

HEALTH AND RETIREMENT STUDY

Cross-wave 2015-2019 Life History Mail Survey (LHMS):

Harmonized and Aggregated Public Data Resource

Data Description

Prepared by

Marina Larkina, Wenqing Qian, Wenshan Yu, and Jacqui Smith

Version 1.0

January 2025

Conditions of Use

By registering for access to HRS Public Release data, the User agrees to all of the following:

- Make no attempts to identify study participants.
- Not to transfer HRS Public Release data to any third party other than staff or students for whom you are directly responsible except as indicated below.
- Not to allow others to use your username and password to access this site.
- To certify the destruction of any downloaded Public Release data file as well as any data files derived from the downloaded file when requested to do so by the Health and Retirement Study.
- To include the following citation in any research reports, papers, or publications based on Public Release data:

In text:

"The HRS (Health and Retirement Study) is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan."

In references:

"Health and Retirement Study, ([insert Product Name]) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740). Ann Arbor, MI, (year)."

- To include the following citation in any research reports, papers, or publications based on any Public Release data file tagged as "Early" or "Preliminary":

"This analysis uses Early Release data from the Health and Retirement Study, ([insert Product Name]), sponsored by the National Institute on Aging (grant number NIA U01AG009740) and conducted by the University of Michigan. These data have not been cleaned and may contain errors that will be corrected in the Final Public Release version of the dataset."

- Please note that although it is necessary to include the above citations to the HRS in any publications, it is not necessary to associate your publications with the HRS grant in [PubMed](#).
- Provide information regarding any research product (dissertation, thesis, journal article, book, book chapter, report, etc.) based on data obtained from the Health and Retirement Study by sending an electronic copy to hrrpublications@umich.edu.
- Report immediately to the Health and Retirement Study at hrrquestions@umich.edu any disclosure of study participant identity as well as any discovery of flaws or errors in the data or documentation files.

Notify the Health and Retirement Study through use of the update function provided at this site or by electronic mail directed to hrrquestions@umich.edu of changes in your electronic mail address, postal address, telephone number, organizational affiliation or organizational status.

Table of Contents

1. Introduction	4
2. General Information	4
2.1. Overview of the LHMS data collection 2015 to 2019	5
3. File Structure	6
3.1. LHMS Wave Indicators	8
4. Data Quality	9
4.1. Sample Updates	9
4.2. Question harmonization in the cross-wave file	9
4.2.1. Harmonization between LHMS 2015 and LHMS 2017	9
4.2.2. Changes to the LHMS 2015 Educational History grids	10
4.2.3. Harmonization between LHMS 2017 Spring and LHMS 2017 Fall, LHMS 2019 Spring, and LHMS 2019 Fall	11
4.3. Editing Data with Inconsistencies or Implausible Values	12
4.4. Coding Specific Question	12
4.4.1. State Coding Scheme	12
4.4.2. LH22 Primary through High School Education	13
4.4.3. Open End Questions	13
4.4.4. “Other: Specify” Questions	14
5. Distribution Files	16
6. Program Statements	16
6.1. Using the Files with SAS	16
6.2. Using the Files with SPSS	17
6.3. Using the Files with Stata	17
7. Linking Respondents across Time	18
8. Documentation: Codebook	18
9. Registration and Downloading the Data	18
9.1. Registration	18
9.2. Conditions of Use	18
9.3. Publications Based on Data	18
10. If You Need to Know More	18
10.1. HRS Website	19
10.2. Contact Information	19

1. Introduction

The Cross-wave 2015 – 2019 Life History Mail Survey (LHMS) file harmonizes and consolidates public data previously released in separate files for the 2015, 2017 and 2019 waves. The LHMS contains questions about residential, educational, employment and health history, and other important childhood and family events (Table 2). The LHMS is part of the Health and Retirement Study (HRS), which is funded under a cooperative agreement between the National Institute on Aging (NIA) and the Survey Research Center at the University of Michigan. The HRS is designed to study labor force, health, and family transitions of the U.S. population aged 51 and older, and the impact of those transitions on economic resources, claims on structured programs such as Social Security, Medicare, and Medicaid, informal assistance and transfers to and from family members as well as health and well-being in later life.

The National Institute on Aging (NIA) provided funding (grants U01AG009740; R01AG051142) for the LHMS, which was conducted by the Survey Research Center (SRC), at the Institute for Social Research (ISR), at the University of Michigan.

By receiving the dataset, you agree to use it for research and statistical purposes only and make no effort to identify respondents. In addition, you agree to send the bibliographic reference (including a PMCID and/or DOI, if available) or a pdf copy to hrrpublications@umich.edu. See Registration and Downloading the Data (at the end of this document) for additional details.

Data alerts, new file notifications, and updates concerning this file will be placed on the HRS Data Alert web page <http://hrsonline.isr.umich.edu/alerts>.

