

---

**NATIONAL CENTER FOR EDUCATION STATISTICS**

---

**User's Manual**

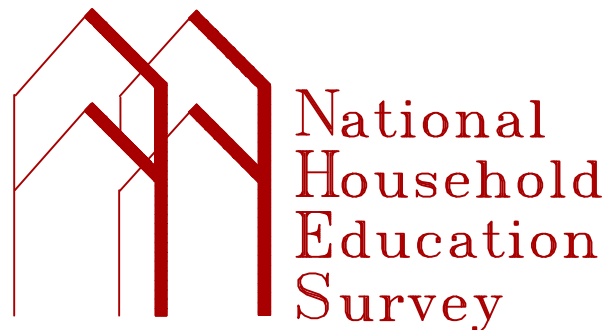
**August 1997**

---

**National Household Education Survey of 1996**

**Data File Users Manual  
Volume IV**

**Youth Civic Involvement Data File**



---

# NATIONAL CENTER FOR EDUCATION STATISTICS

---

User's Manual

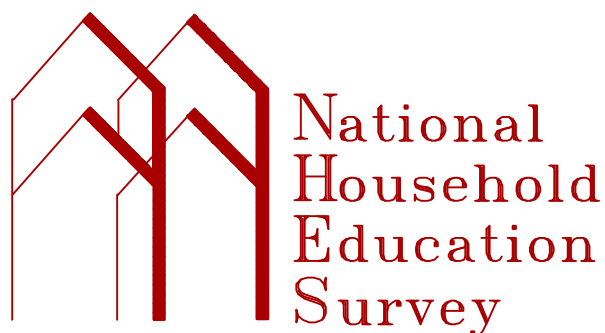
August 1997

---

National Household Education Survey of 1996

## Data File Users Manual Volume IV

### Youth Civic Involvement Data File



Mary A. Collins, Project Director  
J. Michael Brick, Senior Statistician  
Mary Jo Nolin  
Susan Gilmore  
Westat, Inc.

Kathryn Chandler  
National Center for Education Statistics

Chris Chapman  
Education Statistics Services Institute

---

U.S. Department of Education  
Office of Educational Research and Improvement

NCES 97-422

**U.S. Department of Education**

Richard W. Riley  
*Secretary*

**Office of Educational Research and Improvement**

Ramon C. Cortines  
*Acting Assistant Secretary*

**National Center for Education Statistics**

Pascal D. Forgione, Jr.  
*Commissioner*

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

National Center for Education Statistics  
Office of Educational Research and Improvement  
U.S. Department of Education  
555 New Jersey Avenue NW  
Washington, DC 20208-5574

July 1997

The NCES World Wide Web Home Page is  
<http://www.ed.gov/NCES/>

**Suggested Citation**

U.S. Department of Education. National Center for Education Statistics. *National Household Education Survey of 1996: Data File User's Manual, Volume IV: Youth Civic Involvement Data File*, NCES 97-422, by Mary Collins, J. Michael Brick, Mary Jo Nolin, Susan Gilmore, Kathryn Chandler and Chris Chapman. Washington, DC: 1997.

**Contact:**

Kathryn Chandler  
(202) 219-1767  
e-mail: [nhes@ed.gov](mailto:nhes@ed.gov)  
<http://www.ed.gov/NCES/NHES>



**TABLE OF CONTENTS**

**Chapter**

<b>Introduction</b> .....	1
<b>6. Guide to the Data File and Codebook</b> .....	3
6.1 Content and Organization of the Data File.....	3
6.1.1 System Variables .....	3
6.1.2 Household Membership Variables .....	4
6.1.3 Questionnaire Item Variables.....	5
6.1.4 Household Characteristics Variables .....	6
6.1.5 Derived Variables.....	6
6.1.6 Weighting and Variance Estimation Variables .....	12
6.1.7 Imputation Flag Variables.....	12
6.1.8 Numeric and Character Variables.....	13
6.2 Guide to the Codebook .....	13
6.3 Public and Proprietary Data Files .....	13
6.4 Linking the Household & Library File to Other NHES:96 Data Files.....	13
<b>7. Data Considerations and Anomalies</b> .....	15
7.1 Number of Students in Child's Grade .....	15
7.2 Youth Service Activities .....	15
7.3 Type of Community in Which the Household is Located.....	15
7.4 Correspondence Between Age and Grade.....	16
7.5 Income to the Nearest Thousand Dollars .....	16
<b>Appendixes</b>	
Appendix B: Youth CI Public File Layout in Position Order .....	B-1
Appendix C: SAS Code for Derived Variables.....	C-1
Appendix D: Youth CI Codebook .....	D-1
Appendix E: Directions and Code for Linking Data Files .....	E-1
<b>Figure</b>	
6-1 Example of the Codebook Format .....	14

*This page is blank.*

## INTRODUCTION

The 1996 National Household Education Survey (NHES:96) was a random digit dial (RDD) telephone survey of households developed by the National Center for Education Statistics (NCES) and conducted by Westat, Inc. The NHES:96 included two topical survey components:

- Parent and Family Involvement in Education (PFI), in which data were collected about types and frequency of family involvement in children's school, school practices to involve and support families, and learning activities with children outside of school; and
- Civic Involvement (CI), which included sources of information about government, knowledge about government, community service participation, political participation, and attitudes related to democratic values and government.

There were three populations of interest for the NHES:96:

- Children 3 years old through grade 12, whose parents responded to PFI items, and children in 6th through 12th graders, whose parents also responded to CI items;
- Students in grades 6 through 12, who, in addition to their parents, responded to CI items and to a small number of PFI items; and
- Adults, defined as persons 18 years old or older, not enrolled in grade 12 or below, and not on active duty in the military, whose responses to CI items provided estimates representative of all civilian U.S. adults.

In addition to the major topical components, the NHES:96 Screener collected demographic and educational information on all members in every household contacted, whether or not any member of the household was selected for an extended interview. (The term "extended interview" refers to the interview pertaining to the topical component of the study, that is, the Parent PFI/CI, the Youth CI, or the Adult CI interview.)

This manual, the *NHES:96 Data File User's Manual, Volume IV-- Youth Civic Involvement*, provides documentation and guidance for users of the public release data file for the Youth Civic Involvement (Youth CI) component of the 1996 National Household Education Survey (NHES:96). This volume contains a description of the Youth CI data file and a discussion of data considerations and anomalies. Included as appendixes are the public file layout, SAS code for creating derived variables, the codebook for the Youth CI public data file, and directions and sample code for linking NHES:96 data files.

Volume IV is meant to be read in conjunction with the *NHES:96 Data File User's Manual, Volume I*. More information about the purpose of the study, the sample design, the other survey components, the data collection instruments, and data collection and data processing procedures is contained in the *NHES:96 Data File User's Manual, Volume I*. Information about the Household & Library public data file can be found in Volume II, and information about the Parent and Family

Involvement in Education and Civic Involvement (Parent PFI/CI) public data file and the Adult Civic Involvement (Adult CI) public data file can be found in Volumes III and V of the manual, respectively.



## 6. GUIDE TO THE DATA FILE AND CODEBOOK

### 6.1 Content and Organization of the Data File

This section describes the content of the Youth CI public release data file of the NHES:96. This file contains data from all completed Youth CI interviews. There are three records for each Youth interview completed, so the file contains 24,129 records for the 8,043 cases. The file is organized so that logically related sets of variables are grouped together. The data items are listed in the file in the following order: system variables, household membership variables, questionnaire item variables, household characteristics variables, derived variables, weighting and variance estimation variables, and imputation flag variables.

A list of all the variables in the Youth CI data file is shown in appendix B. The VARIABLE NAME column displays the unique identifier for each variable in the data file. The VARIABLE LABEL column displays a short description associated with the variable. The FORMAT column indicates if a variable has a numeric ("N") or a character ("A") format. All of the variables except GRADE, GRADEEQ, SLOW, SHIGH, and ALLGRADE in the Youth CI file have numeric formats. The RECORD NUMBER column indicates whether the variable is located on the first, second, or third record. The LENGTH column indicates the length of the variable by the number of digits. The length descriptor also includes the number of digits found after the decimal point for noninteger numeric variables (e.g., weight variables). The position of the variable is indicated in the START and END columns which indicates the position in the file where the variable begins and ends.

A value of "-1" for any variable on the data file indicates that a case was part of a legitimate skip. For example, if the youth was schooled at home (HOMESCHL, PB2), the question about how often he/she talks with his/her parents about school (FESCHOOL, YA1) would have a value of -1, because the question was inapplicable. This convention of assigning -1 to all legitimate skips applies to all NHES data files.

The NHES:96 data files are provided on CD-ROM and are accessible through an Electronic CodeBook (ECB) that allows data users to view variable frequencies, tag variables for extraction, and create the SAS, SPSS for DOS, or SPSS for Windows code needed to create an extract file for analysis purposes. The ECB contains all NHES:96 data sets: the Household & Library file, the Parent PFI/CI file, the Youth CI file, and the Adult CI file. The ECB also contains all data sets for previous NHES collections and documentation for every file. Instructions for using the CD-ROM and ECB are provided in a separate document, the *National Household Education Survey: NHES:91/93/95/96 Electronic CodeBook (ECB) User's Guide* (Collins and Chandler forthcoming). The sections that follow describe the contents of the Youth CI data file.

#### 6.1.1 System Variables

System variables are created during the conduct of an interview and are instrumental in the successful administration of the interview. Their creation is transparent to the interviewer and to the respondent. System variables fall into two categories: linking variables (record identifiers or ID numbers) and interview status variables. Linking variables are record identifiers that provide a link to other interviews completed in the same household. See appendix E for more information about linking between files.) Status variables are set at the completion of each interview to define interview status.

**BASMID** is the unique 12-digit identifier variable for the interview. It is composed of the eight-digit household identifier, the interview subject's two-digit household-member person number, and the two-digit number, 02, that indicates that the interview was a Youth CI interview.

**ENUMID** is the 10-digit identifier variable for the subject of the interview. It is composed of the eight-digit household identifier and the interview subject's two-digit household-member person number. **ENUMID** can be used to link the Youth CI interview to the Parent PFI/CI interview about that youth. See appendix E for instructions for linking the Parent PFI/CI and Youth CI interviews.

**BASEID** is the eight-digit identifier for the household. This ID number also forms the first eight digits of interview ID numbers for other interviews in the household, providing a means of linking interviews within the same household. See appendix E for instructions for linking the NHES:96 data files.

**ENGLSPAN** is the variable that indicates whether the interview was conducted in English or in Spanish.

The values for **ENGLSPAN** are:

- 1 = Interview was conducted in English
- 2 = Interview was conducted in Spanish

### 6.1.2 Household Membership Variables

All household members were enumerated in the Screener. Data collected included age and sex (S6), educational status (SX7 through SX14), and demographic characteristics (SX15 through SX22 and SX27 through SX33OV). The gender collected during the household enumeration in the Screener (S6) were used to drive the gender-based wording of subsequent questions throughout the Screener and, if appropriate, the extended interview.

