# Overview of the National Longitudinal Survey of Youth 1997 (NLSY97)

Hsueh-Sheng Wu
CFDR Workshop Series
March 14, 2022



## **Outline**

- Introduction to the NLSY97
- Advantages and challenges of using NLSY97
- Navigation through NLSY 97 website
  - > Asterisk tables
  - ➤ Marital history, childcare, and fertility
  - ➤ NLSY 97 sample
  - Using and understanding Data
  - > Other documentations
- How to obtain the NLSY97 data
  - > NLS Investigator website
  - ➤ How to obtain weight variables
- How to weight NLSY97 data
- An example of Using NLSY97 data
- Journal articles using NLSY97
- Conclusions



## Introduction to the NLSY 97

- NLSY97 is one of the five cohort studies conducted by the U.S. Bureau of Labor Statistics (see https://www.bls.gov/nls/cohorts.htm).
- NLSY97 was designed to document how the youth transits from school to work and into adulthood.
- The subject areas of NLSY97 now include:

Education, Training & Achievement Scores	Income, Assets & Program Participation
Employment	Health
Household, Geography & Contextual Variables	Attitudes, Expectations, Non-Cognitive Tests, Activities
Parents, Family Process & Childhood	Crime & Substance Use
Dating, Marriage & Cohabitation; Sexual Activity, Pregnancy & Fertility; Children	



## Introduction to the NLSY97 (Cont.)

- NLSY 97 is a cohort survey whose target population is youths who were 12 to 16 years old as of December 31, 1996.
- NLSY97 was conducted annually between 1997 and 2012 and biennially afterwards. Nineteen rounds of NLSY97 data are currently available. Not every subject topic was examined in every round of NLSY97.
- NLSY97 used a multi-stage stratified sampling method to select eligible households. Blacks and Hispanics were over-sampled. At the first phase, 96,512 households were selected from 1,748 sample segments embedded within 147 primary sampling units (PSUs). At the second phase, screening interviews were used to identify 9,907 households with eligible youths, and 8,984 of them participated in the 1997 survey.
- The NLSY97 data were mainly collected using Computer-Assisted Personal Interview(CAPI), Audio Computer-Assisted Self-Interview (ACASI), and questionnaires.



## Advantages and Challenges of Using NLSY97

#### Advantages:

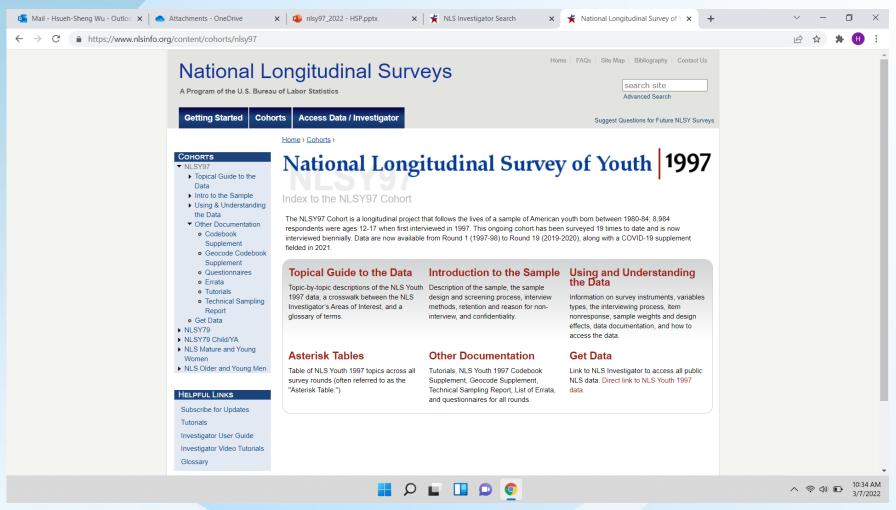
- NLSY97 covers many research tocis.
- NLSY97 oversampled Black and Hispanic respondents, which allows for examining racial differences on many topics.
- NLSY97 data have been collected at a one- or two-year intervals, which provides vast information on when, why, and how youths transited into adulthood.
- Many constucted variables are included in NSLY97 data, and tutorials are provided for researchers to construct additional variables.
- Custom weight variables can be generated for researchers using certain rounds of NLSY97.