2. General Information

The questions in the LHMS expand the content of HRS early-life histories and complement information consolidated in the Cross-Wave Childhood family and Childhood Health and the Cross-Wave Lifecourse Experiences of Adversity and Trauma files (<https://hrs.isr.umich.edu/data-products/life-history>). Items in the LHMS questionnaire are semi-harmonized with similar studies around the world, including the English Longitudinal Study of Aging (ELSA, England), the Survey of Health, Ageing and Retirement in Europe (SHARELIFE), and the China Health and Retirement Longitudinal Study (CHARLS). Those studies collected retrospective life history information in computer-assisted in-person interviews. This strategy was not feasible in the HRS due to cost and the steady-state longitudinal panel protocol followed by the HRS. Instead, the HRS designed a self-administered questionnaire to mail to participants in the off years between the biennial HRS core interviews. The initial LHMS, fielded to a subsample of HRS participants

in 2015, included sections on early-life contexts as well as residential and educational histories. The content was expanded to cover additional topics in 2017, in 2019 and additional HRS sample members were added in those years. In December 2021, HRS released a Cross-Wave 2015-2017 Harmonized and Aggregated data file, including data obtained in Fall 2015, Spring 2017 and Fall 2017. In 2019 HRS collected data from additional respondents who either had not previously completed LHMS in 2017 or were not yet invited to participate in the LHMS before (see details below). The goal of the present aggregation work is to extend harmonized data by adding 2019 data files and also to updating a few measures that were not yet available during the previous 2015-2017 aggregation. As a result, the present cross-wave 2015-2019 file aggregates and harmonizes data obtained in Fall 2015, in Spring and Fall 2017, and in Spring and Fall 2019 from respondents in different subsamples of the HRS panel (see below). The cross-wave LHMS file can be linked using the HRS household and person-specific indicators to other public-use HRS data (e.g., the cross-wave Tracker file, the Cross-wave Childhood Family and Childhood Health file, Cross-Wave Lifecourse Experiences of Adversity and Trauma file as well as biennial HRS core waves).

2.1. Overview of the LHMS data collection 2015 to 2019

The LHMS 2015, 2017 and 2019 data were initially released as separate public data files (see links below). Data collection occurred in different waves for pragmatic reasons.

- In **Fall 2015**, in order to take quick advantage of new funding, a short LHMS survey covering only residential and education histories was mailed to eligible English-speaking panel members who had not been assigned to a different mail survey, the HRS Consumption and Activity Mail Survey (CAMS). (Fall 2015: <https://hrsdata.isr.umich.edu/data-products/2015-life-history-mail-survey>).
- In 2017, data collection was split into separate waves to accommodate the steady-state CAMS longitudinal survey collection that occurs in the fall of each off-year. The **Spring 2017** LHMS was sent to mid Baby Boomers (MBB) and older cohorts in CAMS. (Spring 2017: <https://hrsdata.isr.umich.edu/data-products/2017-spring-life-history-mail-survey>).
- HRS panel members who are not in CAMS received the LHMS in the **Fall of 2017** (MBB and older cohorts). Those participants who responded to the 2015 LHMS received a shortened (supplemental) version of the LHMS in 2017 which excluded questions about residential and education histories that were included in the 2015 survey. All others received the full LHMS. (Fall 2017: <https://hrsdata.isr.umich.edu/data-products/2017-fall-life-history-mail-survey>).
- The sample for the **2019 Spring** LHMS consisted of HRS respondents who are part of the CAMS sample and who had not previously completed the LHMS in 2017. This includes persons who were asked but did not respond to the 2017 CAMS survey and participants from the late-Baby Boomer (LBB) birth cohort who entered HRS in 2016 but

were not assigned to the LHMS in 2017. (<https://hrsdata.isr.umich.edu/data-products/2019-spring-life-history-mail-survey>).

- The sample for the **2019 Fall** LHMS included three groups of HRS respondents. The first group consisted of HRS respondents from the late-Baby Boomer cohort (LBB) who had not previously been invited to participate in the LHMS. The second group were HRS respondents who had been invited previously but had never returned a LHMS questionnaire (<https://hrsdata.isr.umich.edu/data-products/2019-fall-life-history-mail-survey-long-questionnaire>). The third group was made up of HRS respondents who had completed the 2015 LHMS but not the 2017 Supplemental LHMS. Those participants received a shortened (supplemental) version of the LHMS which excluded questions about residential and education histories that were included in the 2015 survey (<https://hrsdata.isr.umich.edu/data-products/2019-fall-life-history-mail-survey-supplemental-questionnaire>).

Table 1 provides summary for each wave about dates of data collection (field period), dates for the latest release and sample sizes (as on December 2024). The simple response rate (RR1) is calculated using the American Association for Public Opinion Research (AAPOR) standards. For Fall 2017 and Fall 2019, RR is calculated across respondents who returned either full or supplemental LHMS. The unweighted RR for the 2015-2019 cross-wave LHMS public resource, which includes partial data from 2015, is 60.8% (AAPOR RR2 definition).

