

README

Overview

This replication package contains all the code necessary to reproduce the results for the paper “*Health Types and Health Inequality*”. Data cleaning and dataset construction are performed in Stata, while the analysis is conducted in R. The workflow is executed through two main scripts: the first generates the data, and the second performs the analysis presented in the paper. The entire process takes approximately 53 minutes to complete.

Data Availability and Provenance Statements

- ☐ This paper does not involve analysis of external data (i.e., no data are used or the only data are generated by the authors via simulation in their code).

Statement about Rights

- I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.
- ☐ I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package. Appropriate permission are documented in the LICENSE.txt file.

Summary of Availability

- All data **are** publicly available.
- ☐ Some data **cannot be made** publicly available.
- ☐ **No data can be made** publicly available.

Details on each Data Source

Data Name	Data File	Location	Provided	Citation
RAND HRS Longitudinal File 2018 (V2)	randhrs1992_2018v2.dta			
Cross-Wave Tracker File	trk2022tr_r.dta			
RAND HRS 1994 Fat File (V1A)	h94f1a.dta			
RAND HRS 1995 Fat File (V2B)	ad95f2b.dta			
RAND HRS 1996 Fat File (V4A)	h96f4a.dta			
RAND HRS 1998 Fat File (V2C)	hd98f2c.dta			
RAND HRS 2000 Fat File (V1D)	h00f1d.dta			
RAND HRS 2002 Fat File (V2C)	h02f2c.dta			
RAND HRS 2004 Fat File (V1C)	h04f1c.dta			
RAND HRS 2006 Fat File (V4A)	h06f4a.dta			
RAND HRS 2008 Fat File (V3A)	h08f3a.dta			
RAND HRS 2010 Fat File (V6A)	hd10f6a.dta			
RAND HRS 2012 Fat File (V3A)	h12f3a.dta			
RAND HRS 2014 Fat File (V2B)	h14f2b.dta			
RAND HRS 2016 Fat File (V2C)	h16f2c.dta	rawdata/	FALSE	RAND (2022)
RAND HRS 2018 Fat File (V2B)	h18f2b.dta			
	randhrsexit1994v1.dta			
	randhrsexit1996v1.dta			
	randhrsexit1998v1.dta			
	randhrsexit2000v1.dta			
	randhrsexit2002v1.dta			
	randhrsexit2004v1.dta			
	randhrsexit2006v1.dta			
	randhrsexit2008v1.dta			
	randhrsexit2010v1.dta			
	randhrsexit2012v1.dta			
	randhrsexit2014v1.dta			
	randhrsexit2016v1.dta			
	randhrsexit2018v2.dta			

RAND HRS Exit/Post-Exit Interview and
Finder Files 2020 (V1)

Data on the Health and Retirement Study (HRS) were downloaded from the RAND Corporation (RAND, 2022) and the HRS. The Longitudinal file, the fat files and exit interviews can all be downloaded from <https://hrsdata.isr.umich.edu/data-products/rand>. The Longitudinal file is available under RAND HRS Archived Data Products, the fat files can be found under Biennial Products, and the exit interviews are accessible via RAND HRS Exit/Post-Exit Interview and Finder Files 2020. The tracker file can be downloaded from <https://hrsdata.isr.umich.edu/data-products/cross-wave-tracker-file>. Please note that the HRS imposes restrictions on data sharing. Per their guidelines, the distribution of HRS data by third parties is strictly prohibited. Access to these publicly available files requires prior registration and approval through the HRS website.