#### Challenges:

- As data collected by a cohort design, the findings of NLSY 97 may not be able to generalized to people of other birth cohorts.
- Changes to designs and questionnaire of NLSY97 over time create some problems in analyzing data longitudinally.
- Researchers should learn to use the NLS Investigator website to select and download the data.
- Special statistical techniques are needed when researchers analying data from multiple rounds of NLSY97.

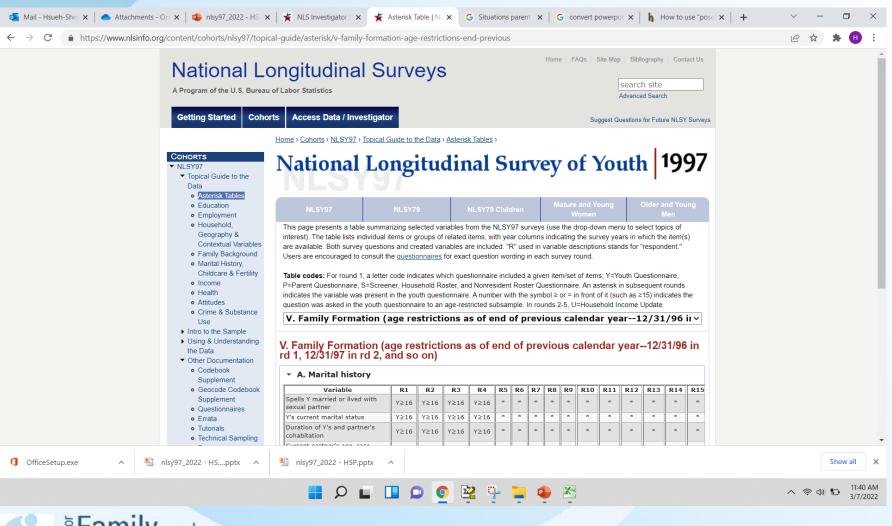
## **NLSY 97 Website**

(https://www.nlsinfo.org/content/cohorts/nlsy97)