Table 1. LHMS Sample Design Summary

Wave	Field Period	Latest Release ^a	N ^b	Response Rate ^c
2015 Fall	Dec 2015 – Aug 2016	Jun 2017 (v1)	6,481	58%
2017 Spring	Jun 2017 – Dec 2017	Feb 2023 (v3)	3,848	74%
2017 Fall	Dec 2017 – Jun 2018	Jun 2023 (v2)	6,660	58%
2019 Spring	Jun 2019 – Dec 2019	Dec 2023 (v3)	1,748	37%
2019 Fall	Oct 2019 – Mar 2020	Apr 2024 (v2)	1,640	24%

Note: a = Latest Release is reported as on December 2024

b = Sample size is based on original unharmonized datasets

c = The simple response rate (RR1)

3. File Structure

Data in this aggregated and harmonized 2015-2019 LHMS file are at the respondent level. The data file contains one record for each respondent ($N = 14,936$) who provided at least one non-missing response to either LHMS 2015, LHMS 2017 Spring, LHMS 2017 Fall Full, LHMS 2019 Spring, LHMS 2019 Fall Full and/or LHMS 2017 / 2019 Fall-Supplement. These data include

proxy responses, in which a spouse or other individual responded to the survey on behalf of the invited HRS sample member. The data file contains 580 variables. Table 2 summarizes the topics covered and their respondent-level source in the different waves, and lists the harmonized variable names related to each topic. Variable names begin with the letters LH and correspond to the question number from the LHMS 2017 Spring questionnaire. Information about the harmonization process is included below.

The data file is called LHMS1519A_R, which stands for Life History Mail Survey 2015-2019, with the “_R” indicating that the file is at the respondent level.

The HRS primary identification variables (IDs) are:

HHID HRS HOUSEHOLD IDENTIFIER

PN HRS PERSON NUMBER IDENTIFIER

Records in the data files are sorted by HHID PN. Identification variables are stored in character format. For further information about HRS identification variables and merging data, see

<https://hrsdata.isr.umich.edu/data-products/cross-wave-tracker-file>.

Table 2. Overview of Harmonized Content in the Cross-Wave File ^a: Data sources from 2015 to 2019

LHMS Sections	Variable Content Groups ^b	File Variable Names	2015 LHMS Fall	2017/19 LHMS Spring	2017/19 LHMS Fall Full	2017/19 LHMS Fall Supp ^c
Residential History	Atypical contexts before/after age 16	LH1, LH2, LH4	✓	✓	✓	
	Age when left parental home	LH3	✓	✓	✓	
	Residences from birth to present	LH5	✓	✓	✓	
	Household composition at age 10	LH6 – LH11	✓	✓	✓	
	Household composition at first job	LH13 – LH17	✓	✓	✓	
	Household composition at age 40	LH18 – LH21	✓	✓	✓	
Education History	K-12 schools	LH22	✓	✓	✓	
	Issues in elementary and high school	LH24 – LH28	✓	✓	✓	
	Extracurricular and sport activities	LH29 – LH33	✓	✓	✓	
	Colleges and professional training	LH35	✓	✓	✓	
Partnership History	Grids of marriages and partnerships	LH36, LH37		✓	✓	✓
Work History	Family-work arrangements	LH38 – LH40		✓	✓	✓
	Industry/job title history ^d	LH41		✓	✓	✓
	Self-defined most important job at age 30-40	LH42 – LH49		✓	✓	✓
	Job Quality of important job age 30-40	LH50		✓	✓	✓
Caregiving History	Start/stop spells of unpaid caregiving	LH51		✓	✓	✓
Health	Health conditions and gynaecological illnesses	LH52, LH53		✓	✓	✓
	Major surgeries before age 50	LH54		✓	✓	✓

	Therapy for drug/alcohol use	LH55		✓	✓	✓
	Injuries in major accidents before 50	LH56		✓	✓	✓
	Spells of ill health/disability after age 16	LH57		✓	✓	✓
Physical activity	Frequency of vigorous and moderate-level activities, ages 18-29, 30-39, and 40-49	LH58, LH59	✓	✓	✓	
Important achievements	Thematic content codes (e.g., education, family, work, health, adversity)	LH60		✓	✓	✓

Note: a = Content is harmonized to the LHMS 2017 Spring

b = Variable Content groups listed in the order questions are in the LHMS 2017 Spring

c = LHMS Fall 2017/19 Supplements were completed by Rs who participated in Fall 2015

d = Categories from 2010 Census, consistent with HRS public employment data

Of the 14,936 total respondents in the 2015-2019 aggregated file, n = 13,877 (92.9%) provided completed data; n = 1,057 (7.1%) provided partial data by completing questions LH1 – LH35 only; and 2 respondents filled only LHMS supplemental questionnaire (LH35 – LH60).

