

Introduction to NLSY

Benedikt Kauf

The **National Longitudinal Surveys (NLS)** are a set of surveys designed to gather information at multiple points in time on the labor market and other significant life events of several groups of men and women.

Topics covered in NLS Surveys

- ▶ **Education, Training and Achievement Scores:** information about high school and college experience, types of training, and cognitive test scores
- ▶ **Employment:** extensive history of employment over time, including wages, hours, occupation, industry, and job search
- ▶ **Health:** Health status, weight, and health conditions
- ▶ **Attitudes, Expectations, and Non-Cognitive Tests:** personality traits, expectations about the future, locus of control
- ▶ And many more...

Survey	Birth Years	Interview Years
Older Men	1906-1921	1966-1990
Mature Women	1922-1937	1967-2003
Younger Men	1941-1952	1966-1981
Younger Women	1943-1953	1968-2003
NLS Youth – 1979 cohort	1957-1964	1979-present
NLS Youth – 1997 cohort	1980-1984	1997-present
NLSY Child and Young Adult	1975-present	1984-present

Table: List of NLS cohorts

NLS Data

- ▶ Based on birth cohorts
- ▶ US nationally representative
- ▶ Oversamples (military, poor, minorities)
- ▶ Interview-based data set focused on labor market experience
- ▶ Periodic interviews

National Longitudinal Survey of Youth, 1979 Cohort (NLSY79)

- ▶ Original sample size of 12,686 individuals born between 1957 and 1964
- ▶ First interviews in 1979 with respondents aged 14 to 22
- ▶ Annual interviewing from 1979 to 1994, biennial interviewing since 1996
- ▶ Most recent published data from interviews in 2016 with respondents aged 51 to 60 and a sample size of 6,912

Types of information gathered in NLSY79

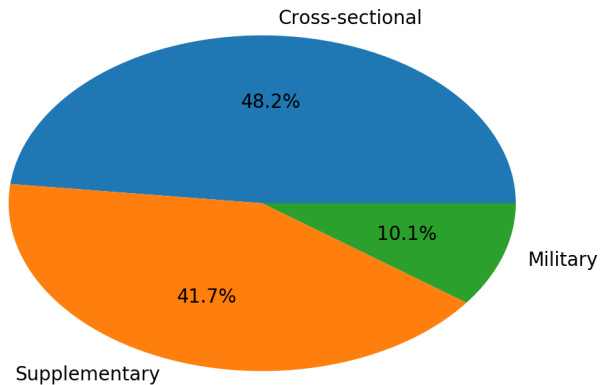
- ▶ Labor market behavior
- ▶ Educational experiences
- ▶ Family background
- ▶ Armed Services Vocational Aptitude Battery (measures knowledge and skills including reading and mathematics)
- ▶ High school information (from respondents' schools and school transcripts)
- ▶ Government program participation
- ▶ Family life
- ▶ Health issues
- ▶ Assets and income

Subsample Design

Three subsamples

- ▶ **Cross-Sectional Sample:** 6,111 individuals that represent the noninstitutionalized civilian segment of people living in the US in 1979
- ▶ **Supplemental Sample:** 5,295 civilian Hispanic or Latino, black, and economically disadvantaged nonblack/non-Hispanic individuals living in the US in 1979
- ▶ **Military Sample:** 1,280 individuals that represent the population serving in the U.S. military as of September 30, 1978

Subsample Design

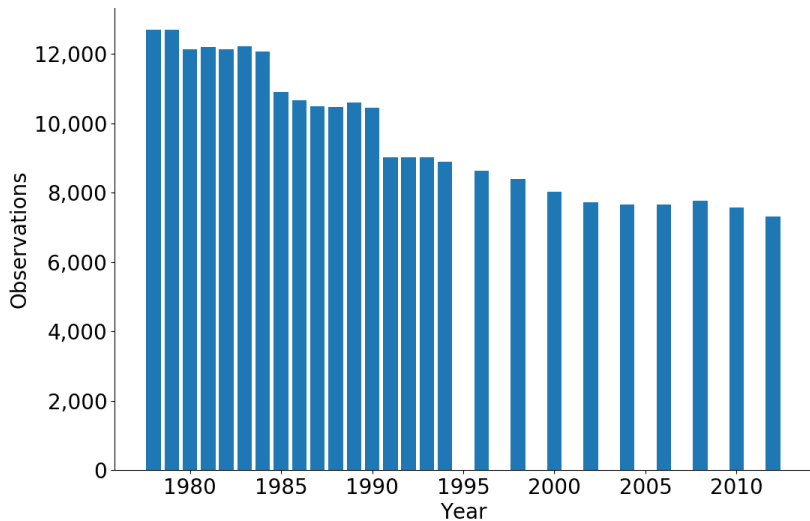


Changes in the subsamples

The subsamples have been subject to significant changes over time:

