

U.S. Department of Labor  
U.S. Bureau of Labor Statistics

## Beginning in 2009: NLS News to Be Online Only!

NLS now offers timely e-mail alerts about changes to the NLS program. Approximately every 2 months, users can receive messages alerting them to updates of NLS products, including the *NLS News* newsletter, user conferences, various data, and important NLS publications. To subscribe, use the form at [www.bls.gov/nls/nls\\_subscribe.htm](http://www.bls.gov/nls/nls_subscribe.htm). □

## 2008 Data Release Schedule

Data releases for all of the active NLS surveys are scheduled in 2008. Data users can expect the following release timetable:

**NLSY79 Main Survey:** The round-22 data set (merged with data sets from the previous rounds) is scheduled for release in spring 2008.

**NLSY79 Child/Young Adult Survey:** The next merged child–young adult round, which will include 1986–2006 data, is scheduled for release in spring 2008, after the release of the NLSY79 main survey.

**NLSY97:** Round 10, merged with previous rounds, is scheduled for release in late summer or early fall of 2008.

Data for all cohorts are available at [www.nlsinfo.org/web-investigator](http://www.nlsinfo.org/web-investigator). □

## Topic Spotlight: Childcare Variables in the NLS

Childcare issues continue to play a major role in labor force participation decisions. The NLS project began collecting infor-

mation on childcare in 1967, providing researchers with historical data as well as current trends. In addition, the NLSY79 and the NLSY Child cohort connection allow a special opportunity to tie maternal childcare and employment decisions to child outcomes. Following is an overview of the childcare variables available in the NLSY79 Main Youth, NLSY79 Child, NLSY79 Young Adult, NLSY97, and Original Cohort surveys.

### NLSY97 Childcare Variables

Childcare usage questions were asked of female NLSY97 respondents in rounds 5 through round 9 (the most currently released round), with round 5 providing a more extensive set of childcare variables specific to each child of the respondent and rounds 6 through 8 providing a shorter list of variables on the childcare arrangements for all children grouped together. In round 9, female respondents born in 1983 and 1984 repeated the longer version found in round 5 and the remaining eligible females received the shorter version.

In round 5 and again in round 9, eligible female respondents provided childcare details for each of their children (all children under the age of 13 in round 5, all children under 7 in round 9). Respondents answered questions about the primary care arrangements (care by a relative, family daycare, care in a childcare center, and so forth) they made for each child during the hours the respondents were at work, in school, in training, or traveling to one of those activities. Information on secondary childcare arrangements also was gathered. If a child was cared for by a relative, the respondent answered a series of questions about (1) the amount of time each relative gave to the care and the childcare location (the child's home, a relative's home, and so on); (2) who provided transportation to

and from the relative's house; (3) the travel time to and from the childcare location; (4) whether the respondent was charged for the relative's help and, if so, the amount charged; and (5) whether an employer, outside agency, or anyone outside the household helped pay for childcare and, if so, the amount of aid given. Similar questions were asked about care by nonrelatives. Users should note that each care-by-relative question is asked for each child about each relative identified as a provider of childcare, so the codebook yields multiple variable titles that are identical except for the child number (Child 01, 02, etc.) and the relative identifier (Relative 01, 02, etc.).

The shorter series of childcare usage questions asked in rounds 6 through 8 (and asked in round 9 of female respondents born during the period 1980–82) included questions about the type of care arrangements for all children together, the cost of the childcare, and the details of any assistance received for childcare (from the Temporary Assistance for Needy Families (TANF) program, the employer, and so forth). These questions were asked separately by type of care, but grouped together all children receiving a given type of care.

In rounds 5 and 9, female respondents born in 1983 or 1984 who had no children and all women with children who did not use relatives as a source of child care were asked a series of questions related to relatives and childcare. These respondents reported the number of relatives within 15 minutes and within 45 minutes of the respondent's home and also reported on the possibility of these relatives caring for the respondent's children (whether actual or hypothetical). Respondents then answered questions on the amount, if any, of pay a relative would require to provide childcare.