Dataset list

Data file	Source	Notes	Provided
rawdata/trk2022tr_r.dta	Health and Retirement Study	As per terms of use	No
rawdata/randhrs1992_2018v2.dta	RAND Corporation	As per terms of use	No
rawdata/h94f1a.dta	All listed	As per terms of use	No
rawdata/ad95f2b.dta	All listed	As per terms of use	No
rawdata/h96f4a.dta	All listed	As per terms of use	No
rawdata/hd98f2c.dta	All listed	As per terms of use	No
rawdata/h00f1d.dta	All listed	As per terms of use	No
rawdata/h02f2c.dta	All listed	As per terms of use	No
rawdata/h04f1c.dta	All listed	As per terms of use	No
rawdata/h06f4a.dta	All listed	As per terms of use	No
rawdata/h08f3a.dta	All listed	As per terms of use	No
rawdata/hd10f6a.dta	All listed	As per terms of use	No
rawdata/h12f3a.dta	All listed	As per terms of use	No
rawdata/h14f2b.dta	All listed	As per terms of use	No
rawdata/h16f2c.dta	All listed	As per terms of use	No
rawdata/h18f2b.dta	All listed	As per terms of use	No
rawdata/randhrsexit1994v1.dta	All listed	As per terms of use	No
rawdata/randhrsexit1996v1.dta	All listed	As per terms of use	No
rawdata/randhrsexit1998v1.dta	All listed	As per terms of use	No
rawdata/randhrsexit2000v1.dta	All listed	As per terms of use	No
rawdata/randhrsexit2002v1.dta	All listed	As per terms of use	No
rawdata/randhrsexit2004v1.dta	All listed	As per terms of use	No
rawdata/randhrsexit2006v1.dta	All listed	As per terms of use	No
rawdata/randhrsexit2008v1.dta	All listed	As per terms of use	No
rawdata/randhrsexit2010v1.dta	All listed	As per terms of use	No
rawdata/randhrsexit2012v1.dta	All listed	As per terms of use	No
rawdata/randhrsexit2014v1.dta	All listed	As per terms of use	No
rawdata/randhrsexit2016v1.dta	All listed	As per terms of use	No
rawdata/randhrsexit2018v2.dta	All listed	As per terms of use	No
dtfiles/CleanPanelBalanced.dta	All listed	Produced by Construct- Panel_replication.do using the files listed above	No
dtfiles/SelfProbDying.dta	All listed	Produced by Construct- Panel_replication.do using randhrs1992_2018v2.dta	No
dtfiles/DeathCause.dta	All listed	Produced by DoFiles/Causes_of_Death.do using randhrs1992_2018v2.dta and all exit interview listed above (randhrsexit'year'vs'.dta)	No

Data file	Source	Notes	Provided
dtfiles/data_clustering.dta	All listed	Produced by Part2_a_clustering.R using trk2022tr_r.dta and CleanPanelBalanced.dta	No
dtfiles/P52_2_Clusters.dta	All listed	Produced by Part2_a_clustering.R using trk2022tr_r.dta and CleanPanelBalanced.dta	No
dtfiles/P52_3_Clusters.dta	All listed	Produced by Part2_a_clustering.R using trk2022tr_r.dta and CleanPanelBalanced.dta	No
dtfiles/P52_4_Clusters.dta	All listed	Produced by Part2_a_clustering.R using trk2020tr_r.dta and CleanPanelBalanced.dta	No
dtfiles/P52_5_Clusters.dta	All listed	Produced by Part2_a_clustering.R using trk2022tr_r.dta and CleanPanelBalanced.dta	No
dtfiles/P52_6_Clusters.dta	All listed	Produced by Part2_a_clustering.R using trk2022tr_r.dta and CleanPanelBalanced.dta	No
dtfiles/P52_7_Clusters.dta	All listed	Produced by Part2_a_clustering.R using trk2022tr_r.dta and CleanPanelBalanced.dta	No
dtfiles/P52_8_Clusters.dta	All listed	Produced by Part2_a_clustering.R using trk2022tr_r.dta and CleanPanelBalanced.dta	No
dtfiles/P52_9_Clusters.dta	All listed	Produced by Part2_a_clustering.R using trk2022tr_r.dta and CleanPanelBalanced.dta	No
dtfiles/P52_10_Clusters.dta	All listed	Produced by Part2_a_clustering.R using trk2022tr_r.dta and CleanPanelBalanced.dta	No

Computational requirements

Software Requirements

- Stata (code was last run with version 18)
- R 4.4.1
 - pacman (0.5.1)
 - tidyverse (2.0.0)
 - haven (2.5.4)
 - stats (4.4.1)
 - ggplot2 (3.5.1)
 - xtable (1.8-4)
 - stargazer (5.2.3)
 - coefplot (1.2.8)
 - gplots (3.2.0)
 - reshape2 (1.4.4)
 - feisr (1.3.0)

- dplyr (1.1.4)
- knitr (1.49)
- kpodclustr (1.1)
- Cairo (1.6-2)
- zoo (1.8-12)
- labelled (2.13.0)
- plot.matrix (1.6.2)
- gcookbook (2.0)
- directlabels (2024.1.21)
- imguR (1.0.3)
- ggrepel (0.9.6)
- data.table (1.16.2)
- foreach (1.5.2)
- DescTools (0.99.58)
- jtools (2.3.0)
- openxlsx (4.2.7.1)
- factoextra (1.0.7)
- kableExtra (1.4.0)
- pscl (1.5.9)
- multiUS (1.2.3)
- caret (6.0-94)
- mefa (3.2-9)
- nnet (7.3-19)
- RColorBrewer (1.1-3)
- caret (6.0-94)
- lmttest (0.9-40)
- mefa (3.2-9)
- latex2exp (0.9.6)
- the file “Install_Libraries.R” will install all dependencies (latest version), and should be run once prior to running other programs.

