NLSY79 Round 19 Data Release

The main file round 19 NLSY79 CD-ROM is now available for distribution to researchers. This latest survey collected information from 8,033 members of the NLSY79 cohort, or almost 81 percent of the eligible respondents. The data are available on compact discs, and include information gathered during all of the 19 interviews conducted since 1979. Since 1994, NLSY79 respondents have been interviewed every other year. Round 19 encompasses the year 2000 interviews.

Also available is the round 19 geocode CD-ROM. This data set includes all of the main file variables and adds more-detailed geographic information. Due to confidentiality restrictions, access to the geocode data is limited; interested researchers should contact the Bureau of Labor Statistics (BLS) for more information. (Contact information may be found on the back page of this newsletter.)

The round 19 NLSY79 work history CD will be available later this spring. The work history data are a special set of created variables tracking respondents’ employment status from January 1, 1978, through the most recent interview date. Users of NLSY79 work history data should be aware that the upcoming distribution of work history CDs for round 19 of the NLSY79 will be different from that of past years. For this round, the work history CD also will include the main file data, and will use the same search and extraction software as the main data release. Work history variables will be located in the work history “area of interest,” but will otherwise be completely integrated into the main data set. (Areas of interest reflect variables that are grouped by common topics. The data CDs include a search function that allows users to view variables by area of interest.) Researchers interested in using work history data should be sure to request the combined work history/main file CD.

Round 19 questionnaire content

The round 19 questionnaire was similar to the instruments used in previous rounds. It asked each NLSY79 respondent about various areas of interest, including labor force behavior, educational attainment, training investments, income and assets, health conditions, workplace injuries, insurance coverage, alcohol and substance abuse, and marital and fertility histories. However, the 2000 questionnaire did include some changes, which are outlined below.

The family background section contained two changes. First, a migration history was added to collect respondent addresses for each move since the date of the last interview. On the main file data CD, researchers will be able to tell if respondents moved, but will not be able to access information on States and counties related to those moves. State and county for each place of residence is available on the NLSY79 geocode CD-ROM. Due to confidentiality restrictions, access to the geocode data is limited; interested researchers should contact BLS for more information. (Contact information may be found on the back page of this newsletter.)

Second, the family background section saw the addition of questions on respondent religious affiliation and attendance at religious services. Respondents were asked to state in what religion, if any, they were raised. Next, they were asked what religion, if any, they presently practiced. Finally, respondents indicated the frequency of their attendance at religious services in the year prior to the date of interview. These same questions were asked about the respondent’s spouse or partner in the marital history section. Religious affiliation and attendance questions were last asked in 1982.

Questions about spouse/partner labor force activity, including the wage module—a series of questions asking about usual earnings and hours worked—were moved in round 19 from the marital history section to a separate section. This section, entitled “spouse/partner labor force activity,” follows the training section.

In the training section of the questionnaire, a response category was added to the question, “Who paid for this training?” The new category allows for the possibility that the training was paid for through welfare or through the Temporary Assistance for Needy Families (TANF) program.

In the health section, several items were removed from the inventory of health conditions in the series addressed to respondents who had reached age 40 since the last interview.

A number of changes were made to the income, recipiency, and assets section for round 19. Most notable are the following: The series on respondent and spouse/partner unemployment, Supplemental Security Income (SSI) benefits, Food Stamp benefits, TANF/Aid to Families with Dependent Children (AFDC) benefits, and Welfare Reform all were streamlined; respondents falling within a rough range of eligibility were asked a new question to determine whether they had filed for an Earned Income Tax Credit in 1999; and several questions in the assets module concerning income-generating assets were modified and augmented to obtain better information on joint ownership of these assets, their valuation, and the income generated from them. Finally, in response to advances in the availability of technology,
a series of questions was added to solicit basic information about computer use and Internet access.

Deleted from the 2000 questionnaire were the drug supplement and the “Current Population Survey” (CPS) module. The CPS module has not been permanently removed from the survey, but has been assigned to recurring status. It will be included every two to three rounds. The CPS section was a snapshot of respondent labor force status and activity. The extensive longitudinal record collected by the survey makes it possible to isolate labor force status information for any week since the beginning of January 1978. Researchers use this information much more frequently than the CPS information.