The household member information is stored on the public release data file in the following order: information about the subject of the interview (the sampled youth), information about the respondent to the parent interview (the most knowledgeable parent/guardian), information about the mother, and information about the father. Please note that the parent respondent information may be repeated in one of two places. If the respondent to the parent interview is the mother or the father, that information will be repeated in the mother or father section. The variables appear on the data file as follows:

**CHILDNUM** is the household member person number of the sampled youth.

**AGE95** is the sampled youth's age as of December 31, 1995.

**SEX** is the sampled youth's sex.

**RACE** indicates the sampled youth's race.

**HISPANIC** indicates whether the sampled youth is Hispanic.

**OTHRAC** indicates the sampled youth's race if "some other race" was reported at RACE (SX21).

**RESPNUM** is the parent interview respondent's household member person number.

**RESPAGE** is the parent interview respondent's age.

**RESPSEX** is the parent interview respondent's sex.

**RESRELN** is the parent interview respondent's relationship to the sampled youth.

**MOMNUM** is the household member person number of the sampled youth's mother.

**MOMAGE** is the mother's age.

**MOMTYPE** is the type of mother (birth, adoptive, step, or foster).

**DADNUM** is the household member person number of the sampled youth's father.

**DADAGE** is the father's age.

**DADTYPE** is the type of father (birth, adoptive, step, or foster).

### 6.1.3 Questionnaire Item Variables

The next variables on the Youth CI data file capture the youth's month and year of birth, the language he or she speaks most at home, school enrollment and grade, and school characteristics. These items were asked of the parent/guardian respondent and have been copied from the Parent PFI/CI data file to aid users. They are on the file in the order they were asked in the Parent PFI/CI interview. The variables from the Youth CI interview appear next on the file in the same order as they were asked. Some variables were excluded from the file for confidentiality reasons. These include the names of household members, verbatim string responses that might identify persons or places, and the individual ZIP Codes. Some of these variables are included in a separate restricted-use data file (see section 6.3 below). The Youth CI questionnaire appears with the Screener, the Parent PFI/CI, and the Adult CI questionnaires, in Volume 1, appendix A; variable names are provided to the left of each question. Those followed by "/R" appear only on a restricted-use data file that may be obtained through a special licensing agreement with NCES.

There are repeating series of questions in the Service Activities section. Variable names and labels reflect the series in this section. For example, the SANOW1 (YC3) variable indicates whether the youth is currently participating in the first service activity reported and the SAREG2 (YC4) variable indicates whether the youth is participating in the second service activity on a regular basis.

#### 6.1.4 Household Characteristics Variables

Household characteristics variables are variables that reflect characteristics of the household as a unit. These questions were asked at the end of the first Parent PFI/CI interview in the household, and they have been copied on the Youth CI file. For example, questions were asked about whether the home was owned or rented and the type and size of community where the household was located. These household items appear on the file in the same order as they were asked. Because they are actually part of the Screener, refer to that instrument in Volume 1, appendix A, for the questions and their order.

#### 6.1.5 Derived Variables

Derived variables were developed and included in the Youth CI public use data file to aid users in their analyses. The derived variables fall into two categories: questionnaire item variables and counter variables. Questionnaire item-derived variables were created by combining two or more items from the questionnaire. Counter-derived variables were created by counting the number of persons with specific characteristics enumerated in the household. Linked-derived variables, created by using the respondent's ZIP Code to extract data from the 1990 Census of Population Summary Tape File 3B (STF3B), are available on a restricted-use file.

The derived variables appear together on the file in their own section in alphabetical order. They are listed below in the same order with an explanation of how they were derived. The actual SAS code to create these variables is found in appendix C, with the exception of counter variables and CENREG (Census region).

**ALLGRADE** is a derived variable that identifies the enrollment status, the grade level of children in graded schools, and the grade level equivalent for children in ungraded schools, special education programs, or home school. ALLGRADE was created using the variables GRADE (PB4) and GRADEEQ (PB5).

The values for ALLGRADE on the Youth CI file are:

- 6 = Sixth grade or equivalent
- 7 = Seventh grade or equivalent
- 8 = Eighth grade or equivalent
- 9 = Ninth grade or equivalent
- 10 = Tenth grade or equivalent
- 11 = Eleventh grade or equivalent
- 12 = Twelfth grade or equivalent
- U = Ungraded/no equivalent

**CENREG** is a derived variable that identifies the Census region in which the subject youth lives. This variable is created by grouping states. The following states and the District of Columbia are in each Census region:

- Northeast: CT, MA, ME, NH, NJ, NY, PA, RI, VT
- South: AL, AR, DC, DE, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN, TX, VA, WV
- Midwest: IA, IL, IN, KS, MI, MN, MO, ND, NE, OH, SD, WI
- West: AK, AZ, CA, CO, HI, ID, MT, NV, NM, OR, UT, WA, WY

The values for CENREG are:

- 1 = Northeast
- 2 = South
- 3 = Midwest
- 4 = West

**COMMUNITY** is a derived variable representing the respondent's report of the size of the community in which the household is located. **COMMUNITY** was created using the variables **HCCOMMUN** (SX31), **HCSUB** (SX31OV), and **HCCITY** (SX31OV2).

The values for **COMMUNITY** are:

- 1 = Very large city (over 500,000 people)
- 2 = Large city (100,000 to 500,000 people)
- 3 = Medium sized city (50,000 to 100,000 people)
- 4 = Suburb of a very large city
- 5 = Suburb of a large city
- 6 = Suburb of a medium city
- 7 = Small city or town of fewer than 50,000 people that is not a suburb of a larger city
- 8 = Rural or farming community

**DADEMPLD** is the derived variable that indicates the employment status of the father (birth/adoptive/step/foster/guardian). **DADEMPLD** was created using the variables **DADWORK** (PM5), **DADLEAVE** (PM6), **DADHOURS** (PM7), **DADLOOK** (PM8), and **DADPUBL**, **DADPRIV**, **DADEMPL**, **DADREL**, **DADANSAD** (all from PM9).

The values for **DADEMPLD** are:

- 1 = Working 35 hours or more per week
- 2 = Working less than 35 hours per week
- 3 = Looking for work
- 4 = Not in the labor force
- 1 = No father for the subject youth in the household

**FAMILY** is the derived variable that describes the family type, based on the presence of parents and siblings. **FAMILY** was created using the derived variables **HHPARN1** and **NUMSIBS**.

The values for **FAMILY** are:

- 1 = Two parents and siblings
- 2 = Two parents and no siblings
- 3 = One parent and siblings
- 4 = One parent and no siblings
- 5 = Other

**HH18OVER** is the counter-derived variable that indicates the number of household members age 18 and older. The screener responses to **HHAGE1** through **HHAGE16** (S6) were counted for this variable. These variables appear in the Household and Library Data file.

**HHDAD** is the derived variable that indicates whether the birth, adoptive, step, or foster father or male guardian of the subject youth resides in the household with him/her. HHDAD was created using the variables DADTYPE (PA7), MOMTYPE (PA6), and RESPSEX (S6).

The values for HHDAD are:

- 1 = Birth or adoptive father
- 2 = Step or foster father
- 3 = Male respondent/no mother or father in household
- 4 = Else

**HHMOM** is the derived variable that indicates whether the birth, adoptive, step, or foster mother or female guardian of the subject youth resides in the household with him/her. HHMOM was created using the variables MOMTYPE (PA6), DADTYPE (PA7), and RESPSEX (S6).

The values for HHMOM are:

- 1 = Birth or adoptive mother
- 2 = Step or foster mother
- 3 = Female respondent/no mother or father in household
- 4 = Else

**HHPARN1** is the derived variable that designates by broad classification the subject youth's parents who reside in the household. It denotes a two-parent family, a one-parent family, or a family with nonparent guardians. HHPARN1 was created using the derived variables HHMOM and HHDAD.

The values for HHPARN1 are:

- 1 = Mother (birth, adoptive, step, or foster) and father (birth, adoptive, step, or foster)
- 2 = Mother (birth, adoptive, step, or foster) only
- 3 = Father (birth, adoptive, step, or foster) only
- 4 = Nonparent guardian(s)

**HHTOTAL** is the counter-derived variable that indicates the total number of household members. The screener responses to HHAGE1 through HHAGE16 (S6), found in the Household & Library file, were counted for this variable.

**HHUNDR6** is the counter-derived variable that indicates the number of household members younger than 6 years old. The screener responses to HHAGE1 through HHAGE16 (S6), found in the Household & Library file, were counted for this variable.

**HHUNDR13** is the counter-derived variable that indicates the number of household members younger than 13 years old. The screener responses to HHAGE1 through HHAGE16 (S6), found in the Household & Library file, were counted for this variable.

**HHUNDR18** is the counter-derived variable that indicates the number of household members younger than 18 years old. The screener responses to HHAGE1 through HHAGE16 (S6), found in the Household & Library file, were counted for this variable.

**HHUNDR21** is a counter-derived variable that indicates the number of household members younger than 21 years old. The screener responses to HHAGE1 through HHAGE16 (S6), found in the Household & Library file, were counted for this variable.

**LANGUAGE** is the derived variable that describes whether the language(s) spoken most often at home by the parent(s)/guardian(s) in the household is English. **LANGUAGE** was created using the variables MOMLANG (PL1), MOMSPEAK (PL2), DADLANG (PM1), and DADSPEAK (PM2), found in the Parent PFI/CI file.

The values for **LANGUAGE** are:

- 1 = Both/only parent(s) main language at home is English
- 2 = One of two parents speaks a non-English language most at home
- 3 = Both/only parent(s) speak a non-English language most at home

**MOMEMPLD** is the derived variable that indicates the employment status of the mother (birth/adoptive/step/foster/guardian). **MOMEMPLD** was created using the variables MOMWORK (PL5), MOMLEAVE (PL6), MOMHOURS (PL7), MOMLOOK (PL9), and MOMPUBL, MOMPRIV, MOMEMPL, MOMREL, and MOMANSAD (all from PL10), found in the Parent PFI/CI file.

The values for **MOMEMPLD** are:

- 1 = Working 35 hours or more per week
- 2 = Working less than 35 hours per week
- 3 = Looking for work
- 4 = Not in the labor force
- 1 = No mom in household

**MOMFTFY** is the derived variable that indicates if the mother (birth/adoptive/step/foster/guardian) of the subject youth currently works full time and has worked 12 months of the past year. **MOMFTFY** was created from the variables MOMWORK (PL5), MOMMTHS (PL8), found in the Parent PFI/CI file, and the derived variable **MOMEMPLD**.

The values for **MOMFTFY** are:

- 1 = Full time (35 hours or more) full year
- 2 = Less than full time or less than full year
- 3 = Not employed
- 1 = No mom in household

**PARGRADE** is the derived variable that indicates the highest level of education for the subject child's parents or nonparent guardians who reside in the household. **PARGRADE** was created using the variables MOMGRADE (PL3), MOMDIPL (PL4), DADGRADE (PM3), and DADDIPL (PM4), found in the Parent PFI/CI file.

The values for PARGRADE are:

- 1 = Less than high school diploma
- 2 = High school diploma or its equivalent
- 3 = Vocational/technical degree or some college
- 4 = Bachelor's degree
- 5 = Graduate/professional school

**RACEETHN** denotes both the race and ethnicity of the youth. If the respondent to the Parent PFI/CI interview designates the youth's ethnicity as Hispanic, RACEETHN is Hispanic regardless of whether RACE was classified as white, black, or another race. RACEETHN was created using the variables RACE (SX21), OTHRAC (SX21A), and HISPANIC (SX22), found in the Household & Library file.

The values for RACEETHN are:

- 1 = White, non-Hispanic
- 2 = Black, non-Hispanic
- 3 = Hispanic
- 4 = All other races (e.g., American Indian or Alaska Native, Asian or Pacific Islander), non-Hispanic

**SCHLGRAD** is a derived variable that classifies the type of school the subject youth attends based on the highest and lowest grades in school. SCHLGRAD was created using the variables SLOW (PD7) and SHIGH (PD8), from the Parent PFI/CI file. Note that although this variable also appears on the Parent PFI/CI file, the values are not the same in both files. Category 1 on the Parent PFI/CI file represents early childhood programs and center-based child care, a category that is not applicable for this 6th through 12th grade population.

The values for SCHLGRAD are:

- 1 = Elementary school (low grade N, K, T, P, 1 to 3; high grade 1 to 8)
- 2 = Middle or junior high school (low grade 4 to 9; high grade 4 to 9)
- 3 = High school (low grade 7 to 12; high grade 10 to 12)
- 4 = Combined grades school
- 1 = Homeschool path

**SCHLTYPE** is a derived variable that classifies the school currently attended as either public or private. Schools that are public are further classified as being chosen or assigned, and schools that are private are also classified as being church-related or not church-related. SCHLTYPE was created using the variables SPUBLIC (PD1), SCHOICE (PD3), and SRELGON (PD4), from the Parent PFI/CI file.



