

## 4.4 Event History

---

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

### Employment

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

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

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

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

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

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

**EMP\_DUAL\_JOB#.** If a respondent holds more than one employee job during a week, the ID number of the second employee job is presented in the dual jobs arrays. These arrays contain only the job number of the overlapping job; labor force status information is only included in the main array. For example, if a respondent held two employee jobs (e.g., the first and third jobs listed on the employer roster), during the 52<sup>nd</sup> week of 1997, the employer number for the first job is recorded in the EMP\_STATUS array and the employer number for the third job is recorded in the EMP\_DUAL\_2 array.

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

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

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

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

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

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

### **Marital Status**

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

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

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

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

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

### **Program Participation**

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

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

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

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

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

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

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

### **Schooling**

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

### Yearly Schooling Variables

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

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

**SCH\_MNTHS\_MISSED.** Not including summer vacation, this array presents the number of months during the school year that the respondent did not attend school.

**SCH\_SUMMER\_SCHOOL.** These variables show whether the respondent attended extra school classes, such as summer school, during an educational break in a given school year.

**SCH\_GRADE\_PROGRESS, SCH\_YEAR\_PROGRESS.** These arrays report whether the respondent was skipped ahead or demoted during a given grade in school or during a given school year.

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

### Monthly Schooling Variables

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

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

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

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

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

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

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

*Related User's Guide Sections*

4.2.2 Educational Status & Attainment

4.3.12 Work Experience

4.8.3 Program Participation

4.9.3 Marital & Marriage-Like Relationships

*Main Area of Interest*

Event History