Since 1994, all job-specific information has been collected in the employer supplements. What used to be termed the “CPS job” is now referred to as “current or most recent job.”

The main indicator of this change is the absence of the employment status recode (ESR) items. However, data users should be able to construct these items from other data. For round 19, the job search questions, which will give users information on a respondent’s job search activity in the preceding 4 weeks, are retained from the CPS module. Other information may be extracted from the employer supplements and gaps modules, which collect information on any periods between employment spells, or from the work history arrays.

The CPS module also contained a couple of questions that established whether anyone in the household owned a business. These questions were used in a few skips in the employer supplements and determined whether the question text contained reference to “working in family business.” Those skips are now determined on the basis of class of worker, where appropriate, or the references to working in a family business are included in the question text for the respondent to choose if applicable.

**NLSY79 data and documentation**

The NLSY79 main file data are available to researchers on CD-ROM. In addition to the data, each CD-ROM contains documentation and search and retrieval software. The NLSY79 CDs include the data collected from each of the 19 rounds of the survey, as well as a number of created variables. The inclusion of data from all rounds allows researchers to easily examine the longitudinal record of a respondent.

The *NLSY79 User’s Guide* is available to researchers at no cost. This guide explains the selection of the NLSY79 sample, describes the contents of the data set, and provides helpful information for researchers using the data. The 1999 version of this guide is currently available on the World Wide Web at [http://www.bls.gov/nls/nlsdoc.htm](http://www.bls.gov/nls/nlsdoc.htm). An updated 2001 edition, containing information about round 19, will be available later this spring.

Researchers interested in purchasing the main file data or any accompanying documentation should contact NLS User Services. (See the back of this newsletter for contact information.)

### Changes to NLSY97

#### Household Roster Data

Eliminate Inconsistencies Discovered in Round 1

Inconsistencies in the household and non-resident rosters from round 1 of the NLSY97 resulted in corrections being made to these data. The changes were made largely in response to the discovery of duplicate persons on and across the household and nonresident rosters. The changes also aimed to eliminate inconsistent sets of information, such as respondents’ being reported as having more than one biological mother.

Users should be aware that the vast majority of the changes involved simply recording missing information from a zero to a valid skip (−4). In most of these cases, various household members were ineligible for certain questions because of their age or because they were sample members who provided information requested in the screener, such as highest grade completed, elsewhere in the survey. At the time the screener was administered, this missing information often was set to zero instead of a −4, valid skip. Changes made to large numbers of cases reflect these types of data cleaning activities, and in no way reflect upon the quality of the data from the survey.

The sections below highlight some of the changes made to the round 1 through 3 rosters between the release of the round 3 main file data and the round 3 event history data. Changes resulting from the zero to −4 conversions were filtered out so that users can focus on the more substantive corrections. Researchers using household and nonresident roster data should obtain the most recent data release before finalizing their analyses. Users interested in viewing a complete list of corrections and changes to the NLSY97 data, including changes to variables not mentioned in this article, should visit [http://www.bls.gov/nls](http://www.bls.gov/nls) and follow the “NLSY97” link for errata and other cohort-specific information.

#### Changes to the household roster

Most cases of invalid duplication occurred on the nonresident roster, but 43 variables in the household roster also were affected. Some of these affected variables required changes to only a handful of cases. As mentioned earlier, the majority of the changes reflect a recoding from a zero to a valid skip (−4). The variables outlined in this article are those requiring changes other than the recoding. They are concentrated in one main area involving the marital status variable.

This variable represents information on whether household members were never-married, married, separated, divorced, or widowed. The largest number of discrepancies for this roster item was for household member number 1 (HH12_/MARSTAT.01), with 92 cases. Household member number 2 (HH12_/MARSTAT.02) had 11 affected cases; number 3 (HH12_/MARSTAT.03) had 2 cases; number 4 (HH12_/MARSTAT.04) had 2 cases; number 5 (HH12_/MARSTAT.05) had 7 cases; and number 6 (HH12_/MARSTAT.06) had 2 cases.