The values for SCHLTYPE are:

- 1 = Public, assigned
- 2 = Public, chosen
- 3 = Private, church-related
- 4 = Private, not church-related
- 1 = Home school path

SCNUMSTU is a derived variable that classifies the estimated number of students in the sampled youth's school. SCNUMSTU was created using the variables SLOW (PD7), SHIGH (PD8), SNUMSTUD (PD9), and SNUMGRAD (PD90V), from the Parent PFI/CI file.

The values for SCNUMSTU are:

- 1 = Under 300
- 2 = 300-599
- 3 = 600-999
- 4 = 1,000 or more
- 1 = Home school path

#### **6.1.6 Weighting and Variance Estimation Variables**

The first variable in this section of the file is FYWT. It is the variable that should be used as the weight variable to estimate the characteristics of youth. This weight contains all of the adjustments for the probabilities of selection, nonresponse, and undercoverage as described in Volume 1, chapter 3.

The 80 replicate weights, FYWTR1 to FYWTR80, are the next variables in this section. These replicate weights can be used with the WesVarPC Windows-based software program to produce estimates of the sampling errors of the estimates. More details on how the replicate weights were created and how they can be used with WesVarPC are given in Volume 1, chapter 3, along with an approximation method that does not involve using the WesVarPC procedure.

The remaining two variables in this section are YSTRATUM and YPSU. These variables are provided to enable users to compute sampling errors using Taylor Series approximations, such as the SUDAAN procedure. The methods used to construct the values for YSTRATUM and YPSU are also discussed in Volume 1, chapter 3.

#### **6.1.7 Imputation Flag Variables**

Item nonresponse occurred when some, but not all, of the responses were missing from an otherwise cooperating respondent. For all the items on the Youth CI public use file except the government knowledge items (Y8a-e and Y9a-e), the missing data were imputed, or "filled in," to help users of the data. For each variable involved in imputation, an imputation flag variable was created. If the imputation flag is equal to 0, then no imputation was performed on that case. This flag can be used to identify imputed values. Volume 1, section 3.8 discusses the meaning of values assigned to the imputation flags.

The naming convention for the imputation flag variables was to drop the last letter of the variable name and replace it with an "f." For example, the imputation flag for SEX is SEF. This naming convention holds true for all Youth CI variables except for variables that originally end in "f," variables that will become confused with other variables when the last letter is dropped, or variables that end in a number. In these cases, the letter before the last letter or last digit is dropped and replaced with an "f." For example, the imputation flag for SAWKS1 (YC5) is SAWKF1. The imputation flags appear on the file in the same order as the item.

### **6.1.8 Numeric and Character Variables**

All of the variables in the Youth CI file have numeric formats except GRADE (PB4), GRADEEQ (PB5), SLOW (PD7), SHIGH (PD8) from the PFI/CI file, and ALLGRADE, a derived variable.

## **6.2 Guide to the Codebook**

The codebook, shown in appendix D, contains complete descriptions of the contents of the data file. The codebook contains system variables, household membership variables, questionnaire variables, household characteristic variables, derived variables, weighting and variance estimation variables, and imputation flag variables. The codebook provides all the pertinent information for the variables in the file, including the variable name, the question wording, the position and format of the variable in the file, and the responses to the item. The unweighted frequency, unweighted percent, and weighted percent are provided with each response. Figure 6-1 provides a description of each of the items appearing in the codebook.

## **6.3 Public and Proprietary Data Files**

This manual is designed to assist users of the public use Youth CI data file. The public use file contains all the variables detailed above but does not contain certain variables excluded from the file for confidentiality reasons. These include the names of household members, verbatim string responses that might identify persons, and respondents' individual ZIP Codes (HZIPCODE). Some of these variables (e.g., verbatim strings of other-specify categories, HZIPCODE) that are excluded from the public file are included on a separate proprietary, or restricted-use, file. These variables are indicated with a "R" on the Youth CI questionnaire in Volume I, appendix A. The Household & Library proprietary data file also contains close to 100 ZIP code variables from the 1990 Census of Population Summary Tape File 3B (STF313), including the median household income of the area, the level of community mobility in the area, and the percentage of owner-occupied households in the area. The proprietary data file may be obtained through a special licensing agreement with NCES. Contact NCES for details on how to become licensed.

## **6.4 Linking the Household & Library File to Other NHES:96 Data Files**

It is possible to link the Household & Library file to the Parent PFI/CI, the Youth CI, and the Adult CI data files. Instructions for doing so are located in appendix E.

**Figure 6-1.—Example of the codebook format**

- (1) PRSTUGOV =           (2) YB1-SCH HAS STUDENT GOVT  
(3) YB1 First, does your (current) school have a student government?

- (4) RECORD: 1 POSITION: 139-140  
(5) FORMAT: N2

(6) RESPONSE	(7) CODES	(8) FREQ	(9) UNWGTD PERCENT	(10) WGTD PERCENT
1 YES	1	6494	80.7%	81.5%
2 NO	2	1446	18.0%	18.5%
RESERVED CODES				
-1 INAPPLICABLE	-1	103	1.3%	(MISS)
TOTALS		8043	100.0%	100.0%

DESCRIPTIONS:

- (1) Variable name: This is the variable name associated with each item. This is the unique identifier present in the SAS or SPSS data file.
- (2) Variable label: A short label, which is associated with each of the variables, is presented here. This label appears in the SAS or SPSS data file. Labels contain the questionnaire item numbers. Labels that begin with the letter "D" indicate a derived variable.
- (3) Question wording: This is the exact question wording as it appeared in the questionnaire.
- (4) Record and position: These provide the record number (1, 2, or 3) and the starting and ending position of the variable in the raw data file on tape.
- (5) Format: This provides the variable type, its width, and the number of positions after the decimal point, if necessary. A data type of "N" represents numeric variables and "A" represents character variables. In this example, PRSTUGOV is a numeric variable with a length of 2.
- (6) Response categories: This column provides the response categories for the variable.
- (7) Response codes: This column provides the actual numeric/alphanumeric codes present in the data files.
- (8) Unweighted frequency counts: This column displays the unweighted frequency counts for this variable. The counts for missing values will also be included for the unweighted values, but not for the weighted values.
- (9) Unweighted percentages: This column displays the unweighted frequency counts from the previous column as percentages. This column will also contain percentages for missing values.
- (10) Weighted percentages: This column displays the percentages of frequency counts weighted up to the population. This column will not include percentages for missing values.

*This page is blank.*

## **7. DATA CONSIDERATIONS AND ANOMALIES**

The purpose of this section is to bring to the user's attention certain data considerations and data anomalies in the NHES:96 Youth CI data, to describe the nature of those anomalies, and, where appropriate, to identify possible means of taking them into account when analyzing the data.

### **7.1 Number of Students in Child's Grade**

For the variable SNUMGRAD (number of students in the child's grade), there were nine parents who reported 2,000 to 4,000 students in their child's grade. Although these numbers are not plausible, they were reported by the parent respondents and thus were left on the data file.

### **7.2 Youth Service Activities**

In the Youth CI interview, students were asked whether they had participated in any community service activities during the current school year. Respondents who indicated that they had participated were asked to name up to three service activities in which they had participated. The purpose of obtaining the names of the activities was to identify them when subsequent questions (whether the activity had been done on a regular basis, how many weeks had the youth done the activity, etc.) were asked. The activities were not intended to be coded, and the youth's own assessment of whether it constituted a type of service was accepted. Review of the names of service activities indicates that the vast majority of youth correctly identified those activities they named as service. However, some of the names given by youth were ambiguous, and it is not possible to ascertain whether the activity in fact constituted a type of service or volunteer work.

### **7.3 Type of Community in Which the Household is Located**

The NHES:96 Youth CI file includes a measure of the urbanicity of the sampled household, COMMUNITY. This variable was copied from the Household & Library data file to the corresponding record on the Youth CI file. The creation of the derived variable COMMUNITY was described in chapter 6. At the household level, the item response rates for the variables used to create COMMUNITY were somewhat lower than most variables in the NHES. HCCOMMUN (SX31, the type of community) had an item response rate of 92.5 percent; HCSUB (SX31OV, size of city to which a suburb belonged, if applicable) had an item response rate of 93.1 percent; and HCCITY (SX31OV2, size of the city, if applicable) had an item response rate of 95.7 percent. When these three variable were combined to create the derived variable COMMUNITY, 85.3 percent had unimputed responses for all three variables. Some respondents simply do not know the size of the community in which they live. This suggests the possibility that some who responded to the questions may have guessed, although this cannot be measured directly. Analysts should keep this in mind when using the variable COMMUNITY, as this variable may contain response error.

#### **7.4 Correspondence Between Age and Grade**

In any survey in which information on people's ages and grades in school (or grade equivalents), some cases appear in which age and grade do not seem to correspond. This is true for each year of the NHES, the CPS October Educational Supplement, and other surveys. In many cases in the NHES, the situation behind these discrepancies are unclear -- it is only known that a CATI edit was tripped and the interviewer had to confirm the information and enter it again. In some cases, interviewers provide more complete explanations. For example, a person may be in a grade far lower than his/her age would indicate, but may be retarded and in a special education program with a low grade equivalent. Some adults long past the modal age of high school completion may report a secondary grade because they are enrolled in adult nighttime high school. Analysts may wish to examine these unusual cases and make their own decisions about how to treat these cases in their analyses.

#### **7.5 Income to the Nearest Thousand Dollars**

In those households whose income category and household size indicated that they may be at or below the poverty line, household income to the nearest thousand dollars was requested. As the values in the data file show, some respondents did not answer in thousands, but gave somewhat more specific answers. Rather than lose this information, the exact response was retained.

