Adult file contains 1994-2000 questionnaire items covering areas such as family background, schooling, training, work and military experiences, relationship history, fertility, health, and drug and alcohol use, as well as a set of created variables for each round. Geographic information for young adults is available on a separate geocode file. Constructed variables drawn from the mothers' records provide information on each mother's household composition, quarterly employment referenced to the birth of each child, and family background. While most information is cross-sectional, many variables link maternal events or behaviors to the child's life cycle—specific points after, or in some instances, before the child's birth. Any item from the complete record of the mother's main Youth record can be linked to the child and young adult files. This chapter discusses the structure of the public data files, the types of documentation available, how to merge files, and where to access user information.

Types of Variables

Variables present in the NLSY79 child and young data files are derived from a number of sources. Some variables represent the unedited respondent answers from the child, mother, or young adult, while others are constructed. The type of variable may affect (1) the

physical placement of the variable within the codebook (its sequence in the reference number list) and (2) the assignment of a variable to a particular area (or areas) of interest. Types of variables that appear in the public releases of the Child and Young Adult files include:

1. Direct (or raw) responses from a questionnaire or assessment or other survey instrument.
2. Recoded or edited variables constructed from raw data according to consistent procedures, e.g., coding of verbatim responses about jobs done for pay or religion other than the precoded categories. Such variables are marked as recode versions of the original.
3. Constructed variables based on responses to more than one data item or multiple reports to the same item, either from cross-sectional or longitudinal information. Some of these created variables are indices or scale summations, such as the assessment scores, and others are individual items edited for consistency where necessary, e.g., child background characteristics such as age, date of birth, and gender.
4. Constructed variables from a non-NLS data source, e.g., the County & City Data Book information present on the NLSY79 Young Adult geocode data files.
5. Variables provided by NORC or another outside organization based on sources not directly available to the user, e.g., the transcript data and test scores from the child school survey.
6. Data collected from or about one universe of respondents reconstructed with a second universe as the unit of observation, e.g., variables on the NLSY79 Child File that are based on inputs from the mother's main Youth record but linked to each child.

Linking the Data & Documentation

Child Question & Variable Names

Variables in the NLSY79 child data set are named according to the source from which they are derived. Items from the Child, Mother, and Child-Self-Administered supplements have question names linked to the location in the instrument (e.g., CS94-14: HAS CHILD EVER BEEN ENROLLED IN HEAD START? - Q.14 asked in the first section of the Child Supplement in 1994). Acronyms are used to identify the child assessments administered at the time of the survey (e.g., *BPI2000* for Behavior Problems Index-Raw Score, 2000). Mnemonic names identify constructed variables that are based on inputs taken from the main NLSY79 survey (e.g., *AGEMOM2000* for Age of mother at interview date, 2000).

Questions from the mother and child assessment instruments are generally identified by “MS” for Mother Supplement or “CS” for Child Supplement and in 2000, “CSAS” for the

Child Self-Administered Supplement. The digits that follow this prefix refer to the year of the survey. Through 1992, the next two numbers refer to the deck number printed at the top of each page in the survey instrument, and finally to the column number printed alongside each question in the survey instrument. For instance, the question name CS861217 refers to deck 12, column 17 of the 1986 Child Supplement. Deck numbers for the Child Self-Administered Supplement follow consecutively those in the regular Child Supplement for each year; questions are numbered in the same manner as the regular Child Supplement.

Starting in 1994, question numbers more closely parallel those of the main youth CAPI. However, the assessment items from the Child Supplement have been assigned mnemonic variable names. The current Child documentation assigns question names, the unique identifiers given to identical questions across CAPI survey years. They replace the deck-and-column numbering system used in previous surveys. *Subsequent CAPI surveys that field this question will use the same question name to help identify comparable questions across years.*

Question names for items from the child questionnaires generally indicate the section of the instrument from which each question was derived. For example, child question names beginning with SCHL come from the schooling section of the Child Supplement questionnaire; items prefixed with “HLTH” designate the child health questions in the supplement. The question name also indicates the order in which the question was administered within each section of a questionnaire. Constructed variables are usually named according to the topic of their content. The item called “FSTYRAFT” (C00052.) in the CHILD BACKGROUND area of interest refers to the “first survey year of mother after the child’s birth.” Such items on NLSY79 mothers and their children are usually based on multiple inputs from the main youth and child survey instruments. Codebook entries for these items may include a reference to one or more main file reference number to give the user an example of the main Youth question on which the child-specific item is based. In the PRE/POST NATAL CARE area of interest the item PRE0009 (C03201.) refers the user to R13284. to see the exact inputs to the variable.