The section on extended childcare in

rounds 5 and 9 also asked these women how much they would expect to pay for someone to come to their home to care for their child and how much they would be willing to pay for that care. Similar questions were asked about family daycare and childcare centers. For both of these types of care, respondents reported the travel time to the closest center, the amount they would expect to pay, and the amount they would be willing to pay to use the services.

Beginning in round 10, scheduled for release later this year, childcare usage questions also were asked of male respondents with resident children ages 7 and under.

#### NLSY79 Childcare Variables

For the NLSY79, the main data collection on childcare arrangements occurred during the 1982–86, 1988, 1992, and 1994–2004 surveys. (The 2006 survey, scheduled for release this year, also will include childcare variables.) In addition, a limited number of childcare questions, fielded within the 1987, 1989, and 1990 fertility series, obtained information from select universes on whether respondents made use of a regular childcare arrangement and whether respondents encountered childcare problems that affected their employment. Information on company-provided childcare is available in the fringe benefits section of the survey. There also was a special experimental *Childcare Supplement*, administered to 347 NLSY79 mothers who were interviewed during the first month of the 1989 fielding, that collected a wide range of information, including data on every childcare arrangement used for at least 10 hours per week since the date of previous interview. Because childcare variables in the NLSY79 cohort have varied a great deal between survey years, users are encouraged to study the *NLSY79 Users Guide* for a more detailed understanding of these variables.

Data on types and locations of childcare arrangements are available for the 1982–1986, 1988, 1992, and 1994–2006 NLSY79 survey years. In addition, during certain survey years, supplementary information was collected on the number of hours that childcare services were required or provided, the nature of the payments (cash or noncash), the total cost per child per provider, and the effect of available childcare services on such activities as

employment, job searches, training, and so forth.

Universes of respondents vary widely both within and across survey years, from respondents—both male and female—engaged in some educational or labor market pursuit, to all women with a child in the household, to not-employed respondents with an employed spouse. The focus during the initial survey years was on collecting information on childcare arrangements used over the previous month for only the youngest child(ren) in the household. In 1986 and 1988, the previous-month time-frame was extended to all children in the household. In addition, retrospective data were gathered during these same 2 interview years, and in 1992 and 1994–2006 retrospective data were collected on up to three childcare arrangements used by NLSY79 mothers during each of the first 3 years of the child’s life.

Typical categories of childcare arrangements include self-care; care by relatives (the child’s other parent, a stepparent, siblings, or grandparents); care by non-relatives; and care provided by institutions such as daycare centers, nurseries, or preschools. Information on care provided by individuals usually differentiates between that occurring in the child’s home and that occurring in another private home. Information was gathered during select survey years (1983–1986 and 1988) on both primary and secondary childcare arrangements.

NLSY79 data on childcare are unique because they can be linked with the NLSY79 child survey. Retrospective information on childcare arrangements for the first 3 years of each child’s life appears in the child data files (C03564.–C03590.), while the other cross-sectional childcare information can be extracted from the main Youth files and merged. (See the section of the *NLSY79 Child and Young Adult Data Users Guide* titled “Child Care” for more information.) The NLSY79 main survey and the NLSY79 child survey together provide researchers with the opportunity to tie a parent’s childcare and employment decisions, documented in the NLSY79 surveys, directly to child outcomes, recorded in the NLSY79 child surveys.

To extend this knowledge even further, beginning in 1994 questions on childcare are also contained in the NLSY79 young adult surveys, which focus on the older

children of NLSY79 mothers. These questions allow researchers to trace childcare across three generations.

#### NLSY79 Young Adult Cohort

##### Childcare Variables

From 1994 to 1998, young adult respondents answered questions about childcare for each of their children living with them, including questions about the type of childcare arrangement, the location of the care, the usual hours per week, and the usual cost per week. In 2000 and 2002, the questions were asked only about the youngest child in the household. A limited number of questions about general childcare were asked in the 2004–06 survey years. For the 1994–2006 survey years, respondents were asked whether they or their partner had lost time from work or had to turn down a job offer because of difficulties in childcare arrangements.

##### Original Cohort Childcare Variables

The first mature women’s survey, in 1967, included questions about whether childcare was used and, if so, the location and cost of the care. The survey continued to ask these questions in 1969, 1971, 1972, and 1977. The 1971 and 1972 surveys also asked about the number of hours per week that childcare was required. A final, extended series of questions on childcare issues was asked of the mature women in 1995.