Respondents with partial data (LH1 – LH35 only) have special codes (97, 997, 99997) for LH35 – LH60 to indicate that they were not asked those questions (e.g., 97. DID NOT COMPLETE 2017/2019 FALL SUPP).

3.1. LHMS Wave Indicators

There are two indicators for the LHMS waves (see Table 3): “LHMS” provides a code to indicate 2015, 2017 Spring, 2017 Fall, 2019 Spring or 2019 Fall waves. A more detailed variable “LHMSIWIND” indicates the LHMS interview completion.

Table 3. Indicators for LHMS Waves in Harmonized Data File

LHMS	LHMSWIND	n	Q1-Q35	Q36-Q61
1	1	5212	2015 (Q1-Q35)	2017 Fall Supplement (Q36-Q61)
1	2	205	2015 (Q1-Q35)	2019 Fall Supplement (Q36-Q61)
1	3	1057	2015 (Q1-Q35)	
1	4	2		2017 Fall Supplement (Q36-Q61)
2	5	3836	2017 Spring (Q1-Q61)	
3	6	1442	2017 Fall Full (Q1-Q61)	
4	7	1748	2019 Spring (Q1-Q61)	
5	8	1434	2019 Fall Full (Q1-Q61)	
	Total	14936		

Note. LHMS = "LHMS STUDY"

1 = 2015 and/or 2017/2019 Fall Supplement

2 = 2017 Spring

3 = 2017 Fall Full

4 = 2019 Spring

5 = 2019 Fall Full

LHMSIWIND = "LHMS WAVE INDICATOR"

1 = completed 2015 and 2017 Fall Supplement

2 = completed 2015 and 2019 Fall Supplement

3 = completed 2015 but not 2017 or 2019 Fall Supplement

4 = completed 2017 Fall Supplement but was not in the 2015 interviewed sample

5 = completed 2017 Spring

6 = completed 2017 Fall Full

7 = completed 2019 Spring

8 = completed 2019 Fall Full

4. Data Quality

4.1. Sample Updates

During the harmonization process of the initial LHMS data across 2015, 2017/2019 Spring, and 2017/2019 Fall waves, we made a few changes in the samples, compared to the unharmonized datasets. The changes were made to account for cases when respondents filled in more than one questionnaire across waves. In most cases, we selected the most recent questionnaire to include in the harmonized data and removed other copies ($n = 15$). Two respondents who completed the LHMS 2017 Fall Supplement but did not complete the LHMS 2015 questionnaire were added to the sample (those respondents have LHMSWIND = 4, see Table 3). In addition, we included in the harmonized data those participants that were available at the time of aggregation (December 2024). Respondents that are added to the LHMS individual data files in later versions (i.e., late-returned questionnaires cases) will be added at the follow-up updates for the harmonized data when next the LHMS waves are released (i.e., 2023 waves).

Respondents' identification variables (HHID and PN) are based on the HRS 2022 Early Release Tracker file (v2.0; November 2024) and reflect the current respondents' status. Two respondents, who belong to "household merge overlap" (see Tracker data description <https://hrsdata.isr.umich.edu/data-products/cross-wave-tracker-file>), have HHID/PN in the aggregated data file that are different from their identification variables in the original unharmonized LHMS data files. They are: HHID = 906543, PN = 020 (was HHID=906126, PN=010; first year of new HHID/PN is 2018); and HHID = 541110, PN = 020 (was HHID = 541986, PN = 010; first year of new HHID/PN is 2020).

4.2. Question harmonization in the cross-wave file

4.2.1. Harmonization between LHMS 2015 and LHMS 2017

For the harmonized data, we used the LHMS 2017 Spring questionnaire for the question order, variable names and labels, answer choices and coding categories (with few exceptions, see below). Because the LHMS 2015 questionnaire could be considered as a pilot questionnaire, there were some differences between the LHMS 2015 and LHMS 2017 questionnaires. Those differences are found in questions Q1 – Q35, because the LHMS 2015 questionnaire covered only Residential History and Educational History (see Table 4 for the list of issues and corrections).

Table 4. List of Question Harmonization Issues and Corrections

	Issue	Correction
1	Questions are in the LHMS 2015 but not in the LHMS 2017	These questions are not included in the harmonized data (data from these questions are available in the public data for the LHMS 2015)
2	Questions are not in the LHMS 2015 but are in the LHMS 2017	Responses for the respondents from the LHMS 2015 were coded as missing and indicated in the codebook as “Blank. MISSING/NOT ASKED IN 2015”
3	Questions are in the LHMS 2015 and 2017 but wording, response scale, or instructions are different	LHMS 2017 wording is used in the harmonized data
4	Questions are in the LHMS 2015 and 2017 but order is different	LHMS 2017 question order is used in the harmonized data