Researchers should make note of these changes because of the importance of roster information in the NLSY97. For some variables, the roster information may be more accurate, because some rosters are updated during the interview if the initial report was inaccurate. When survey staff prepare the data for release, they clean up the rosters if necessary but do not necessarily clean the corresponding raw data. Whenever a roster variable and a regular question appear to store the same information, the roster variable always should be used. Starting with the most recent release of data, many of the raw question variables have been deleted from the data when a corresponding roster variable exists. This
Changes to the nonresident roster

The majority of the changes to the NLSY97 rosters were made to the nonresident roster. The changes were intended to achieve three main purposes:

- to eliminate duplicate persons from the roster
- to correct information that was wrong in earlier versions
- to add information that was previously missing from the roster but that did exist in the original "raw" screener questions.

On the nonresident roster, 476 variables were affected. As in the household roster, however, most of the variables required changes to only a handful of cases, and the majority of the changes to the nonresident roster were grouped among a few main sets of variables. Again, the zero to -4 conversions have been filtered out so that the numbers of cases below reflect other types of changes.

Roster items NONHHI_RACE.01 through NONHHI_RACE.18, which report, for 1997, the race of nonresident members, had a combined 773 cases affected. Items NONHHI_RACE.01 through NONHHI_RACE.16, which provide the same information for 1998, had 658 cases requiring changes.

Similarly, the roster items reporting nonresident ethnicity based on the question “Is nonresident member Hispanic?” were affected for both the 1997 (questions NONHHI_ETHNICITY.01 through NONHHI_ETHNICITY.18, 168 cases) and 1998 (questions NONHHI_ETHNICITY.01 through NONHHI_ETHNICITY.16, 180 cases) fieldings.

Variables NONHHI_HIGHGRADE.01 through NONHHI_HIGHGRADE.18 list the highest grade completed by each nonresident member. A total of 1,505 cases from 1997 needed cleaning on this series.

Employment status of nonresident members in 1997 was recorded in roster items NONHHI_EMPLOYED.01 through NONHHI_EMPLOYED.16, with 702 cases affected.

In 1998, NONHHI_RELY.01 through NONHHI_RELY.16 sought the nonresident’s relationship to the respondent. For these items, 1,056 cases were changed.

Gender was reported in NONHHI_SEX.01 through NONHHI_SEX.19. For these 19 roster items using 1998 data, there were 1,075 changed cases.

Items NONHHI_HHFLAG.01 through NONHHI_HHFLAG.22 checked the status of nonresidents in 1998 by noting whether each had been a household resident in round 1, a nonresident in round 1, or newly reported in 1998. Of these, 1,562 cases were checked and cleaned.

The age of nonresidents section also required some adjustment. There were 1,034 cases fixed for items NONHHI_AGE.01 through NONHHI_AGE.26 in the 1998 data.

Marital status and UID of nonresidents were the final major roster groups affected. The marital status items for 1998, NONHHI_MARSTAT.01 through NONHHI_MARSTAT.23, had 188 changed cases, while the UID variables from 1998, NONHHI_UID.01 through NONHHI_UID.23, had 1,572 cases.

Created variables affected by changes to the household roster and nonresident roster

Updates to the household and nonresident rosters led to changes in eight created variables. Listed below are the affected variables and the number of cases. The corrected data are available on the round 3 event history CD.

- CV_BIO_MOM_AGE_CHILD1, age of biological mother at first birth, 17 cases
- CV_BIO_MOM_AGE_YOUTH, age of biological mother when respondent was born, 6 cases
- CV_YTH_REL_HH_CURRENT, respondent’s relationship to household parent figure 1997, 4 cases
- CV_HGC_BIO_DAD, biological father’s highest grade completed, 466 cases
- CV_HGC_BIO_MOM, biological mother’s highest grade completed, 93 cases
- CV_HGC_RES_DAD, residential father’s highest grade completed, 2 cases
- CV_YTH_REL_HH_CURRENT, respondent’s relationship to household parent figure 1998, 40 cases
- CV_YTH_REL_HH_CURRENT, respondent’s relationship to household parent figure 1999, 72 cases

Methodological Experiments in the NLS

Throughout the years of the NLS, survey administrators have performed several methodological experiments that offer researchers a number of opportunities to investigate the outcomes of differing survey approaches. This article looks at many of these experiments.