Memory and Runtime Requirements

Summary Approximate time needed to reproduce the analyses on a standard (2024) desktop machine:

- ☐ <10 minutes
- ☒ 10-60 minutes
- ☐ 1-2 hours
- ☐ 2-8 hours
- ☐ 8-24 hours
- ☐ 1-3 days
- ☐ 3-14 days
- ☐ > 14 days
- ☐ Not feasible to run on a desktop machine, as described below.

Details Both the R and the Stata codes were last run on a **11-core Apple M3 laptop with MacOS version Sonoma 14.3**. The Stata code (cleaning codes) takes 98 seconds to run. The R code (analysis codes) takes 51 minutes to run. These times do not include the time required to install any dependencies needed to execute the code. The final output folder, after running all results, occupies approximately 10 GB.

Description of programs/code

- The replication package is organized into five main folders:
 - **acodes/**: Contains all analysis codes that generate the main results of the paper.
 - **ccodes/**: Contains all cleaning codes used to process raw data files and produce the clean datasets for analysis.
 - **dtfiles/**: Stores the cleaned datasets generated by the cleaning codes, as well as additional datasets generated by the analysis codes.
 - **output/**: Contains all results, including tables and figures used in the paper.
 - **rawdata/**: Contains the raw data files used as inputs for the analysis.
- Programs in **ccodes/** produce the clean datasets needed for the analysis. Running **ccodes/Master_Stata.do** will execute all cleaning scripts in the folder, generating and storing the resulting datasets in **dtfiles/**.

- Programs in `acodes/` generate all the results of the paper. Running `acodes/Master_Code.R` will execute all analysis scripts, saving outputs such as tables and figures in `output/` and additional datasets in `dtfiles/`.
- The file `acodes/Master_Code.R` calls subroutines that are divided in three main parts:
 - `acodes/Part1`: Installs and imports all necessary packages.
 - `acodes/Part2`: Generates all tables and figures, except for those related to Section 5 of the paper.
 - `acodes/Part3`: Produces all tables and figures specific to Section 5 of the paper.
- Output files are stored in the folder corresponding to the programs used to produce them. For example, files produced by `acodes/Part2/Part2_b_graphs.R` will be saved in the folder `output/Part2_output/Part2_b_graphs`.
- Output files are called (with a few exceptions to be listed) appropriate names (`tab5.tex`, `figure12.pdf`) and should be easy to correlate with the manuscript. The exceptions are:
 - Tables 3 and 11 are both pulled directly from the tables in `Cause_of_death_tables.xlsx`. See cell A18 for further instructions.

Instructions to Replicators

- Download the data files listed in “Data Sources”. Each should be stored in the prepared subdirectories of `rawdata/`, in the format that you download them in. No further action is needed on the replicator’s part.
- Edit `ccodes/Master_Stata.do` to adjust the default paths.
- Run `ccodes/Master_Stata.do` once in Stata to generate datasets. Datasets will be stored in `dtfiles/`.
- Edit `acodes/Master_Code.R` to adjust the default paths.
- Run `acodes/Master_Code.R` to run all other steps in sequence. Output will be stored in `output/`, and additional datasets will be stored in `dtfiles/`.

Details

- `ccodes/Master_Stata.do`: will create cleaned datasets from the raw data. Originally used in Russo, McGee, De Nardi, Borella, and Abram (2024)
 - The global ‘mainpath’ at line 4 should be edited to match the directory of the replication package.
- `acodes/Part1`: Contains two files, one which installs the necessary packages, one which imports them.
 - The indicators `install_lib` and `import_lib` in `acodes/Master_Code.R` determine whether to install and import the necessary libraries, respectively. It is suggested to set both to 1 the first time the code is run, then set `install_lib` to 0 for later runs.
- `acodes/Part2`.
 - The contents of this folder are the files to produce all figures and graphs in our paper, and appendixes except for those in Section 5.
 - The file `acodes/Part2/Part2_main_code.R` runs everything in order, and it is called by the file `acodes/Main_code.R`.
 - If running programs individually, note that ORDER IS IMPORTANT.
 - The programs were last run top to bottom on January 2, 2025.
- `acodes/Part3`.
 - The contents of this folder are the files to produce all figures and graphs in Section 5.
 - The file `acodes/Part2/Part3_main_code.R` runs everything in order, and it is called by the file `acodes/Main_code.R`.
 - If running programs individually, note that ORDER IS IMPORTANT.
 - The programs were last run top to bottom on January 2, 2025.
- Tables 3 and 11: The tables can be reproduced using the excel file `Cause_of_death_tables.xlsx`, which starts with the contents of `Cause_of_death.xlsx`, the direct output of `Part2_h_cause_of_death.R`, and produces Tables 3 and 11.