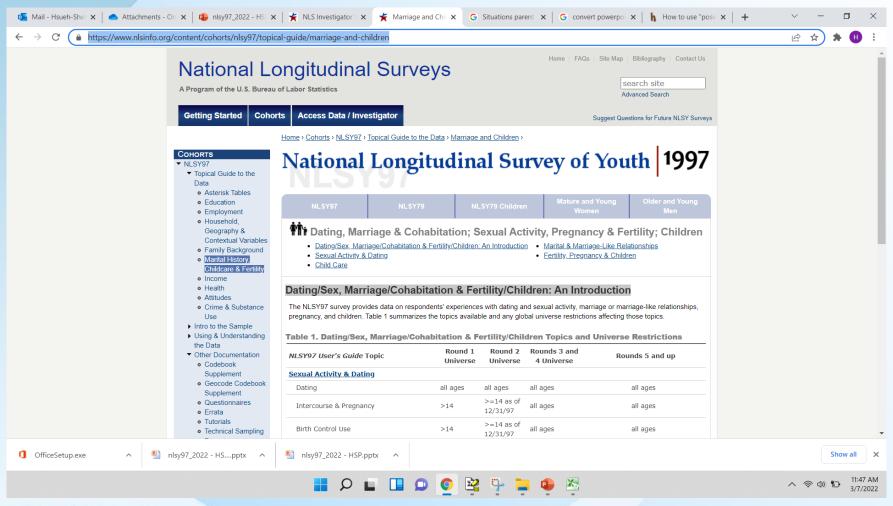


#### Asterisk tables



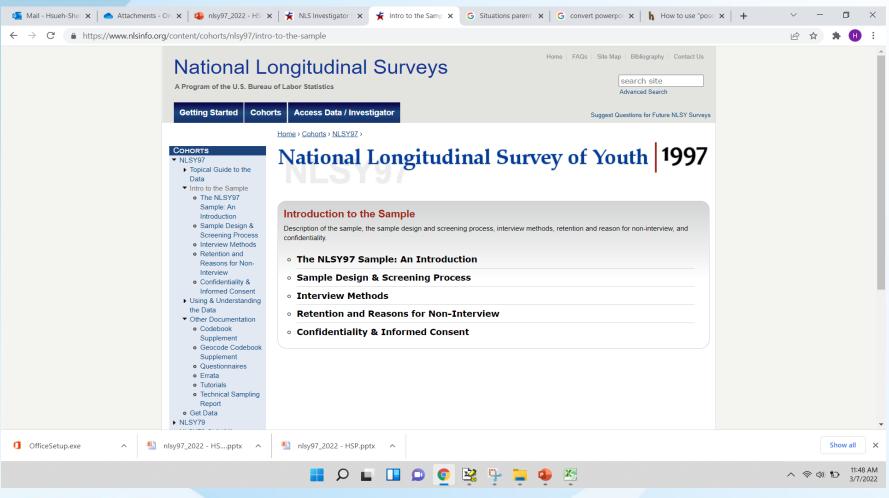


### Marital history, childcare, and fertility



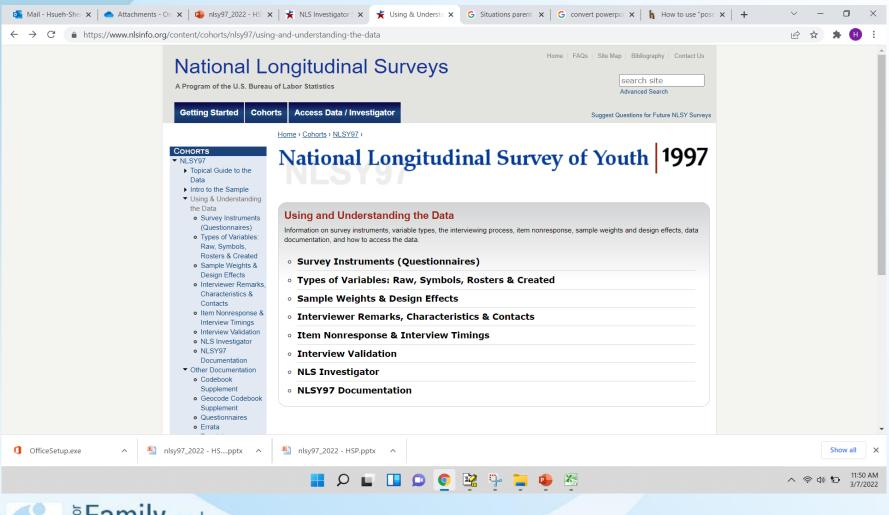


## NLSY 97 sample



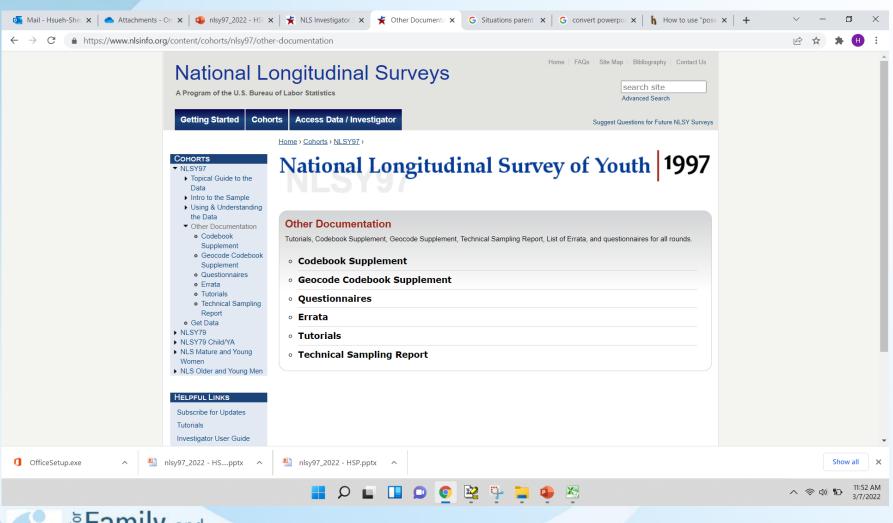


