

CenHRS Crosswalk User Guide

Updated June 2025

The development of the CenHRS is supported by a grant from the Alfred P. Sloan Foundation (2011-6-21) with additional support from the National Science Foundation NCRN project (SES 1131500). 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.

Contents

1 Acronyms and Abbreviations	2
2 Project Overview	3
3 CenHRS Data	3
3.1 Data Inputs to the CenHRS	3
3.1.1 The HRS	4
3.1.2 The Business Register (BR)	5
3.2 CenHRS Record Linkage Procedure	5
3.3 CenHRS Crosswalk Variables	6
4 Using the CenHRS Crosswalk	7
4.1 Linking to HRS Datasets	7
4.2 Linking to Census Employer and Establishment Files	8
5 Data Access	11
5.1 Application Procedures	11
5.2 Reporting Project Progress and Modifications	12
6 Data Citation and Acknowledgement of Funding	12
6.1 Data Citation	12
6.2 Acknowledgment of Funding	12
7 Contact Information	12
8 References	13
9 Additional References	14

1 Acronyms and Abbreviations

Acronym/ Abbreviation	Description
AHEAD	Assets and Health Dynamics Among Oldest Old
BR	Business Register
CenHRS	Census-Enhanced Health and Retirement Study
CFN	Census File Number
EIN	Employer Identification Number
FSRDCs	Federal Statistical Research Data Centers
HRS	Health and Retirement Study
IRS	Internal Revenue Service
JCF	HRS Cross-Wave Harmonized Job Characteristics File, or simply Job Characteristics File
LBD	Longitudinal Business Database
LEHD	Longitudinal Employer-Household Dynamics
ML-MI	Machine learning and multiple imputation
MI	Multiple imputation
SAS	Formerly Statistical Analysis System, a statistical software package
SSA	Social Security Administration
SSS	Special Sworn Status

2 Project Overview

The Census-Enhanced Health and Retirement Study (CenHRS) project links Health and Retirement Study (HRS) respondents to data on a broad range of characteristics of their employers. The purpose of integrating the HRS with Census employer data is to enhance understanding of the labor market dynamics of older Americans. To create this data infrastructure, the CenHRS project compiled and integrated reports by HRS respondents about their employers to create job histories for HRS respondents from the time they enter the HRS survey. Using these job histories and Employer Identification Numbers (EINs) from Social Security Administration (SSA) earnings records for respondents who consented to link them to the HRS, the CenHRS project used machine learning methods to match HRS respondents to their employers in the Census Bureau's Business Register (BR). The project then used this model to develop a probabilistic crosswalk between the HRS and the BR, incorporating uncertainty from the machine learning matching process. The crosswalk consists of two files: one linking HRS respondents to firms, and the other linking HRS respondents to establishments.

The crosswalk is available to the research community through [Federal Statistical Research Data Centers \(FSRDCs\)](#) with permission from the Census Bureau and the Health and Retirement Study at the University of Michigan, so that researchers with Census-approved projects may conduct research using Census Bureau economic and business data.

In addition to employer information available in the BR, the crosswalk permits linkage of the HRS to other datasets linked to the BR. These include Census-held datasets such as Longitudinal Employer-Household Dynamics (LEHD) data and Census Bureau business surveys. The crosswalk also permits linkage of the HRS to non-Census datasets that are themselves linkable to the BR, such as Compustat or Form 5500 data, for users authorized to access that data. More information on Census-held datasets can be found on [ResearchDataGov.org](#).

The CenHRS creates a distinctive data infrastructure combining business data from the Census Bureau and household data from the HRS. The HRS includes data on wealth, income, work, demographics, family structure, expectations, health, and cognition; modules on a wide range of economic and health topics; links to other administrative data including Social Security earnings records and company pension plans; and biomarkers. The Census data include information on organizational structure, firm size, employment, industry, location, factor inputs, and international transactions, as well as data from specialized Census Bureau surveys. They also include characteristics of the firm's workforce, including, for example, earnings, age, race, gender, immigrant status, and employment tenure.