**APPENDIX B**

**YOUTH CI PUBLIC FILE LAYOUT IN POSITION ORDER**

*This page is blank.*



## Youth CI Public File Layout in Position Order

VARIABLE NAME	VARIABLE LABEL	FORMAT	RECORD NUMBER	LENGTH	START COLUMN	END COLUMN
BAS MID	INTERVIEW ID NUMBER	N	1	12	1	12
ENUMID	SUBJECT ID NUMBER	N	1	10	13	22
BASEID	HOUSEHOLD ID NUMBER	N	1	8	23	30
ENGLSPAN	EXTENDED IN ENGLISH OR SPANISH	N	1	2	31	32
CHILDNUM	PERSON'S EMUMERATION NUMBER	N	1	2	33	34
AGE95	CHILD'S AGE AS OF 12/31/95	N	1	2	35	36
SEX	S6-SEX	N	1	2	37	38
RACE	SX21-RACE	N	1	2	39	40
HISPANIC	SX22-HISPANIC	N	1	2	41	42
OTHRAC	SX21A-OTHER RACE CATEGORY	N	1	2	43	44
RESPNUM	SX23-PARENT RESPONDENT PERSON NUMBER	N	1	2	45	46
RESPAGE	PARENT RESPONDENT'S AGE	N	1	2	47	48
RESPSEX	PARENT RESPONDENT'S SEX	N	1	2	49	50
RESRELN	PARENT R'S RELATIONSHIP TO CHILD	N	1	2	51	52
MOMNUM	ENUM NUMBER OF CHILD'S MOTHER	N	1	2	53	54
MOMAGE	MOTHER'S AGE	N	1	2	55	56
MOMTYPE	SPECIFIC RELATIONSHIP OF MOTHER TO CHILD	N	1	2	57	58
DADNUM	ENUM NUMBER OF CHILD'S FATHER	N	1	2	59	60
DADAGE	FATHER'S AGE	N	1	2	61	62
DADTYPE	SPECIFIC RELATIONSHIP OF FATHER TO CHILD	N	1	2	63	64
CDOBMM	PA1-MONTH OF BIRTH	N	1	2	65	66
CDOBY	PA1-YEAR OF BIRTH	N	1	2	67	68
CSPEAK	PA3-LANG CHLD SPEAKS MOST AT HOME	N	1	2	69	70
ENROLL	PB1-CHILD ENROLLED/ATTENDING SCHOOL	N	1	2	71	72
HOMESCHL	PB2-CHILD BEING SCHOOLED AT HOME	N	1	2	73	74
GRADE	PB4-GRADE/YR CHLD IS ATTENDING	A	1	2	75	76
GRADEEQ	PB5-GRADE EQUIV/HOME SCH/SP ED/UNGRD	A	1	2	77	78
SPUBLIC	PD1-CHLD ATTNDS PUBL/PRIV SCH	N	1	2	79	80
SCHOICE	PD3-SCH ASSIGNED OR CHOSEN	N	1	2	81	82
SRELGN	PD4-CHLD ATTNDS CHURCH RELATED SCH	N	1	2	83	84
SCATHLIC	PD5-CHLD ATTNDS CATHOLIC SCH	N	1	2	85	86
SLOW	PD7-LOWEST GRADE AT CHLD'S SCH	A	1	2	87	88
SHIGH	PD8-HIGHEST GRADE AT CHLD'S SCH	A	1	2	89	90
SNUMSTUD	PD9-# OF STDTS AT CHLD'S SCH	N	1	2	91	92
SNUMGRAD	PD9OV-# OF STDTS IN CHLD'S GRADE	N	1	4	93	96
SETHNIC	PD10-PERCENTAGE STDTS OF CHLD'S RACE/ETH	N	1	2	97	98
FESCHOOL	YA1-FREQ CHLD TALKS W/FAM RE SCH	N	1	2	99	100
FEFUTURE	YA2-DISCUSSES FUTURE PLANS W/FAM	N	1	2	101	102
FESCHINV	YA3-LEVEL OF INVOLVEMENT IN SCH	N	1	2	103	104
FENOTICE	YA4-SCH GIVES WRN NOTICE TO TAKE HOME	N	1	2	105	106
FENOTGIV	YA5-FREQ NOTICES ARE TAKEN HOME	N	1	2	107	108
FERBED	YA6A-RULES ABT BEDTIME/SCH NIGHTS	N	1	2	109	110
FERSCHNT	YA6B-RULES ABT TIME HOME/SCH NIGHTS	N	1	2	111	112
FERHMWRK	YA6C-RULES ABT DOING HOMEWORK	N	1	2	113	114
FERTVTIM	YA6D-RULES ABT TV VIEWING TIME	N	1	2	115	116
FERTVPRG	YA6E-RULES ABT TV PRGMS WATCHED	N	1	2	117	118
FEFAMDEC	YA7A-FAM DISCUSSES DECISIONS W/CHLD	N	1	2	119	120
FEYRSIDE	YA7B-FAM LISTENS CHLDS SIDE/ARGUMNT	N	1	2	121	122
FERULES	YA7C-FAM LETS CHLD HAVE SAY IN RULES	N	1	2	123	124
FECHALNG	YA8A-CHLD IS CHALLENGED AT SCH	N	1	2	125	126
FEENJOY	YA8B-CHLD ENJOYS SCHOOL	N	1	2	127	128
FETEADIS	YA8C-TCHRS MAINTAIN DISCIPLINE	N	1	2	129	130
FERESPCT	YA8D-STDTS/TCHRS RESPECT EACH OTHR	N	1	2	131	132
FEPRIDIS	YA8E-PRINCIPAL MAINTAINS DISCIPLINE	N	1	2	133	134
FEWATCH	YA8F-FAM MONITORS SCH PROGRESS	N	1	2	135	136
FELISTEN	YA8G-STDY OPINIONS COUNT AT SCH	N	1	2	137	138
PRSTUGOV	YB1-SCH HAS STUDENT GOVT	N	1	2	139	140
PRREP GOV	YB2-SERVED/WORKED IN STUDENT GOVT	N	1	2	141	142
PRSCHACT	YB3-PARTICIPATED IN SCH ACTIVITIES	N	1	2	143	144
PRGRPACT	YB4-PARTICIPATED OUT-OF-SCH ACTIVITIES	N	1	2	145	146
PRWORK	YB5-WORKS FOR PAY	N	1	2	147	148
PRWRKHRS	YB6-HRS/WK WORKS	N	1	2	149	150
PRLOOK	YB7-LOOKED FOR JOB THIS SCH YR	N	1	2	151	152
SACTY	YC1-DOES COMMTY SERVICE ACTY	N	1	2	153	154
SANOW1	YC3-PARTICIPATING IN ACTIVITY #1 NOW	N	1	2	155	156
SAREG1	YC4-SERVICE ACTIVITY #1 SCHEDULE	N	1	2	157	158
SAWKS1	YC5-FREQ OF SERVICE ACTIVITY #1	N	1	2	159	160
SAWKSNU1	YC5OV-NUM WKS FOR SERV ACTY #1	N	1	2	161	162
SAHRS1	YC6-HRS/WK FOR SERV ACTY #1	N	1	2	163	164
SAHRSNU1	YC6OV-NUM HRS/WK FOR SERV ACTY #1	N	1	2	165	166
SANOW2	YC3-PARTICIPATING IN ACTIVITY #2 NOW	N	1	2	167	168

VARIABLE NAME	VARIABLE LABEL	FORMAT	RECORD NUMBER	LENGTH	START COLUMN	END COLUMN
SAREG2	YC4-SERVICE ACTIVITY #2 SCHEDULE	N	1	2	169	170
SAWKS2	YC5-FREQ OF SERVICE ACTIVITY #2	N	1	2	171	172
SAWKSNU2	YC5OV-NUM WKS FOR SERV ACTY #2	N	1	2	173	174
SAHRS2	YC6-HRS/WK FOR SERV ACTY #2	N	1	2	175	176
SAHRSNU2	YC6OV-NUM HRS/WK FOR SERV ACTY #2	N	1	2	177	178
SANOW3	YC3-PARTICIPATING IN ACTIVITY #3 NOW	N	1	2	179	180
SAREG3	YC4-SERVICE ACTIVITY #3 SCHEDULE	N	1	2	181	182
SAWKS3	YC5-FREQ OF SERVICE ACTIVITY #3	N	1	2	183	184
SAWKSNU3	YC5OV-NUM WKS FOR SERV ACTY #3	N	1	2	185	186
SAHRS3	YC6-HRS/WK FOR SERV ACTY #3	N	1	2	187	188
SAHRSNU3	YC6OV-NUM HRS/WK FOR SERV ACTY #3	N	1	2	189	190
SAARRYOU	YC7-SCH ARR THIS STDT SERV ACTY	N	1	2	191	192
SAARRSER	YC8-SCH ARRANGES SERV ACTIVITIES	N	1	2	193	194
SAREQSER	YC9-SCH REQUIRES SERV ACTY	N	1	2	195	196
SAREQYOU	YC10-SCH REQD THIS STDT SERV ACTY	N	1	2	197	198
SATALK	YC11-TALK IN CLASS/GRP ABT SERV ACTY	N	1	2	199	200
SAJOURNL	YC12-REQUIRED TO WRITE ABT SERV ACTY	N	1	2	201	202
SAGRADE	YC13-ACTIVITY FOR A GRADE IN CLASS	N	1	2	203	204
SASCHLYR	YC14-WILL DO SERV ACTY LATER THIS SCH YR	N	1	2	205	206
SANEXTYR	YC15-WILL DO SERV ACTY NEXT YR	N	1	2	207	208
PSPEACE	YC16A-HEARD OF THE PEACE CORPS	N	1	2	209	210
PSVISTA	YC16B-HEARD OF VISTA	N	1	2	211	212
PSAMCORP	YC16C-HEARD OF AMERICORPS	N	1	2	213	214
SASERVC	YC17-FAM PARTICIPATES COMMTY SERV	N	1	2	215	216
CYRDNEWU	YD1-FREQ YOUTH READS NATL NEWS	N	1	2	217	218
CYWATCHU	YD2-FREQ YOUTH WATCH/LSTN NATL NEWS	N	1	2	219	220
CYNEWSHH	YD3-WATCH/LSTN NATL NEWS W/FAM PST WK	N	1	2	221	222
CYISTALK	YD4-FREQ TALK ABT NATL NEWS W/FAM	N	1	2	223	224
CYCOMPLI	YD5A-CAN'T UNDERSTAND POLITICS/GOVT	N	1	2	225	226
CYFAMSAY	YD5B-FAM HAS NO SAY IN WHAT GOVT DOES	N	1	2	227	228
CYAGNST	YD5C-ALLOW FREEDOM TO SPEAK AGNST RELGN	N	1	2	229	230
CYBOOK	YD5D-SOME BKS SHLD BE KPT OUT/PUB LIB	N	1	2	231	232
CYLETTER	YD6-COULD WRITE LETTER TO GOVT OFCL	N	1	2	233	234
CYMTG	YD7-COULD MAKE STATEMENT AT PUBLIC MTG	N	1	2	235	236
CYVP	YD8A-JOB/POL OFC HELD BY AL GORE	N	1	2	237	238
CYLAW	YD8B-WHO DETERMINES LAW CONSTITUTIONAL	N	1	2	239	240
CYHOUSE	YD8C-PARTY W/MOST MEMBRS IN HOUSE	N	1	2	241	242
CYVETO	YD8D-MAJORITY NEEDED TO OVERRIDE VETO	N	1	2	243	244
CYCONSRV	YD8E-PARTY MORE CONSERV/NATL LEVEL	N	1	2	245	246
CYSPKR	YD9A-JOB/POL OFC HELD BY NEWT GINGRICH	N	1	2	247	248
CYJUDGE	YD9B-WHO NOMINATES FED COURT JUDGES	N	1	2	249	250
CYSENATE	YD9C-PARTY W/MOST MEMBRS IN SENATE	N	1	2	251	252
CYCONST	YD9D-1ST 10 AMENDMENTS TO CONSTIT	N	1	2	253	254
CYDFENS	YD9E-PARTY FAVORS LRGR DEFENSE BUDGET	N	1	2	255	256
CYCRSE	YD10-COURSE REQS ATTN TO GOVT ISSUES	N	1	2	257	258
CYCRSLST	YD11-LST YR COURSE REQD ATTN TO GOVT ISS	N	1	2	259	260
CYINTRST	YD12-CLASS INCREASED INT/GOVT ISSUES	N	1	2	261	262
CYSCHLET	YD13A-IN CLASS WROTE LTR TO UNKNOWN PERS	N	1	2	263	264
CYSCHSPE	YD13B-IN CLASS GAVE SPEECH/ORAL REPT	N	1	2	265	266
CYSCHDEB	YD13C-IN CLASS TOOK PART IN DEBATE	N	1	2	267	268
HOWNHOME	SX27-OWN, RENT HOME/OTHR ARRNGMNT	N	1	2	269	270
HCCOMMUN	SX31-COMMUNITY DESCRIPTION	N	1	2	271	272
HCSUB	SX31OV-SIZE OF SUBURB	N	1	2	273	274
HCCITY	SX31OV2-SIZE OF CITY	N	1	2	275	276
HWIC	SX32A-FAMILY RECD WIC PAST 12 MO	N	1	2	277	278
HFOODST	SX32B-FAMILY RECD FOOD STMPs PAST 12 MO	N	1	2	279	280
HAFDC	SX32C-FAMILY RECD AFDC PAST 12 MO	N	1	2	281	282
HINCMRNG	SX33- TOTAL HH INCOME RANGE	N	1	2	283	284
HINCOME	SX33-TOTAL HH INCOME RANGE 2	N	1	2	285	286
HINCMEXT	SX33OV-EXACT HH INC NEAREST \$1000	N	1	5	287	291
ALLGRADE	D-CHILLD'S ENROLLMENT AND GRADE/EQUIV	A	1	2	292	293
CENREG	D-CENSUS REGION	N	1	2	294	295
COMMUNTY	D-SIZE OF COMMUNITY CHILD RESIDES	N	1	2	296	297
DADEMPLD	D-WORK STATUS-DAD/STEP/FOSTER DAD/GUARD	N	1	2	298	299
FAMILY	D-FAMILY TYPE	N	1	2	300	301
HH18OVER	D-NUMBER OF HH MMBRS AGE 18 AND OLDER	N	1	2	302	303
HHDAD	D-FATHER LIVES IN HOUSEHOLD	N	1	2	304	305
HHMOM	D-MOTHER LIVES IN HOUSEHOLD	N	1	2	306	307
HHPARN1	D-PARENTS IN HH, GENERAL	N	1	2	308	309
HHTOTAL	D-TOTAL NUMBER OF HH MEMBERS	N	1	2	310	311