Researchers interested in looking at the results of these experiments in-depth are invited to claim a topic and write an NLS user’s guide section on it. Following review by survey staff, these sections will be published in the guide. This opportunity is offered on a first-come, first-serve basis. Researchers are encouraged to contact NLS User Services early to claim a topic. User Services also can answer questions about these experiments.

PAPI to CAPI

The switch in the administration of the NLSY79 from a paper-and-pencil interview (PAPI) to computer-assisted personal interviews (CAPI) provides a number of opportunities for comparison. The gradual implementation of CAPI includes the 1989
fielding of the NLSY79, in which 600 Ohio cases were randomly split between PAPI and CAPI administration; and the 1990 fielding of the NLSY79, with one-quarter of the surveys administered via CAPI and three-quarters administered with PAPI. The 1993 survey of the NLSY79 marked the first time that computer-assisted interviews were administered to the full sample of NLSY79 respondents. Computer-assisted interviewing software also has been used in the interviews with the NLS of mature women and young women since 1995, as well as in all rounds of the NLSY97. While many papers already have been written on this switch from PAPI to CAPI, there is still a wealth of areas open to investigation. Researchers interested in reviewing the papers already written on this transition are encouraged to read:


Additional papers may be found at http://www.nlsbibliography.org.

The NLSY79 recall experiment
The NLSY79 was switched from an annual to a biennial fielding schedule beginning with round 17. Prior to that, respondents were asked to recall information about their labor force and educational experiences, health and disability, marital status, and other topics only for the past year. With the start of the new schedule, respondents would be asked to recall 2 years’ worth of information. The recall experiment in round 16 was aimed at learning more about the possible effects that an extended gap between surveys might have on recall.

During the round 16 survey, 900 NLSY79 sample members, who were interviewed in round 15, were randomly chosen to be interviewed about the period since their round 14 interviews, which had occurred approximately 2 years earlier. During their round 16 interviews, these respondents were treated as if their round 15 interviews had never occurred. Their responses in round 16 were compared with the existing data these respondents had reported for the same period in the round 15 interview with only a 1-year recall.

The recall experiment encompassed all retrospective information collected in the interview. In the paper, “Event History Data and Survey Recall: An Analysis of the National Longitudinal Survey of Youth 1979 Recall Experiment” (Journal of Human Resources 36,3 (Summer 2001): 439-466), Charles R. Pierre focused on several labor force and recipiency variables included in the experiment. He found that the longer recall period did produce data of lesser quality in some circumstances. Respondents often failed to report short spells of recipiency, employment, and non-employment when faced with the longer recall period.

Incentive fee experiment
In round 4 of the NLSY79, survey administrators offered different levels of incentives to survey respondents in an effort to study the effects of incentive level on survey participation. In the NLSY79 experiment, three levels of compensation ($10, $15, and $20) were offered. In addition, half of the respondents at each dollar level were paid in advance, and half were paid upon completion of the interview.

In “Evaluation of a Monetary Incentive Payment Experiment in the National Longitudinal Survey of Youth, 1997 Cohort,” researchers A. Rupa Datta, Michael W. Horrigan, and James R. Walker examined the impact of response rates for differing levels of incentive and differing categories of panel members. They did not find “any discernable evidence of a strong and consistent impact from either the timing or the amount of the incentive payment on response rates.” Researchers are encouraged to use the administration of this incentive experiment as a springboard to other findings.

Split administration of the drug supplement in the NLSY79
In 1988, the administration of the drug supplement in the NLSY79 was split, with respondents being randomly sent down one of two paths. On one path, respondents self-administered the supplement. On the other path, the interviewer administered the supplement. One opportunity this split administration offers researchers is the chance to examine any differences that might occur between the self-administration and interviewer administration of sensitive questions.

Presence of others during interview
A series of questions across the years in the mature and young women surveys has asked whether another person was present during all or part of the interview.