3 CenHRS Data

3.1 Data Inputs to the CenHRS

The CenHRS project links HRS respondents to data on a broad range of characteristics of their employers by linking to the BR, which covers the universe of private sector employers in the United States. This section details how HRS and BR data are used to build the CenHRS crosswalk.

3.1.1 The HRS

The [HRS](#) is a nationally representative longitudinal household survey that collects data from more than 22,000 Americans over the age of 50 every two years. It is a large-scale longitudinal project that studies the labor force participation and health transitions that individuals undergo toward the end of their work lives and in the years that follow. It includes the “original” HRS study (data collection in 1992, 1994, and 1996) and the Assets and Health Dynamics Among the Oldest Old (AHEAD) study (data collections in 1993 and 1995). These studies were merged in 1998 and now represent the U.S. population over age 50 in 1998. Two new cohorts were added in 1998: the Children of the Depression (born 1924-1930) and the War Babies (1942-1947). A fourth cohort, the Early Baby Boomers (1948-1953), was added in 2004; a fifth cohort, the Mid Baby Boomers (1954-1959), was added in 2010; and in 2016, the Late Baby Boomers cohort (1960-1965) became the sixth (Health and Retirement Study, 2020; Sonnega et al., 2014). HRS data are collected via various modes of data collection, including face-to-face and phone interviews and mail questionnaires. For more information on HRS data collection, see Sonnega et al. (2014).

The HRS collects information on respondents’ main jobs at the time of interview, prior jobs, and last jobs. However, the CenHRS crosswalk includes only respondents’ current main jobs at the time of interview.

To create the CenHRS crosswalk, it is first necessary to compile and harmonize job information across survey waves. This is needed because, in light of the longitudinal nature of the HRS, the characteristics of main jobs are collected at different survey waves: some characteristics are collected only at the first time the job is reported, some are collected every wave, and some are collected after the job ended. The Health and Retirement Study Cross-Wave Harmonized Job Characteristics File (JCF) achieves this by linking main jobs across waves and imputing relevant variables that are missing or inconsistent. The file contains harmonized information on HRS respondents’ current main jobs at the time of the interview at the person-wave level. For more information on the JCF, see Abramowitz et al. (2023).

Only jobs in which respondents report working for someone else are eligible for linkage (i.e., respondents reporting that they are self-employed are excluded). The table that follows summarizes the number of HRS respondents eligible to be included in the CenHRS crosswalk in each survey wave:

Number of HRS Respondents Reporting Working for Someone Else by Survey Wave

Wave	1992	1994	1996	1998	2000	2002	2004	2006	2008	2010	2012	2014	2016	2018
Obs.	6,845	6,141	5,293	6,523	5,643	4,630	6,068	5,079	4,447	7,028	6,284	5,396	6,971	5,408

3.1.2 The BR

The BR is a comprehensive database that contains a record for each known legal establishment that is located in the U.S. or its territories and has employees, except employers classified as private households. Accordingly, the BR includes the universe of private sector employers in the

United States. It is updated using information from several federal agencies, including the Internal Revenue Service, the Bureau of Labor Statistics, the Social Security Administration, and the Census Bureau. Variables available on the BR include payroll, employment, sales and revenue, industry, and firm structure. The BR also includes geographic information and firm and establishment identifiers, which are vital to the creation of the CenHRS crosswalk. The firm and establishment identifiers on the BR can also be used to link it to select other firm- and establishment-level datasets. For more information on the BR see DeSalvo, Limehouse, and Klimek (2016).

3.2 CenHRS Record Linkage Procedure

This section draws extensively from Abowd et al. (2024), which describes the record linkage methodology underpinning the CenHRS.