VARIABLE NAME	VARIABLE LABEL	FORMAT	RECORD NUMBER	LENGTH	START COLUMN	END COLUMN
HHUNDR6	D-NUMBER OF HH MMBRS AGE 5 AND YOUNGER	N	1	2	312	313
HHUNDR13	D-NUMBER OF HH MMBRS AGE 12 AND YOUNGER	N	1	2	314	315
HHUNDR18	D-NUMBER OF HH MMBRS AGE 17 AND YOUNGER	N	1	2	316	317
HHUNDR21	D-NUMBER OF HH MMBRS AGE 20 AND YOUNGER	N	1	2	324	325
LANGUAGE	D-IS ENGLISH SPOKEN BY PRNTS	N	1	2	318	319
MOMEMPLD	D-WORK STATUS-MOM/STEP/FOSTER MOM/GUARD	N	1	2	320	321
MOMFTFY	D-NOT EMPLOYED DURING YEAR	N	1	2	322	323
PARGRADE	D-HIGHEST LEVEL OF PRNT/GUARD EDUCATION	N	1	2	326	327
RACEETHN	D-RACE-ETHNICITY	N	1	2	328	329
SCHLGRAD	D-CLASSIFICATION OF CHILD'S SCHOOL	N	1	2	330	331
SCHLTYPE	D-TYPE OF SCHOOL CHILD ATTENDS	N	1	2	332	333
SCNUMSTU	D-ESTIMATED NUMBER STDTS IN CHILD'S SCH	N	1	2	334	335
FYWT	FINAL (RAKED) YOUTH INTERVIEW WEIGHT	N	1	10.3	336	345
BAS MID	INTERVIEW ID NUMBER	N	2	12	1	12
FYWTR1	FINAL (RAKED) YOUTH INTV. REPL. WGT-1	N	2	10.3	13	22
FYWTR2	FINAL (RAKED) YOUTH INTV. REPL. WGT-2	N	2	10.3	23	32
FYWTR3	FINAL (RAKED) YOUTH INTV. REPL. WGT-3	N	2	10.3	33	42
FYWTR4	FINAL (RAKED) YOUTH INTV. REPL. WGT-4	N	2	10.3	43	52
FYWTR5	FINAL (RAKED) YOUTH INTV. REPL. WGT-5	N	2	10.3	53	62
FYWTR6	FINAL (RAKED) YOUTH INTV. REPL. WGT-6	N	2	10.3	63	72
FYWTR7	FINAL (RAKED) YOUTH INTV. REPL. WGT-7	N	2	10.3	73	82
FYWTR8	FINAL (RAKED) YOUTH INTV. REPL. WGT-8	N	2	10.3	83	92
FYWTR9	FINAL (RAKED) YOUTH INTV. REPL. WGT-9	N	2	10.3	93	102
FYWTR10	FINAL (RAKED) YOUTH INTV. REPL. WGT-10	N	2	10.3	103	112
FYWTR11	FINAL (RAKED) YOUTH INTV. REPL. WGT-11	N	2	10.3	113	122
FYWTR12	FINAL (RAKED) YOUTH INTV. REPL. WGT-12	N	2	10.3	123	132
FYWTR13	FINAL (RAKED) YOUTH INTV. REPL. WGT-13	N	2	10.3	133	142
FYWTR14	FINAL (RAKED) YOUTH INTV. REPL. WGT-14	N	2	10.3	143	152
FYWTR15	FINAL (RAKED) YOUTH INTV. REPL. WGT-15	N	2	10.3	153	162
FYWTR16	FINAL (RAKED) YOUTH INTV. REPL. WGT-16	N	2	10.3	163	172
FYWTR17	FINAL (RAKED) YOUTH INTV. REPL. WGT-17	N	2	10.3	173	182
FYWTR18	FINAL (RAKED) YOUTH INTV. REPL. WGT-18	N	2	10.3	183	192
FYWTR19	FINAL (RAKED) YOUTH INTV. REPL. WGT-19	N	2	10.3	193	202
FYWTR20	FINAL (RAKED) YOUTH INTV. REPL. WGT-20	N	2	10.3	203	212
FYWTR21	FINAL (RAKED) YOUTH INTV. REPL. WGT-21	N	2	10.3	213	222
FYWTR22	FINAL (RAKED) YOUTH INTV. REPL. WGT-22	N	2	10.3	223	232
FYWTR23	FINAL (RAKED) YOUTH INTV. REPL. WGT-23	N	2	10.3	233	242
FYWTR24	FINAL (RAKED) YOUTH INTV. REPL. WGT-24	N	2	10.3	243	252
FYWTR25	FINAL (RAKED) YOUTH INTV. REPL. WGT-25	N	2	10.3	253	262
FYWTR26	FINAL (RAKED) YOUTH INTV. REPL. WGT-26	N	2	10.3	263	272
FYWTR27	FINAL (RAKED) YOUTH INTV. REPL. WGT-27	N	2	10.3	273	282
FYWTR28	FINAL (RAKED) YOUTH INTV. REPL. WGT-28	N	2	10.3	283	292
FYWTR29	FINAL (RAKED) YOUTH INTV. REPL. WGT-29	N	2	10.3	293	302
FYWTR30	FINAL (RAKED) YOUTH INTV. REPL. WGT-30	N	2	10.3	303	312
FYWTR31	FINAL (RAKED) YOUTH INTV. REPL. WGT-31	N	2	10.3	313	322
FYWTR32	FINAL (RAKED) YOUTH INTV. REPL. WGT-32	N	2	10.3	323	332
FYWTR33	FINAL (RAKED) YOUTH INTV. REPL. WGT-33	N	2	10.3	333	342
FYWTR34	FINAL (RAKED) YOUTH INTV. REPL. WGT-34	N	2	10.3	343	352
FYWTR35	FINAL (RAKED) YOUTH INTV. REPL. WGT-35	N	2	10.3	353	362
FYWTR36	FINAL (RAKED) YOUTH INTV. REPL. WGT-36	N	2	10.3	363	372
FYWTR37	FINAL (RAKED) YOUTH INTV. REPL. WGT-37	N	2	10.3	373	382
FYWTR38	FINAL (RAKED) YOUTH INTV. REPL. WGT-38	N	2	10.3	383	392
FYWTR39	FINAL (RAKED) YOUTH INTV. REPL. WGT-39	N	2	10.3	393	402
FYWTR40	FINAL (RAKED) YOUTH INTV. REPL. WGT-40	N	2	10.3	403	412
FYWTR41	FINAL (RAKED) YOUTH INTV. REPL. WGT-41	N	2	10.3	413	422
FYWTR42	FINAL (RAKED) YOUTH INTV. REPL. WGT-42	N	2	10.3	423	432
FYWTR43	FINAL (RAKED) YOUTH INTV. REPL. WGT-43	N	2	10.3	433	442
FYWTR44	FINAL (RAKED) YOUTH INTV. REPL. WGT-44	N	2	10.3	443	452
FYWTR45	FINAL (RAKED) YOUTH INTV. REPL. WGT-45	N	2	10.3	453	462
FYWTR46	FINAL (RAKED) YOUTH INTV. REPL. WGT-46	N	2	10.3	463	472
FYWTR47	FINAL (RAKED) YOUTH INTV. REPL. WGT-47	N	2	10.3	473	482
FYWTR48	FINAL (RAKED) YOUTH INTV. REPL. WGT-48	N	2	10.3	483	492
FYWTR49	FINAL (RAKED) YOUTH INTV. REPL. WGT-49	N	2	10.3	493	502
FYWTR50	FINAL (RAKED) YOUTH INTV. REPL. WGT-50	N	2	10.3	503	512
FYWTR51	FINAL (RAKED) YOUTH INTV. REPL. WGT-51	N	2	10.3	513	522
FYWTR52	FINAL (RAKED) YOUTH INTV. REPL. WGT-52	N	2	10.3	523	532
FYWTR53	FINAL (RAKED) YOUTH INTV. REPL. WGT-53	N	2	10.3	533	542
FYWTR54	FINAL (RAKED) YOUTH INTV. REPL. WGT-54	N	2	10.3	543	552
FYWTR55	FINAL (RAKED) YOUTH INTV. REPL. WGT-55	N	2	10.3	553	562
FYWTR56	FINAL (RAKED) YOUTH INTV. REPL. WGT-56	N	2	10.3	563	572
FYWTR57	FINAL (RAKED) YOUTH INTV. REPL. WGT-57	N	2	10.3	573	582
FYWTR58	FINAL (RAKED) YOUTH INTV. REPL. WGT-58	N	2	10.3	583	592
FYWTR59	FINAL (RAKED) YOUTH INTV. REPL. WGT-59	N	2	10.3	593	602
FYWTR60	FINAL (RAKED) YOUTH INTV. REPL. WGT-60	N	2	10.3	603	612

