USC Dornsife
Center for Economic and Social Research
Program on Global Aging, Health, and Policy

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

UKD Universitätsklinikum Düsseldorf

Harmonized CHARLS Life History Documentation

VERSION A, JULY 2022

Morten Wahrendorf, Christian Deindl, Drystan Phillips, and Jinkook Lee

We greatly appreciate support from the National Institute on Aging (R01 AG030153, RC2 AG036619, R03 AG043052)

g2aging.org

Preface

The China Health and Retirement Longitudinal Study (CHARLS) is a longitudinal household survey for the study of retirement and health among the elderly in China. It is modeled after the Health and Retirement Study (HRS) and its sister studies (e.g. the English Longitudinal Study of Ageing (ELSA), Survey of Health, Ageing and Retirement in Europe (SHARE)) which are similar longitudinal surveys in the United States and Europe. Similar to HRS, SHARE, and ELSA, CHARLS conducted a separate life history interview. This life history interview comprises retrospective information on previous histories, specifically, detailed data on previous partnership, children, residential, and work histories. The data collection on the life history interview took place in 2014.

In order to make the CHARLS life history data more accessible to researchers and to facilitate comparisons with HRS, SHARE, and ELSA, we, Christian Deindl and Morten Wahrendorf from the TU Dortmund University and the University of Duesseldorf in Germany and the USC Gateway to Global Aging team, created the Harmonized CHARLS Life History, a user-friendly subset of the CHARLS life history interviews. This initiative is part of a larger set of projects. With funding and support from the National Institute on Aging, we have also created harmonized life history data for other studies, harmonized data from longitudinal core interviews, and harmonized end of life data. The other harmonized life history data include the Harmonized ELSA Life History (England) and the Harmonized SHARE Life History (Europe + Israel). The harmonized longitudinal core interview data include Harmonized HRS (United States), Harmonized ELSA (England), Harmonized SHARE (Europe + Israel), Harmonized KLoSA (South Korea), Harmonized JSTAR (Japan), Harmonized CHARLS (China), Harmonized LASI (India), Harmonized MHAS (Mexico), Harmonized TILDA (Ireland), and Harmonized CRELES (Costa Rica). The harmonized end of life data includes the Harmonized HRS End of Life (United States). Further information about these harmonized data files including questionnaires and other metadata is available on our searchable website, https://g2aging.org/.

The Harmonized CHARLS Life History consists of variables with a similar data structure and naming conventions to other Harmonized Life History variables.

We are grateful for the continuing support of and funding from NIA.

Requested Acknowledgment

We ask all users of the Harmonized Life History CHARLS to please inform our team of any written analysis using data from the Harmonized Life History CHARLS or information from the Harmonized Life History CHARLS Codebook by sending an email to papers@g2aging.org. We also ask users to include the following acknowledgement in their written work:
"This analysis uses data or information from the Harmonized CHARLS Life History dataset and Codebook, Version A as of July 2022 developed by the Gateway to Global Aging Data in collaboration with the University of Duesseldorf and the TU Dortmund University. The development of the Harmonized CHARLS was funded by the National Institute on Aging (R01 AG030153, RC2 AG036619, R03 AG043052). For more information, please refer to https://g2aging.org/."

CHARLS Version and Acknowledgment

This document used CHARLS Waves 1 through 4 as of June 2021. CHARLS is supported by Peking University, the National Natural Science Foundation of China, the National Institute on Aging, and the World Bank.

Contents

1. Introduction and Overview

This report documents the harmonized sequence data from the life history interviews of the China Health and Retirement Longitudinal Study (CHARLS), as provided by the Gateway to Global Aging platform.

CHARLS started data collection for its first wave in 2011 and continued with surveys in 2013, 2014, 2015, and 2018. In the first wave (2011/12) 17,708 respondents had been interviewed. Each following wave contains a refresher sample (see Zhao et al 2014 for details on CHARLS). Similar to HRS, SHARE, and ELSA, CHARLS conducted a life history survey in 2014. This life history interview collected details of respondents' lives before they entered the study. This interview includes retrospective information on previous histories, as for example detailed data on previous partnership, children, residential, and work histories. CHARLS also includes detailed information about childhood conditions, characteristics of their parents and some information about grandparents. Not all information contained in CHARLS life history interview was harmonized due to country-specific themes and larger differences between surveys that made harmonization impossible. We encourage users to use our code and to modify the data fitting their research needs. As similar life history interviews have been conducted in SHARE, ELSA and HRS, this data proves a great opportunity to combine life history interviews across three continents.

1.1. Gateway to Global Aging Data

CHARLS is part of the family of Health and Retirement Studies (HRS) around the world. These surveys – and the life history interviews – were designed with harmonization as a goal to allow for cross-country studies. Yet, despite the richness of the available information, there are only a limited number of empirical studies exploiting such opportunities. This is partly due to the difficulty associated with learning multiple surveys, but – specifically in case of the life history data – also due to the complex data structure resulting from life history interviews. In detail, data of the life history interview are usually stored in a "wide spell-data format" with repeated information on spells (e.g. information for each job spell). Another way of storing this information is to rearrange the format in a so called "state sequence format" with one state definition for each age (e.g. work state at age 20, 21, 22, 23, 24, …..). The latter seems more intuitive and user-friendly. Furthermore, it is also known to be the standard format for sequence analysis, as an advanced statistical tool to investigate life courses. Cross-country studies using the life history data thus require learning multiple surveys and managing complex data structures associated with life history data. Harmonized life history data from the Gateway are provided in a rearranged "state sequence format", to facilitate an in-depth study of individual life-courses in a cross-national perspective.