The CenHRS linkage matches HRS respondents' employers and establishments for their current main jobs at the time of their HRS interview to the BR. The establishment is the workplace, and a given firm may operate many establishments. The linkage uses distinct models and creates distinct crosswalk files for employers and establishments.

The majority of HRS respondents' jobs are linked to the BR using EINs. HRS respondents gave permission to the Social Security Administration (SSA) to provide earnings records, which include EINs, to the HRS for purposes of enhancing the HRS data infrastructure. Over the 1992-2020 waves, 75.2 percent of HRS respondents gave permission to the SSA to provide earnings records, and 68.4 percent of HRS respondents are matched to their earnings records. For more information on the consent process, see Abramowitz, Fang, and Hyde (2024).

For respondents that are not linked to their earnings records, their jobs are linked to the BR using employer names and establishment addresses that they provide as part of the survey. The HRS elicits information about employer name, establishment address, and telephone number for the respondent's main job. Respondent reports of employer identity and address are obtained at the survey baseline (i.e., when new respondents are enrolled in the study, generally every six years when a new cohort is added to the study) and in each subsequent wave if the respondent reports having changed jobs. The job spell is defined using the job spell start year and job spell end year variables in the JCF.

To create the CenHRS employer and establishment linkages, two procedures are employed. First, for each crosswalk file, a deterministic match based on EIN is used with respondents for whom HRS has valid EINs that can be matched to an establishment in the BR. This case happens when the respondent has just one job (and therefore just one EIN) in a given year and that EIN corresponds to exactly one establishment in the BR.

Second, a machine learning and multiple imputation (ML-MI) procedure is used to produce a dataset that contains multiply-imputed matches. This approach is used when respondents consent to SSA linkage, but cannot be deterministically matched to an establishment in the BR. This case happens either because the respondent has multiple jobs in a given year (and consequently has multiple EINs), or because the respondent's EIN is associated with multiple establishments in the

BR, or both. This approach is also used for respondents for whom HRS does not have valid EINs. Further details about the linkage procedure can be found in Abowd et al. (2024).

Each CenHRS crosswalk file is created independently, and should be used independently. That is, no relationship should be assumed between the firms and establishments to which a given person is linked. Each file is structured with 10 implicate links associated with each respondent-year to properly account for the additional variability introduced by probabilistic matching. For a given respondent-year, multiple implicates may link to the same firm or establishment, depending on the outcome of the linkage procedures. Valid statistical inference requires that multiple imputation methods be used to obtain point and variance estimates of any parameter of interest (for example, sample means or regression coefficients). Documentation on how to implement MI with SAS and Stata software is provided in Klauth (2023).

3.3 CenHRS Crosswalk Variables

The products of the linkage procedure are the CenHRS crosswalk files, one for firms (F file) and one for establishments (E file), which provide HRS and BR identifiers for each matched HRS-BR pair. To incorporate uncertainty in the matching process, each of the crosswalk files consists of 10 implicates per HRS respondent-year. The files contain the following variables:

Variables in the Crosswalk	Source	Description	File
hhidpn	HRS	HRS household person identifier	E and F
linkage_type	CenHRS	Type of linkage: 1 = EIN-based deterministic link 2 = EIN-based probabilistic link 3 = Name and address-based probabilistic link	E and F
mainjobid	CenHRS/JCF	Equal to 1 for first reported main job in the HRS, and incrementing by 1 for each subsequent report of a new current main job	E and F
Replicate	CenHRS	Implicate number ranging from 1 to 10 for each hhipdn/year, with the number itself having no meaning besides identifying specific implicates	E and F
st_respondent	HRS	The state where the respondent works as reported in the HRS	E and F
year	CenHRS	Year in YYYY format (See section 4.1 below for guidance on definition and usage)	E and F
br_mu_indicator	BR	Indicator: 0 = firm has single unit/establishment 1 = firm has multiple units/establishments	E and F
estabid	BR	Establishment identifier: =CFN in 2001 and earlier =empunit_id_char in 2002 and later	E
firmid	BR	Firm identifier	F
lbdnum	BR	Longitudinal establishment identifier	E

Note: HRS = Health and Retirement Study, E = establishment, F = firm, JCF = Job Characteristics File, BR = Business Register, CFN = Census File Number, estabid takes its values from the BR variables CFN in years before 2002 and empunit_id_char in years 2002 and later (Chow et al., 2021). For information on the Health and Retirement Study, visit <https://hrs.isr.umich.edu/>. For information on the Business Register, see DeSalvo, Limehouse, and Klimek (2016).

