National Longitudinal Surveys

A Program of the U.S. Bureau of Labor Statistics

National Longitudinal Survey of Youth | 1997
NLSY97

NLSY97	NLSY79	NLSY79 Children	Mature and Young Women	Older and Young Men
(/content/cohorts/nlsy97/topical-guide/employment)	(/content/cohorts/nlsy79/topical-guide/employment)	(/content/cohorts/nlsy79-children/topical-guide/employment)	(/content/cohorts/mature-and-young-women/topical-guide/employment)	(/content/cohorts/older-and-young-men/topical-guide/employment)

💼 Employment

Employment: An Introduction

Linking Job Information with Employers

To associate job information with the correct employer, researchers need to understand how employment information is collected during the interview. The following paragraphs describe how the data are gathered and how employers can be identified in different types of questions and across survey rounds.

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

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

Starting in round 2, the employer information was collected in a similar manner. Respondents reported all new employers since the last interview date in no particular order. As employers were reported, the CAPI program included a check for whether each employer had been reported in a previous interview. If the respondent reported a new employer, then the YEMP_UID.xx variables contain a new number, as shown in Figure 1. If the employer had been previously reported, the employer kept the same ID number (9701-9707 for round 1 employers, 9801-9809 for round 2 employers, and so on) as it had in previous rounds. This system permits users to link employers across survey rounds, even if there was a break in employment, and to identify the round in which an employer was first reported. After the ID numbers were either continued from a previous round or newly assigned, the roster was sorted according to the stop date of each job. Therefore, employers from different rounds may be mingled on the roster; previous round employers do **not** necessarily precede current round employers. Note that old employers for whom the respondent has not worked since the last interview do not appear on the current round's roster.

Figure 1. NLSY97 Unique Employer ID Numbers

Round	Maximum # Jobs	Unique ID # Range[1]
1	7	9701-9707
2	9	9801-9809
3	9	199901-199909

4	9	200001-200009, 199998, 199999
5	8	200101-200109, 200099
6	11	200201-200209, 200199
7	10	200301-200310
8	7	200401-200407
9	9	200501-200509
10	9	200601-200609
11	8	200701-200708
12	8	200801-200808
13	9	200901-200909
14	9	201001-201009
15	13	201101-201113
16	10	201301-201310
17	12	201501-201512
18	7	201701-201707
19	11	201901-201911
20	10	202101-202110

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

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

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

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

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

When the data are being prepared for public release, the original missing values are inserted into the employer roster. At this point the employer roster reflects the actual responses given during the interview and not the temporary imputed values. Therefore, researchers can use the original answers in their analyses. However, they may wish to know what imputed values were substituted so that they can follow the correct question paths and understand the respondent's answers. A complete, detailed explanation of the imputation process is contained in Appendix 6 (/content/cohorts/nlsy97/other-documentation/codebook-supplement/appendix-6-event-history-creation-and) in the *NLSY97 Codebook Supplement*.

Event history data. The created event history variables (see Employers & Jobs (/content/cohorts/nlsy97/topical-guide/employment/employers-jobs)) can be used in conjunction with the main file information about the respondent's employment. Like the main file variables, the event history variables use two systems of identification for a respondent's employers. First, the event history variables contained in the week-by-week status (e.g., EMP_STATUS_1997.01, where "01" indicates the first week of the year "97") and dual job (e.g., EMP_DUAL_2_1997.01) arrays use the unique ID numbers (UID) for each employer; to associate these employers with characteristic information collected during the interview, researchers must use the YEMP_UID.xx crosswalk variables. A second set of event history variables, those providing start and stop date information (e.g., EMP_START_WEEK_1997.01, EMP_END_WEEK_1997.01, where "01" indicates job #01), use the employer roster line numbers to identify the jobs. The number in the title of these variables refers to the same job as the variables in the main data set with the same number, so users can compare all information about job #02, for example, without any additional ID variables. However, to compare event history start and stop date information about job #02, for example, with information in the event history week-by-week status arrays, researchers must first use the YEMP_UID.xx crosswalk variables to identify the employer ID (9701-9707, 9801-9809, etc.) that matches job #02. See the example below to understand how this process works.