## Using and understanding data





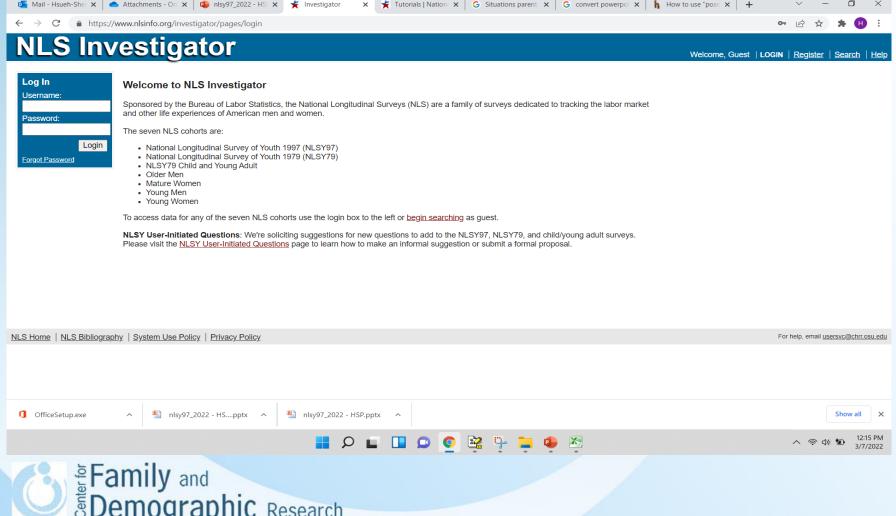
#### Other documentations





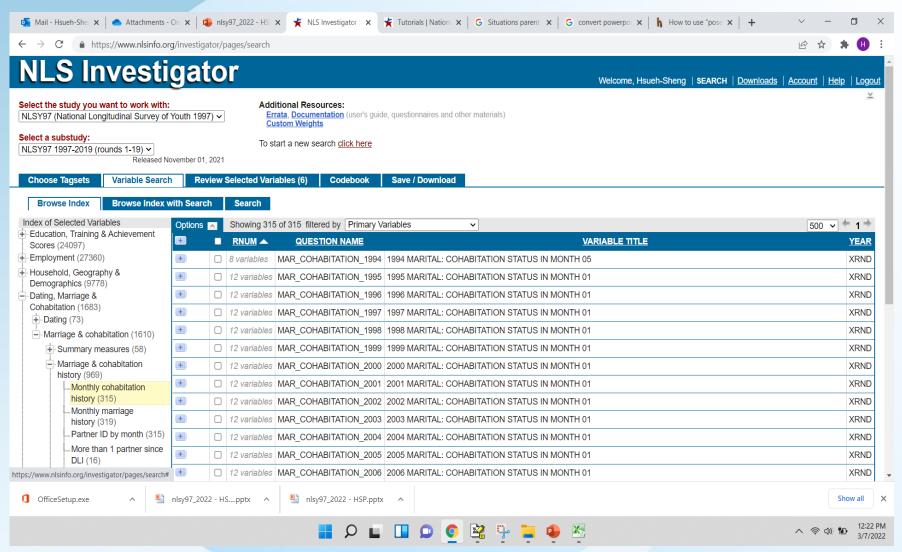
## How to Obtain NLSY97 Data

**NLS Investigator** (https://www.nlsinfo.org/investigator/pages/login)



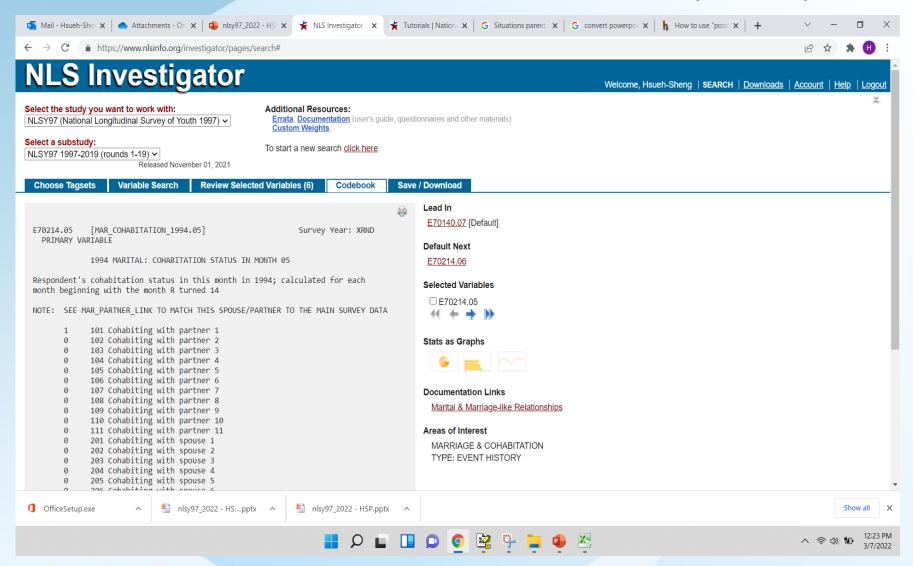