4 Using the CenHRS Crosswalk

Each CenHRS crosswalk file provides identifiers for merging HRS datasets with data on HRS respondents' employers and establishments in the BR and other Census employer-and establishment-level data. The crosswalk files are thus the backbones on which administrative information about the firm or survey information on the worker can be added depending on the research purpose.

Each of the crosswalk files contain 10 implicates for each respondent in each year that the respondent has a valid linked *firmid*. HRS respondents are not included in the crosswalk in a given year if (1) they did not report a current main employer in that year or (2) they reported an employer, but were classified as not matched owing to the probabilistic match criteria in Abowd (2024). Rows in the crosswalk are uniquely identified by the variables *hhidpn* (a respondent identifier in the HRS), *year* (identifying reference year of questions regarding current main job at time of interview in the HRS), and *Replicate* (identifying each of the implicates for a given respondent-year).

4.1 Linking to HRS Datasets

For each implicate, the crosswalk variables *year* and *hhidpn* can be used to merge with HRS datasets. The variable *hhidpn* uniquely identifies an HRS respondent longitudinally. The variable *year* is the reference year for current employment questions in the HRS.¹ Year concepts in the HRS are complex, so care must be taken to ensure HRS records are merged to the crosswalk as intended. Several HRS and CenHRS year concepts, and considerations for merging HRS data with the CenHRS crosswalk, are detailed in the table below.

¹ The calendar year variable in the CenHRS crosswalk (and the crosswalk as a whole) is created using the Cross-Wave Harmonized Job Characteristics File (JCF), the purpose of which is to harmonize information about respondents' current jobs across HRS interviews. As a result, the crosswalk year can be thought of as roughly equivalent to the reference year for current employment questions in the HRS (in turn equivalent to the HRS interview year). For more information on the JCF, see Abramowitz et al. (2023).

Concept	Description	Example	Implications for HRS-to-CenHRS merge
HRS Wave year	HRS wave for which a given HRS record was collected	Waves are mainly named for even-numbered years since 1992 (e.g. 1998, 2014), as well as a few additional years (see section 3.1.1)	Depending on the vintage or variety of HRS data, HRS records may be organized into files by wave year
HRS Interview year (<i>iwave</i>)	Calendar year in which an HRS record was collected	A record for a respondent interviewed in 2019 has an interview year of 2019, even if the interview was performed for the 2018 survey wave	The CenHRS crosswalk is constructed using current main job employment as of the interview year
HRS Reference year	Calendar year a given HRS record refers to for a given variable	HRS questions about the last calendar year asked in interview year 2012 have reference year 2011	Reference year is the relevant concept if using the crosswalk to link an HRS variable to a contemporaneous BR variable
CenHRS Crosswalk year	Calendar year in the CenHRS crosswalk	A respondent linked to their main job in 2018 will have a record in the CenHRS crosswalk year 2018 for that job	The CenHRS crosswalk year variable can be merged with the HRS wave, interview, or reference year concepts, as appropriate

4.2 Linking to Census Employer and Establishment Files

The crosswalk contains identifiers to facilitate merging to several Census employer and establishment datasets, including the Business Register (BR), Longitudinal Business Database (LBD), and Longitudinal Employer-Household Dynamics (LEHD). Consisting of the universe of U.S. business establishments, the BR includes information on establishment industry, employment, and sales among other information (DeSalvo et al. 2016). The LBD links the Business Register longitudinally (Chow et al. 2021). The LEHD is a quarterly database of linked employer-employee data derived from state unemployment insurance records covering over 95 percent of employment in the United States and includes demographic information on all employees at a given employer including age, sex, race, job tenure, and quarterly earnings among other information (Graham et al. 2022).