4.2.2. Changes to the LHMS 2015 Educational History grids

Educational History data (K-12 school grids) in the LHMS 2015 were collected separately for Primary/Elementary schools (Q20) and for Junior/Middle and High schools (Q21). We harmonized LHMS 2015 educational grids with LHMS 2017 data by merging LHMS 2015 Q20 and Q21 into one grid to match the 2017 variables. During this process, we removed extra empty rows and renamed variables. The question SCHOOL END GRADE was not asked in LHMS 2015 and was coded as Blank. MISSING/NOT ASKED IN 2015. Three questions from the LHMS 2015, which are not in the LHMS 2017, were not included in the harmonized data file: LH20_1E_15 – LH207E_15 ELEMENTARY SCHOOL NUMBER IN CLASS; LH21_1E_15 - LH21_7E_15 HIGH SCHOOL HOW FAR, and LH21_1F_15 - LH21_7F_15 HIGH SCHOOL HOW TRAVEL.

For College / Community Colleges and Professional or Technical schools or programs (Q35 in the LHMS 2017 and Q23 in LHMS 2015), college degrees (2017 LH35_1G_17 - LH35_7G_17 and LH23_1F_15 - LH23_7F_15) were coded using different categories in the previously released public data files. During harmonization, the LHMS 2015 college degree categories were recoded to match the LHMS 2017 categories. After recoding, respondents from 2015 did not contribute any observations to the categories 9, 10, or 11 in the harmonized data (9. SPECIALIST DEGREE BEYOND MASTER'S; 10. CURRENTLY ENROLLED; 11. CONTINUING EDUCATION).

4.2.3. Harmonization between LHMS 2017 Spring and LHMS 2017 Fall, LHMS 2019 Spring, and LHMS 2019 Fall

Starting with 2017 Spring, LHMS questionnaires are the same with few exceptions. First, in LHMS 2019 Spring and LHMS 2019 Fall 23 additional questions were included: 9 questions about childhood experience of adversity (LH4_AA – LH4_AI), and 14 questions about lifetime traumas and major experiences of discrimination (LH4_BA – LH4_BN). Those questions are not included in the cross-wave 2015-2019 harmonized LHMS data file but are available in the separate public data files (LHMS 2019 Spring and LHMS 2019 Fall) and in the Cross-wave Lifecourse Experiences of Adversity and Trauma aggregated data (2024, v1.0, <https://hrsdata.isr.umich.edu/data-products/cross-wave-lifecourse-experiences-adversity-and-trauma-aggregated-data>).

Second, starting with LHMS 2017 Spring (released 2023, v3), HRS included additional variables to indicate whether a respondent listed more than one job for the same time period:

LH41_xFLM. LAST JOB TITLE MASKED (it is called “PROMOTION TITLE MASKED” in the original LHMS 2017 Spring public data). Starting with LHMS 2017 Fall (released 2023, v2), HRS included another variable to describe the additional reported jobs:

LH41_xELM. TYPE BUSINESS, LAST JOB MASKED. Because this measure is not available for LHMS 2017 Spring respondents (n = 3836), they received a code 996. NOT CODED IN 2017 SPRING. Those two measures were not available in the 2015-2017 cross-wave aggregated data and earlier releases of the LHMS 2017 separate public data files, but they are included in the cross-wave 2015-2019 harmonized LHMS data file.

Other changes are related to additional codes for some variables that were introduced starting with LHMS 2019 (Spring and Fall). Those codes are available only for respondents who filled LHMS in 2019 waves (2019 Spring and Fall). Specifically:

1. One new code for State variables: “-7. MORE THAN ONE STATE/COUNTRY MENTIONED” for: LH5_xB. RESIDENCE STATE; LH22_xC. SCHOOL STATE; LH35_xC. POST HIGH SCHOOL STATE; LH44. JOB STATE. Multiple states were not coded for 2015 and 2017 waves.

2. Two new codes for Type of Residence variables: “4. MILITARY HOUSING/BARRACKS;” and “5. ROOMING HOUSE/BOARDING HOUSE/YMCA/DORMITORY” for LH9. TYPE OF RESIDENCE AGE 10 and LH20. TYPE OF RESIDENCE AGE 40. For respondents who filled LHMS in 2015 and 2017 waves, those types of residency were included in 97. OTHER category and not released as separate codes.

3. One new code 12. MILITARY for LH18Mx. LIVED WITH AGE 40. For respondents who filled LHMS in 2015 and 2017 waves, this category was part of the 97. OTHER category and not released as a separate code.