Young Adult Question & Variable Names

Question names in the Young Adult surveys follow one of three general patterns. Individual items from the Young Adult CAPI questionnaire are usually prefixed by “Q” and the section number. The suffix generally indicates the order in which questions appear. Because questions asked in multiple rounds retain the same question name, many suffixes also include letters (e.g., Q4-31B). Thus Q2-19 represents the following question (asked in each survey year) of the Family Background portion (Section 2) of the YA CAPI questionnaire:

Q2-19 IS THIS THE FIRST TIME R STOPPED LIVING WITH MOTHER?

Questions from the Young Adult self-report booklet (in 1994 and 1996) are prefixed with “YA” followed by the survey year, and then a designation of the deck and column number:

YA940939 YA SELF-REPORT: EVER BEEN ON PROBATION?

YA960839 YA SELF-REPORT: EVER BEEN ON PROBATION?

In 1998, question names for these self-report items no longer incorporate the deck and column numbers and are simply named according to question number sequence in the booklet:

YA98100 YA SELF-REPORT: EVER BEEN ON PROBATION?

Starting in 2000, self-report questions appear in the self-report section of the YA CAPI questionnaire and are prefixed by “YASR”:

YASR-67 EVER BEEN ON PROBATION?

Constructed variables on the Young Adult file are identified by mnemonic names related to their content. For example, the series of variables that identifies the cohabitation status in each survey round are named “COHAB” followed by a year designation.

Child Question Naming Conventions—Details

Mother Supplement Question Names. From 1986-1996 question names for items in the Mother Supplement have the following three components: (1) “MS” to designate the source of the item, (2) a 2-digit number to indicate the year of the survey round, and (3) deck and column number. For example in the item listed below, MS960329, the “MS” indicates the Mother Supplement, the “92” means the item is from the 1992 survey, and “0329” shows that the item was in Deck 3, Column 29 in the booklet.

1992: MS920323 HOME PART B (3-5 YRS): HOW OFTEN MOTHER
READS TO CHILD
1994: MS940331 HOME PART B (3-5 YRS): HOW OFTEN MOTHER
READS TO CHILD
1996: MS960329 HOME PART B (3-5 YRS): HOW OFTEN MOTHER
READS TO CHILD

In 1998 a similar system was used in naming Mother Supplement questions except that section number and question numbers replace the deck and column numbers, which were no longer relevant. In the example below, MS981B01 means that this item is from the Mother Supplement, 1998, Section 1-B (the HOME), question 1.

MS981B01 HOME PART B (3-5 YRS): HOW OFTEN MOTHER READS
TO CHILD

In 2000 a new naming system was introduced that more closely parallels that used for the main Youth and young adult surveys. As seen in the example below, questions from the Mother Supplement still use the MS designation and a reference to the section and item number within the section but do not include a reference to the year of the survey.

HOME-B01 HOME (AGE 3 YRS): HOW OFTEN MOTHER READS TO
CHILD
MS1-A01 HOME (AGES 4-5 YRS): HOW OFTEN MOTHER READS TO
CHILD

Note: In 2000 the HOME and Temperament items have different question naming conventions depending on the age of the child. Items for children under age 4 were administered in the Child CAPI Supplement are named according to the name of the assessment (e.g., HOME-B01 above). HOME and Temperament questions for children age 4 and older appear in the Mother Supplement and are therefore prefixed with “MS” followed by the section and item number (MS1-A01 means Mother Supplement, Section 1-A, question 01). Notes have been entered into the codebook for these items to alert users to this change in 2000.

Child Supplement Question Names. Through 1992 questions that were administered in the Child Supplement include CS for the supplement, a year number, and deck and column. When CAPI was introduced in 1994 questions were no longer identified by deck and column but simply by item number.