This section provides an overview of how to link the CenHRS crosswalk to these datasets. The variable *firmid* identifies firms in both the firm crosswalk file and the BR. The variable *estabid* identifies establishments in the establishment crosswalk file, and can be linked with the BR using the establishment identifiers *CFN* (before 2002) and *empunit_id_char* (in 2002 and later waves). Note that due to the probabilistic nature of the crosswalk and the fact that multiple HRS participants may work for the same firm in the same year, firmid-year pairs are not uniquely identified in the CenHRS crosswalk. The graphic that follows indicates the relevant identifiers for merging the crosswalk to these datasets.

Key Identifiers of CenHRS, BR, and Select Linkable Data Files

CenHRS	Business Register	Longitudinal Business Database	Longitudinal Employer-Household Dynamics (LEHD)																																																														
<table border="1" style="width: 100%; border-collapse: collapse;"> <tr><th style="background-color: #e1f5fe;">CenHRS Crosswalk-Firm</th></tr> <tr><td style="background-color: #e8f5e9;">year</td></tr> <tr><td style="background-color: #ffe0b2;">firmid</td></tr> <tr><td>state_respondent</td></tr> <tr><td>Replicate</td></tr> <tr><td>...</td></tr> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><th style="background-color: #e1f5fe;">CenHRS Crosswalk-Establishment</th></tr> <tr><td style="background-color: #e8f5e9;">year</td></tr> <tr><td style="background-color: #fff9c4;">estabid</td></tr> <tr><td>state_respondent</td></tr> <tr><td>Replicate</td></tr> <tr><td>...</td></tr> </table>	CenHRS Crosswalk-Firm	year	firmid	state_respondent	Replicate	...	CenHRS Crosswalk-Establishment	year	estabid	state_respondent	Replicate	...	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr><th style="background-color: #e1f5fe;">Business Register (BR)</th></tr> <tr><td style="background-color: #fff9c4;">CFN/empunit_id_char</td></tr> <tr><td style="background-color: #ffe0b2;">ALPHA</td></tr> <tr><td style="background-color: #ffe0b2;">EIN</td></tr> <tr><td style="background-color: #ffe0b2;">firmid</td></tr> <tr><td style="background-color: #e8f5e9;">year</td></tr> <tr><td>State Abbreviation</td></tr> <tr><td>...</td></tr> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><th style="background-color: #e1f5fe;">Combined Business Register and County Business Pattern Microdata (cbpbr)</th></tr> <tr><td style="background-color: #fff9c4;">id</td></tr> <tr><td style="background-color: #ffe0b2;">ALPHA</td></tr> <tr><td style="background-color: #ffe0b2;">EIN</td></tr> <tr><td style="background-color: #ffe0b2;">firmid</td></tr> <tr><td style="background-color: #e8f5e9;">year</td></tr> <tr><td>State Abbreviation</td></tr> <tr><td>...</td></tr> </table>	Business Register (BR)	CFN/empunit_id_char	ALPHA	EIN	firmid	year	State Abbreviation	...	Combined Business Register and County Business Pattern Microdata (cbpbr)	id	ALPHA	EIN	firmid	year	State Abbreviation	...	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr><th style="background-color: #e1f5fe;">Firm-Level LBD File (lbdfirm)</th></tr> <tr><td>state (fips code)</td></tr> <tr><td style="background-color: #ffe0b2;">firmid</td></tr> <tr><td style="background-color: #e8f5e9;">year</td></tr> <tr><td>...</td></tr> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><th style="background-color: #e1f5fe;">Establishment-Level LBD File (lbd)</th></tr> <tr><td style="background-color: #ffe0b2;">firmid</td></tr> <tr><td style="background-color: #e8f5e9;">year</td></tr> <tr><td style="background-color: #fff9c4;">estabid</td></tr> <tr><td>...</td></tr> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><th style="background-color: #e1f5fe;">Firm-Level Revenues</th></tr> <tr><td style="background-color: #e8f5e9;">year</td></tr> <tr><td style="background-color: #ffe0b2;">firmid</td></tr> <tr><td>...</td></tr> </table>	Firm-Level LBD File (lbdfirm)	state (fips code)	firmid	year	...	Establishment-Level LBD File (lbd)	firmid	year	estabid	...	Firm-Level Revenues	year	firmid	...	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr><th style="background-color: #e1f5fe;">State-Firm-Level (SEIN) Employer Characteristics File (ECF SEIN T13)</th></tr> <tr><td>state (fips code)</td></tr> <tr><td style="background-color: #e1f5fe;">SEIN</td></tr> <tr><td style="background-color: #e8f5e9;">year</td></tr> <tr><td style="background-color: #e8f5e9;">quarter</td></tr> <tr><td>...</td></tr> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><th style="background-color: #e1f5fe;">National-Firm-Level Employer Characteristics Files (ECF T26)</th></tr> <tr><td style="background-color: #ffe0b2;">firmid</td></tr> <tr><td style="background-color: #e1f5fe;">SEIN</td></tr> <tr><td style="background-color: #e8f5e9;">year</td></tr> <tr><td style="background-color: #e8f5e9;">quarter</td></tr> <tr><td>...</td></tr> </table>	State-Firm-Level (SEIN) Employer Characteristics File (ECF SEIN T13)	state (fips code)	SEIN	year	quarter	...	National-Firm-Level Employer Characteristics Files (ECF T26)	firmid	SEIN	year	quarter	...	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr><th style="background-color: #e1f5fe;">(SEINUNIT) Employer Characteristics File (ECF SEINUNIT T13)</th></tr> <tr><td>state (fips code)</td></tr> <tr><td style="background-color: #e1f5fe;">SEIN</td></tr> <tr><td style="background-color: #ffe0b2;">SEINUNIT</td></tr> <tr><td style="background-color: #e8f5e9;">year</td></tr> <tr><td style="background-color: #e8f5e9;">quarter</td></tr> <tr><td>...</td></tr> </table>	(SEINUNIT) Employer Characteristics File (ECF SEINUNIT T13)	state (fips code)	SEIN	SEINUNIT	year	quarter	...