4.3. Editing Data with Inconsistencies or Implausible Values

While cleaning the data, when extreme, unlikely or conflicting values were encountered, electronic images of the questionnaires were checked directly 1) to determine if errors were introduced during the data entry/scanning process and 2) to possibly correct those values. The instances of conflicting values include where a start year is greater than an end year, or a start age is greater than an end age. Inconsistencies were detected for thematically related questions, especially for the questions with a skip pattern instruction. For example, a respondent could select NO for Q1. LIVED OUTSIDE US but answered either Q1A. LIVED IN HOW MANY COUNTRIES or any of the Q1B. LIVED OUTSIDE US BETWEEN questions. We reviewed respondents' electronic images of the questionnaires and, when possible, corrected detected inconsistencies. In cases of inadequate or ambiguous information, responses were not changed. We did not include a skip question Q29 "Did you go to high school?" in the harmonized data file because respondents often marked "no" but answered questions that followed (e.g., Q29a – Q33) that indicated that they did (and the HRS Tracker file reports number of school years consistent with having attended high school).

We also identified and edited values that were outside of expected or plausible value ranges. For example, for the question Q3A. AGE ESTABLISHED OWN HOME all values that were smaller than 16 were reviewed and, when possible, corrected based on relevant information provided in the questionnaires. In cases of inadequate or ambiguous information, values were recoded to 96. IMPLAUSIBLE VALUE.

Please note, that for Q22 (Educational history) Start/End age variables, there are some out-of-expected range values of 0 or younger than age 4. Users are advised to refer to the HRS restricted data (<https://hrs.isr.umich.edu/data-products/restricted-data>) where data were comprehensively checked and corrected and included additional information about school city and state.

4.4. Coding Specific Question

4.4.1. State Coding Scheme

Respondents were asked to provide the state for their residency (LH5), education (LH22, LH35), and employment history (LH44). The state variable in these series is coded as follows: beginning with LH5, the first state mentioned is coded as 1, the second state is coded as 2, and so on for each unique state mentioned. If the first state mentioned is repeated on a later line in LH5, it is coded as 1, if the second state mentioned is repeated on a later line it is coded as 2, etc. Mentions of foreign countries are coded as 96. The same code frame is maintained across LH22, LH35 and LH44 within each individual's data. For example, a 1 in the state variables in the education and employment series represents the same state as the 1 in the LH5 residential series. Please note, that depending whether respondents filled full or supplemental versions of LHMS, the different coding was applied for the question Q44, "*In what city and state or country did you work for this employer or business?*". For respondents who filled a full version of LHMS, state information is

provided in Q44. JOB STATE and coded as an order mentioned in LH5 (see above). This code frame is not available for respondents who filled the 2017 and 2019 supplemental version of the LHMS, they received a code 999. STATE - MASKED (see LH44M). For respondents who filled the 2017 and 2019 supplemental version of the LHMS, state information of employer is provided in LH44M. STATE – MASKED and coded as Census defined regions where states categories have been collapsed to protect participant confidentiality (codes 1 – 9: 1. Northeast Region: New England Division; 2. Northeast Region: Middle Atlantic Division, etc.,). Mentions of foreign countries are coded as 11. This code frame is not available for respondents who filled a full version of LHMS; they received a code 999. STATE CODED IN ORDER MENTIONED (see LH44).

4.4.2. LH22 Primary through High School Education

For the Educational History grid, respondents were instructed to write the start grade, end grade, start age and end age for each school they attended from Kindergarten through high school. In the LH22 series, a number of respondents provided information for post high school education or training (rather than writing those in the LH35 series). In these cases, if the “School Name” field in LH22 clearly indicated a post-high school institution, the response was moved to LH35; otherwise it was left in LH22.

4.4.3. Open End Questions

Responses were reviewed for six open end questions:

- LH35f, major field of study in post high school education
- LH35g, degree earned in post high school education
- LH35h, certificate earned in post high school education
- LH54, major surgeries
- LH57a, periods of ill health or disability
- LH60, most important accomplishments

Responses to the LH35 series, post high school training or education, were coded based on the Classification of Instructional Programs (CIP) categories, provided by the National Center for Educational Statistics (<https://nces.ed.gov>). The codes for LH35FM, major field of study, were then grouped into larger categories to protect respondent confidentiality, as shown below:

LHMC code	NCES CIP code	NCES Description
1	52	BUSINESS, MANAGEMENT, MARKETING, AND RELATED SUPPORT SERVICES
2	9	COMMUNICATION, JOURNALISM
3	13	EDUCATION
4	51	HEALTH PROFESSIONS, RESIDENCY PROGRAMS, AND RELATED PROGRAMS
4	60	RESIDENCY PROGRAMS
5	16	FOREIGN LANGUAGES, LITERATURES, AND LINGUISTICS
5	23	ENGLISH LANGUAGE AND LITERATURE/LETTERS