In 1967, a question to the mature women cohort asked if another person was present when the respondent answered the marital status and family questions (R00220.), and if that person influenced the respondent’s answers (R00221.). This portion of the survey also asked whether another person was present when the respondent answered non-labor-market-activity questions (R00350.) and, if so, whether that person influenced the respondent’s answers (R00351.).

In 1968, the young women survey included questions on whether someone else was present when the respondent answered questions about education (R00157.), work attitudes (R00322.), expected occupation at age 35 (R00336.), and assets and income (R00375.). In 1969, the young women’s survey asked if another person was present when the respondent answered knowledge of work world questions (R01115.). In 1993, the question asked whether the respondent’s husband/partner was present during the interview (R15697.).
other person during the interview and whether that person was the respondent’s spouse/partner were asked of both women’s cohorts in 1995 (R34893., R34895.), 1997 (R42333., R42335.), and 1999 (R54040.).

Similar questions about the presence of others have been asked in most surveys of the NLSY79. These questions are answered by the interviewer at the end of the entire interview and also after sensitive sections, including those concerning fertility and income.

**Addition of neutral category in young women job satisfaction question**

The young women sample was split in 1993 on a question about the respondent’s attitude toward her current or last job. Approximately half of the interviewed sample chose from the following answers: They liked their job very much, liked it fairly well, disliked it somewhat, or disliked it very much (R15782.). The other part of the sample answered the same question but had an added, neutral-response category that allowed them to say they neither liked nor disliked their job (R15781.).

**Transfer questions in the mature and young women**

The 1997 surveys of mature and young women contain several questions about transfers of time and money to and from mothers and daughters. Administered only to respondents with a mother or daughter in the other women’s cohort, these questions give researchers the unique ability to compare mothers’ perceptions about their level of receipt or giving with the perceptions of their daughters.

**NLSY79 young adult dating/marijuana questions at time of survey redesign**

When the young adult survey was redesigned for 2000, survey administrators deliberately re-asked two series of questions. One series concerned whether the respondent had begun dating and the age at which dating had begun. The other series concerned whether the respondent had ever smoked marijuana and the age at which the respondent first did so. The questions were re-asked for two main reasons: 1) The survey was switching from primarily in-person to primarily telephone as mode of interview, and 2) the self-report booklet had been integrated into the CAPI instrument so that it was no longer a paper-and-pencil supplement.

In addition to the questions asked above, the young adult survey redesign in 2000 offers ample potential for analyzing mode effects due to the combination of switching from primarily in-person to primarily telephone administration and having the sensitive questions moved from PAPI to CATI/CASI. The self-report areas in 2000 were designed to be interviewer-administered for telephone interviews and self-administered for in-person interviews.

**Religious affiliation and attendance**

Round 19 of the NLSY79 saw the addition of questions on respondent religious affiliation and attendance at religious services. Respondents were asked to state in what religion, if any, they were raised. Next they were asked what religion, if any, they presently practiced. These attendance and affiliation questions were last asked in 1982. They also had been asked in round 1 of the survey. This is an interesting area for investigation because of the number of people who changed answers to the question on the denomination in which they were raised between rounds 1 and 19.

**Interviewer characteristics**

Every NLSY79 CD-ROM contains an interviewer characteristic database, which shows the age, race, and sex of the interviewers and indicates their pay rates, experience, and education. Information on how many times a particular interviewer has administered the survey to each respondent also is available. The wealth of information in this untapped resource provides almost unlimited opportunity for research.

Researchers wanting to explore any of these experiments in detail and to write a user guide chapter are encouraged to contact NLS User Services to claim their topic. (Contact information is provided on the back of this newsletter.)

**Frequently Asked Questions**

NLS User Services encourages researchers to contact them with questions and problems they have encountered while accessing and using NLS data and/or documentation. Every effort is made to answer these inquiries. Some recently asked questions that may be of general interest to NLS users are listed below with their answers.

**Q1:** In the NLSY79, I found that there are zeros for the 1996 and 1998 county and State FIPS codes. Are they missing values or something else?