VARIABLE NAME	VARIABLE LABEL	FORMAT	RECORD NUMBER	LENGTH	START COLUMN	END COLUMN
FYWTR61	FINAL (RAKED) YOUTH INTV. REPL. WGT-61	N	2	10.3	613	622
FYWTR62	FINAL (RAKED) YOUTH INTV. REPL. WGT-62	N	2	10.3	623	632
FYWTR63	FINAL (RAKED) YOUTH INTV. REPL. WGT-63	N	2	10.3	633	642
FYWTR64	FINAL (RAKED) YOUTH INTV. REPL. WGT-64	N	2	10.3	643	652
FYWTR65	FINAL (RAKED) YOUTH INTV. REPL. WGT-65	N	2	10.3	653	662
FYWTR66	FINAL (RAKED) YOUTH INTV. REPL. WGT-66	N	2	10.3	663	672
FYWTR67	FINAL (RAKED) YOUTH INTV. REPL. WGT-67	N	2	10.3	673	682
FYWTR68	FINAL (RAKED) YOUTH INTV. REPL. WGT-68	N	2	10.3	683	692
FYWTR69	FINAL (RAKED) YOUTH INTV. REPL. WGT-69	N	2	10.3	693	702
FYWTR70	FINAL (RAKED) YOUTH INTV. REPL. WGT-70	N	2	10.3	703	712
FYWTR71	FINAL (RAKED) YOUTH INTV. REPL. WGT-71	N	2	10.3	713	722
FYWTR72	FINAL (RAKED) YOUTH INTV. REPL. WGT-72	N	2	10.3	723	732
FYWTR73	FINAL (RAKED) YOUTH INTV. REPL. WGT-73	N	2	10.3	733	742
FYWTR74	FINAL (RAKED) YOUTH INTV. REPL. WGT-74	N	2	10.3	743	752
FYWTR75	FINAL (RAKED) YOUTH INTV. REPL. WGT-75	N	2	10.3	753	762
FYWTR76	FINAL (RAKED) YOUTH INTV. REPL. WGT-76	N	2	10.3	763	772
FYWTR77	FINAL (RAKED) YOUTH INTV. REPL. WGT-77	N	2	10.3	773	782
FYWTR78	FINAL (RAKED) YOUTH INTV. REPL. WGT-78	N	2	10.3	783	792
FYWTR79	FINAL (RAKED) YOUTH INTV. REPL. WGT-79	N	2	10.3	793	802
FYWTR80	FINAL (RAKED) YOUTH INTV. REPL. WGT-80	N	2	10.3	803	812
YPSU	FOR USE IN TAYLOR SERIES VARIANCE	N	2	4	813	816
YSTRATUM	FOR USE IN TAYLOR SERIES VARIANCE	N	2	2	817	818
SEF	IMPUTATION FLAG	N	2	2	819	820
RACF	IMPUTATION FLAG	N	2	2	821	822
HISPANIF	IMPUTATION FLAG	N	2	2	823	824
OTHRAF	IMPUTATION FLAG	N	2	2	825	826
RESRELF	IMPUTATION FLAG	N	2	2	827	828
MOMTYPF	IMPUTATION FLAG	N	2	2	829	830
DADTYPF	IMPUTATION FLAG	N	2	2	831	832
CSPEAF	IMPUTATION FLAG	N	2	2	833	834
ENROLF	IMPUTATION FLAG	N	2	2	835	836
HOMESCHF	IMPUTATION FLAG	N	2	2	837	838
GRADF	IMPUTATION FLAG	N	2	2	839	840
GRADEEF	IMPUTATION FLAG	N	2	2	841	842
SPUBLIF	IMPUTATION FLAG	N	2	2	843	844
SCHOICF	IMPUTATION FLAG	N	2	2	845	846
SRELGOF	IMPUTATION FLAG	N	2	2	847	848
SCATHLIF	IMPUTATION FLAG	N	2	2	849	850
SLOF	IMPUTATION FLAG	N	2	2	851	852
SHIGF	IMPUTATION FLAG	N	2	2	853	854
SNUMSTUF	IMPUTATION FLAG	N	2	2	855	856
SNUMGRAF	IMPUTATION FLAG	N	2	2	857	858
SETHNIF	IMPUTATION FLAG	N	2	2	859	860
FESCHOOF	IMPUTATION FLAG	N	2	2	861	862
FEFUTURF	IMPUTATION FLAG	N	2	2	863	864
FESCHINF	IMPUTATION FLAG	N	2	2	865	866
FENOTICF	IMPUTATION FLAG	N	2	2	867	868
FENOTGIF	IMPUTATION FLAG	N	2	2	869	870
FERBEF	IMPUTATION FLAG	N	2	2	871	872
FERSCHNF	IMPUTATION FLAG	N	2	2	873	874
FERHMWRF	IMPUTATION FLAG	N	2	2	875	876
FERTVTIF	IMPUTATION FLAG	N	2	2	877	878
FETVPRF	IMPUTATION FLAG	N	2	2	879	880
FEFAMDEF	IMPUTATION FLAG	N	2	2	881	882
FEYRSIDF	IMPUTATION FLAG	N	2	2	883	884
FERULEF	IMPUTATION FLAG	N	2	2	885	886
FECHALNF	IMPUTATION FLAG	N	2	2	887	888
FEENJOF	IMPUTATION FLAG	N	2	2	889	890
FETEADIF	IMPUTATION FLAG	N	2	2	891	892
FERESPCF	IMPUTATION FLAG	N	2	2	893	894
FEPRIDIF	IMPUTATION FLAG	N	2	2	895	896
FEWATCF	IMPUTATION FLAG	N	2	2	897	898
FELISTEF	IMPUTATION FLAG	N	2	2	899	900
PRSTUGOF	IMPUTATION FLAG	N	2	2	901	902
PRREPGOF	IMPUTATION FLAG	N	2	2	903	904
PRSCHACF	IMPUTATION FLAG	N	2	2	905	906
PRGRPACF	IMPUTATION FLAG	N	2	2	907	908
PRWORF	IMPUTATION FLAG	N	2	2	909	910
PRWRKHRF	IMPUTATION FLAG	N	2	2	911	912
PRLOOF	IMPUTATION FLAG	N	2	2	913	914
SACTF	IMPUTATION FLAG	N	2	2	915	916
SANOF1	IMPUTATION FLAG	N	2	2	917	918
SAREF1	IMPUTATION FLAG	N	2	2	919	920
SAWKF1	IMPUTATION FLAG	N	2	2	921	922

VARIABLE NAME	VARIABLE LABEL	FORMAT	RECORD NUMBER	LENGTH	START COLUMN	END COLUMN
SAWKSNF1	IMPUTATION FLAG	N	2	2	923	924
SAHRF1	IMPUTATION FLAG	N	2	2	925	926
SAHRSNF1	IMPUTATION FLAG	N	2	2	927	928
SANOF2	IMPUTATION FLAG	N	2	2	929	930
SAREF2	IMPUTATION FLAG	N	2	2	931	932
SAWKF2	IMPUTATION FLAG	N	2	2	933	934
SAWKSNF2	IMPUTATION FLAG	N	2	2	935	936
SAHRF2	IMPUTATION FLAG	N	2	2	937	938
SAHRSNF2	IMPUTATION FLAG	N	2	2	939	940
SANOF3	IMPUTATION FLAG	N	2	2	941	942
SAREF3	IMPUTATION FLAG	N	2	2	943	944
SAWKF3	IMPUTATION FLAG	N	2	2	945	946
SAWKSNF3	IMPUTATION FLAG	N	2	2	947	948
SAHRF3	IMPUTATION FLAG	N	2	2	949	950
SAHRSNF3	IMPUTATION FLAG	N	2	2	951	952
SAARRYOF	IMPUTATION FLAG	N	2	2	953	954
SAARRSEF	IMPUTATION FLAG	N	2	2	955	956
SAREQSEF	IMPUTATION FLAG	N	2	2	957	958
SAREQYOF	IMPUTATION FLAG	N	2	2	959	960
SATALF	IMPUTATION FLAG	N	2	2	961	962
SAJOURNF	IMPUTATION FLAG	N	2	2	963	964
SAGRADF	IMPUTATION FLAG	N	2	2	965	966
SASCHLYF	IMPUTATION FLAG	N	2	2	967	968
SANEXTYF	IMPUTATION FLAG	N	2	2	969	970
PSPEACF	IMPUTATION FLAG	N	2	2	971	972
PSVISTF	IMPUTATION FLAG	N	2	2	973	974
PSAMCORF	IMPUTATION FLAG	N	2	2	975	976
SASERVF	IMPUTATION FLAG	N	2	2	977	978
CYRDNEWF	IMPUTATION FLAG	N	2	2	979	980
CYWATCHF	IMPUTATION FLAG	N	2	2	981	982
CYNEWSHF	IMPUTATION FLAG	N	2	2	983	984
CYISTALF	IMPUTATION FLAG	N	2	2	985	986
CYCOMPLF	IMPUTATION FLAG	N	2	2	987	988
CYFAMSAF	IMPUTATION FLAG	N	2	2	989	990
CYAGNSF	IMPUTATION FLAG	N	2	2	991	992
CYBOOF	IMPUTATION FLAG	N	2	2	993	994
CYLETTEF	IMPUTATION FLAG	N	2	2	995	996
CYMTF	IMPUTATION FLAG	N	2	2	997	998
CYVF	IMPUTATION FLAG	N	2	2	999	1000
CYLAF	IMPUTATION FLAG	N	2	2	1001	1002
BASMID	INTERVIEW ID NUMBER	N	3	12	1	12
CYHOUSF	IMPUTATION FLAG	N	3	2	13	14
CYVETF	IMPUTATION FLAG	N	3	2	15	16
CYCONSRF	IMPUTATION FLAG	N	3	2	17	18
CYSPKF	IMPUTATION FLAG	N	3	2	19	20
CYJUDGF	IMPUTATION FLAG	N	3	2	21	22
CYSENATF	IMPUTATION FLAG	N	3	2	23	24
CYCONSF	IMPUTATION FLAG	N	3	2	25	26
CYDFENF	IMPUTATION FLAG	N	3	2	27	28
CYCRSF	IMPUTATION FLAG	N	3	2	29	30
CYCRSLSF	IMPUTATION FLAG	N	3	2	31	32
CYINTRSF	IMPUTATION FLAG	N	3	2	33	34
CYSCHLEF	IMPUTATION FLAG	N	3	2	35	36
CYSCHSPF	IMPUTATION FLAG	N	3	2	37	38
CYSCHDEF	IMPUTATION FLAG	N	3	2	39	40
HOWNHOMF	IMPUTATION FLAG	N	3	2	41	42
HCCOMMUF	IMPUTATION FLAG	N	3	2	43	44
HCSUF	IMPUTATION FLAG	N	3	2	45	46
HCCITF	IMPUTATION FLAG	N	3	2	47	48
HWIF	IMPUTATION FLAG	N	3	2	49	50
HFOODSF	IMPUTATION FLAG	N	3	2	51	52
HAFDF	IMPUTATION FLAG	N	3	2	53	54
HINCMRNF	IMPUTATION FLAG	N	3	2	55	56
HINCOMF	IMPUTATION FLAG	N	3	2	57	58
HINCMEXF	IMPUTATION FLAG	N	3	2	59	60

NOTE: The variables RECNUM is located in the last column of each record (column 1,024). The value of RECNUM varies with the record number of a given case. RECNUM is set to one on the first record of every case, 2 for the second record, and 3 for the third record. Each case on the Youth Civic Involvement data set has three records of data.

*This page is blank.*

## **APPENDIX C**

### **SAS CODE FOR DERIVED VARIABLES**

*This page is blank.*



```

/*--ALLGRADE--*/
IF GRADE = '-1' & GRADEEQ = '-1' THEN ALLGRADE = '0';
ELSE IF GRADE IN('T','K','P') OR GRADEEQ IN('T','K','P')
THEN ALLGRADE = 'K';
ELSE IF GRADE IN('N','1','2','3','4','5','6','7','8','9',
'10','11','12') THEN ALLGRADE = GRADE;
ELSE IF (GRADE IN('U','S','-1') & GRADEEQ IN('U',' '))
THEN ALLGRADE = 'U';
ELSE IF (GRADE IN('U','S','-1') & GRADEEQ NE ' ')
THEN ALLGRADE = GRADEEQ;
ELSE ALLGRADE = '-1';

/*--COMMUNITY--*/
IF HCCOMMUN = 4 & HCCITY = 1 THEN COMMUNITY = 1;
ELSE IF HCCOMMUN = 4 & HCCITY = 2 THEN COMMUNITY = 2;
ELSE IF HCCOMMUN = 4 & HCCITY = 3 THEN COMMUNITY = 3;
ELSE IF HCCOMMUN = 2 & HCSUB = 1 THEN COMMUNITY = 4;
ELSE IF HCCOMMUN = 2 & HCSUB = 2 THEN COMMUNITY = 5;
ELSE IF HCCOMMUN = 2 & HCSUB = 3 THEN COMMUNITY = 6;
ELSE IF HCCOMMUN = 3 THEN COMMUNITY = 7;
ELSE IF HCCOMMUN = 1 THEN COMMUNITY = 8;

/*--DAEMPLD--*/
IF ((DADWORK = 1 OR (DADWORK = 2 & DADLEAVE = 1))
& DADHOURS GE 35) THEN DAEMPLD = 1;
ELSE IF ( (DADWORK = 1 OR (DADWORK = 2 & DADLEAVE = 1))
& DADHOURS < 35) THEN DAEMPLD = 2;
ELSE IF (DADWORK = 2 & DADLEAVE = 2 & (DADLOOK = 1 &
(DADPUBL = 1 OR DADPRIV = 1 OR DAEMPL = 1 OR
DADREL = 1 OR DADANSAD = 1))) THEN DAEMPLD = 3;
ELSE IF DADWORK = -1 THEN DAEMPLD = -1;
ELSE DAEMPLD = 4;

/*--FAMILY--*/
IF HHPARN1 = 1 & NUMSIBS > 0 THEN FAMILY = 1;
ELSE IF HHPARN1 = 1 & NUMSIBS = 0 THEN FAMILY = 2;
ELSE IF HHPARN1 IN(2,3) & NUMSIBS > 0 THEN FAMILY = 3;
ELSE IF HHPARN1 IN(2,3) & NUMSIBS = 0 THEN FAMILY = 4;
ELSE FAMILY = 5;

/*--HHDAD--*/
IF DADTYPE IN(1,2) THEN HHDAD = 1;
ELSE IF DADTYPE IN(3,4) THEN HHDAD = 2;
ELSE IF (DADTYPE = -1 & MOMTYPE = -1) & RESPSEX = 1 THEN HHDAD = 3;
ELSE HHDAD = 4;

/*--HHMOM--*/
IF MOMTYPE IN(1,2) THEN HHMOM = 1;
ELSE IF MOMTYPE IN(3,4) THEN HHMOM = 2;
ELSE IF (MOMTYPE = -1 & DADTYPE = -1) & RESPSEX = 2 THEN HHMOM = 3;
ELSE HHMOM = 4;

/*--HHPARN1--*/
IF (HHMOM IN(1,2) & HHDAD IN(1,2)) THEN HHPARN1 = 1;
ELSE IF (HHMOM IN(1,2) & HHDAD IN(3,4)) THEN HHPARN1 = 2;
ELSE IF (HHMOM IN(3,4) & HHDAD IN(1,2)) THEN HHPARN1 = 3;
ELSE HHPARN1 = 4;

