



**National Social Life, Health and Aging
Project (NSHAP)**

Wave 2 Public Use File v1



May 19, 2014

Contents

Contents

Introduction	8
Unrestricted versus Restricted W2 Data	8
Sample and Weights	9
Variable Coding	9
Acknowledgements and Appropriate Citations	11
Sample Text for Acknowledging NIH Support	11
Suggested Citation for NSHAP Data	11
Reporting a Problem	12

Introduction

This document is intended to accompany the unrestricted, downloadable public use file (PUF) of the second wave of the National Social Life, Health and Aging Project (NSHAP) dataset.

NSHAP is a population-based study of health, social life, and well-being among older Americans. In Wave 1, a nationally- representative probability sample of community-dwelling individuals ages 57-85 was selected from households across the U.S. screened in 2004. African-Americans, Latinos, men and the oldest-old (75-84 years at the time of screening) were over-sampled. In-home interviews were conducted in English and Spanish by professional interviewers from the National Opinion Research Center (NORC) between July 2005 and March 2006, yielding a total of 3,005 respondents (1,455 men and 1,550 women).

A second wave was conducted from August 2010 through May 2011, during which Wave 1 respondents were reinterviewed. An attempt was also made to interview individuals who were sampled in Wave 1 but declined to participate (referred to as *Wave 1 non-respondents*). In addition, spouses or co-resident partners were also interviewed using the same instruments as the main respondent. The resulting Wave 2 dataset contains 3,377 total respondents, including 957 couples.

The overall unconditional response rate for Wave 2 was 74%; the conditional response rate of Wave 1 respondents was 89%; the conditional response rate of partners was 84%; and the conversion rate for Wave1 non-respondents was 26%.

Unrestricted versus Restricted W2 Data

This unrestricted dataset is designed for public dissemination and includes Wave 2 data only. All identifiers have been removed to protect respondents' confidentiality. This file is appropriate for preliminary analysis and exploration of the health and functioning of older Americans.

These data were subjected to a deductive disclosure review before release in order to ensure the

confidentiality of the respondents to the survey while maintaining the quality of the data and the conclusions drawn from them. The data were thoroughly reviewed by NORC and NSHAP staff. To prevent the risk of deductive disclosure, some variables are excluded in this release, including, but not limited to:

- all direct identifiers (including date of birth and names from network roster and marital/partner history)
- actual date and location of interview (month and year are included), language in which interview was conducted, and description of exactly who was present during interview (presence of 3rd party is included)
- all references (either direct or indirect) to geographical location
- all open-ended responses (e.g., further specification of “Other”)
- interviewer’s physical description of the respondent
- assays conducted on biological samples
- any specific dates, including marital and partner history dates
- socio-demographic variables that, when cross-tabbed with other variables, produced cell sizes smaller than 5
- disease-specific variables that reduced cell sizes when cross-tabbed with other easily identifiable variables
- medication use
- network roster

The authoritative results of the NSHAP data are based on the complete data set that is housed at NACDA at ICPSR (<http://www.icpsr.umich.edu/icpsrweb/NACDA/>). These data are available upon request and researchers should contact NACDA directly.

Sample and Weights

NSHAP used a complex, multi-stage area probability sample. The sample design was integrated with that of the Health and Retirement Study (HRS) in order to share the costs of frame development and sample screening between the two projects.

A weight incorporating the inverse probability of selection and a non-response adjustment based on age and urbanicity is contained in the variable `weight_adj`. Researchers wishing to compute design-based variance estimates may use the variables `stratum` and `cluster`. These variables were constructed from the original sampling units for the purpose of variance estimation; the former may be treated as (pseudo) strata and the latter as (pseudo) Primary Sampling Units (PSUs).

Variable Coding

A consistent coding scheme has been used for all variables. In assigning numeric codes, the following rules were used:

1. no/yes responses are always coded 0/1
2. Whenever an item has a natural directionality, this is reflected in the coding even if the responses appeared differently in the actual questionnaire. For example, although the responses to the questions on self-rated health were presented in the order “Excellent”, “Very Good”, etc., the variables in the dataset are coded as:

- 1 poor
- 2 fair
- 3 good
- 4 very good
- 5 excellent

This has been done to facilitate analysis. In all cases, the definitive source for the item as it was presented to the respondent is the questionnaire.