CS921746 CHILD BACKGROUND: HAS CHILD EVER BEEN
ENROLLED IN HEAD START?

CS94-14 CHILD BACKGROUND: HAS CHILD EVER BEEN ENROLLED
IN HEAD START?
CS96-14 CHILD BACKGROUND: HAS CHILD EVER BEEN ENROLLED
IN HEAD START
CS98-14 CHILD BACKGROUND: HAS CHILD EVER BEEN ENROLLED
IN HEAD START?

Starting in 2000 questions in the Child CAPI Supplement are generally identified by the section of the questionnaire. In the example that follows BKGN stands for the Child Background section of the questionnaire:

2000: BKGN-12 CHILD BACKGROUND: HAS CHILD EVER BEEN
ENROLLED IN HEAD START?

Child Self-Administered Supplement Question Names. Questions in the Child Self-Administered Supplement for children 10 and older are prefixed by “CS” until the 2000 survey when they begin with “CSAS.” Through 1996 deck and column numbers were used to designate the question location in the CSAS booklet. Starting in 1998, item numbers are used.

CS960113 CHILD SELF-ADMIN: GONE TO MOVIES WITH PARENTS
IN LAST MONTH
CS98001A CHILD SELF-ADMIN: GONE TO MOVIES WITH PARENTS
IN LAST MONTH
CSAS001A CHILD SELF-ADMIN: GONE TO MOVIES WITH PARENTS
IN LAST MONTH

Child & Young Adult Reference Numbers

All variables on the NLSY79 Child and Young Adult files are assigned unique identifiers called reference numbers, which determine the relative position of each variable within the codebook. Reference numbers that start with “C” are used for data items on the Child file, and “Y” numbers for the Young Adult. The “C” numbers (for the Child) and the “Y” numbers (for the Young Adult) appear in each codebook entry and also form the basis of the variable names on the SAS and SPSS control cards that are generated by the extraction procedures on the data set.

Through 1994, reference numbers assigned to child variables were specific to a given data release and changed in subsequent releases. Different prefixes identified different releases: variables from the first 1986 release of the NLSY79 child data were preceded by the letter Cxxxxx.xx; second release 1988 variables were preceded by a Dxxxxx.xx, and so forth.

Since 1996, all child variables have been numbered Cxxxxx.xx. Reference numbers for Young Adult variables are prefixed with the letter “Y.” Decimals in the Young Adult reference numbers generally indicate that more than one variable has been derived from a single question. Users of the main Youth data will recognize these reference numbers as conceptually equivalent to the “R” numbers used in the NLSY79 main Youth documentation.

Contents of the Child & Young Adult Public Data Files

NLSY79 Child and Young Adult data files and documentation materials are available on CD-ROM or online for download via FTP. The current public release contains the complete set of Child and Young Adult files that can be linked with the entire record for all NLSY79 females. The 1979-2000 NLSY79 main Youth file is also on the Child/Young Adult CD as a separate file and is also available for download via FTP. See below for specific instructions about each type of file.

The current Child and Young Adult data release supersedes all the files created in connection with the previous releases of the data. The current version of this file includes nearly all the child-based variables created for the last data release, updating information for children already born by 1998, as well as adding information for those born between 1998 and 2000. The file includes the individual items from the Child assessments as well as all the derived raw and normed assessment scores for each survey round. Child-specific information on the file includes demographic and family background, pre-and postnatal health history, home environment reports, childcare experiences, and all items and scores from the biennial child assessments. The Young Adult file contains 1994-2000 questionnaire items, covering areas such as work experience, military service, training, schooling, family background, relationship history, income, health, fertility, childcare, and drug and alcohol use, as well as a set of created variables for each round. Constructed mother-specific variables include information on each mother’s household composition, quarterly employment referenced to the birth of each child, and family background.

The Child and Young Adult files are grouped in topical areas of interest (formerly referred to as “record types”) but may be accessed using a variety of search strategies, including “any word in context,” year, reference number, and question number. The search and extraction software facilitates the process of finding and using the appropriate variables.

Users should refer to the current *NLS Database Investigator Guide*, which accompanies the data, both on the CD and via download.