CenHRS Crosswalk-Firm																																																																	
year																																																																	
firmid																																																																	
state_respondent																																																																	
Replicate																																																																	
...																																																																	
CenHRS Crosswalk-Establishment																																																																	
year																																																																	
estabid																																																																	
state_respondent																																																																	
Replicate																																																																	
...																																																																	
Business Register (BR)																																																																	
CFN/empunit_id_char																																																																	
ALPHA																																																																	
EIN																																																																	
firmid																																																																	
year																																																																	
State Abbreviation																																																																	
...																																																																	
Combined Business Register and County Business Pattern Microdata (cbpbr)																																																																	
id																																																																	
ALPHA																																																																	
EIN																																																																	
firmid																																																																	
year																																																																	
State Abbreviation																																																																	
...																																																																	
Firm-Level LBD File (lbdfirm)																																																																	
state (fips code)																																																																	
firmid																																																																	
year																																																																	
...																																																																	
Establishment-Level LBD File (lbd)																																																																	
firmid																																																																	
year																																																																	
estabid																																																																	
...																																																																	
Firm-Level Revenues																																																																	
year																																																																	
firmid																																																																	
...																																																																	
State-Firm-Level (SEIN) Employer Characteristics File (ECF SEIN T13)																																																																	
state (fips code)																																																																	
SEIN																																																																	
year																																																																	
quarter																																																																	
...																																																																	
National-Firm-Level Employer Characteristics Files (ECF T26)																																																																	
firmid																																																																	
SEIN																																																																	
year																																																																	
quarter																																																																	
...																																																																	
(SEINUNIT) Employer Characteristics File (ECF SEINUNIT T13)																																																																	
state (fips code)																																																																	
SEIN																																																																	
SEINUNIT																																																																	
year																																																																	
quarter																																																																	
...																																																																	