3. Whenever a response set contains a natural zero, this is also reflected in the coding. For example, the variable `hlthvis` (corresponding to the question “During the past 12 months, how many times have you seen a doctor or other health care professional...”) is encoded like this:

- 0 none
- 1 one
- 2 2-3
- 3 4-9
- 4 10-12 (about once a month)
- 5 13-20
- 6 20-30 (about twice a month)
- 7 30 or more

This has also been done to facilitate analysis. In cases where a natural zero is not present, integers beginning with the number one are used.

4. No system missing values (i.e., `sysmiss` or `.`) appear in the data files. Instead, specific reasons for missingness have been identified and labeled using Stata’s extended missing values (i.e., `.a`, `.b`, etc.). The following missing value codes are used consistently throughout all data files:

- `.a` refused
- `.b` don’t know
- `.c` not applicable
- `.d` no answer
- `.e` not returned
- `.f` missing in error
- `.g` insufficient data
- `.h` incomplete interview

All cases where an item was skipped as specified by the questionnaire are marked as “not applicable”. The categories “no answer” and “not returned” apply only to responses obtained from the leave-behind. One respondent broke off the interview and did not complete it; all variables from that point forward for that individual are coded “incomplete interview”. Finally, “missing in error” is used whenever a value should be available but is not (e.g., due to interviewer error, technical error, or some other type of problem); in these cases a note usually appears in the file explaining what happened.

Very few derived variables have been included in this release. However, when a derived variable cannot be computed because of missing information among the constituent items, this is usually coded as “not applicable”. Note also that whenever a derived variable is provided, the Stata code used to derive that variable is also provided in a note attached to the variable.

Acknowledgements and Appropriate Citations

The National Social Life Health and Aging Project is supported by the National Institutes of Health. Each publication, press release, or other document about research supported by an NIH award must include an acknowledgment of NIH award support.

When submitting a publication using NSHAP data, please keep in mind the following conditions of NSHAP data use:

- 1. Acknowledge NIH support**
- 2. Appropriately cite the data used in the methods section of your manuscript**
- 3. Comply with the NIH Public Access Policy.** Compliance with this policy is solely the responsibility of the authors and not NSHAP staff. For more information, see NOT-OD-08-033 and the Public Access website: <http://publicaccess.nih.gov/>.

Sample Text for Acknowledging NIH Support

NSHAP Wave 1:

The National Social Life, Health, and Aging Project is supported by the National Institutes of Health, including the National Institute on Aging, the Office of Women's Health Research, the Office of AIDS Research, and the Office of Behavioral and Social Sciences Research (R01AG021487).

NSHAP Wave 2 and Partners:

The National Social Life, Health and Aging Project is supported by the National Institute on Aging of the National Institutes of Health (R37AG030481; R01AG033903). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Suggested Citation for NSHAP Data from NACDA

NSHAP Wave 1:

Waite, Linda J., Edward O. Laumann, Wendy Levinson, Stacy Tessler Lindau, and Colm A. O'Muircheartaigh. National Social Life, Health, and Aging Project (NSHAP): Wave 1. ICPSR20541-v6. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2014-04-30. doi:10.3886/ICPSR20541.v6

NSHAP Wave 2:

Waite, Linda J., Kathleen Cagney, William Dale, Elbert Huang, Edward O. Laumann, Martha McClintock, Colm A. O'Muircheartaigh, L. Phillip Schumm, and Benjamin Cornwell. National Social Life, Health, and Aging Project (NSHAP): Wave 2 and Partner Data Collection. ICPSR34921-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2014-04-29. doi:10.3886/ICPSR34921.v1

Reporting a Problem

We regret that we cannot respond to questions regarding how to manipulate or analyze the data (if, however, you believe the supplied documentation is inaccurate, incomplete or contains ambiguous information, please report this and we shall fix it). Users with such questions are referred to the user's mailing list corresponding to the software they are using (e.g., Stata users can consult the Stata user's mailing list at <http://www.stata.com/statalist>).

Please report any problems with the dataset or documentation by sending email to NSHAP_PUF@norc.org. Make sure to describe the problem clearly, including specifying exactly which variables and cases are involved. If necessary, someone may contact you for further clarification. Problems will be fixed in a subsequent update of the dataset.