Constructed Child- and Mother-Specific Variables. In addition to the child questionnaire items and constructed assessment scores discussed in greater detail elsewhere in this document, the NLSY79 Child and Young Adult data sets contain a number of other constructed variables. Also, some constructed variables, such as pre- and postnatal care and child usual residence, are drawn from child-specific information collected in the mother's main Youth interview. Other elements are constructed from mother-based information that does not vary across children, for example, maternal household composition and family background. Constructed variables are generally found in the following Areas of Interest: CHILD BACKGROUND, FAMILY BACKGROUND, MATERNAL HEALTH, MATERNAL HOUSEHOLD COMPOSITION, PRE/POST NATAL CARE, and MATERNAL WORK HISTORY. These created items include sibling identifiers, maternal family background, maternal household composition at each interview, maternal work histories, family educational background, and maternal health. Details on these constructed variables can be found in Chapter 2 of this guide and in Chapter 4 of the *NLS Handbook*. Mother-specific information present on the NLSY79 main data file and on special data sources such as the work history and geocode main youth files can be linked with the child data by case ID.

Constructed Young Adult Variables. In addition to the questionnaire items from the 1994-2000 Young Adult surveys, several constructed variables for Young Adults are available in the current data release (see Chapter 3). Some of these are available for all YA respondents who were interviewed in *any* survey year, while others are specific to a particular survey round. These variables are all located in the Area of Interest on the CD called YA COMMON KEY VARIABLES. The following variables are constructed for *all* young adults: young adult ID (Y00001.00), date of birth, gender, race, and the ID code of the mother.

Child & Young Adult Data on CD-ROM

The Child and Young Adult data, together with search and extraction software, are available on compact disc (CD-ROM) media for use in a Windows environment. The CD contains the 1986-2000 Child file, the 1994-2000 Young Adult file, the complete NLSY79 main Youth 1979-2000 file, and documentation for the Round 19 (2000) Child & Young

Adult. The Young Adult questionnaire, along with an overview and associated help screens, is available in HTML format on the CD. The Young Adult questionnaire is fully hyperlinked, allowing users to easily follow the flow of respondents through the questionnaire. The CD contains all the child-based created variables, all the assessment variables (both “raw” items and created scores), all the Young Adult CAPI and self-report data, and selected constructed variables drawn from the NLSY79 main Youth 1979-2000 database³. One set of current Child questionnaires and the 2000 Child Assessment Tables report are included in PDF format on the CD. Users who wish to retrieve information from the mother’s record beyond the variables included in the child files will need to access the main Youth file, which is on the CD as a separate zipped file. A copy of the *NLSY79 User’s Guide* is automatically available on the CD when accessing the main Youth file.

Child & Young Adult Data On-line for Download

The Child-Young Adult data file, as well as the main Youth file, is available for download via FTP. Instructions for obtaining the data can be found on the *NLS Information* Web page <<http://www.nlsinfo.org>>. At this site, click on Order Data, and select NLSY79-Children-Young Adults. You will see item number:

DCYA-R10	NLSY79 Child/Young Adult 1979-2000 Data Release R10.0 (CD-ROM includes NLSY79 main file data R11.2)
----------	--

Clicking the box on the far right (Size/Download) will make the Child-YA file available for download. Users who access the child and young adult data in this manner will receive the 1986-2000 Child data and codebook, the 1994-2000 Young Adult file, and documentation for the Round 19 (2000) Child & Young Adult that includes one set of current questionnaires and the 2000 Child Assessment Tables report. Unlike the CD, the complete Youth 1986-2000 file is NOT automatically included as part of the Child-YA download. The NLSY79 main file can be accessed separately from the same website. Users should note that an electronic copy of the *NLSY79 User’s Guide* comes with the download of the main NLSY79 data set. Users who are unfamiliar with the NLS extraction software and wish to use

³ NOTE: Users should note that the Young Adults are represented in all non-round specific Areas of Interest on the Child-YA CD, such as PRE/POST NATAL CARE, MATERNAL WORK HISTORY, and CHILD BACKGROUND. The Young Adults are also included in the variables in MATERNAL HOUSEHOLD COMPOSITION, which are based on data for all mothers and thus attached to each of her children regardless of age.

the file via FTP should review Chapter 8 of the current *NLS Handbook*, which is found on-line at <<http://www.bls.gov/nls/handbook/nlshndbk.htm>>.

