

ICPSR 21600

**National Longitudinal Study of  
Adolescent Health (Add Health),  
1994-2008**

Kathleen Mullan Harris

*University of North Carolina-Chapel Hill*

J. Richard Udry

*University of North Carolina-Chapel Hill*

Wave 2, Grand Sample Weights, Public Use  
Data, Original Codebook

Inter-university Consortium for  
Political and Social Research  
P.O. Box 1248  
Ann Arbor, Michigan 48106  
[www.icpsr.umich.edu](http://www.icpsr.umich.edu)

# Terms of Use

The terms of use for this study can be found at:  
<http://www.icpsr.umich.edu/cocoon/ICPSR/TERMS/21600.xml>

## Information about Copyrighted Content

Some instruments administered as part of this study may contain in whole or substantially in part contents from copyrighted instruments. Reproductions of the instruments are provided as documentation for the analysis of the data associated with this collection. Restrictions on "fair use" apply to all copyrighted content. More information about the reproduction of copyrighted works by educators and librarians is available from the United States Copyright Office.

### NOTICE

#### WARNING CONCERNING COPYRIGHT RESTRICTIONS

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use," that user may be liable for copyright infringement.

*National Longitudinal Study of  
Adolescent Health*

*Public-use  
Grand Sample Weights*



Roger Tourangeau  
Hee-Choon Shin  
National Opinion Research Center

Carolina Population Center  
University of North Carolina at Chapel Hill

August 1998

This research was supported by grant P01-HD31921 from the National Institute of Child Health and Human Development. Further information may be obtained by contacting Jo Jones, PhD, Project Manager, 919/ 962-8412 (email: jo\_jones@unc.edu) at the Carolina Population Center, CB# 8120 University Square, Chapel Hill, NC 27516-3997.

# Weights for the Public-Use Sample

## 1. Wave I Weights

The public-use sample includes data from 6,504 respondents, who were originally selected for the core sample, the high education black supplement sample, or both. For the purpose of weighting the public-use data, respondents were divided into two groups. One group consisted of those who were eligible only for the core sample (and who had, therefore, only one chance of selection for the public-use sample); the second included those who were eligible for selection into both samples (thus, giving them two chances of selection). The weights for the latter group had to take into account the respondents' selection probabilities for both samples.

For the respondents eligible only for the core sample, the calculation of initial public-use weights was straightforward. The public-use cases are a random subsample that includes one-half of the core respondents. Thus, their initial public-use weights were twice their final core weights. The core weights already incorporated adjustments for In-home Questionnaire nonresponse, but they did not include post-stratification adjustments; we intended to post-stratify the public-use weights as the final stage in computing the new weights, after we had computed preliminary weights for both groups of public-use cases.

For the respondents eligible for selection into both samples, however, a more elaborate weighting procedure was needed. This procedure consisted of two main steps. In the first, a base weight was calculated that reflected the student's probability of selection into either sample ( $P_{ijk}$ ). That probability, in turn, reflected the school's selection probability ( $P_i$ ), the conditional probability for selection into the core sample (given that school  $i$  was selected), and the conditional probability for selection into the black supplement ( $P_B$ ):

$$P_{ijk} = P_i \left( 1 - \left[ \left( 1 - \frac{n_{ij}}{N_{ij}} \right) (1 - P_B) \right] \right) .$$

in which  $n_{ij}$  is the number of selections taken from stratum  $i$  at school  $j$  (within each school, separate core selections were made within each grade-sex cell, or stratum),  $N_{ij}$  is the total number of students in that stratum and school, and  $P_B$  is a constant (equal to .3077) representing the sampling fraction for high education black students at all the sample schools. In computing the weights, we used the inverse of the final school weight ( $W_3$ ) rather than the school's selection probability; the school weight took into account school-level nonresponse and school-level post-stratification adjustments. In addition, we multiplied the base probability by one-half (since the public-use file included data for only half of the eligible respondents). Thus, the initial weight for cases eligible for both the core and black supplement samples was:

$$W_{1ijk} = \frac{2W_{3i}}{\left( 1 - \left[ \left( 1 - \frac{n_{ij}}{N_{ij}} \right) (1 - P_B) \right] \right)} . \tag{1}$$

The next step in computing weights for the cases eligible for both samples was to calculate a nonresponse adjustment. We calculated weighted nonresponse rates for each grade-sex combination (using  $W_1$  defined in equation [1] above as the weight). Altogether there were

1,318 cases who were selected for the high education black supplement (including 218 cases who were also selected for the core) and an additional 410 core selections who were eligible for selection into the supplement but not actually selected. Of these 1,728 cases, 1,375 completed the Wave I in-home interview and an additional 13 turned out to be ineligible for the interview. The adjusted weight  $W_2$  for each respondent in grade-sex cell  $i$  was simply the base weight divided by the response rate ( $R_i$ ) for that cell:

$$W_{2ijk} = \frac{W_{1ijk}}{R_i}.$$

For nonrespondents, this weight was set to zero.

The final step in computing the Wave I public-use weights was to put the two groups of respondents back together again and to adjust the sum of the weights for the entire sample to agree with estimates from the Census Bureau of the size of each grade-sex-race (black vs. non-black) population. This post-stratification procedure removed any differences between the composition of the sample and the estimated composition of the population.

## **2. Wave II Weights**

Some 5,469 of the Wave I public-use respondents were deemed eligible for follow-up in Wave II. (The remainder were seniors who had graduated after Wave I). Of those fielded during Wave II, 4,834 completed the Wave II questionnaire. Because there was no subsampling of eligible Wave I respondents, the weight for Wave II is simply the final Wave I public-use weight adjusted for additional nonresponse during second wave of data collection. Separate nonresponse adjustments were calculated for each grade-sex cell. The final Wave II weight for a respondent in cell  $i$  was simply the Wave I weight divided by the weighted Wave II response rate for that cell; for the Wave II nonrespondents, the weight is zero.

*Public-use Grand Sample Weights*

Frequency Code	Response	Variable Name	Type/Length
	Respondent Identifier	<b>AID</b>	char 8
	Sample cluster	<b>CLUSTER2</b>	num 3
	Grand Sample Weight - Wave II	<b>GSWGT2</b>	num 8, with decimal places
4834	range 282.4468 to 21107.10		