5	24	LIBERAL ARTS AND SCIENCES, GENERAL STUDIES AND HUMANITIES
5	38	PHILOSOPHY AND RELIGIOUS STUDIES
5	39	THEOLOGY AND RELIGIOUS VOCATIONS
5	50	VISUAL AND PERFORMING ARTS
5	54	HISTORY
6	22	LEGAL PROFESSIONS AND STUDIES
6	28	MILITARY SCIENCE, LEADERSHIP, OPERATIONAL ART
6	29	MILITARY TECHNOLOGIES AND APPLIED SCIENCES
6	43	HOMELAND SECURITY, LAW ENFORCEMENT, FIREFIGHTING AND RELATED PROTECTIVE SERVICES
7	1	AGRICULTURE, AGRICULTURE OPERATIONS, AND RELATED SCIENCES
7	3	NATURAL RESOURCES AND CONSERVATION
8	4	ARCHITECTURE AND RELATED SERVICES
8	14	ENGINEERING
8	15	ENGINEERING TECHNOLOGIES AND ENGINEERING-RELATED FIELDS
8	26	BIOLOGICAL AND BIOMEDICAL SCIENCES
8	27	MATHEMATICS AND STATISTICS
8	40	PHYSICAL SCIENCES
9	5	AREA, ETHNIC, CULTURAL, GENDER, AND GROUP STUDIES
9	19	FAMILY AND CONSUMER SCIENCES/HUMAN SCIENCES
9	42	PSYCHOLOGY
9	44	PUBLIC ADMINISTRATION AND SOCIAL SERVICE PROFESSIONS
9	45	SOCIAL SCIENCES
10	10	COMMUNICATIONS TECHNOLOGIES/TECHNICIANS AND SUPPORT SERVICES
10	11	COMPUTER AND INFORMATION SCIENCES AND SUPPORT SERVICES
10	25	LIBRARY SCIENCE
11	12	PERSONAL AND CULINARY SERVICES
11	31	PARKS, RECREATION, LEISURE, AND FITNESS STUDIES
11	36	LEISURE AND RECREATIONAL ACTIVITIES
12	46	CONSTRUCTION TRADES
12	47	MECHANIC AND REPAIR TECHNOLOGIES/TECHNICIANS
12	48	PRECISION PRODUCTION
12	49	TRANSPORTATION AND MATERIALS MOVING
12	53	HIGH SCHOOL/SECONDARY DIPLOMAS AND CERTIFICATES
13	96	OTHER

Responses to the employment series beginning at LH41 were coded based on the same 2010 US Census code frames used to code J166M and J168M (Section J) in the core interview. These codes were then grouped into larger categories to protect respondent confidentiality, also using the same masking code used in the HRS core study.

Responses to the health questions (LH54 and LH57), and to major accomplishments (LH60), were coded based on a content analysis of response frequencies.

4.4.4. “Other: Specify” Questions

Responses were reviewed for all questions with an “Other: Specify” option, and were either coded into existing code frames, or new codes were created. For the six questions below, the following new code frames were created:

LH14 *Who lived with you when you started your first full-time job?*

- 11. Children, unspecified
- 12. In the military

LH16 *What type of residence did you live in when you started first full-time job?*

- 4. Military housing/barracks
- 5. Rooming house, boarding house/YMCA/dormitory

LH18 *Who lived with you when you were 40?*

- 11. Children, unspecified

LH33b *In what sports did you participate (in junior/middle and high school)?*

- 12. Softball
- 13. Wrestling
- 14. Cheerleading, drill team

LH41d *What did you do after leaving this job?*

- 6. Retired, took buy out package
- 7. Pursued further education
- 8. Started own business/became self-employed
- 9. Moved/relocated
- 10. Joined military
- 11. Still work at this job

LH49 *Why did you leave or stop working for this employer?*

- 11. The company closed, went out of business, or relocated
- 12. Started own business/became self-employed
- 13. Did not like it; problems with the owners; problems with coworkers; hours were cut
- 14. Moved to a different job, reason unspecified
- 15. Divested ownership in the business/sold the business

5. Distribution Files

The *Cross-wave 2015-2019 Life History Mail Survey* data are packaged for distribution in a .ZIP file. Extract the data file, the program statement file that matches your analysis environment, the data description (this file), and the codebook file. If you have problems when downloading this data set or in extracting its contents, please contact the HRS Help Desk (<mailto:hqsquestions@umich.edu>).

The following extensions are used for the six different types of distribution files:

- .DA for data files
- .SAS for SAS program statements
- .SAS7BDAT for “ready-to-use” SAS files
- .SPS for SPSS program statements
- .SAV for “ready-to-use” SPSS files
- .DO for Stata DO statements
- .DCT for Stata dictionary statements
- .DTA for “ready-to-use” Stata files
- .TXT for codebook files

For example,

- LHMS1519A_R.DA contains ASCII data
- LHMS1519A_R.SAS contains corresponding SAS program statements
- LHMS1519A_R.sas7bdat contains “ready-to-use” data in SAS format,
- LHMS1519A_R.SPS contains corresponding SPSS program statements
- LHMS1519A_R.sav contains “ready-to-use” data in SPSS format,
- LHMS1519A_R.DO contains corresponding Stata DO statements
- LHMS1519A_R.DCT contains corresponding Stata dictionary statements
- LHMS1519A_R.da contains “ready-to-use” data in Stata format,
- LHMS1519A_R.TXT contains the ASCII codebook

6. Program Statements

The data are provided as SAS (.sas7bdat), SPSS (.sav) and Stata (.dta) “ready-to-use” files and in ASCII format. Each ASCII data file comes with associated SPSS, SAS, or STATA program statements to read the data. Files containing SPSS statements are named with a .sps extension, those with SAS statements with a .sas extension, and those with STATA statements with .do and .dct extensions.