Identifying comparable questions across surveys is the first step toward cross-country analyses. The Gateway to Global Aging Data (Gateway) helps users understand and use these large-scale

population surveys on health and retirement. The Gateway includes several tools to facilitate cross-national health and retirement research. It includes a digital library of survey questions for all participating surveys. Its search engine enables users to find relevant survey questions. The Gateway also includes a concordance with information comparing measures within and across surveys over time. Using these tools, researchers can identify all questions related to particular key words or within a domain. The Gateway also includes population and sub-population estimates for key harmonized variables and present them in graphs and tables that can be downloaded, and provides access to harmonized datasets and codebooks.

The Gateway can be accessed at https://g2aging.org. For more information about using the Gateway visit the Help page. For more information about creating the Harmonized CHARLS Life History using the Gateway see "Chapter 4. Distribution and Technical Notes."

1.2. Unit of Observation

The harmonized life history of CHARLS includes all individuals that participated in the life history interview. To enable merging with data from other waves, the data includes the person identifier ("ID"), the household and community identifier ("householdID" "communityID").

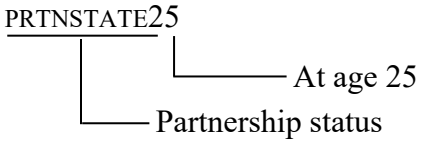
1.3. Data File Structure

Data is provided as individual data where each observation represents one individual. The harmonized sequence data from the life history interview is arranged in a state sequence format. Therefore, the data contains a variable for each age covered in the life history interview to describe the state at a particular age (e.g. work situation or total number of children at the age of 20, 21, 22, 23, 24, …etc.). This format has been created by rearranging the original data, which are provided in a "wide spell-data format". The original format provides information on all spells to which respondents gave information (e.g. details on job spells, as well as information on the beginning and the end of the job spell). The state sequence format, in contrast, appears more intuitive and user-friendly than a spell format and has several advantages. For example, it directly allows to describe the circumstance at a specific age (without need of extensive data management) and fully exploits the longitudinal nature of the data by providing information for each age covered in the life history interview. Further, the state sequence format allows to derive summary measures of interest (e.g. years spent in a specific state) and to apply methods of sequence analysis (Gabadinho, Ritschard, Müller, & Studer, 2011; Halpin, 2017) for an in-depth study of histories.

Sequence data of the Gateway are available for four domains: employment, partnership, children, and hukou information. Additionally, childhood information that was covered in the life history questionnaire is included in the dataset. Details are described below for each domain separately.

1.4. Variable Naming Convention

With few exceptions, variable names in the Harmonized CHARLS Life History data follow a consistent pattern. The first characters indicate to which domain the state variable refers to: : "cc" and "ch" = childhood, "empl" = employment history, "prtn" = partner history, "chld" & "ychd" = children history, and "hk" = hukou history. The last character indicates the age to which the state variable refers (e.g. 20, 21, 22, 22, 23, …). For example:



Variable PRTNSTATE25 captures the partnership status of the individual at age 25 as reported by the respondent during the CHARLS Life History interview.

In the text below, we may refer to variables such as PRTNSTATE15-80 for example, without specifying an age. This reference points at the group of variables PRTNSTATE15, PRTNSTATE16, …, PRTNSTATE79, PRTNSTATE80.

Variable labels also follow a consistent pattern. The first characters denote the name of the variable, followed by a colon. Then follows the unit to which the variable refers ("R" for reference person). The remainder of the label describes the concept that the variable captures and the age to which it refers. For example, the variable label of PRTNSTATE25 is:

```
prtnSTATE25:r partnership state at age 25
```

It may seem duplicative to include the name of the variable in the variable label. However, statistical packages often suppresses the variable name and instead uses its label in the presentation of results.

Variable names in the Harmonized CHARLS Life History are meant to be identical where possible with variable names in other Harmonized Life History datasets. Measures which are exactly or near-exactly comparable between the Harmonized CHARLS Life History and Harmonized ELSA Life History or Harmonized SHARE Life History use the exact same name. If the Harmonized CHARLS Life History measure is deemed only somewhat comparable with the Harmonized ELSA Life History or Harmonized SHARE Life History version of that measure, the variable name in the Harmonized CHARLS Life History will often start or end in "_C." This variable name prefix or suffix indicates some CHARLS-specific difference with the Harmonized ELSA Life History or Harmonized SHARE Life History version of this measure.

1.5. Missing Values, and Nonresponse

Variables may contain missing values for several reasons. Stata, SAS, and SPSS offer the capability to distinguish multiple types of missing values, and we have attempted to record as much information as possible. Generally, the codes adhere to the classification in Table 1.