- ▶ After the 1984 interview, 1,079 members of the military sample were no longer eligible for interview; 201 respondents randomly selected from the military sample remained in the survey
- ▶ After the 1990 interview, none of the 1,643 members of the economically disadvantaged, nonblack/non-Hispanic sample were eligible for interview

Number of Observations



Pros & Cons

Pros:

- ▶ Covers a long time period
- ▶ Large part of the data set is freely available
- ▶ Lots of helpful internet resources (use them!)

Cons:

- ▶ Significant reduction of the sample size over the years
- ▶ Not the whole data set is freely available (might be a problem for replication)

Time availability of variables

- ▶ Note that not all variables are available for all points in time.
- ▶ Go to <https://www.nlsinfo.org/content/cohorts/nlsy79/topical-guide/asterisk/i-labor-market-experience-variables> to check for which timepoints a certain variable is available.
- ▶ Helpful to do this quite early in order to avoid data availability problems at later stages.

NLSY79 Child and Young Adult

- ▶ Co-resident children of women in the NLSY79 have been followed with their mothers since 1986
- ▶ Not based on a certain birth cohort
- ▶ Psychosocial assessments through age 10
- ▶ Children graduate to "Young Adults" at age 15
- ▶ Significant overlap of the children sample with NLSY97

National Longitudinal Survey of Youth, 1997 Cohort

- ▶ Original sample size of 8,984, including oversamples of Blacks and Hispanics
- ▶ Annual interviewing, beginning in 1997
- ▶ Good for comparison with earlier cohorts
- ▶ Disadvantage: less (non-)cognitive info and more attrition

Access NLSY data

- ▶ To access NLSY data, visit <https://www.nlsinfo.org/investigator/pages/login.jsp>
- ▶ You can either create an account for free or search for data as a guest.
- ▶ Setting up an account only requires an email address and a username, not necessarily your real name.

Selecting a data source

NLS Investigator

Select the study you want to work with:

(Choose One)	▼
(Choose One)	
NLSY79 (National Longitudinal Survey of Youth 1979)	
NLSY79 Child & Young Adult	
NLSY97 (National Longitudinal Survey of Youth 1997)	
Original Cohorts - Mature Women & Young Women	
Original Cohorts - Older Men & Young Men	

Select the study that you want to use.

For this exercise we will work with NLSY79.

Tagsets tab

NLS Investigator

Select the study you want to work with:

NLSY79 (National Longitudinal Survey of Youth 1979) ▾

Released March 01, 2017

Additional Resources:

[Errata](#), [Documentation](#) (user's guide, questionnaires and other materials)
[Custom Weights](#)

To start a new search [click here](#)

Choose Tagsets

Variable Search

Review Selected Variables (5)

Codebook

Save / Download

Required / Recommended Variables:

- ☒ Required ID Variable - CASEID will always be selected (1 variable)
- ☒ Recommended Demographic Variables (4 variables)

Saved Tagsets (on server):

None Available

Upload Tagset (from PC):

Durchsuchen... Keine Datei ausgewählt.

Upload

[NLS Home](#) | [NLS Bibliography](#) | [Privacy Policy](#)

Tagsets tab

After selecting the study you will see a series of tabs. The first one is to choose a *tagset*. Tagsets are basically saved searches.

Note that by default six variables will be added to your data: id (can't remove this) and the following demographics (optional): gender, age, race/ethnicity and birthday (month and year).