6.1. Using the Files with SAS

To create a SAS system file for a particular dataset, two file types must be present for that dataset -- .SAS program statement files and .DA data files.

To create a SAS system file, load the *.SAS file into the SAS Program Editor. If the *.SAS file is located in "c:\ LHMS1519 \sas" and the data file is located in "c:\ LHMS1519\data", you can run the file as is. A SAS system file (*.SD2 or *.SAS7BDAT) will be saved to directory "c:\LHMS1519\sas".

If the files are not located in the specified directories, you will need to edit the *.SAS file to reflect the proper path names prior to running the file.

6.2. Using the Files with SPSS

To create an SPSS system file for a particular dataset, two file types must be present for that dataset -- .SPS program statement files and .DA data files.

To create an SPSS system file, open the *.SPS file in SPSS as an SPSS Syntax File.

If the *.SPS file is located in "c:\ LHMS1519\spss" and the data file is located in "c:\ LHMS1519\data", you can run the file as is. An SPSS system file (*.SAV) will be saved to directory "c:\ LHMS1519\spss".

If the files are not located in the specified directories, you will need to edit the *.SPS file to reflect the proper path names prior to running the file.

6.3. Using the Files with Stata

To use Stata with a particular dataset, the following three file types must be present for that dataset -- .DCT files, .DO files, and .DA data files.

Files with the suffix .DA contain the raw data for Stata to read. Files with the suffix .DCT are Stata dictionaries used by Stata to describe the data. Files with the suffix .DO are short Stata programs ("do files") which you may use to read in the data. Load the .DO file into Stata and then submit it.

If the *.DO and *.DCT files are located in "c:\ LHMS1519\stata" and the data file is located in "c:\ LHMS1519\data", you can run the .DO file as is.

If the files are not located in these directories, you must edit the *.DO and *.DCT files to reflect the proper path names before you run the files.

Note that the variable names provided in the .DCT files are uppercase. If you prefer lower case variable names, you may wish to convert the .DCT files to lower case prior to use. You may do this by reading the .DCT file into a text or word processing program and changing the case. For instance, in Microsoft Word, Edit, Select All, Format, Change Case, lowercase.

7. Linking Respondents across Time

Respondent records in the *Cross-wave 2015-2019 Life History Mail Survey* file can be linked to respondent records in the cross-wave Tracker file and biennial core HRS waves by HHID and PN. The core sub-household identifiers can be used to link household data with the cross-sectional respondent level data.

8. Documentation: Codebook

The *Cross-wave 2015-2019 Life History Mail Survey* codebook is provided as an ASCII text file. Each variable has its own codebook entry that includes variable name, label, variable type (e.g., character, numeric). Variable names begin with the letters LH to designate the Life History Mail Survey and correspond to the question number found in the LHMS questionnaire. For example:

LH1A Q1A. LIVED IN HOW MANY COUNTRIES

9. Registration and Downloading the Data

9.1. Registration

The HRS data are available for free to researchers and analysts on the HRS web site. In order to obtain public release data, you must first register <https://hrsdata.isr.umich.edu/user/>. Once you have completed the registration process, a username and password will be sent to you via e-mail. Your username and password are required to download data files.

Registered users receive user support, information related to errors in the data, future releases, workshops, and publication lists. The information you provide will not be used for any commercial use, and will not be redistributed to third parties.

9.2. Conditions of Use

By registering, you agree to the [Conditions of Use](#) governing access to Health and Retirement public release data.

9.3. Publications Based on Data

As part of the data registration process, you agree to include specified citations and to inform HRS of any papers, publications, or presentations based on HRS data. Please send a copy and full citation of any publications you produce based on HRS data to hrspublications@umich.edu.

10. If You Need to Know More

This document is intended to serve as a brief overview and to provide guidelines to using the *Cross-wave 2015-2019 Life History Mail Survey* data. Additional information about the HRS and the LHMS can be obtained from the HRS web site. If you have questions or concerns that are not adequately covered here or on our web site, please contact us at hqsquestions@umich.edu. We will do our best to provide answers.

10.1. HRS Website

Public release data from the Health and Retirement Study and additional information about the study are available at <http://hrs.isr.umich.edu/>.

10.2. Contact Information

For specific questions, please contact the helpdesk at hqsquestions@umich.edu.

The HRS mailing address and fax are:

Health and Retirement Study
The Institute for Social Research
426 Thomson Street
Ann Arbor, MI 48106-1248
FAX: (734) 647-1186