List of tables and programs

The provided code reproduces:

- ☐ All numbers provided in text in the paper
- ☐ All tables and figures in the paper
- ☒ Selected tables and figures in the paper, as explained and justified below.

Output files are stored in the folder corresponding to the programs used to produce them. For example, files produced by `acodes/Part2/Part2_b_graphs.R` will be saved in the folder `output/Part2_output/Part2_b_graphs`.

Figure/Table #	Program	Output file	Note
Table 1	n.a. (no data)		From Russo et al (2024)
Table 2	Part2/Part2_g_predictive_power.R	tab2.tex	

Figure/Table #	Program	Output file	Note
Table 3	Part2/Part2_h_cause_of_death.R	Cause_of_death_tables.xlsx	See details above.
Table 4	Part2/Part2_d_cluster_statistics.R	tab4.tex	tab4.tex is compiled using tab4-header.tex, tab4-a.tex, tab4-b.tex, tab4-c.tex, tab4-d.tex and tab4-e.tex that are also produced and stored.
Table 5	Part2/Part2_j_mlogit.R	tab5.tex	
Table 6	Part3/Part3_b_testing.R	tab6.tex	
Table 7	Part3/Part3_b_testing.R	tab7.tex	
Table 8	Part2/Part2_a_clustering.R	tab8.tex	
Table 9	Part2/Part2_a_clustering.R	tab9.tex	
Table 10	Part2/Part2_i_frailty_distribution.R	tab10.tex	
Table 11	Part2/Part2_h_cause_of_death.R	Cause_of_death_tables.xlsx	See details above.
Table 12	Part2/Part2_g_predictive_exc.R	tab12.tex	
Table 13	Part2/Part2_g_predictive_exc.R	tab13.tex	
Table 14	Part2/Part2_j_mlogit.R	tab14.tex	
Table 15	Part2/Part2_f_deficit_prevalence.R	tab15.tex	
Figure 1	Part2/Part2_i_frailty_distribution.R	fig1.pdf	
Figure 2	Part2/Part2_b_graphs.R	fig2-1.pdf, fig2-2.pdf, fig2-3.pdf	
Figure 3	Part2/Part2_e_heatmaps.R	fig3-1.pdf, fig3-2.pdf, fig3-3.pdf, fig3-4.pdf, fig3-5.pdf	
Figure 4	Part3/Part3_c_simulations.R	fig4-1.pdf, fig4-2.pdf	
Figure 5	Part3/Part3_c_simulations.R	fig5-1.pdf, fig5-2.pdf	
Figure 6	n.a. (no data)		Pulled from the Health and Retirement Study (2024).
Figure 7	Part2/Part2_c_number_cluster.R	fig7-1.pdf, fig7-2.pdf	
Figure 8	Part2/Part2_c_number_cluster.R	fig8-1.pdf, fig8-2.pdf	
Figure 9	Part2/Part2_a_clustering.R	fig9-1.pdf, fig9-2.pdf	
Figure 10	Part2/Part2_g_predictive_exc.R	fig10-1.pdf, fig10-2.pdf	
Figure 11	Part2/Part2_g_predictive_exc.R	fig11-1.pdf, fig11-2.pdf	
Figure 12	Part2/Part2_b_graphs.R	fig12-1.pdf, fig12-2.pdf, fig12-3.pdf	
Figure 13	Part2/Part2_b_graphs.R	fig13-1.pdf, fig13-2.pdf, fig13-3.pdf	
Figure 14	Part2/Part2_b_graphs.R	fig14-1.pdf, fig14-2.pdf, fig14-3.pdf	
Figure 15	Part2/Part2_b_graphs.R	fig15-1.pdf, fig15-2.pdf, fig15-3.pdf	

References

Health and Retirement Study, (RAND HRS Longitudinal File 2018 (V2)) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant numbers NIA U01AG009740 and NIA R01AG073289). Ann Arbor, MI, (July 2022).

RAND HRS Longitudinal File 2018 (V2). Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA (July 2022).

Health and Retirement Study, (Cross-Wave Tracker File) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant numbers NIA U01AG009740 and NIA R01AG073289). Ann Arbor, MI, (November 2024).

Health and Retirement Study, (RAND HRS 1994 Fat File (V1A)) public use dataset. Produced and distributed by the

University of Michigan with funding from the National Institute on Aging (grant numbers NIA U01AG009740 and NIA R01AG073289). Ann Arbor, MI, (June 2004).