Childcare questions in the young women’s survey began in 1968 and continued regularly through 1995. Basic information obtained each survey year included information on the location of care and details about the provider of the care. In select survey years, respondents also provided information about the cost of care and the hours of care required, with varying reference periods. In some years, the survey collected data on only the youngest child; in others, several children in different age groups were included. In 1983, 1991, 1993, and 1995, young women also were asked about missing work because of childcare issues.

The 1971 survey of the young and mature women’s cohorts included an extended series of childcare questions that asked respondents their attitudes toward their current childcare arrangements and their preferences regarding what was their ideal arrangement. Unemployed respondents also provided information about what their

likely childcare arrangements would be if they took a job.

In 1999, both young women and mature women respondents answered questions about time spent the previous year helping with the care of grandchildren. These questions were part of a series of questions about intergenerational transfers of time and money.

#### Accessing Childcare Variables

Childcare variables can be accessed by type of cohort through the NLS Web Investigator, available at [www.nlsinfo.org/web-investigator](http://www.nlsinfo.org/web-investigator). Use the “area of interest” search option to search “Child Care.” For the young adult cohort, users should search “YA Child Care.” □

### The Survey Round: From Development to Release

What steps take place prior to the release of an NLS data round? This article provides survey details that may be invisible to the average user, but that are necessary in any successful NLS data release.

#### Development and approval of survey content:

As part of the development process, the previous round’s survey instrument is evaluated to decide what updates will be made for the next round. Although many core questions remain constant from survey year to survey year, some series of questions are asked on a rotating basis. Other questions are added or retired on the basis of the maturation of the cohort. If new questions were added in the previous round and they proved problematic (respondents misunderstood a question or there was a high “don’t know” or refusal rate, for example), these questions would be reevaluated and possibly removed or reworded. In addition, as different issues become more pertinent in the changing U.S. society, new series of questions might be considered. Current examples of questions addressing such issues are bankruptcy questions and questions about computer usage.

Whatever the changes might be, a balance must be achieved: the survey developers try to gather as much significant longitudinal information as possible each round, while still keeping the respondent burden at a reasonable level. At times, researchers have contacted NLS User Serv-

ices to ask why a particular question that would be helpful to their research area was never asked. Often, their question was in fact considered, but space may have been too limited for its insertion. While not every good question may appear in a survey for a particular year, users are encouraged to send NLS User Services their ideas on the content of questions, because such ideas provide a barometer for user interest.

Once a draft of the newly revised questionnaire is agreed upon, the draft must be submitted to, and approved by, the Office of Management and Budget (OMB) and Institutional Review Boards (IRBs). OMB, an office within the Executive Office of the President of the United States, evaluates all Federal surveys. An IRB reviews the questionnaire’s content and associated data collection procedures to ensure that adequate protections for human subjects are in place.

#### Survey strategy and computerization:

The NLSY surveys are complex, because not every question in a survey round is asked of every respondent; some questions may instead be asked of subsets of respondents, such as females only or those who are employed. There are also several looped sections of the questionnaire, whereby a respondent may go through certain sets of questions more than once. For example, if a respondent reports having two jobs since the last interview, he or she will go twice through the questions asking about employer details. Another situation adding to the survey’s complexity occurs when a respondent may have skipped one or more interviews. If the respondent’s previous interview did not take place in the last survey round fielded, additional or differently worded survey questions may be used to recover or fill in information not gathered since his or her last interview. A key feature of the NLSY interviews is *bounded interviewing*, which uses information from earlier rounds to refresh the respondent’s memory and to increase the likelihood that ongoing events are reported consistently. All of these factors add up to a survey that takes time and effort to program correctly.

In the early years of the NLS project, paper questionnaires for new survey rounds were literally cut and pasted, with skips determined by hand. Since the mid-1990s, the process has been computerized and a cen-

tral, integrated database now drives both the execution of the instrument and the creation of documentation, which together make creating a survey a more efficient process. However, it still takes considerable labor to program new questions and the proper skip patterns.