## How to Obtain the NLSY97 Data (Cont.)



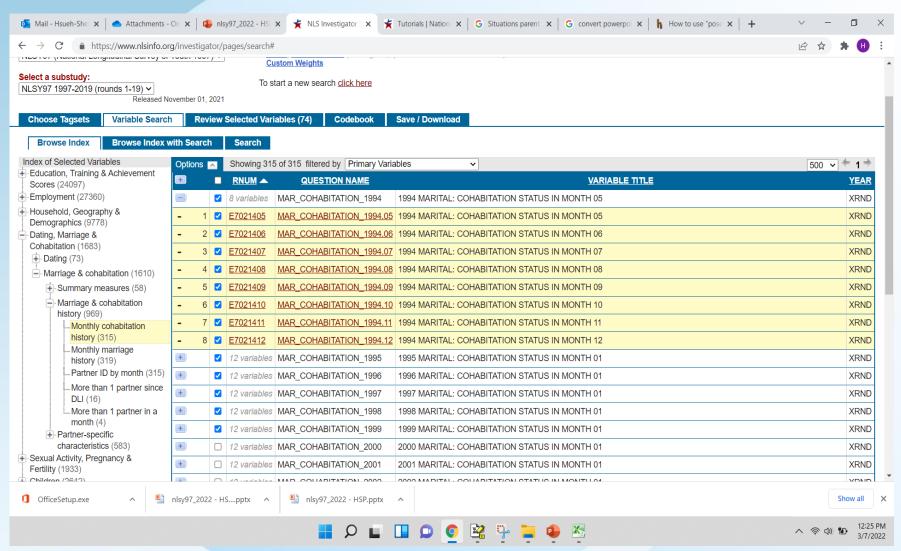


## How to Obtain NLSY97 Data (Cont.)



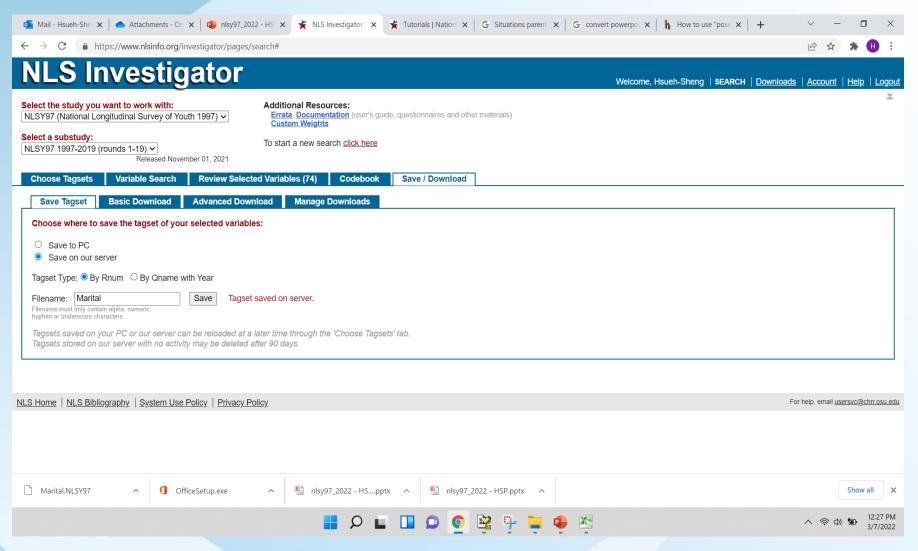


## How to Obtain the NLSY97 Data (Cont.)



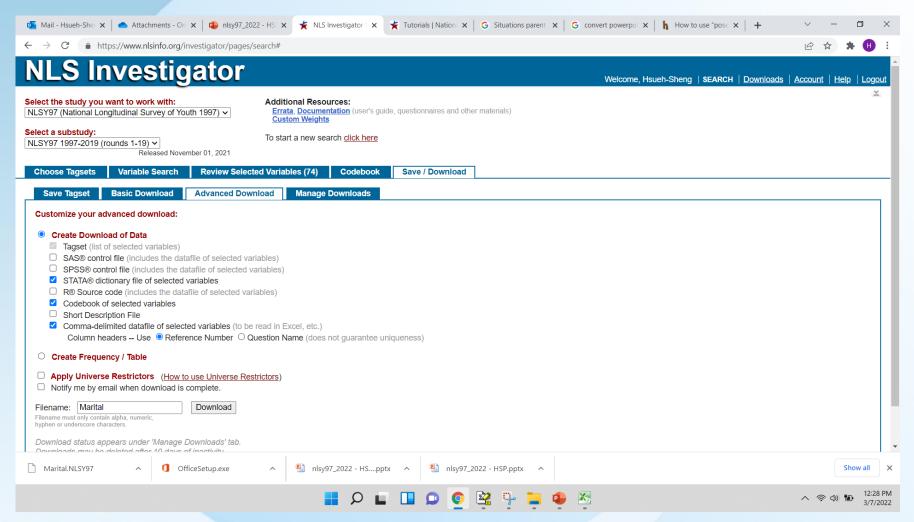


## How to Obtain NLSY97 Data (Cont.)





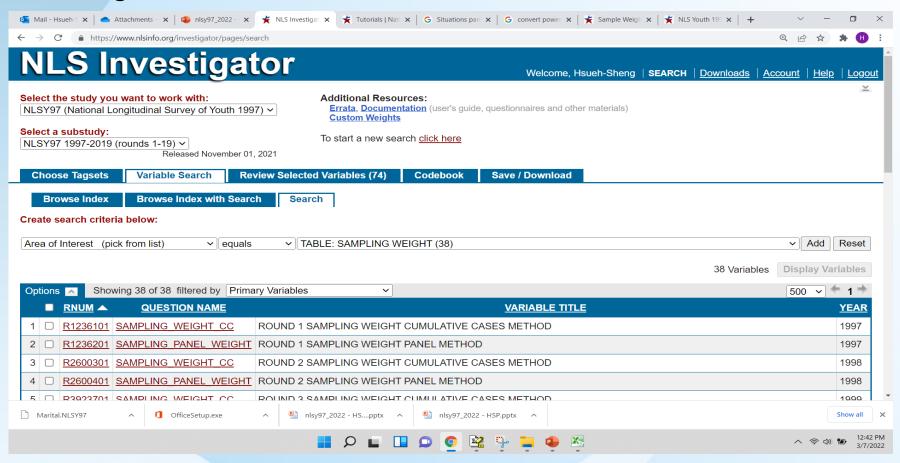