Table 1. Missing Codes

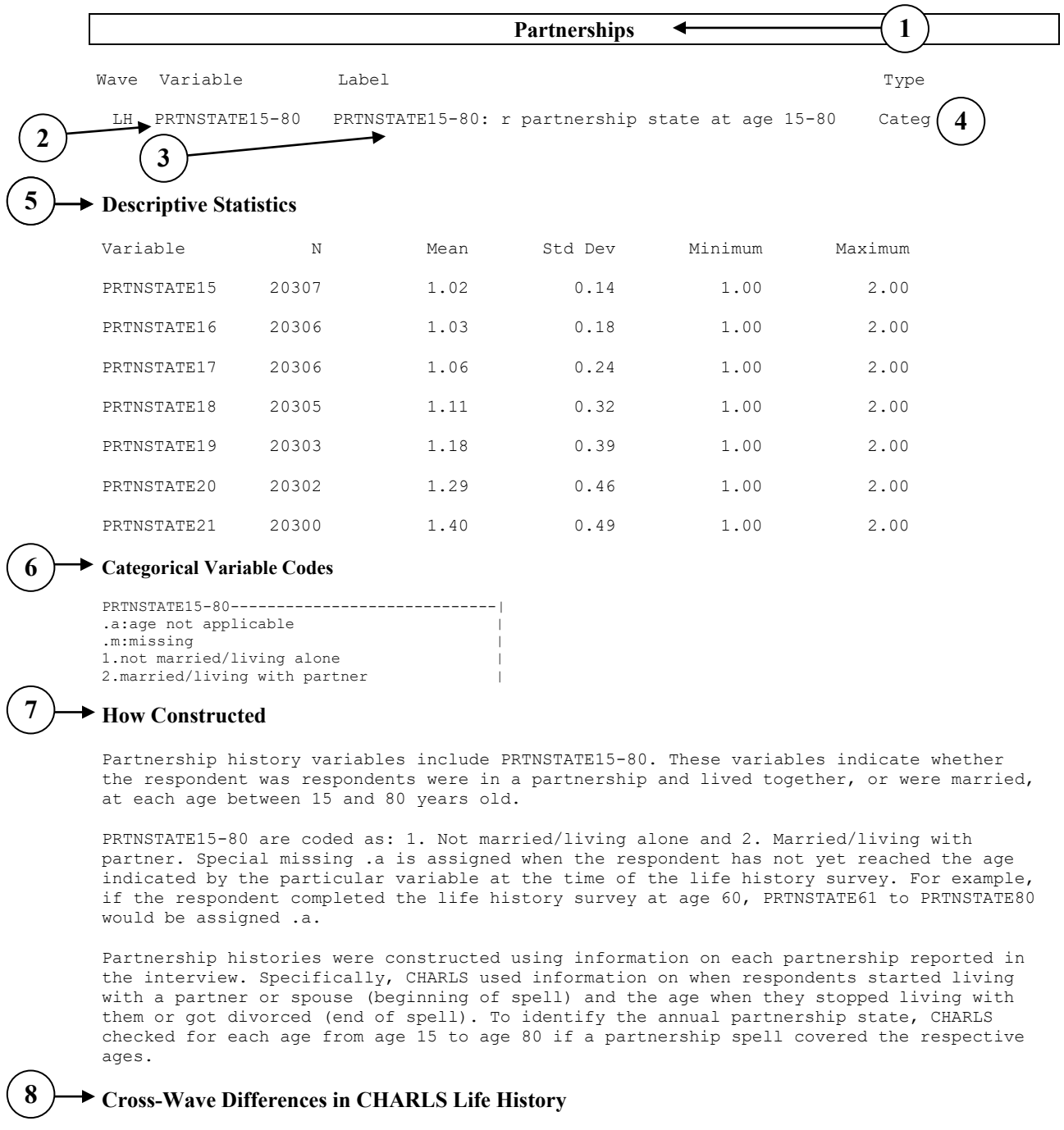
Code	Reason for missing
.d	Don't know
.r	Refused
.a	Not yet reached this age at the time of the survey
.m	Other missing

The coding scheme varies across variables. Consult the Data Codebook for details on individual variables.

Overall, the amount of missing state information due to individuals who do not know or refuse to answer is very low. In that case we assigned simple missing values denoted ".m". In addition, we assigned special missing values in case respondents have not yet reached the age to which the state-variable refers to, denoted as ".a". As such, missing information is gradually higher for variables that refer to older ages, as these state variables could only be identified for older respondents. For example, respondents aged 70 at the life history interview will have missing information for all workstates at higher ages (emplstate71…emplstate80).

2. Structure of Codebook

The Data Codebook contains the codebook documenting all variables in the Harmonized CHARLS Life History Data. This section explains how to interpret the codebook entries. The figure below shows a typical codebook page; the numbers in circles correspond to comments below.



```
There is currently only one wave of CHARLS Life History.
```

9 → **Differences with SHARE/ELSA Life History**

```
CHARLS ask for married and unmarried partnerships separately (different loops according
to marital status), while ELSA asks for partnerships in general (same loops regardless of
marital status) in addition to information on marital status as part of each partnership-
loop.
```

10 → **ELSA Variables Used**

```
Life History:
    P001A                   Have You Ever Been Married?
    P002_1_                 The Year of Marriage
    P002_2_                 The Year of Marriage
    P002_3_                 The Year of Marriage
    P002_4_                 The Year of Marriage
    P002_5_                 The Year of Marriage
    P007_1_                 Still Living with Spouse
    P007_2_                 Still Living with Spouse
    P007_3_                 Still Living with Spouse
    P007_4_                 Still Living with Spouse
    P007_5_                 Still Living with Spouse
    P010_1_                 The Year of Stopping Living with Spouse
    P010_2_                 The Year of Stopping Living with Spouse
    P010_3_                 The Year of Stopping Living with Spouse
    P010_4_                 The Year of Stopping Living with Spouse
    P013A_1_                The Year of Starting Living with Partner
    P013A_2_                The Year of Starting Living with Partner
    P013A_3_                The Year of Starting Living with Partner
    P013A_4_                The Year of Starting Living with Partner
    P017_1_                 Still Living with Partner
    P017_2_                 Still Living with Partner
    P017_3_                 Still Living with Partner
    P017_4_                 Still Living with Partner
    P020_1_                 The Year of Stopping Living with Partner
    P020_2_                 The Year of Stopping Living with Partner
    P020_3_                 The Year of Stopping Living with Partner
```

---

1 *Title*: The variables are documented in groups according to the concept that they measure. The title is often followed by a short description of the concept that is captured.