"Variable search" tab

NLS Investigator

Select the study you want to work with:

NLSY79 (National Longitudinal Survey of Youth 1979) ▾

Released March 01, 2017

Additional Resources:

[Errata](#), [Documentation](#) (user's guide, questionnaires and other materials)
[Custom Weights](#)

To start a new search [click here](#)

Choose Tagsets

Variable Search

Review Selected Variables (5)

Codebook

Save / Download

Browse Index

Browse Index with Search

Search

Index of Selected Variables

- + Education, Training & Achievement Scores (896)
- + Employment (47873)
- + Health (2513)
- + Survey Methodology (180)
- + Household, Geography & Contextual Variables (2696)
- + Marriage & Cohabitation (664)
- + Sexual Activity, Pregnancy & Fertility (1169)
- + Children (3508)
- + Parents, Family Process & Childhood (540)
- + Income, Assets & Program Participation (3351)
- + Attitudes, Expectations & Non-cognitive Tests (847)
- + Crime & Substance Use (365)

Options ▾

Please browse the index on the left to display variables.
This index contains a set of NLSY79 variables commonly used in research and is not the full data set.

[NLS Home](#) | [NLS Bibliography](#) | [Privacy Policy](#)

"Variable search" tab

"Variable Search" offers three search modes: by topics (index), by searching the index and a general search.

Note that the site does not offer access to the full data sets.

Searching for data (1)

NLS Investigator

Select the study you want to work with:

NLSY79 (National Longitudinal Survey of Youth 1979) ▾

Released March 01, 2017

Additional Resources:

[Errata](#), [Documentation](#) (user's guide, questionnaires and other materials)
[Custom Weights](#)

To start a new search [click here](#)

Choose Tagsets

Variable Search

Review Selected Variables (5)

Codebook

Save / Download

Browse Index

Browse Index with Search

Search

Create search criteria below:

(Choose One) ▾	▾	
(Choose One)		
Area of Interest (pick from list)		
Word in Title (pick from list)		
Word in Title (enter search term)		
Question Text (enter search term)		play variables.
Question Name (pick from list)		
Question Name (enter search term)		
Reference Number (pick from list)		
Reference Number (enter search term)		
Survey Year (pick from list)		
Codebook (enter search term)		
Variable Type (pick from list)		

The general search offers a variety of options. From the dropdown menu select the type of search you want.

Searching for data (2)

NLS Investigator

Welcome, benedikt.hauf@gmail.com | SEARCH | Downloads | Account | Help | Logout

Select the study you want to work with:

NLSY79 (National Longitudinal Survey of Youth 1979) ▾

Additional Resources:

[Ezra's Documentation](#) (user's guide, questionnaires and other materials)
[Custom Weights](#)

Released March 01, 2017

To start a new search [click here](#)

Choose Targets

Variable Search

Review Selected Variables (15)

Codebook

Save / Download

Browse Index

Browse Index with Search

Search

Create search criteria below:

Include only intersecting (AND) ▾

Area of Interest (pick from list)	equals	BETWEEN JOB GAPS	Remove
Word in Title (enter search term)	contains	month	Remove
(Choose One) ▾	▾	▾	Add Reset

126 Variables

Display Variables

Options: <input checked="" type="checkbox"/> Showing 126 of 126 filtered by Primary Variables ▾			
	RNUM	QUESTION NAME	VARIABLE TITLE
1	<input checked="" type="checkbox"/> B0279300	Q7-10_1_M	MONTH BEGAN 1ST PERIOD NOT EMPLOYED SINCE LAST INT
2	<input checked="" type="checkbox"/> B0279300	Q7-11_1_M	MONTH ENDED 1ST PERIOD NOT EMPLOYED SINCE LAST INT
3	<input checked="" type="checkbox"/> B0280100	Q7-10_2_M	MONTH BEGAN 2ND PERIOD NOT EMPLOYED SINCE LAST INT
4	<input checked="" type="checkbox"/> B0280200	Q7-11_2_M	MONTH ENDED 2ND PERIOD NOT EMPLOYED SINCE LAST INT
5	<input checked="" type="checkbox"/> B0281000	Q7-10_3_M	MONTH BEGAN 3RD PERIOD NOT EMPLOYED SINCE LAST INT
6	<input checked="" type="checkbox"/> B0281100	Q7-11_3_M	MONTH ENDED 3RD PERIOD NOT EMPLOYED SINCE LAST INT
7	<input checked="" type="checkbox"/> B0281900	Q7-10_4_M	MONTH BEGAN 4TH PERIOD NOT EMPLOYED SINCE LAST INT
8	<input checked="" type="checkbox"/> B0282000	Q7-11_4_M	MONTH ENDED 4TH PERIOD NOT EMPLOYED SINCE LAST INT
9	<input checked="" type="checkbox"/> B0282800	Q7-10_5_M	MONTH BEGAN 5TH PERIOD NOT EMPLOYED SINCE LAST INT
10	<input checked="" type="checkbox"/> B0282900	Q7-11_5_M	MONTH ENDED 5TH PERIOD NOT EMPLOYED SINCE LAST INT
11	<input type="checkbox"/> B0456500	Q7-10_1_M	MONTH BEGAN 1ST PERIOD NOT EMPLOYED SINCE LAST INT
12	<input type="checkbox"/> B0456600	Q7-11_1_M	MONTH ENDED 1ST PERIOD NOT EMPLOYED SINCE LAST INT
13	<input type="checkbox"/> B0456800	Q7-10_2_M	MONTH BEGAN 2ND PERIOD NOT EMPLOYED SINCE LAST INT
14	<input type="checkbox"/> B0466100	Q7-11_2_M	MONTH ENDED 2ND PERIOD NOT EMPLOYED SINCE LAST INT