# How to Obtain NLSY97 Data (Cont.)





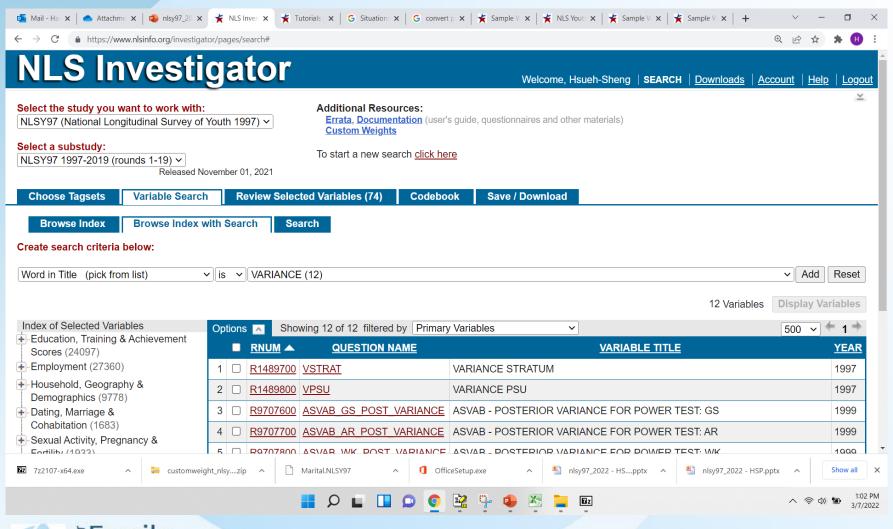
# How to Obtain Weight Variables of NLSY 97

Sampling weights at each round, sampling panel weights, and design effect variables at 1997 are available via NLS Investigator