2 *Variable Names*: This entry shows the names of variables in the group.

3 *Variable Labels*: This entry shows the Stata variable labels. As discussed above, the labels typically include the name of the variable, the file on which it is present, and a description of its contents.

4 *Variable Type*: This entry indicates the type of variable. It may be continuous (Cont), categorical (Categ), or character (Char).

5 *Descriptive Statistics*: This entry shows descriptive statistics on each variable. They include the number of nonmissing values, the mean, standard deviation, minimum value, and maximum value.

6 *Categorical Value Codes*: This entry shows the value label codes. These are

only relevant for categorical variables. The first character(s) of the value labels indicate the value to which each label has been assigned. For example, value "1" is mapped into "1.not married/living alone" (not just "not married/living alone"). These values will be used for all variables in the group referenced (for example PRTNSTATE15-80).

7 *How Constructed*: This entry provides background on the manner in which variables were constructed.

8 *Cross-Wave Differences in CHARLS Life History*: This entry briefly describes differences in question wording or contents between life history interview waves. Note that currently there has only been one CHARLS Life History interview.

9 *Differences with the Harmonized SHARE/ELSA Life History*: This entry describes any differences between the Harmonized CHARLS Life History version of the variable and the Harmonized SHARE Life History or the Harmonized ELSA Life History version of the variable. It is imperative these differences are understood when using harmonized measures.

10 *CHARLS Variables Used*: This entry provides the names and labels of raw CHARLS variables that were used to construct the new variables.

3. Distribution and Technical Notes

The Harmonized CHARLS Life History data file is distributed by the CHARLS team. The Harmonized CHARLS Life History data file is made available free of charge to users who register China Health and Retirement Longitudinal Study webiste http://charls.pku.edu.cn/en and agree to the standard conditions.

For more information on obtaining access to the CHARLS data visit: http://charls.pku.edu.cn/en.

This is version ***A*** of the Harmonized CHARLS Life History data.

A copy of this Harmonized CHARLS Life History codebook can be obtained on the Gateway to Global Aging Data (https://g2aging.org/) under the Download tab.

4. Data Codebook

Section A: Life History

Identifiers

Wave	Variable	Label	Type
LH	ID	Individual ID	Char
LH	HOUSEHOLDID	Household ID	Char
LH	COMMUNITYID	Community ID	Char

How Constructed

ID is the 12-digit character identifier that uniquely identifies each respondent starting in Wave 2 and CHARLS Life History. HouseholdID is a 10-digit character identifier that that uniquely identifies each household starting in Wave 2 and CHARLS Life History.

ID consists of two separate parts, including:

(1) HouseholdID, which is a 10-digit character household identifier to indicate the household each individual belonged to when entering the panel.

(2) PNC, which is a 2-digit character person identifier used by CHARLS to indicate each participant within the household. PN is the numeric version of the person identifier.

Please note that ID and HouseholdID are specific to the CHARLS Wave 2 forward and CHARLS Life History. Starting in wave 2, CHARLS revised the ID and HouseholdID formats to account for splitting of households due to divorce. Wave 2, following waves, and CHARLS Life History employ this 12-digit character individual identifier and this 10-digit character household identifier.

CommunityID is a 7-digit character identifier that identifies which community each household belonged to.

Cross Wave Differences in CHARLS Life History

There is currently only one wave of CHARLS Life History. ID does not merge back with Wave 1.

Differences with the Harmonized SHARE/ELSA Life History

No differences known.

CHARLS Variables Used

Life History:

COMMUNITYID	Community ID
HOUSEHOLDID	Household ID
ID	Individual ID

Demographics

Wave	Variable	Label	Type
LH	RLHAGEY	rLHagey: age in years at life history interview	Cont
LH	RABYEAR	rabyear:r birth year	Cont
LH	RAFBYEAR	rafyear:r flag birth year	Categ
LH	RABMONTH	rabmonth:r birth month	Cont
LH	RAFBMONTH	rafmonth:r flag birth month	Categ
LH	RAIDBYEAR	raidbyear:r ID birth year	Cont
LH	RAIDBMONTH	raidbmonth:r ID birth month	Cont
LH	RAGENDER	ragender:r gender	Categ

Descriptive Statistics

Variable	N	Mean	Std Dev	Minimum	Maximum
RLHAGEY	20590	58.30	10.94	13.00	104.00
RABYEAR	20637	1955.26	10.95	1910.00	2000.00
RAFBYEAR	20638	3.87	0.96	1.00	9.00
RABMONTH	20635	6.57	3.52	0.00	12.00
RAFBMONTH	20638	3.88	1.00	1.00	9.00
RAIDBYEAR	20592	1955.28	10.90	1910.00	2000.00
RAIDBMONTH	20583	17.56	145.66	0.00	1975.00
RAGENDER	20655	1.53	0.50	1.00	2.00

Categorical Variable Codes

Value	RAFBYEAR
.m:missing	18
1.Wave 1	650
2.Wave 2	210
3.Wave 3	2452
4.Wave 4	16913
9.Life History Wave	413

Value	RAFBMONTH
.m:missing	18
1.Wave 1	672
2.Wave 2	206
3.Wave 3	2407
4.Wave 4	16890
9.Life History Wave	463

Value	RAGENDER
.m:missing	1
1.man	9811
2.woman	10844

How Constructed

RLHAGEY indicates respondent's age in years at life history interview. It is derived from respondent's life history interview year and month and respondent's birth year and month. Other missing responses are assigned as .m.