#### Survey pretesting:

Before going to the field, a new survey is pretested so that what appears to be correct and viable on the computer monitor actually works in practice. Pretesting is done in layers. First, the survey is tested with a software utility. Then it is tested by staff who are familiar with the instrument. Finally, a formal pretest is conducted with a special sample of individuals. Computer programming errors may be caught at this time. Sometimes a new question will need to be reworded slightly because pretest respondents were not consistent in their interpretation of the question or had difficulty with it in some way.

#### Interviewer training:

When a survey has been approved and programmed, NLS interviewers receive training on the new instrument. The NLS project retains many of its interviewers from previous survey years, but in each round there are new interviewers who need fundamental training in survey and administrative procedures. All interviewers receive training on the new instrument.

#### Fielding the Interview:

Over a period of several months, interviewers make efforts to contact eligible respondents and interview them in person or by phone. For the NLSY79, the field period starts with a phone effort and is followed up with a more traditional face-to-face effort. For the NLSY97, most cases are completed face-to-face. Sometimes, respondents prove difficult to contact due to locational problems (the respondent may have moved, or his or her phone number may have changed), reluctant gatekeepers (household members may refuse the interview on behalf of the respondent, or they may be reluctant to pass along messages), availability issues (the respondent may be at work during the hours the interviewer seeks an interview), or changes in the respondents’ lives (the respondent may come down with an illness, for instance). These circumstances make it more of a challenge for an interviewer to take

place. Frequently, it takes many attempts to complete interviews with each of the respondents in the sample, resulting in a field period that lasts several months.

**Data cleaning:** Within each survey round of each cohort are thousands of data variables that show the direct responses of the respondents. Generally, most variables are released to the public without any postinterview editing, because the computer-generated survey precisely controls branching of the questionnaire and restricts the range of valid answers, greatly minimizing the need for data cleaning. Occasionally, however, data archivists find it necessary to delete or change the answers originally recorded in the data file. The two main reasons for cleaning data are instrument error, which can take the form of either design errors or technical errors, and interviewer or respondent confusion, which can lead to the wrong information being collected or the correct information being collected in the wrong part of the survey. As an example of a design error, a group of respondents may be asked a set of questions that they are supposed to skip. When this happens, the archivist will generally remove the answers to those questions from the data file prior to its release. As another example, respondents might list multiple periods working for the same employer in separate employer supplements, even though they are supposed to report the information in only a single supplement. Data archivists will then combine the answers into a single employer supplement so that users of the data will recognize that only one employer is being reported. These two types of problems are relatively rare and generally affect only a very small percentage of the answers that are collected.

Another major data-cleaning task is ensuring that no identifying data are released as part of the data file. For this purpose, verbatim information, such as information about the respondent's occupation and industry, is released in coded form rather than as reported by the respondents themselves.

**Variable creation:** In addition to the direct-response variables, variables are constructed or created on the basis of responses to more than one data item. For example, a created wage variable (CV\_HRLY\_PAY.xx in the NLSY97, for

instance) combines information on the respondent's rate of pay (such as whether it is hourly, monthly, or yearly), amount earned, number of hours worked per week, and number of weeks worked per year (when the salary is reported on an annual basis) into an hourly wage. This variable is created for each job the respondent reports. Because variables created for income or assets may contain identifying information on the respondent's annual salary (or the value of his or her house) when viewed in conjunction with the respondent's other characteristics, many created variables are topcoded to prevent that information from being revealed. The NLSY97 variable CV\_INCOME\_GROSS\_YR, for example, provides a single numerical value that represents the total income (the sum of the wage or salary income, business income, interest income, and any other income) reported by the respondent in separate questions during the survey round. After this variable is created, the amounts reported by the 2 percent of respondents with the highest income are averaged, and that mean is reported in the created income variable for those respondents. The asset values reported for respondents at age 25 also are topcoded, although a different method is used. Topcodes for these variables (for instance, CVC\_HH\_NET\_WORTH\_25, CVC\_ASSETS\_FINANCIAL\_25, and CVC\_ASSETS\_DEBTS\_25) were determined from asset values reported in other surveys. These topcodes were applied to the variables created for assets and remain stable from round to round.

Beginning in the 2000 survey year, respondents were asked a short series of questions about the degree of financial strain their household was facing, including how often the household put off buying something necessary, the level of difficulty the household had in paying bills, and whether there was any money left over at the end of each month.