```

```

/*--LANGUAGE--*/
IF ((MOMLANG IN(1,3) OR MOMSPEAK IN(1,3)) &
    (DADLANG IN(-1,1,3) OR DADSPEAK IN(-1,1,3)))
    THEN LANGUAGE = 1;
ELSE IF (MOMLANG = -1 & (DADLANG IN(1,3) OR DADSPEAK IN(1,3)))
    THEN LANGUAGE = 1;
ELSE IF ((MOMLANG IN(1,3) OR MOMSPEAK IN(1,3)) & DADSPEAK IN(2,91))
    THEN LANGUAGE = 2;
ELSE IF (MOMSPEAK IN(2,91) & (DADLANG IN(1,3) OR DADSPEAK IN(1,3)))
    THEN LANGUAGE = 2;
ELSE IF (MOMSPEAK IN(2,91) & (DADSPEAK IN(2,91) OR DADLANG = -1))
    THEN LANGUAGE = 3;
ELSE IF (MOMLANG = -1 & DADSPEAK IN(2,91)) THEN LANGUAGE = 3;
ELSE LANGUAGE = -1;

```

```

/*--MOMEMPLD--*/
IF ((MOMWORK = 1 OR (MOMWORK = 2 & MOMLEAVE = 1))
    & MOMHOURS GE 35) THEN MOMEMPLD = 1;
ELSE IF ((MOMWORK = 1 OR (MOMWORK = 2 & MOMLEAVE = 1))
    & MOMHOURS < 35) THEN MOMEMPLD = 2;
ELSE IF (MOMWORK = 2 & MOMLEAVE = 2 & (MOMLOOK = 1 &
    (MOMPUBL = 1 OR MOMPRIV = 1 OR MOMEMPL = 1 OR
    MOMREL = 1 OR MOMANSAD = 1))) THEN MOMEMPLD = 3;
ELSE IF MOMWORK = -1 THEN MOMEMPLD = -1;
ELSE MOMEMPLD = 4;

```

```

/*--MOMFTFY--*/
IF MOMWORK = -1 THEN MOMFTFY = -1;
ELSE IF (MOMEMPLD = 1 & MOMMTHS = 12) THEN MOMFTFY = 1;
ELSE IF (MOMEMPLD = 1 & (0 <= MOMMTHS <= 11)) THEN MOMFTFY = 2;
ELSE IF MOMEMPLD = 2 THEN MOMFTFY = 2;
ELSE IF ((MOMEMPLD = 3 OR MOMEMPLD = 4) & MOMMTHS > 0)
    THEN MOMFTFY = 2;
ELSE IF (MOMEMPLD = 3 OR MOMEMPLD = 4) THEN MOMFTFY = 3;
ELSE MOMFTFY = -1;

```

```

/*--PARGRADE--*/
IF (MOMGRADE >= 10 OR DADGRADE >= 10) THEN PARGRADE = 5;
ELSE IF (MOMGRADE = 9 OR DADGRADE = 9) THEN PARGRADE = 4;
ELSE IF ((5 <= MOMGRADE <= 8) OR (5 <= DADGRADE <= 8))
    THEN PARGRADE = 3;
ELSE IF (MOMGRADE = 4 OR (MOMGRADE IN(1,2, 3) & MOMDIPL = 1)) OR
    (DADGRADE = 4 OR (DADGRADE IN(1,2,3) & DADDIPL = 1))
    THEN PARGRADE = 2;
ELSE IF (MOMGRADE IN(1,2,3) OR DADGRADE IN(1,2,3))
    THEN PARGRADE = 1;
ELSE IF MOMGRADE = -1 & DADGRADE = -1 THEN PARGRADE = 0;

```

```

/*--RACEETHN--*/
IF HISPANIC = 1 THEN RACEETHN = 3;
ELSE IF RACE = 1 THEN RACEETHN = 1;
ELSE IF RACE = 2 THEN RACEETHN = 2;
ELSE IF RACE IN(3,4) OR (RACE = 5 & OTHRAC IN(2,91))
    THEN RACEETHN = 4;

```

NOTE: Following creation of this variable, using the above code, a recode was performed for the Youth CI file. Because category 1 (early childcare) was not applicable to youth, that category was dropped and the remaining categories were recoded as shown on page 10.

```

                /*--SCHLGRAD--*/

IF SOTHGRAD = 1 THEN SLOW = 'N';
ELSE IF SOTHGRAD = 2 THEN DO;
    SLOW = 'N';
    SHIGH = 'N';
END;

IF SLOW = '-1' & SHIGH = '-1' THEN SCHLGRAD = -1;
ELSE IF SLOW IN('N','K','T','P') & SHIGH IN('N','K','T','P')
    THEN SCHLGRAD = 1;
ELSE IF SLOW IN('N','K','T','P','1','2','3') &
    SHIGH IN('1','2','3','4','5','6','7','8')
    THEN SCHLGRAD = 2;
ELSE IF SLOW IN('4','5','6','7','8','9') &
    SHIGH IN('4','5','6','7','8','9') THEN SCHLGRAD = 3;
ELSE IF SLOW IN('7','8','9','10','11','12') &
    SHIGH IN('10','11','12') THEN SCHLGRAD = 4;
ELSE SCHLGRAD = 5;

                /*--SCHLTYPE--*/

IF (SPUBLIC = 1 & SCHOICE = 1) THEN SCHLTYPE = 1;
ELSE IF SPUBLIC = 1 & SCHOICE IN(2,3) THEN SCHLTYPE = 2;
ELSE IF SRELGON = 1 THEN SCHLTYPE = 3;
ELSE IF SRELGON = 2 THEN SCHLTYPE = 4;
ELSE SCHLTYPE = -1;

                /*--SCNUMSTU--*/

IF (MAINRSLT IN('CH','CN') OR SHIGH = '-1' OR SLOW = '-1')
    THEN SCNUMSTU = -1;
ELSE DO;
IF SLOW IN('N','T','K','P') THEN TSLOW = '0';
    ELSE TSLOW = SLOW;
IF SHIGH IN('N','T','K','P') THEN TSHIGH = '0';
    ELSE TSHIGH = SHIGH;
SHIGHN = TSHIGH * 1;
TSLOWN = TSLOW * 1;
IF SNUMSTUD LE 4 & SNUMSTUD GE 1 THEN SCNUMSTU = SNUMSTUD;
ELSE DO;
    IF SNUMGRAD GE 1 THEN DO;
        IF (TSLOWN GE 0 & SHIGHN GE 0) THEN
            NUMSCHL = ((SHIGHN - TSLOWN) + 1) * SNUMGRAD;
        ELSE IF TSLOWN LT 0 OR SHIGHN LT 0 THEN NUMSCHL = -1;
        IF NUMSCHL LT 300 THEN SCNUMSTU = 1;
            ELSE IF (300 LE NUMSCHL LT 600) THEN SCNUMSTU = 2;
            ELSE IF (600 LE NUMSCHL LT 1000) THEN SCNUMSTU = 3;
            ELSE IF NUMSCHL GE 1000 THEN SCNUMSTU = 4;
        END;
    ELSE SCNUMSTU = -1;
END;
END;

```

*This page is blank.*

## **APPENDIX E**

### **DIRECTIONS AND CODE FOR LINKING DATA FILES**

*This page is blank.*

## Linking the NHES:96 Data Files

It is possible to link information from the four distinct NHES:96 files. This process is clear once the structure of the file identifiers is understood. First, the types of identifiers found on the NHES:96 data files are discussed. In the NHES:96, there are **household** identification numbers (**BASEID**), interview **subject** identification numbers (**ENUMID**), and **interview** or case identification numbers (**BASMID**). In addition, two-digit **person numbers** are provided for household members within households. The household, interview, and subject identification numbers are first discussed and then the person numbers.

- **BASEID**, is the **household** identification number. This eight-digit identification number is the same for every data record within a household and is the case identification number for the Household & Library data file. It is also provided on the Parent PFI/CI file, the Youth CI file, and the Adult CI file to permit data users to form linkages between the files.
- **ENUMID** is the **interview subject** identification number and is composed of 10 digits. "Interview subject" identification means that this number is unique to the person who is the subject of the interview. For example, in the Parent PFI/CI data file, **ENUMID** is the ID number of the child or youth who is the *subject* of the interview. Thus, **ENUMID** is the same in the Parent PFI/CI interview about a given youth and in that youth's own Youth CI interview record. **ENUMID** appears in the Parent PFI/CI and Youth CI data files only.
- **BASMID** is the unique **interview** or case identification number and is composed of 12 digits. Each Parent PFI/CI, Youth CI, or Adult CI interview has this unique interview ID. The first eight digits of **BASMID** are the same as **BASEID** for the household to which the interview belongs. The first 10 digits of **BASMID** are the same as the **ENUMID** of the subject of the interview. Therefore, a Parent PFI/CI interview record and a Youth CI record about the same youth would have the same value for **ENUMID**, but each interview would have its own unique **BASMID**. For the Parent PFI/CI interview, the last two digits are 01; for the Youth CI interview, the last two digits are 02; and for the Adult CI interview, the last two digits are 03. **BASMID** is the variable specified as the case identification number when creating a WesVarPC analysis from the Parent PFI/CI, Youth CI, or Adult CI files (see chapter 3 of this volume for a discussion of variance estimation and WesVarPC.)

Two-digit person identification numbers are provided on each of the NHES:96 data files in order to permit data users to copy information about certain individuals across interview records. In developing the public use data files, care was taken to include those person-level variables that were most likely to be needed by analysts. For example, the educational attainment of and languages spoken by the child's parents are included on the Parent PFI/CI file, since these parent characteristics are likely to be of interest to many analysts. These characteristics of subject adults are also included on the Adult CI file. In addition, since household characteristics (such as own/rent and income) are likely to be of interest to many analysts, these variables are contained on all four of the NHES:96 data sets and it is not necessary to copy them from the Household & Library file to the Parent PFI/CI, Youth CI, or Adult CI data files.

However, there may be circumstances in which an analyst would like to copy data items about a household member from one file to another. In order to facilitate linkages between the NHES:96 data files for the purpose of merging person characteristics, individual person numbers are provided on the data files. These two-digit numbers represent the number assigned to the person during enumeration of the household, i.e., 01 for the first person listed by the Screener respondent, 02 for the second person, etc.

- **PNUM(n)** is the two-digit person number variable in the Household & Library file. As noted above, many individual characteristics appear on this file for each household member, and these characteristics have sequential numbers, e.g., AGE1, AGE2, AGE3, and so on. Similarly, PNUM(n) carries the same sequential numbers in the variable name, so that it appears as PNUM1, PNUM2, PNUM3, etc. PNUM(n) contains the number assigned to each household member at the time of enumeration,

**NOTE:** *The value of PNUM(n) is not the same as the sequential number it carries in its variable name.* That is, PNUM1 is not necessarily equal to '01,' PNUM2 is not always equal to '02,' etc. There are two reasons for this. First, during data collection and data preparation, household members who were originally enumerated may have been deleted because they were later determined not to be household members according to the study definition. The enumeration numbers assigned to the household members were not changed when this occurred, because doing so would have disrupted linkages between segments in the hierarchical CATI data base. Second, after imputation of person-level records, the household members were sorted by age, oldest to youngest, before constructing the rectangular Household & Library file.