RABYEAR indicates the respondent's reported birth year. RABMONTH indicates the respondent's reported birth month. RABYEAR and RABMONTH are taken from the Harmonized CHARLS dataset. In cases where the Life History respondent did not participate in any core CHARLS wave and therefore they are not included in the Harmonized CHARLS dataset, RABYEAR and RABMONTH are taken from the respondent's report in the Life History interview.

RAFBYEAR and RAFBMONTH are flag variables that indicate the wave at which the respondent reported the value used in RABYEAR and RABMONTH respectively. RAFBYEAR and RAFBMONTH are coded as follow, 1. Wave 1, 2. Wave 2, 3. Wave 3, 4. Wave 4, and 9. Life History Wave. Other missing responses are assigned as .m.

RAIDBYEAR indicates the respondent's birth year from their ID card. RAIDBMONTH indicates the respondent's birth month from their ID card. RAIDBYEAR and RAIDBMONTH are taken from the Harmonized CHARLS dataset. In cases where the Life History respondent did not participate in any core CHARLS wave and therefore they are not included in the Harmonized CHARLS dataset, RAIDBYEAR and RAIDBMONTH are taken from the respondent's report in the Life History interview. Special missing value .q is assigned if the respondent has not participated in an interview since Wave 2 so was never asked to report the birth date listed on their ID card. Don't know, refused, and missing responses are assigned special missing values .d, .r, and .m, respectively. RAIDBYEAR and RAIDBMONTH are coded as blank missing . if the respondent did not respond to this wave.

RAGENDER indicates the respondent's gender. RAGENDER is set to 1 for man and 2 for woman. RAGENDER is taken from the Harmonized CHARLS dataset. In cases where the Life History respondent did not participate in any core CHARLS wave and therefore they are not included in the Harmonized CHARLS dataset, RAGENDER is taken from the respondent's report in the Life History interview. In wave 1, gender is reported in the Demographic module. The gender reported in the demographic module is taken as the respondent's gender in wave 1. Starting in Waved 2, gender is reported in the Demographic module and the Health module. In the Life History wave, gender is reported in the Demographic module and the Residence module. We first take the reported gender values from the Demographic module and in the case that is missing we take reported gender values from Residence module. There are cases where respondents report different genders during the same interview at different modules and cases in which gender reports change between waves. In cases that gender is not consistent between waves or gender is inconsistent between modules, RAGENDER determines gender using the most consistent gender reported.

Cross Wave Differences in CHARLS Life History

There is currently only one wave of CHARLS Life History.

Differences with the Harmonized SHARE/ELSA Life History

Unlike SHARE, CHARLS does not provide birthdate information beyond the birth year. Thus, age is calculated using birth year and month and interview year and month.

CHARLS Variables Used

Life History:

RBIRTHID	Is Birthday Same as ID Card
RBIRTHID_1	Birth Year in ID Card
RBIRTHID_2	Birth Month in ID Card
RBIRTHMONTH	Respondent's Birth Month
RBIRTHYEAR	Respondent's Birth Year

Interview Dates

Wave	Variable	Label	Type
LH	RLHIWM	rLHiwm: life history interview month	Cont
LH	RLHIWY	rLHiwy: life history interview year	Cont

Descriptive Statistics

Variable	N	Mean	Std Dev	Minimum	Maximum
RLHIWM	20656	7.13	0.51	6.00	12.00
RLHIWY	20656	2014.00	0.00	2014.00	2014.00

How Constructed

RLHIWM and RLHIWY indicate the life history interview month and year, respectively.

Cross Wave Differences in CHARLS Life History

There is currently only one wave of CHARLS Life History.

Differences with the Harmonized SHARE/ELSA Life History

No differences known.

CHARLS Variables Used

Life History:

IMONTH	Interview Month
IYEAR	Interview Year

Work Histories

Wave	Variable	Label	Type
LH	C_WRKSTATE15-80	c_wrkstate:r employment state at age 15-80 (CHARLS-specific)	Categ

Descriptive Statistics

Variable	N	Mean	Std Dev	Minimum	Maximum
C_WRKSTATE15	20025	5.91	4.17	1.00	10.00
C_WRKSTATE16	20037	4.88	4.03	1.00	10.00
C_WRKSTATE17	20041	4.11	3.73	1.00	10.00
C_WRKSTATE18	20042	3.54	3.37	1.00	10.00
C_WRKSTATE19	20029	3.19	3.06	1.00	10.00
C_WRKSTATE20	20019	2.92	2.75	1.00	10.00
C_WRKSTATE21	20000	2.78	2.55	1.00	10.00
C_WRKSTATE22	19992	2.70	2.39	1.00	10.00
C_WRKSTATE23	19978	2.66	2.29	1.00	10.00
C_WRKSTATE24	19968	2.64	2.20	1.00	10.00
C_WRKSTATE25	19964	2.64	2.13	1.00	10.00
C_WRKSTATE26	19951	2.65	2.09	1.00	10.00
C_WRKSTATE27	19950	2.66	2.06	1.00	10.00
C_WRKSTATE28	19952	2.69	2.04	1.00	10.00
C_WRKSTATE29	19946	2.71	2.01	1.00	10.00
C_WRKSTATE30	19943	2.74	1.98	1.00	10.00
C_WRKSTATE31	19941	2.77	1.96	1.00	10.00
C_WRKSTATE32	19931	2.79	1.94	1.00	10.00
C_WRKSTATE33	19931	2.81	1.93	1.00	10.00
C_WRKSTATE34	19932	2.83	1.91	1.00	10.00
C_WRKSTATE35	19931	2.85	1.89	1.00	10.00
C_WRKSTATE36	19918	2.87	1.88	1.00	10.00
C_WRKSTATE37	19911	2.89	1.87	1.00	10.00
C_WRKSTATE38	19886	2.92	1.86	1.00	10.00
C_WRKSTATE39	19857	2.95	1.87	1.00	10.00
C_WRKSTATE40	19814	2.98	1.86	1.00	10.00