**Documentation update:** With each survey round, a cohort-specific *Users Guide* is updated with new variable information and edited tables and figures. The updated *Users Guides* are made available online when the data round is released.

**Merging the data sets:** After the variables are cleaned and new variables created, the new round of data is merged with the other

data rounds so that researchers can take advantage of the longitudinal aspect of the data. Once the merge has been deemed successful, the data set is released for public access at [www.nlsinfo.org/web-investigator](http://www.nlsinfo.org/web-investigator). Apart from a preliminary review, all researchers have access to the main-file data at the same time. □

### Errata Notice: NLSY97 Created Variable Omitted

A set of collapsed created variables was included for the first time in the NLSY97 round-9 release. These variables provide users with the most up-to-date created variables for selected topics of a respondent, regardless of his or her interview status in the current round. The topics covered by the collapsed created variables (QNAME=CVC\_\*) include schooling, assets, and program participation. The variable that indicates the respondent's most recent interview round (QNAME=CVC\_RND\_2005) was inadvertently omitted from the round-9 data release, but is available by downloading the file (cvc\_rnd.xls) from the errata page at [www.nlsinfo.org/nlsy97/nlsy97\\_errata.php3](http://www.nlsinfo.org/nlsy97/nlsy97_errata.php3). This variable should be used in conjunction with the CVC\_\* variables to determine from which round the information in the created variable was collected. □

### Frequently Asked Questions

The NLS staff encourages researchers to contact NLS User Services with questions and problems encountered while accessing and using NLS data or documentation. Every effort is made to answer these inquiries. Following are some recent questions and answers that may be of general interest to NLS users:

*Q1. Are there any NLSY variables that describe educational expectations or aspirations?*

A1. In its initial year (1979), the NLSY79 main survey included three questions related to a respondent's educational aspirations. Respondents were asked to name the highest year or grade of school (from elementary school to graduate school) they wanted to complete

(R00234.00), what year or grade they expected to complete as things stood at the time of the survey interview (R00235.00), and what year or grade their closest friend would like to complete (R00236.00). These questions also were asked of the NLSY79 young adult respondents in their first interview as a young adult.

Information on educational expectations is available as well for the children of the NLSY79 main respondents. Beginning in 1988, mothers of children who were at least school age were asked to estimate how far they thought each child would go in school. In turn, the children themselves were asked to estimate how far they would get. A related question asking the mother to rate each child's prospects for the future was introduced in 1992.

Respondents in the NLSY97 cohort were asked in round 1 (1997) to state the percent chance they thought they had of getting a high school diploma by age 20 (R05147.00), in round 5 (2001) to name the highest year or grade they expected to complete (R58281.00), and in both rounds 1 (R05151.00) and 5 (R68844.00) to state the percent chance they thought they had of getting a college degree by age 30.

*Q2. In one NLSY79 question, respondents are asked the year they arrived in the United States if they were foreign born. However, when the information garnered from this question is cross-tabulated with whether or not a person was born in the United States, one learns that a few people report being born in the United States, but also report a year of arrival. What could be happening here?*

A2. In nearly all surveys, some respondents are inconsistent. Whenever a similar question is asked at different points in the survey, there are some respondents whose answers are not consistent. Sometimes a question is misunderstood, sometimes the respondent is not paying attention and answers incorrectly, and sometimes a respondent, for whatever reason, chooses to give an incorrect answer.

Users who find inconsistencies should first check the skip patterns of the question to understand the universe of respondents receiving the question. For larger-than-expected inconsistencies, a query to NLS User Services would certainly be warranted, because it occasionally indicates

a data error.