A1: A code of zero in either the county or State variable means the respondent is not living in the United States. Specifically:

- If the county code is 0, then the State code variable represents residence in a U.S. territory or other outlying area. For example, if a respondent has a county code of 0 but a State code of 72, that respondent resides in Puerto Rico (State code 72).
- If the State code is 0, then the county code variable represents residence in another country. If a respondent has a State code of 0, for example, but a county code of 272, that respondent resides in Germany.

**Q2:** In the NLSY79 data, is the “total amount AFDC, food stamps, or other welfare/SSI received during the calendar year 1980” for the family or for the respondent?

A2: Question R04993.14 is the total received by the respondent and his/her spouse. Receipt by other family members is not included in this variable.

**Q3:** This is in regard to the measurement of height and weight for the children of the NLSY79 in 1986. In other survey years, there is recorded information regarding whether the child was physically measured or whether the mother reported the child’s height/weight. In 1986, there is no such information in the data. I would assume this is because in 1986 everyone was physically measured. Is this true?

A3: The 1986 Child Supplement did not contain a question asking the interviewer to code how the measurements were taken. So, for 1986 it is not known how many measurements were physically taken and how many were reported by the mother.
Q4: Were questions about wife’s education and age asked in 1973 and 1975 in the NLS of Young Men? I looked for these spousal characteristics in these years and have been unable to find the information.

A4: The 1973 and 1975 young men interviews were conducted by telephone, as opposed to a personal interview. In general, telephone interviews (which occurred for years 1973, 1975, 1978, and 1980) were considerably shorter and asked fewer questions. You should be able to get the wife’s age by locating the line on the household record card for the wife and getting her age from that.

Q5: I am interested in obtaining data about the educational attainment of the parents of NLSY97 respondents. I have consulted the codebook, and I am confused as to which variables contain parental education information. Which are the correct variables to use?

A5: As of the round 3 event history data release, there are created variables that contain the educational attainment of respondents’ biological parents (CV_HGC_BIO_MOM and CV_HGC_BIO_DAD) and residential parents (CV_HGC_RES_MOM and CV_HGC_RES_DAD). For cases in which respondents live with their biological parent or parents, these data will overlap.

Completed NLS Research

The following is a listing of recent research based on data from the NLS cohorts that has not appeared in its current form in a previous issue of the NLS News. See the NLS Annotated Bibliography at www.nlsbibliography.org for a comprehensive listing of NLS-related research.


Parcel, Toby L. and Dufur, Mikaela J. “Capital at Home and at School: Effects on Student Achievement.” Social Forces 79,3 (March 2001): 881-911. [NLSY79 Children]


Are You Working With NLS Data?

If you are, we are interested in your work!

- Have you received funding to sponsor a project using NLS data?
- Are you working on a paper that uses NLS data?
- Have you published a recent paper using NLS data?

If you have received funding on a project, are working on a paper, or published a recent paper that uses NLS data, please contact: NLS User Services, Center for Human Resource Research, 921 Chatham Lane, Suite 100, Columbus, OH 43221; (614) 442-7366; e-mail: usersvc@postoffice.chrr.osu.edu.
NLS News is published quarterly by the Bureau of Labor Statistics. It is distributed both nationwide and abroad without charge to researchers using NLS data, as well as to other interested persons.

NLS Contact Information

NLS User Services: Center for Human Resource Research
The Ohio State University
921 Chatham Lane, Suite 100
Columbus, Ohio 43221-2418
usersvc@postoffice.chrr.ohio-state.edu
(614) 442-7366
(614) 442-7329 (Fax)

NLS Program Office: National Longitudinal Surveys
2 Massachusetts Avenue, NE
Room 4945
Washington, DC 20212-0001
Attention: Rita Jain
NLS_INFO@bls.gov
(202) 691-7405
(202) 691-6425 (Fax)

NLS documentation, data, and data updates: usersvc@postoffice.chrr.ohio-state.edu

NLS web site: http://www.bls.gov/nls

BLS-NLS publications: Jain_Rita@bls.gov
(202) 691-7405

NLS News Editor: Donna S. Rothstein
Rothstein_D@bls.gov

NLS Program Director and Media Contact: Charles R. Pierret
Pieirut_C@bls.gov
(202) 691-7519