C_WRKSTATE41	19728	3.00	1.86	1.00	10.00
C_WRKSTATE42	19574	3.03	1.86	1.00	10.00
C_WRKSTATE43	19236	3.05	1.86	1.00	10.00
C_WRKSTATE44	18725	3.07	1.87	1.00	10.00
C_WRKSTATE45	18198	3.11	1.89	1.00	10.00
C_WRKSTATE46	17567	3.13	1.91	1.00	10.00
C_WRKSTATE47	16991	3.16	1.93	1.00	10.00
C_WRKSTATE48	16379	3.20	1.96	1.00	10.00
C_WRKSTATE49	15661	3.24	1.99	1.00	10.00
C_WRKSTATE50	14944	3.36	2.07	1.00	10.00
C_WRKSTATE51	14147	3.41	2.11	1.00	10.00
C_WRKSTATE52	13278	3.44	2.14	1.00	10.00
C_WRKSTATE53	12673	3.48	2.16	1.00	10.00
C_WRKSTATE54	12303	3.53	2.18	1.00	10.00
C_WRKSTATE55	11784	3.62	2.23	1.00	10.00
C_WRKSTATE56	11317	3.68	2.25	1.00	10.00
C_WRKSTATE57	10704	3.73	2.27	1.00	10.00
C_WRKSTATE58	10062	3.79	2.30	1.00	10.00
C_WRKSTATE59	9414	3.86	2.32	1.00	10.00
C_WRKSTATE60	8663	4.10	2.41	1.00	10.00
C_WRKSTATE61	7942	4.19	2.43	1.00	10.00
C_WRKSTATE62	7295	4.25	2.44	1.00	10.00
C_WRKSTATE63	6614	4.29	2.46	1.00	10.00
C_WRKSTATE64	6021	4.35	2.47	1.00	10.00
C_WRKSTATE65	5432	4.44	2.48	1.00	10.00
C_WRKSTATE66	4892	4.51	2.48	1.00	10.00
C_WRKSTATE67	4405	4.58	2.48	1.00	10.00
C_WRKSTATE68	3949	4.67	2.48	1.00	10.00
C_WRKSTATE69	3557	4.79	2.47	1.00	10.00
C_WRKSTATE70	3200	4.89	2.46	1.00	10.00
C_WRKSTATE71	2856	4.99	2.44	1.00	10.00

C_WRKSTATE72	2552	5.11	2.41	1.00	10.00
C_WRKSTATE73	2219	5.19	2.39	1.00	10.00
C_WRKSTATE74	1928	5.26	2.37	1.00	9.00
C_WRKSTATE75	1674	5.34	2.34	1.00	9.00
C_WRKSTATE76	1462	5.44	2.31	1.00	9.00
C_WRKSTATE77	1248	5.57	2.25	1.00	9.00
C_WRKSTATE78	1046	5.68	2.19	1.00	9.00
C_WRKSTATE79	861	5.79	2.13	1.00	9.00
C_WRKSTATE80	724	5.90	2.06	1.00	9.00

Categorical Variable Codes

Value
.a:age not applicable
.m:missing
1.Agri employed
2.Agri self-employed
3.Non-agri employed
4.Non-agri self-employed
5.Non-agri family business
6.Unemployed
7.Retired
8.Home
9.Education
10.Other

How Constructed

Work history variables include C_WRKSTATE15-80. These variables indicate the employment status of the respondent by age from age 15 to age 80.

C_WRKSTATE15-80 are constructed using reported schooling periods, work periods with details on periods when respondents were not working, and timing of processing retirement. These multiple sources of information were used to identify respondents into ten defined employment states. These variables are coded as: 1.Agri employed, 2.Agri self-employed, 3.Non-agri employed, 4.Non-agri self-employed, 5.Non-agri family business, 6.Unemployed, 7.Retired, 8.Home, 9.Education, and 10.Other. Special missing .a is assigned when the respondent has not yet reached the age indicated by the particular variable at the time of the life history survey. For example, if the respondent completed the life history survey at age 60, EMPLSTATE61 to EMPLSTATE80 would be assigned .a. Special missing .m is assigned if it was not possible to assign an employment status to the respondent at that age.

Determining formal education status at a given age first considers the middle school, high school, vocational school, three-year college, four-year college, master's program, and Ph.D. program periods. Besides middle school and high school, the educational period is only counted if the respondent identified the schooling as part of a formal school education and not if the schooling was identified as evening classes, self-study examination, correspondence education, radio and television university, or other.