*Q3. Is there a simple way to determine whether an NLSY97 participant is a parent or a child?*

A3. The unit of observation in the NLSY97 is the youth respondent, and the PUBID uniquely identifies each one. The parent completed a parent interview that contained questions specific to the parent and loops specific to each youth respondent. The parent-specific data are attached to data pertaining to each of the parent's children who are youth respondents, and only the child-specific parent interview data are attached to data pertaining to that youth respondent. □

### 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](http://www.nlsbibliography.org) for a comprehensive listing of NLS-related research):

Arkes, Jeremy. "Does the Economy Affect Teenage Substance Use?" *Health Economics* 16,1 (January 2007): 19–36. [NLSY97]

Bacolod, Marigee Ponla. "Do Alternative Opportunities Matter? The Role of Female Labor Markets in the Decline of Teacher Quality." *Review of Economics and Statistics* 89,4 (November 2007): 737–751. [NLSY79, Young Men, Young Women]

Berger, Lawrence M. "Socioeconomic Factors and Substandard Parenting." *Social Service Review* 81,3 (September 2007): 485–522. [Children of the NLSY79, NLSY79, NLSY79 Young Adult]

Caputo, Richard K. "Perceived Work-Related Discrimination by Women: Implications for Social Justice and Affirmative Action." *Journal of Policy Practice* 6,2 (Summer 2007): 5–22. [Young Women]

Gabriel, Paul E., and Schmitz, Susanne. "Gender Differences in Occupational Distributions Among Workers." *Monthly*

*Labor Review* 130,6 (June 2007): 19–24. [NLSY79]

Gius, Mark Paul. "Impact of Provider Availability and Legal Restrictions on the Demand for Abortions by Young Women." *Social Science Journal* 44,3 (July 2007): 495–506. [NLSY79]

Hoffmann, John P., Dufur, Mikaela, and Huang, Lynn. "Drug Use and Job Quits: A Longitudinal Analysis." *Journal of Drug Issues* 37,3 (Summer 2007): 569–596. [NLSY79]

Li, Chaoyang, Goran, Michael I., Kaur, Harsohena, Nollen, Nicole, and Ahluwalia, Jasjit S. "Developmental Trajectories of Overweight During Childhood: Role of Early Life Factors." *Obesity* 15,3 (March 2007): 760–771. [Children of the NLSY79, NLSY79]

McFarlin, Isaac, Jr. "Do School Teacher Parents Make a Difference?" *Economics of Education Review* 26,5 (October 2007): 615–628. [Children of the NLSY79, NLSY79]

McLeod, Jane D., and Fettes, Danielle L. "Trajectories of Failure: The Educational Careers of Children with Mental Health Problems." *American Journal of Sociology* 113,3 (November 2007): 653–701. [Children of the NLSY79]

McNamara, Justine M. "Long-Term Disadvantage among Elderly Women: The Effects of Work History." *Social Service Review* 81,3 (September 2007): 423–452. [Mature Women]

Sen, Bisakha. "Frequency of Family Dinner and Adolescent Body Weight Status: Evidence from the National Longitudinal Survey of Youth, 1997." *Obesity* 14 (2006): 2266–2276. [NLSY97]

Smith, Jacqueline, and Boone, Anniglo. "Future Outlook in African American Kinship Care Families." *Journal of Health and Social Policy* 22,3/4 (Spring/Summer 2007): 9–30. [NLSY97]

Teachman, Jay D. "Race, Military Service, and Marital Timing: Evidence from the NLSY-79." *Demography* 44,2 (May 2007): 389–404. [NLSY79] □

Official Business  
Penalty for Private Use, \$300  
Address Service Requested

FIRST CLASS MAIL  
Postage and Fees Paid  
U.S. Department of Labor  
Permit No. G-738

U.S. Bureau of Labor Statistics  
Postal Square Building, Rm. 2850  
2 Massachusetts Ave., NE.  
Washington, DC 20212-0001

# NLS News



## NLS Contact Information

NLS News is published quarterly by the U.S. 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 User Services:	Center for Human Resource Research The Ohio State University 921 Chatham Lane, Suite 100 Columbus, Ohio 43221-2418 usersvc@chrr.osu.edu (614) 442-7366 (614) 442-7329 (Fax)	NLS documentation, data, and data updates:	usersvc@chrr.osu.edu
		NLS Web site:	www.bls.gov/nls
		BLS-NLS publications:	Jain_Rita@bls.gov (202) 691-7405
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)	<i>NLS News</i> Editor:	Donna S. Rothstein Rothstein_D@bls.gov
		NLS Program Director and Media Contact:	Charles R. Pierret Pierret_C@bls.gov (202) 691-7519