Searching for data (2)

This is a nested search in which we are looking for "month" in "between job gaps". You can add more terms if needed.

Results will be shown in the lower box. Sometimes you will see a "+" next to RNUM which means that there are more variables available (as a subset of responses). Check the square next to RNUM to select all variables or check the ones you need.

If you hover the cursor over the question name, a window will pop up with detailed information about that variable.

Codebook tab

Choose Tagsets

Variable Search

Review Selected Variables (15)

Codebook

Save / Download

R02792.00 [q7-10_1_M] Survey Year: 1980

PRIMARY VARIABLE

MONTH BEGAN 1ST PERIOD NOT EMPLOYED SINCE LAST INT

ORIGINAL QUESTION NAME: S09D1837

I: DRAW LINES ON ROW C TO REPRESENT PERIODS DURING WHICH THERE ARE NO LINES IN ROW A OR B. USE DATES ENTERED IN ROWS A & B TO INDICATE IN ROW C DATES R BEGAN AND ENDED EACH PERIOD OF NON-EMPLOYMENT. ENTER THE DATES FOR EACH PERIOD INTO BOX A. MOST RECENT PERIOD FIRST.

1 THRU 12 ACTUAL MONTH

UNIVERSE: R >= 16; period not employed since DLI

0	0: < 1
391	1
1174	2
1580	3
698	4
506	5
321	6
286	7
941	8
563	9
272	10
283	11
475	12
0	13
0	14
0	15
0	16
0	17 TO 99999: 17*

7490

Refusal (-1)	0
Don't Know (-2)	0
Invalid Skip (-3)	2
TOTAL =====>	7492
VALID SKIP (-4)	4649
NON-INTERVIEW (-5)	545




Min:	1	Max:	12	Mean:	5.45
------	---	------	----	-------	------

Lead In: [R02791.00](#) [Default]
Default Next Question: [R02792.10](#)

Lead In
[R02791.00](#) [Default]

Default Next
[R02792.10](#)

Selected Variables
☒ R02792.00
◀ ▶ ⏪ ⏩

Stats as Graphs
  

Questionnaire Links
[1980 Questionnaire](#)

Area of Interest
BETWEEN JOB GAPS

Here you can get some additional information about your variables along with some basic stats

25 / 32

Downloading data (1)

NLS Investigator

Select the study you want to work with:
NLSY79 (National Longitudinal Survey of Youth 1979) ▾

Released March 01, 2017

Additional Resources:
[Errata, Documentation](#) (user's guide, questionnaires and other materials)
[Custom Weights](#)

To start a new search [click here](#)

Choose Tagsets

Variable Search

Review Selected Variables (18)

Codebook

Save / Download

Save Tagset

Basic Download

Advanced Download

Manage Downloads

Choose where to save the tagset of your selected variables:

☒ Save to PC
☐ Save on our server

Tagset Type: ☒ By Rnum ☐ By Qname with Year

Filename:

Filename must only contain alpha, numeric, hyphen or underscore characters.

Tagsets saved on your PC or our server can be reloaded at a later time through the 'Choose Tagsets' tab.
Tagsets stored on our server with no activity may be deleted after 90 days.