Child & Young Adult Documentation

Codebook. The NLSY79 Child and Young Adult (and main Youth) data files are each documented by a codebook that lists each variable and briefly describes its content. Each codebook entry contains the variable's unique reference number, coding categories for the variable, frequency distributions, minimum and maximum values, and any special notes. Entries for questionnaire items also contain links to the previous and next question in the instrument. The hyperlinked reference numbers that appear in codebook entries for created variables simply take the user to the next item in the list, generally within the area of interest.

Areas of Interest. The main Youth and Child/Young Adult data sets are also assigned to topical "Areas of Interest" (formerly called Record Types) in order to group the variables by topic. Since the Child and Young Adult cases reside in the data set as one merged file, users are urged to use caution when searching or selecting "Child" versus "YA" variables. To help users distinguish the child from young adult items, the Areas of Interest for Young Adults are prefixed with "YA." For example, users browsing the topical "Areas of Interest" will find both YA CHILD CARE and CHILD CARE. The former contains all the questions from the Young Adult CAPI questionnaire related to use of childcare for their children. The CHILD CARE index holds information drawn from the mother's main Youth record related to the childcare experiences of the NLSY79 children, including the early experiences of the children who are now Young Adults.

Child and Young Adult Survey Instruments. Since most of the primary variables found within the child data set are derived directly from one or more survey instruments, users should have access to at least one set of questionnaires. Copies of the following Child and Young Adult questionnaires and interview aids are available from NLS User Services at cost or on-line in PDF format:

- Child Supplement (CS)—CAPI 1994-present
- Mother Supplement (MS)
- Child Self-Administered Supplement (CSAS)
- CSAS Confidential Card
- Child Face Sheet
- Young Adult CAPI Questionnaire

Users can access the Round 18 (1998) and Round 19 (2000) child and young adult questionnaires at the BLS-NLS website: <<http://www.bls.gov/nls/nlsy79ch.htm>>. This site has links to the following questionnaires in PDF format:

Round 19 (2000)

- Child CAPI Supplement (abbreviated paper version of Child CAPI questionnaire)
- Child Self-Administered Supplement (paper self-report questionnaire for children age 10-14)
- Child Confidential report (13 and 14, sexual experience report)
- Mother Supplement (paper maternal report for each child age 4-14)
- *The 2000 NLSY79 Child Assessments: Selected Tables*
- Young Adult CAPI Supplement (abbreviated paper version of the Young Adult questionnaire)

Round 18 (1998)

- Child CAPI Supplement (abbreviated paper version of questionnaire)
- Child Self-Administered Supplement (paper self-report questionnaire completed by children age 10-14)
- Mother Supplement (paper maternal report for each child under age 15)
- *The 1998 NLSY79 Child Assessments: Selected Tables*
- Young Adult Self-Report booklet (paper self-administered questionnaire containing sensitive questions)

Copies of child and young adult questionnaires for all survey rounds can also be downloaded via FTP at the <<ftp://ftp.chrr.ohio-state.edu/usersvc/>> CHRR site or by visiting the <http://www.nlsinfo.org/ordering/display_db.php3> Product Availability page.

The NLS Handbook. The *NLS Handbook*, updated periodically, provides an introduction to and overall picture of all the cohorts that comprise the National Longitudinal Surveys: NLSY97, NLSY79 and children, Mature Women, Young Women, Older Men, and Young Men. It is particularly useful for those who are unfamiliar with the surveys and their data. Each cohort-specific chapter of this handbook is accompanied by detailed tables that provide users with information about many of the variables contained in each of the surveys over time.

NLS Bibliographies. NLS User Services maintains an annotated archive of NLS research on-line at <<http://www.nlsbibliography.org>>. Records of more than 4,000 journal articles, working papers, monographs, reports, presentations, and dissertations are available for users to search, retrieve, and print customized listings of NLS research. CHRR has also

compiled a comprehensive listing of research based on the child, maternal, and fertility data from the NLSY79. This bibliography, *Research Using NLSY79 Data on Fertility, Child Care, & Child Development*, can be downloaded at <<http://www.bls.gov/nls/nlsdoc.htm>> or requested from NLS User Services (see contact information at the end of this chapter).