# How to Obtain NLSY 97 Weight Variables (Cont.)

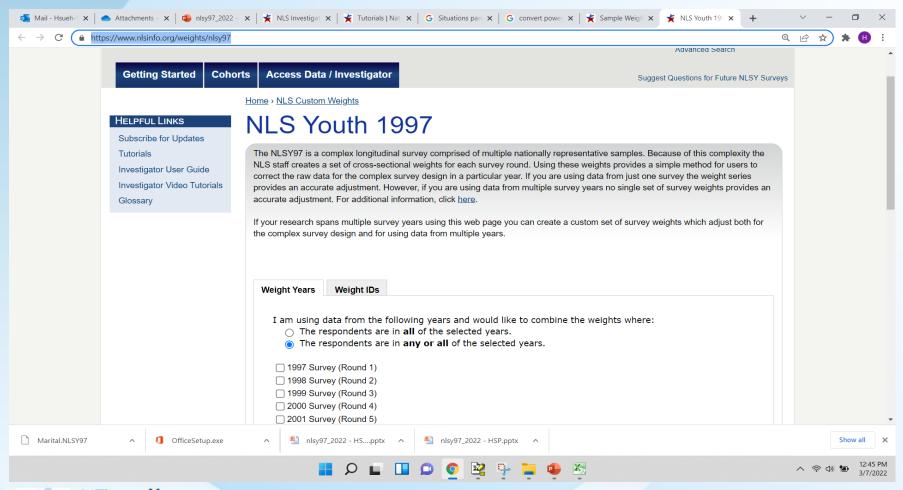




# How to Obtain NLSY 97 Weight Variables (Cont.)

Customized Panel Weights are available at:

https://www.nlsinfo.org/weights/nlsy97





# How to Obtain NLSY97 Weight Variables (Cont.)

- NLSY97 panel weight variables can be customized, depending on how many rounds of surveys are used and whether the focus is on respondents participating in all or any of these rounds.
- The customized panel weight variables were created by adjusting for differences in base weight, nonresponse to the screener, different response rates to the screener, different cross-sectional weights for Black and Hispanic respondents, nonresponse to the main interview.
- NLSY97 customized panel weight variables are post-stratification person weight variable.



## How to Weight NLSY97 Data

Source: https://www.nlsinfo.org/content/cohorts/nlsy97/using-and-understanding-the-data/sample-weights-design-effects/page/0/1

```
use NLSY97DATAFILE.dta,clear
svyset [pweight=WTVAR], strata(vstrat) psu(vpsu) singleunit(scaled)
svy: mean ANALYSISVAR //mean for continuous variables
svy: proportion ANALYSISVAR //proportion for categorical variables
svy, subpop (if SUBGROUP==1): mean ANALYSISVAR // mean within a
subpopulation
svy: tabulate ANALYSISVAR //one way table
```



# An example of Using NLSY97 data

- Research questions: Do age and education predict the respondent's marital status on December 2019?
- Select variables and weight variables
- Download the data
- Use the Stata to read in the data
- Examine the data to make sure the data are read in correctly.
- Code the missing values of variables
- Identify the analytic sample
- Specify survey design characteristics
- Conduct the descriive and inferential analysis.



## Journal Articles Using NLSY 97

- Manning, Wendy D. (2015) Family Formation Processes: Assessing the Need for a New Nationally Representative Household Panel Survey in the United States. Journal of Economic and Social Measurement, 40(1-4), 197-219.
- Guzzo, K.B. (2015). Young parents and co-residence with their own parents (2015 Working Paper Series). Retrieved from California Center for Population Research: <a href="http://papers.ccpr.ucla.edu/papers/PWP-BGSU-2015-008/PWP-BGSU-2015-008.pdf">http://papers.ccpr.ucla.edu/papers/PWP-BGSU-2015-008/PWP-BGSU-2015-008.pdf</a>
- Sandberg-Thoma, S.S.; Snyder, A.R. & Jang. BJ. (2015). Exiting and Returning to the Parental Home for Boomerang Kids, Journal of Marriage and Family, 77(3), 806-818. http://DX.DOI: 10.1111/jomf.12183
- Tumin, D.; Siqi Han, S. & Qian, Z. (2014). Estimates and Meanings of Marital Separation. Journal of Marriage and Family, 77(1), 312-322. http://DX.DOI: 10.1111/jomf.12149



## Conclusions

- NLSY97 provides an excellent opportunity to examine how the youth transited into the adulthood. With measures of various subject areas, NLSY may shed light on why some youths made the transition while some did not.
- NLSY97 already includes many constructed variables on the timing and history of many events, so researchers do not need to start everyting from scrach.
- Excelletn tutorials are provided at NLSY website and document how to link different rounds of NLSY data and constuct additional variables.
- When analyzing NLSY97 data, researchers need to correctly specify the survey design characteristics to obtain the correct estimate.
- If you have any questions about using NLSY97 data. Please feel free to contact Hsueh-Sheng Wu at wuh@bgsu.edu.