[NLS Home](#) | [NLS Bibliography](#) | [Privacy Policy](#)

Once you are satisfied with your search, you can save it as a "tagset". You can keep the defaults, select a name (in this case we choose "Month") and click on "Save".

Downloading data (2)

NLS Investigator

Select the study you want to work with:

NLSY79 (National Longitudinal Survey of Youth 1979) ▾

Additional Resources:

[Errata](#), [Documentation](#) (user's guide, questionnaires and other materials)
[Custom Weights](#)

Released March 01, 2017

To start a new search [click here](#)

Choose Tagsets

Variable Search

Review Selected Variables (16)

Codebook

Save / Download

Save Tagset

Basic Download

Advanced Download

Manage Downloads

Customize your advanced download:

☒ **Create Download of Data**

- ☒ Tagset (list of selected variables)
- ☐ SAS® control file (includes the datafile of selected variables)
- ☐ SPSS® control file (includes the datafile of selected variables)
- ☒ STATA® dictionary file of selected variables
- ☐ R® Source code (includes the datafile of selected variables)
- ☒ Codebook of selected variables
- ☐ Short Description File
- ☒ Comma-delimited datafile of selected variables (to be read in Excel, etc.)
Column headers -- Use ☒ Reference Number ☐ Question Name (does not guarantee uniqueness)

☐ **Create Frequency / Table**

- ☐ **Apply Universe Restrictors** ([How to use Universe Restrictors](#))
- ☐ Notify me by email when download is complete.

Filename:

Filename must only contain alpha, numeric, hyphen or underscore characters.

Download status appears under "Manage Downloads" tab.
Downloads may be deleted after 10 days of inactivity.

Downloading data (2)

1. Select the data format you want. For this example we will select Stata and comma-delimited along with the codebook.
2. Select a name for the file and click "Download"

Downloading data (3)

NLS Investigator

Select the study you want to work with:

NLSY79 (National Longitudinal Survey of Youth 1979) ▾

Released March 01, 2017

Additional Resources:

[Errata, Documentation](#) (user's guide, questionnaires and other materials)
[Custom Weights](#)

To start a new search [click here](#)

Choose Tagsets

Variable Search

Review Selected Variables (16)

Codebook

Save / Download

Save Tagset

Basic Download

Advanced Download

Manage Downloads

Download Status:

All downloads are available. Please click a download link below to begin downloading.

All Available Downloads:

		Date	Study	Name	Size	Download
1	<input checked="" type="checkbox"/>	2017-11-22 09:43:21	NLSY79	Month	166.4K	download
2	<input type="checkbox"/>	2017-11-19 06:01:13	NLSY79	default	118.4K	download
3	<input type="checkbox"/>	2017-11-19 05:36:40	NLSY79	default	106.8K	download

Delete Selected Files

Öffnen von Month.zip

Sie möchten folgende Datei öffnen:

 Month.zip

Vom Typ: Compressed (zipped) Folder

Von: <https://www.nlsinfo.org>

Wie soll Firefox mit dieser Datei verfahren?

☐ Öffnen mit Windows-Explorer (Standard) ▾

☒ Datei speichern

☐ Für Dateien dieses Typs immer diese Aktion ausführen

OK

Abbrechen

Downloading data (3)

If you set your browser to ask you where to save, a pop-up window will prompt you to select a folder location for the zip file, click "OK" and select the folder.

If it gets downloaded automatically, then the file should be in the "Downloads" folder or the default place for downloads.

Further information

- ▶ For detailed information on all NLS cohorts:
<https://www.bls.gov/nls/handbook/nlshndbk.htm>
- ▶ For an extensive user guide to the NLSY79:
<https://www.bls.gov/nls/79guide/2001/nls79g0.pdf>
- ▶ For an extensive user guide to the NLSY97:
<https://www.bls.gov/nls/97guide/rd5/nls97ugall.pdf>

Further information

- ▶ For information on how to use NLSY in STATA;
<http://dss.princeton.edu/training/NLS101.pdf>
- ▶ For a tutorial on how to construct comparable sample across NLSY79 and NLSY97:
<https://www.nlsinfo.org/content/getting-started/intro-to-the-nls/tutorials/tutorial-constructing-comparable-samples-across>
- ▶ For information of sample weights/clustering adjustments:
<https://www.nlsinfo.org/content/cohorts/nlsy79/using-and-understanding-the-data/sample-weights-clustering-adjustments>