NLS News. The Bureau of Labor Statistics publishes a quarterly NLS newsletter that includes information about new data releases, error notices, completed NLS research, and other information of general interest to the NLS research community. Issues can be downloaded in PDF at <<http://www.bls.gov/nls/nlsnews.htm>>.

Other Documentation. In addition to the codebooks, variable descriptions, search indices, and items described above, CHRR provides the following materials to users as additional documentation for the NLSY79 Child and Young Adult data:

- NLS CD-ROM Guide (for CD users)
- NLSY79 Child Handbooks 1986-1992

Comprehensive documentation and bibliographies are available at no charge from CHRR. Selected reports, user updates, and current questionnaires can be accessed on-line on the BLS Child-YA website. Paper versions of original field instruments can also be purchased from CHRR.

User Notes & Errata

Periodically CHRR provides users with additional updates to the existing data or documentation. Notes are sometimes prepared that clarify certain elements of the current files. This information is compiled into a list of “User Notes,” a copy of which is provided with each data release. Updates are also posted to the Child & Young Adult pages on the BLS-NLS website (<<http://www.bls.gov/nls/nlsy79ch.htm>>). Current known updates to the Child and Young Adult files are listed below.

Cases deleted from file. The following child cases have been deleted from the Round 19 (2000) NLSY79 Child-Young Adult file: Child Public ID = 318802, 318803, 567001, 864903, 902102,1031303. A review of the longitudinal main Youth record has shown that these cases are duplicates, nonbiological children, or non-live births previously reported by NLSY79 respondents.

Case edit for CINTRV2000 (C00115.07). There are two Child cases in the current file for which the Round 19 (2000) child interview status flag (CINTRV2000, C00115.07,

“INTERVIEW STATUS OF CHILD”) and the child sampling weight for 2000 (CSAMWT2000, C24955., “CHILD SAMPLING WEIGHT”) disagree. For case CPUBID=256503, the child interview status flag C00115.07 should be recoded to 0; the child sampling weight (C24955.) value of 0 for this case is correct. For case CPUBID=745801, a telephone interview was completed for the child, even though the child was part of the excluded over sample in 2000. Thus, the child interview status flag C00115.07 has a value of 1 (indicating that the Child Supplement was administered) and a child sampling weight of 0 because the child was in the deleted over sample. Users may decide to recode the interview status value for this case depending on their analysis. See below for an explanation of the deleted over sample in 2000.

Over samples reduced in 2000. In Round 19 (2000), due to funding constraints, children of the black and Hispanic over samples were reduced by about 38 percent. These excluded cases were randomly selected and will be eligible for re-interview in Round 20 (2002). Users can obtain a file that identifies the excluded oversample from NLS User Services or by checking the <<ftp://ftp.chrr.ohio-state.edu/usersvc>> FTP site.

HOME Recode short titles. Several of the variable titles for the HOME recode items (e.g., C19028., C19055., C19082., C25091., C25119., and C25146.) do not have a space between the “&” and the next word (e.g., &DAD). When running SAS this produces a warning, but does not affect the use of the title.

Child cases with no mother interview. In Round 19 (2000) two children (Public ID=11202, 592604) were interviewed whose mothers were not. Users will notice that variables tied to the mother’s interview date, such as AGECH2000 (C00047.45) or AGEMOM2000 (C00382.37) are missing. Users may want to default to the child’s age at the date of assessment (CSAGE2000, C00070.47) for AGECH2000.

Interview status flags reduced and simplified. In Round 19 (2000), the set of child interview status flags has been reduced to 7 from 8 for 1994-98. The key child interview status flag (CINTRV2000) has been created as a dichotomous rather than a 7-category variable as in 1998.

Users should make sure that their supply of documentation includes any errata sheets or users notes that contain corrections and updates to the data files. The full set of user

documentation is listed in a document on the CD. User update information is available from NLS User Services (see the contact information at the end of this chapter).