Details regarding job periods, CHARLS contains information about type of work, start and end time on up to 19 job periods. CHARLS allowed respondents to identify their work using the following work types: own agricultural production and business activities, agricultural employment, non-agricultural employment, non-agricultural self-employment, unpaid household business help, and army. In CHARLS only jobs lasting 6 months or more were recorded. Periods where respondents were not working between two jobs (gaps) were also recorded. CHARLS measured if the gap was because of being "unemployed" (both searching and not searching for a job), "home/family work" (caretaking, gave birth), or a remaining group of "other" (incl.

being sick or disabled, travelling, being in prison, etc.). Respondents could also report that during the gap they were in training or in school.

CHARLS also asked all respondents whether they have processed retirement procedures and if so, they are asked what year they processed retirement procedures. This reported year is used to determine when retirement started.

Formal education periods, employment periods, employment gaps, and retirement timing were combined into one final work history. In the case that respondents reported more than one employment status at the same age, we prioritized non-agricultural and employment jobs over agricultural and self-employment jobs. In the case that respondents reported both an employment period and a job gap or having entered retirement at the same age, we prioritized information regarding job gaps and retirement. Education was only assigned based on schooling periods in the case that the respondent was not working or did not have any job gap information at the same age. All ages before the respondent first started working when the respondent did not report being in school were categorized as "other".

In sum, the harmonized data account for different forms of labor market situation, in terms of work sequences with annual state-information for each year of age between 15 and 80 (66 years). When creating the histories, we aimed to harmonize with the Harmonized CHARLS labor force variables and to incorporate most relevant and frequent states. In case users are interested in more complex, though less comparable, sequences we suggest to modify the provided do-files of the Gateway to create other state-definitions.

Cross Wave Differences in CHARLS Life History

There is currently only one wave of CHARLS Life History.

Differences with the Harmonized SHARE/ELSA Life History

There are many differences between C_WRKSTATE15-80 and the work states provided as part of the Harmonized SHARE and ELSA Life History data. C_WRKSTATE15-80 was designed to be CHARLS-specific and so to more accurately reflect how labor status was asked during the CHALRS Life History interview. Some notable differences are the separation of agricultural and non-agricultural jobs in C_WRKSTATE15-80.

Periods before the first job and after the last job are not considered in CHARLS. Gaps between two jobs episodes were recorded if they lasted three months or longer in ELSA and six months or longer in CHARLS. To describe situations where respondents were not working, ELSA allows for multiple responses and only one response for CHARLS. In the unlikely case that respondents gave multiple responses in ELSA (less than 10% of all gaps between two jobs), information was set to "Other".

CHARLS Variables Used

Life History:

E000S1	Educational Level
E000S10	Educational Level
E000S11	Educational Level
E000S12	Educational Level
E000S2	Educational Level
E000S3	Educational Level
E000S4	Educational Level
E000S5	Educational Level
E000S6	Educational Level
E000S7	Educational Level
E000S8	Educational Level
E000S9	Educational Level
E024B_5_1_	Ending Time: Year
E024B_5_2_	Ending Time: Year
E024B_5_3_	Ending Time: Year
E024B_5_4_	Ending Time: Year
E035B_5_1_	Ending Time: Year
E035B_5_2_	Ending Time: Year
E035B_5_3_	Ending Time: Year
E050B_5_1_	Ending Time: Year

E050B_5_2_	Ending Time: Year
E050B_5_3_	Ending Time: Year
E050B_TYPE_1_	Way
E050B_TYPE_2_	Way
E050B_TYPE_3_	Way
E057B_5_1_	Ending Time: Year
E057B_5_2_	Ending Time: Year
E057B_TYPE_1_	Way
E057B_TYPE_2_	Way
E065B_5_1_	Ending Time: Year
E065B_5_2_	Ending Time: Year
E065B_5_3_	Ending Time: Year
E065B_TYPE_1_	Way
E065B_TYPE_2_	Way
E065B_TYPE_3_	Way
E074B_5_1_	Ending Time: Year
E074B_TYPE_1_	Way
E092B_5_1_	Ending Time: Year
E092B_TYPE_1_	Way
F151	Year Took Retirement Procedures
F_1	Have Ever Worked?
F_1_0	Answer Worked in Former Survey
F_1_1	Are You Sure You Never Worked?
F_2_1_0_	Starting Time for 0_th Work: Year
F_2_1_10_	Starting Time for 10_th Work: Year
F_2_1_11_	Starting Time for 11_th Work: Year
F_2_1_12_	Starting Time for 12_th Work: Year
F_2_1_13_	Starting Time for 13_th Work: Year
F_2_1_14_	Starting Time for 14_th Work: Year
F_2_1_15_	Starting Time for 15_th Work: Year
F_2_1_16_	Starting Time for 16_th Work: Year
F_2_1_17_	Starting Time for 17_th Work: Year
F_2_1_18_	Starting Time for 18_th Work: Year
F_2_1_1_	Starting Time for 1_th Work: Year
F_2_1_2_	Starting Time for 2_th Work: Year
F_2_1_3_	Starting Time for 3_th Work: Year
F_2_1_4_	Starting Time for 4_th Work: Year
F_2_1_5_	Starting Time for 5_th Work: Year
F_2_1_6_	Starting Time for 6_th Work: Year
F_2_1_7_	Starting Time for 7_th Work: Year
F_2_1_8_	Starting Time for 8_th Work: Year
F_2_1_9_	Starting Time for 9_th Work: Year
F_2_5_0_	Ending Time for 0_th Work: Year
F_2_5_10_	Ending Time for 10_th Work: Year
F_2_5_11_	Ending Time for 11_th Work: Year
F_2_5_12_	Ending Time for 12_th Work: Year
F_2_5_13_	Ending Time for 13_th Work: Year
F_2_5_14_	Ending Time for 14_th Work: Year
F_2_5_15_	Ending Time for 15_th Work: Year
F_2_5_16_	Ending Time for 16_th Work: Year
F_2_5_17_	Ending Time for 17_th Work: Year
F_2_5_18_	Ending Time for 18_th Work: Year
F_2_5_1_	Ending Time for 1_th Work: Year
F_2_5_2_	Ending Time for 2_th Work: Year
F_2_5_3_	Ending Time for 3_th Work: Year
F_2_5_4_	Ending Time for 4_th Work: Year
F_2_5_5_	Ending Time for 5_th Work: Year
F_2_5_6_	Ending Time for 6_th Work: Year
F_2_5_7_	Ending Time for 7_th Work: Year
F_2_5_8_	Ending Time for 8_th Work: Year
F_2_5_9_	Ending Time for 9_th Work: Year
F_3_0_	Category of the 0_th Job
F_3_10_	Category of the 10_th Job