Since the links that underlie the crosswalk are constructed using data from the BR, there is a direct mapping from the *firmid* identifiers in the crosswalk to the *firmid* identifiers in the LBD. The *CFN* and *empunit_id_char* identifiers in the establishment-level crosswalk can also be linked to the *estabid* identifier in the establishment-level LBD files.

However, in the case of the LEHD, linking requires an additional step, as we do not have a direct link between workers in the HRS and the *SEIN/SEINUNIT* identifiers that uniquely identify employers in the LEHD. As the LEHD's SEIN identifiers represent all establishments of a firm aggregated at the state level, users may merge the CenHRS crosswalk to the LEHD SEIN-level files using a combination of the *firmid* and the state in which the respondent reports working. Note that many SEINs may be associated with one *firmid* in each state, since one firm may have multiple SEINs at a point in time. At the moment, we do not have a way of mapping *CFNs* or *estabids* to the establishment-level *SEINUNIT* files in the LEHD.

5 Data Access

5.1 Application Procedures

The CenHRS crosswalk is available exclusively in FSRDCs. To access the CenHRS crosswalk in the FSRDCs, researchers must first apply for access to the crosswalk as well as any HRS and Census data they would like to link. Researchers are advised to follow the below process for preparing a proposal:

1. Before beginning the application process, researchers should reach out to the CenHRS team at CenHRSinfo@umich.edu to discuss the feasibility of their research question.
2. After consulting with the CenHRS team, researchers should contact their local FSRDC administrator to begin a discussion to develop a proposal.
3. Researchers must first prepare and submit an application to the HRS requesting access to use the CenHRS crosswalk, which is a restricted HRS data product, as well as any other requested restricted HRS data products, via the HRS restricted data agreement (RDA) portal.
4. Once the application has been approved by the HRS, the researcher must submit a proposal to the Census Bureau via the Standard Application Portal including documentation of HRS approval. Once submitted, the proposal will be reviewed first by the Census Bureau and then by the Internal Revenue Service (IRS).
5. Once the proposal is approved by the Census Bureau, researchers will need to obtain Special Sworn Status (SSS). Researchers should work with their FSRDC administrator to apply for SSS.
6. Once the project is approved by the Census Bureau, CenHRS, the IRS, and HRS and the researcher(s) have obtained SSS, the researcher(s) will work with their FSRDC administrator to begin accessing the data.

5.2 Reporting Project Progress and Modifications

After a project is approved, it is the researcher's responsibility to provide documentation of progress and modifications during the data access period to the Census Bureau and HRS. This may include the following actions:

1. Adding or Removing Research Staff
2. Requesting Additional Data
3. Uploading updated IRB Documentation when the project's IRB approval has expired
4. Submitting Annual Reports to the Census Bureau and HRS
5. Submitting project completion documentation when the data access period has ended

6 Data Citation and Acknowledgement of Funding

6.1 Data Citation

Publications and other scholarly work based on the CenHRS crosswalk should cite the resource, including the following information: the project name “Census-Enhanced Health and Retirement Study,” the version number of the crosswalk file being used, the year of publication of the crosswalk file, the name that version of the crosswalk including HRS wave years included in that version, acknowledgement of the the University of Michigan as publisher, and a link to the CenHRS project website, <https://cenhrs.isr.umich.edu/>. An example citation is provided below:

Census-Enhanced Health and Retirement Study. 2023. Census-Enhanced Health and Retirement Study Crosswalk for 1992-2018 HRS Wave Years, V2023-08-01 [dataset]. Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan.
<https://cenhrs.isr.umich.edu/>.