- In the Parent PFI/CI and Youth CI data files, four two-digit person numbers are provided -- **MOMNUM** for the child's mother (if any), **DADNUM** for the child's father (if any), **RESPNUM**, the person number of the respondent to the Parent PFI/CI interview, and **CHILDNUM** for the subject child or youth. If the mother or the father was the respondent to the Parent PFI/CI interview, MOMNUM (or DADNUM) will have the same value as RESPNUM.
- In the Adult CI data file, the person number for the sampled adult is contained in **PERSNUM**.

In order to effectively approach linkages between NHES:96 data files, it is important to remember the structure of the NHES:96 sample. Every household with a completed Screener interview has a household record in the Household & Library file. Therefore, every Parent PFI/CI, Youth CI, and Adult CI data record belongs to a household that is also represented in the Household & Library file. Because the Youth CI interview was only attempted if the corresponding Parent PFI/CI interview was completed, every Youth CI interview has a corresponding Parent PFI/CI record. As noted earlier, the sample of telephone numbers for the NHES:96 was split, and 95 percent was assigned to Parent/Youth interviewing and 5 percent was assigned to Adult CI interviewing. As a result, there are no Parent PFI/CI or Youth CI interviews for which there is an Adult CI interview in the same household.

The following examples are provided for the general populations for each component. Data users should consider the following tips regarding the length of time required to run a program and use of disk space:



- The data files are provided in ID order -- all of the following examples present code for sorting data files prior to linking (merging). Sorting the files can take up considerable time and disk space. If the files are already in the order required by the analyst, sorting is unnecessary.
- Keep only the variables required for your analysis -- specifying only the variables needed for the analysis will significantly improve the speed of the linking and the created data file will use less disk space. The use of a KEEP option, demonstrated in some of the following examples, can be used for this purpose.
- Keep only the relevant records -- when linking, for example, the Parent PFI/CI file with the Household & Library file, a match for every parent can be found that will bring together the Parent PFI/CI variables with the parents' related Household & Library variables. However, there are additional records in the Household & Library file unrelated to the Parent PFI/CI file, i.e., household information on respondents found in the Adult CI and Youth CI files. The example on **Linking between Parent PFI/CI and Youth CI files** demonstrates a technique for dropping unwanted records resulting from such a merge; in the SAS example note the use of the ONPARENT variable and in the SPSS examples the INPARENT recode.

**Linking between Parent PFI/CI and Youth CI files** is straightforward. The common identifier (key) is the ENUMID, and is available on both files. Sample SAS code to bring together these two files follows:

```
DATA TEMP;
MERGE parent_filename (IN=ONPARENT) youth_filename (IN=ONYOUTH);
BY ENUMID;
RUN;
```

Sample SPSS for DOS code is:

```
SET MORE = OFF

GET FILE = 'youth file'.
SORT CASES BY ENUMID.
SAVE OUTFILE = 'temp'.

GET FILE = 'parent file'.
SORT CASES BY ENUMID.

JOIN MATCH FILE = */FILE = 'temp'
  /BY ENUMID
  /MAP.
SAVE OUTFILE = 'merged file'.
```

Sample SPSS for Windows code is:

```
GET FILE = 'youth file'.
  /KEEP = ALL.
SORT CASES BY ENUMID.
SAVE OUTFILE = 'temp'.

GET FILE = 'parent file'.
  /KEEP = ALL.
SORT CASES BY ENUMID.

MATCH FILES FILE = */FILE = temp
  /BY ENUMID
```

```
/MAP.  
SAVE OUTFILE = 'merged file'.
```

**Linking between the Household & Library file and the Parent PFI/CI, Youth CI, or Adult CI files** requires using the key common to both files. This is accomplished using BASEID, which appears on all data files. The following example shows how to join selected library items from the Household & Library file with the Parent PFI/CI file and retain only records from the Parent PFI/CI file. Similarly, the Youth CI or Adult CI file can be substituted where references to the Parent PFI/CI file are made to allow joining library items with that file. Example SAS code to bring these files together follows:

```
DATA TEMP;  
MERGE parent_filename (IN=ONPARENT) household_and_library_filename (KEEP = BASEID LCOMP LCONSUME  
LDISTANC LJOBHELP);  
BY BASEID;  
IF ONPARENT;  
RUN;
```

Sample SPSS for DOS code is:

```
SET MORE = OFF
```

```
GET FILE = 'household & library file'  
/KEEP = BASEID LCOMP LCONSUME LDISTANC LJOBHELP.  
SORT CASES BY BASEID.  
SAVE OUTFILE = 'temp'.
```

```
GET FILE = 'parent file'.  
COMPUTE INPARENT = 1.  
SORT CASES BY BASEID.  
SAVE OUTFILE = 'temp2'.
```

```
JOIN MATCH TABLE='temp'  
/FILE='temp2'  
/BY BASEID  
/MAP.
```

```
SELECT IF (INPARENT = 1).  
SAVE OUTFILE = 'merged file'.
```

Sample SPSS for Windows code is:

```
GET FILE = 'household & library file'  
/KEEP = BASEID LCOMP LCONSUME LDISTANC LJOBHELP.  
SORT CASES BY BASEID.  
SAVE OUTFILE = 'temp'.
```

```
GET FILE = 'parent file'  
/KEEP = ALL.  
COMPUTE INPARENT = 1.  
SORT CASES BY BASEID.  
SAVE OUTFILE = 'temp2'.
```

```
MATCH FILES TABLE = 'temp'  
/FILE='temp2'  
/BY BASEID  
/MAP.
```

```
SELECT IF (INPARENT = 1).  
SAVE OUTFILE = 'merged file'.
```

**Linking between selected household member characteristics and the Parent PFI/CI file** requires the use of household member person numbers. (The same approach can be used to link person-level variables in the Household & Library file with the Youth CI and Adult CI data files.) As noted above, the household member person numbers on the Parent PFI/CI file are stored in the variables CHILDNUM, MOMNUM, DADNUM, and RESPNUM. On the Household & Library file, the characteristics of household members have been stored in arrayed fields that number from 1 to 16, for example MARITL1-MARITL16 gives the marital status of each household member. Household member numbers have been stored in the fields PNUM1-PNUM16 on the Household & Library file. To determine the marital status of the child's father, each PNUM(*n*) field must be compared to the value of the DADNUM field and the corresponding arrayed MARITL(*n*) field contains the father's marital status, for example, if DADNUM equals the value in PNUM5, then the MARITL5 field contains the father's marital status. Building on the preceding code, the following code demonstrates a way this can be accomplished in SAS. Note that characteristics can be determined for the child, mother, or interview respondent by substituting CHILDNUM, MOMNUM, or RESPNUM for references to DADNUM in the following code.

```
DATA TEMP;
MERGE parent_filename (IN=ONPARENT) household_and_library_filename (KEEP = BASEID PNUM1-PNUM16
MARITL1-MARITL16);
BY BASEID;
IF ONPARENT;
/*determine which element contains dads info and assign dads marital status to DADSMAR*/
IF DADNUM = PNUM1 THEN DADSMAR = MARITL1;
ELSE IF DADNUM = PNUM2 THEN DADSMAR = MARITL2;
ELSE IF DADNUM = PNUM3 THEN DADSMAR = MARITL3;
ELSE IF DADNUM = PNUM4 THEN DADSMAR = MARITL4;
ELSE IF DADNUM = PNUM5 THEN DADSMAR = MARITL5;
ELSE IF DADNUM = PNUM6 THEN DADSMAR = MARITL6;
ELSE IF DADNUM = PNUM7 THEN DADSMAR = MARITL7;
ELSE IF DADNUM = PNUM8 THEN DADSMAR = MARITL8;
ELSE IF DADNUM = PNUM9 THEN DADSMAR = MARITL9;
ELSE IF DADNUM = PNUM10 THEN DADSMAR = MARITL10;
ELSE IF DADNUM = PNUM11 THEN DADSMAR = MARITL11;
ELSE IF DADNUM = PNUM12 THEN DADSMAR = MARITL12;
ELSE IF DADNUM = PNUM13 THEN DADSMAR = MARITL13;
ELSE IF DADNUM = PNUM14 THEN DADSMAR = MARITL14;
ELSE IF DADNUM = PNUM15 THEN DADSMAR = MARITL15;
ELSE IF DADNUM = PNUM16 THEN DADSMAR = MARITL16;
RUN;
```

Sample SPSS for DOS code is:

```
SET MORE = OFF
GET FILE = 'household & library file'
/KEEP = BASEID PNUM1 TO PNUM16 MARITL1 TO MARITL16.
SORT CASES BY BASEID.
SAVE OUTFILE = 'temp'.

GET FILE = 'parent file'.
COMPUTE INPARENT = 1.
SORT CASES BY BASEID.
SAVE OUTFILE = 'temp2'.

JOIN MATCH TABLE='temp'
/FILE='temp2'
/BY BASEID
/MAP.

SELECT IF (INPARENT = 1).

IF (DADNUM = PNUM1) DADSMAR = MARITL1.
IF (DADNUM = PNUM2) DADSMAR = MARITL2.
IF (DADNUM = PNUM3) DADSMAR = MARITL3.
```

```
IF (DADNUM = PNUM4) DADSMAR = MARITL4.
IF (DADNUM = PNUM5) DADSMAR = MARITL5.
IF (DADNUM = PNUM6) DADSMAR = MARITL6.
IF (DADNUM = PNUM7) DADSMAR = MARITL7.
IF (DADNUM = PNUM8) DADSMAR = MARITL8.
IF (DADNUM = PNUM9) DADSMAR = MARITL9.
IF (DADNUM = PNUM10) DADSMAR = MARITL10.
IF (DADNUM = PNUM11) DADSMAR = MARITL11.
IF (DADNUM = PNUM12) DADSMAR = MARITL12.
IF (DADNUM = PNUM13) DADSMAR = MARITL13.
IF (DADNUM = PNUM14) DADSMAR = MARITL14.
IF (DADNUM = PNUM15) DADSMAR = MARITL15.
IF (DADNUM = PNUM16) DADSMAR = MARITL16.
```

```
SAVE OUTFILE = 'merged file'.
```

### Sample SPSS for Windows code is:

```
GET FILE = 'household & library file'
  /KEEP = BASEID PNUM1 TO PNUM16 MARITL1 TO MARITL16.
SORT CASES BY BASEID.
SAVE OUTFILE = 'temp'.
```

```
GET FILE = 'parent file'
  /KEEP = ALL.
COMPUTE INPARENT = 1.
SORT CASES BY BASEID.
SAVE OUTFILE = 'temp2'.
MATCH FILES TABLE = 'temp'
  /FILE='temp2'
  /BY BASEID
  /MAP.
```

```
SELECT IF (INPARENT = 1).
```

```
IF (DADNUM = PNUM1) DADSMAR = MARITL1.
IF (DADNUM = PNUM2) DADSMAR = MARITL2.
IF (DADNUM = PNUM3) DADSMAR = MARITL3.
IF (DADNUM = PNUM4) DADSMAR = MARITL4.
IF (DADNUM = PNUM5) DADSMAR = MARITL5.
IF (DADNUM = PNUM6) DADSMAR = MARITL6.
IF (DADNUM = PNUM7) DADSMAR = MARITL7.
IF (DADNUM = PNUM8) DADSMAR = MARITL8.
IF (DADNUM = PNUM9) DADSMAR = MARITL9.
IF (DADNUM = PNUM10) DADSMAR = MARITL10.
IF (DADNUM = PNUM11) DADSMAR = MARITL11.
IF (DADNUM = PNUM12) DADSMAR = MARITL12.
IF (DADNUM = PNUM13) DADSMAR = MARITL13.
IF (DADNUM = PNUM14) DADSMAR = MARITL14.
IF (DADNUM = PNUM15) DADSMAR = MARITL15.
IF (DADNUM = PNUM16) DADSMAR = MARITL16.
```

```
SAVE OUTFILE = 'merged file'.
```