RAND HRS 1994 Fat File (V1A). Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA (June 2004).

Health and Retirement Study, (RAND HRS 1995 Fat File (V2B)) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant numbers NIA U01AG009740 and NIA R01AG073289). Ann Arbor, MI, (February 2013).

RAND HRS 1995 Fat File (V2B). Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA (February 2013).

Health and Retirement Study, (RAND HRS 1996 Fat File (V4A)) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant numbers NIA U01AG009740 and NIA R01AG073289). Ann Arbor, MI, (April 2006).

RAND HRS 1996 Fat File (V4A). Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA (April 2006).

Health and Retirement Study, (RAND HRS 1998 Fat File (V2C)) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant numbers NIA U01AG009740 and NIA R01AG073289). Ann Arbor, MI, (April 2013).

RAND HRS 1998 Fat File (V2C). Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA (April 2013).

Health and Retirement Study, (RAND HRS 2000 Fat File (V1D)) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant numbers NIA U01AG009740 and NIA R01AG073289). Ann Arbor, MI, (February 2021).

RAND HRS 2000 Fat File (V1D). Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA (February 2021).

Health and Retirement Study, (RAND HRS 2002 Fat File (V2C)) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant numbers NIA U01AG009740 and NIA R01AG073289). Ann Arbor, MI, (September 2017).

RAND HRS 2002 Fat File (V2C). Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA (September 2017).

Health and Retirement Study, (RAND HRS 2004 Fat File (V1C)) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant numbers NIA U01AG009740 and NIA R01AG073289). Ann Arbor, MI, (October 2019).

RAND HRS 2004 Fat File (V1C). Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA (October 2019).

Health and Retirement Study, (RAND HRS 2006 Fat File (V4A)) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant numbers NIA U01AG009740 and NIA R01AG073289). Ann Arbor, MI, (July 2022).

RAND HRS 2006 Fat File (V4A). Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA (July 2022).

Health and Retirement Study, (RAND HRS 2008 Fat File (V3A)) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant numbers NIA U01AG009740 and NIA R01AG073289). Ann Arbor, MI, (November 2017).

RAND HRS 2008 Fat File (V3A). Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA (November 2017).

Health and Retirement Study, (RAND HRS 2010 Fat File (V6A)) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant numbers NIA U01AG009740 and NIA R01AG073289). Ann Arbor, MI, (July 2022).

RAND HRS 2010 Fat File (V6A). Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA (July 2022).

Health and Retirement Study, (RAND HRS 2012 Fat File (V3A)) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant numbers NIA U01AG009740 and NIA R01AG073289). Ann Arbor, MI, (February 2021).

RAND HRS 2012 Fat File (V3A). Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA (February 2021).

Health and Retirement Study, (RAND HRS 2014 Fat File (V2B)) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant numbers NIA U01AG009740 and NIA R01AG073289). Ann Arbor, MI, (February 2021).

RAND HRS 2014 Fat File (V2B). Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA (February 2021).

Health and Retirement Study, (RAND HRS 2016 Fat File (V2C)) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant numbers NIA U01AG009740 and NIA R01AG073289). Ann Arbor, MI, (March 2023).

RAND HRS 2016 Fat File (V2C). Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA (March 2023).

Health and Retirement Study, (RAND HRS 2018 Fat File (V2B)) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant numbers NIA U01AG009740 and NIA R01AG073289). Ann Arbor, MI, (March 2023).

RAND HRS 2018 Fat File (V2B). Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA (March 2023).

Health and Retirement Study, (RAND HRS Exit/Post-Exit Interview and Finder Files 2020 (V1)) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant numbers NIA U01AG009740 and NIA R01AG073289). Ann Arbor, MI, (May 2024).

RAND HRS Exit/Post-Exit Interview and Finder Files 2020 (V1). Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA (May 2024).

The Health and Retirement Study (2024), “HRS survey design and methodology: Longitudinal Cohort Sample Design.” URL <https://hrs.isr.umich.edu/documentation/survey-design>. Accessed: 2024-04-10.

Russo, Nicolò, Rory McGee, Mariacristina De Nardi, Margherita Borella, and Ross Abram (2024), “Health inequality and economic disparities by race, ethnicity, and gender.[Stata code]” Working Paper 32971, National Bureau of Economic Research, URL <http://www.nber.org/papers/w32971>.

Acknowledgements

Some content was adapted from Russo et al (2024), data cleaning and construction of the frailty index, with the author’s permission.

We also thank Johanna Torres Chain for her contributions to the data cleaning and construction of the cause of death data (*DeathCause.dta*) using the exit interviews file of the HRS.