6.2 Acknowledgment of Funding

Researchers who have used CenHRS data, including the crosswalk files or the JCF, must acknowledge the funding source for CenHRS in their publications or any types of scholarly work (e.g., poster, presentation, etc.) as follows:

The development of the CenHRS is supported by a grant from the Alfred P. Sloan Foundation ((2011-6-21). The Health and Retirement Study is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan.

7 Contact Information

For questions about the CenHRS, please email CenHRSinfo@umich.edu.

8 References

- Abowd, J. M., Abramowitz, J., Levenstein, M. C., McCue, K., Patki, D., Raghunathan, T., Rodgers, A. M., Shapiro, M. D., Wasi, N., & Zinsser, D. (2024). *Finding Needles in Haystacks: Multiple-Imputation Record Linkage Using Machine Learning*.
- Abramowitz, J., Fang, C., & Hyde, J. S. (2024). An Overview of Linked Social Security Administration Data in the Health and Retirement Study. University of Michigan. Available at: <https://hrs.isr.umich.edu/publications/biblio/14167>.
- Abramowitz, J., Klauth, B., & Wasi, N. (2023) Health and Retirement Study Cross-Wave Harmonized Job Characteristics File for 1992-2018 HRS Wave Years, V2023-05-05, Data Description. Available at: https://hrsdata.isr.umich.edu/sites/default/files/documentation/data-descriptions/1684339744/hrs_jcf_public_92to18_v20230505_datadescription.pdf.
- Chow, M., Fort, T. C., Goetz, C., Goldschlag, N., Lawrence, J., Perlman, E. R., Stinson, M., & White, T. K. (2021). *Redesigning the Longitudinal Business Database*. Available at: <https://www2.census.gov/ces/wp/2021/CES-WP-21-08.pdf>.
- DeSalvo, B., Limehouse, F. F., & Klimek, S. D. (2016). Documenting the Business Register and Related Economic Business Data. U.S. Census Bureau Center for Economic Studies Working Paper Number CES-16-17. Available at: <https://www.census.gov/library/working-papers/2016/adrm/ces-wp-16-17.html>.
- Graham, M., McEntarfer, E., McKinney, K., Tibbets, S., & Tucker, L. (2022). *LEHD Snapshot Documentation*. Census Bureau Center for Economic Studies Working Paper Number 22-51. Available at: <https://ideas.repec.org/p/cen/wpaper/22-51.html>.
- Health and Retirement Study. (2020). *2020 HRS Early Core public use dataset. Version 2.0*. Available at: <https://hrsdata.isr.umich.edu/sites/default/files/documentation/data-descriptions/1660571337/h20dd.pdf>.
- Klauth, B. (2023). *How to Combine Inferences from Multiply-Imputed Data Sets: SAS and STATA Examples*. Available at: https://cenhrs.isr.umich.edu/wp-content/uploads/2023/02/sas_stata_examples_v20230202.pdf.
- Sonnega, A., Faul, J. D., Ofstedal, M. B., Langa, K. M., Phillips, J. W., & Weir, D. R. (2014). Cohort Profile: the Health and Retirement Study (HRS). *International Journal of Epidemiology*, 576–585. <https://doi.org/10.1093/ije/dyu067>.

9 Additional References

Rubin, D. (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley.

SAS Institute Inc. (2017). *SAS Help Center: Combining Inferences from Imputed Data Sets*.

Available at:

https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.3/statug/statug_mianalyze_detail_s08.htm.

SAS Institute Inc. (2017). SAS Help Center: The MIANALYZE Procedure. Available at:

https://documentation.sas.com/?cdcId=pgmsascdc&cdcVersion=9.4_3.3&docsetId=statug&docsetTarget=statug_mianalyze_toc.htm&locale=en.