File and Variable Linking Procedures

The Child/Young Adult data files contain many constructed variables drawn from multiple sources, including both cross-sectional and longitudinal information in both the child and mother records. Users of the Child/Young Adult data who wish to construct variables not found in the Child/ Young Adult files may access the mothers' records in the main Youth files in order to obtain the necessary inputs. A useful variable for linking the mother's longitudinal record to the child is C52., which defines the first survey year (of the mother) following the child's date of birth. (All children born prior to the 1979 survey date are assigned 1979 as their first post-birth survey point.) In general, a researcher interested in creating a variable that references a particular post-birth survey point for all children would need to follow a two-step procedure. For example, to create a variable that defines the first post-birth survey maternal school enrollment status for all children from information collected at the mother's survey dates, one would take the following steps: (1) use C52.00 to define the appropriate survey year following each child's date of birth, and (2) use this variable to choose a maternal enrollment status variable, depending on the date of the birth of the child. If the child's first post-birth survey year is 1982, the maternal enrollment status variable as of the 1982 interview (C570.) would be selected. Similarly the appropriate enrollment status information could be combined across the relevant surveys into one created "First post-birth survey enrollment status" variable.

For more refined linking procedures, the user may need to access maternal interview dates as well as the dates of many relevant events, such as school leaving, job changes, and marital changes. Variables on the file that are keyed to the child's date of birth or to the mother's date of interview can be used to (1) precisely link many social, economic, and demographic maternal behaviors with a child's birth or age at a particular point or to (2) define the survey point closest to a particular child's birth or attained age. The first of these two objectives can be facilitated by using variable C2700., which defines the birth of every child in terms of the number of weeks that have elapsed between January 1, 1978 and the current maternal interview date. This variable can be linked with the NLSY79 main Youth

Work History File, a data set available from CHRR that constructs a week-by-week profile of the respondent's employment experience.

The pre- and postnatal variables (C3200.-C3392.10) are already defined in an appropriate life cycle context, since they reference either the pregnancy leading up to a particular birth or an event/behavior during the first year following a birth. A missing value in this series of constructed child variables means that at least one input from the main Youth file was missing (refer to the note in the Child codebook for C2700.).

Linking Data Files

The child sequential identification code (C1.) allows the user to link children with their siblings on the NLSY79 Child files and with their mothers on the main NLSY79 files. The child ID is composed of the first 5 digits of the mother's sequential ID plus a 2-digit code (01-10) generally *but not always* indicating the child's order of birth. Children with the same first 5 digits in their IDs are siblings. Since the child ID contains the first 5 digits of the mother ID, either the child ID or the mother ID (C2.) enables the users wishing to create a mother-based file to extract maternal characteristics and write them out only for mothers.

Child ID vs. Young Adult ID. Users will note that identification codes are provided for all children as CPUBID (C00001.) in the CHILD BACKGROUND area of interest and again for young adults as CASEID (Y00001.) in YA COMMON KEY VARIABLES. The only distinction between these two variables is that CASEID is only available for children who have been interviewed as young adults. The CPUBID is a comprehensive ID code for all children in the child file, regardless of age or interview status. They are both included for convenience, particularly if users are accessing only the Young Adult file.

Linking Children and Young Adults to Mothers. Space constraints require that the current release of the Child and Young Adult data be stored in a file *separate* from the main Youth files. The Child/Young Adult file does *not* contain software that will enable users to automatically create a merged extract of children or young adults and data selected from the mother's record. Such a merge can be accomplished by means of Child case ID (C1.) and Mother case ID (C2.) saved on a Child/Young Adult extract and mother case ID (R1.) saved on an extract from the main Youth CD files. Respondent identification code should always be among the variables saved in creating any specification file or extract. When selecting variables from the NLSY79 main Youth file, users should include R1. (Identification Code) in

the list of saved variables. Before exiting the Child/Young Adult data set, users should verify that Child ID and mother ID are among the saved variables. The user can employ any of a number of methods for merging files. Appendix F of this document offers sample programs designed to assist users in merging files.

Missing Data

On the current release, missing values other than those reserved for Don't Know (-2), Refused (-1), and Invalid Skip (-3) have been collapsed into a "-7." This value includes noninterviews and valid skips. Users should rely on the Interview/Assessment flags in the CHILD BACKGROUND area of interest to determine child or young adult interview status. The relevant Child interview status variables, available starting in 1998, are identified by reference numbers C00115.01-C00115.15:

CINTRV98 (C00115.01) INTERVIEW STATUS OF CHILD
CASSMT98 (C00115.02) ASSESSMENT STATUS OF CHILD
MINTRV98 (C00115.03) INTERVIEW STATUS OF MOTHER
INCSUP98 (C00115.04) DOES CHILD HAVE A CHILD SUPPLEMENT?
INMSUP98 (C00115.05) DOES CHILD HAVE A MOTHER
SUPPLEMENT RECORD?
INCSAS98 (C00115.06) DOES CHILD 10-14 YEARS OLD HAVE A CSAS
RECORD?
CINTRV2000 (C00115.07) INTERVIEW STATUS OF CHILD
MINTRV2000 (C00115.09) INTERVIEW STATUS OF MOTHER
INCSUP2000 (C00115.10) DOES CHILD HAVE A CHILD
SUPPLEMENT?
INMSUP2000 (C00115.11) DOES CHILD HAVE A MOTHER
SUPPLEMENT RECORD?
INCSAS2000 (C00115.12) DOES CHILD 10-14 YEARS OLD HAVE A
CSAS RECORD?

The series of Young Adult interview status flags, available for 1994-2000, is also in the CHILD BACKGROUND area of interest and can be identified by reference numbers C00112.00 - C00112.07:

YASAMP [1994-2000] IS CHILD ELIGIBLE FOR YOUNG ADULT
INTERVIEW?
YAINTV [1994-2000] WAS CHILD INTERVIEWED AS YOUNG ADULT?
YASAMP2000 IS CHILD ELIGIBLE FOR YA INTERVIEW IN CURRENT
ROUND?
YAINTV2000 WAS CHILD INTERVIEWED AS YOUNG ADULT IN
CURRENT ROUND?

These interview status flags are cross-sectional and thus only reflect valid values for cases from a particular survey round. The “MINTRV” variables in this series indicate if the mother was interviewed as part of the main Youth survey. More details on mothers’ interview status are provided by a series of variables titled “REASON FOR NONINTERVIEW” in the main Youth record COMMON VARIABLES.

Users may also employ the Child and Young Adult sampling weight variables to delineate their analysis sample. The Child sampling weights, in the ASSESSMENT areas of interest, have values greater than “0” for any child interviewed in a particular year. The Young Adult sampling weight for each year is specific to young adults interviewed in that year; any young adult not interviewed or any child who is not a young adult in that year is assigned a value of “0.” Therefore, restricting the sample to cases greater than “0” on either of these types of variables will result in cases interviewed in that year.

Contact and Ordering Information

The NLSY79 Child-YA Web Page

Data users can view information about the NLSY79 Child and Young Adult Surveys on the NLSY79 pages of the BLS web site: <<http://www.bls.gov/nls/nlsy79ch.htm>>. The web site features downloadable public use data and documentation for the current NLSY79 main and Child files, as well as information about any modifications or updates to the data and/or documentation. Published reports from previous years’ surveys are also available.

Users wishing more detailed information about the NLSY79 Child assessments as well as supplementary information about sample constraints in previous survey rounds should consult the *NLSY Child Handbook: 1986-1990* (Baker et al., 1993) and *The NLSY Children 1992* (Mott et al., 1995). Tabular reports for the assessments scores, starting with the 1994 survey round, are available in print format from NLS User Services or online at the CHRR site: <<ftp://ftp.chrr.ohio-state.edu/usersvc/>>.

Information on the NLSY79 Child & Young Adult surveys is available on-line at:
<http://www.bls.gov/nls/nlsy79ch.htm>

The NLS website, managed by The Bureau of Labor Statistics, can be accessed at:
<http://www.bls.gov/nls/home.htm>

Searches of the NLS annotated bibliography can be made via:
<http://www.nlsbibliography.org/>

NLS User Services at CHRR. Please direct all questions and comments related to ordering or using the NLSY79 Child and Young Adult data and documentation to:

NLS User Services
Center for Human Resources Research (CHRR)
921 Chatham Lane, Suite 100
Columbus, OH 43221-2418 USA
Phone: (614) 442-7366
Fax: (614) 442-7329
E-mail: usersvc@postoffice.chrr.ohio-state.edu