F_3_11_	Category of the 11_th Job
F_3_12_	Category of the 12_th Job
F_3_13_	Category of the 13_th Job
F_3_14_	Category of the 14_th Job
F_3_15_	Category of the 15_th Job
F_3_16_	Category of the 16_th Job
F_3_17_	Category of the 17_th Job
F_3_18_	Category of the 18_th Job
F_3_1_	Category of the 1_th Job
F_3_2_	Category of the 2_th Job
F_3_3_	Category of the 3_th Job
F_3_4_	Category of the 4_th Job
F_3_5_	Category of the 5_th Job
F_3_6_	Category of the 6_th Job
F_3_7_	Category of the 7_th Job
F_3_8_	Category of the 8_th Job
F_3_9_	Category of the 9_th Job
F_6_10_S1	Check Status Before 10_th Job:Unemployed but
F_6_10_S10	Check Status Before 10_th Job:In Prison
F_6_10_S11	Check Status Before 10_th Job:In School
F_6_10_S12	Check Status Before 10_th Job:Other
F_6_10_S2	Check Status Before 10_th Job:Unemployed but
F_6_10_S3	Check Status Before 10_th Job:Illness
F_6_10_S4	Check Status Before 10_th Job:Give Birth
F_6_10_S5	Check Status Before 10_th Job:Take Care of Ch
F_6_10_S6	Check Status Before 10_th Job:Take Care of El
F_6_10_S7	Check Status Before 10_th Job:Take Care of Ot
F_6_10_S8	Check Status Before 10_th Job:Leisure,Traveli
F_6_10_S9	Check Status Before 10_th Job:Training
F_6_11_S1	Check Status Before 11_th Job:Unemployed but
F_6_11_S10	Check Status Before 11_th Job:In Prison
F_6_11_S11	Check Status Before 11_th Job:In School
F_6_11_S12	Check Status Before 11_th Job:Other
F_6_11_S2	Check Status Before 11_th Job:Unemployed but
F_6_11_S3	Check Status Before 11_th Job:Illness
F_6_11_S4	Check Status Before 11_th Job:Give Birth
F_6_11_S5	Check Status Before 11_th Job:Take Care of Ch
F_6_11_S6	Check Status Before 11_th Job:Take Care of El
F_6_11_S7	Check Status Before 11_th Job:Take Care of Ot
F_6_11_S8	Check Status Before 11_th Job:Leisure,Traveli
F_6_11_S9	Check Status Before 11_th Job:Training
F_6_12_S1	Check Status Before 12_th Job:Unemployed but
F_6_12_S10	Check Status Before 12_th Job:In Prison
F_6_12_S11	Check Status Before 12_th Job:In School
F_6_12_S12	Check Status Before 12_th Job:Other
F_6_12_S2	Check Status Before 12_th Job:Unemployed but
F_6_12_S3	Check Status Before 12_th Job:Illness
F_6_12_S4	Check Status Before 12_th Job:Give Birth
F_6_12_S5	Check Status Before 12_th Job:Take Care of Ch
F_6_12_S6	Check Status Before 12_th Job:Take Care of El
F_6_12_S7	Check Status Before 12_th Job:Take Care of Ot
F_6_12_S8	Check Status Before 12_th Job:Leisure,Traveli
F_6_12_S9	Check Status Before 12_th Job:Training
F_6_14_S1	Check Status Before 14_th Job:Unemployed but
F_6_14_S10	Check Status Before 14_th Job:In Prison
F_6_14_S11	Check Status Before 14_th Job:In School
F_6_14_S12	Check Status Before 14_th Job:Other
F_6_14_S2	Check Status Before 14_th Job:Unemployed but
F_6_14_S3	Check Status Before 14_th Job:Illness
F_6_14_S4	Check Status Before 14_th Job:Give Birth
F_6_14_S5	Check Status Before 14_th Job:Take Care of Ch
F_6_14_S6	Check Status Before 14_th Job:Take Care of El
F_6_14_S7	Check Status Before 14_th Job:Take Care of Ot

F_6_14_S8	Check Status Before 14_th Job:Leisure,Traveli
F_6_14_S9	Check Status Before 14_th Job:Training
F_6_2_S1	Check Status Before 2_th Job:Unemployed but L
F_6_2_S10	Check Status Before 2_th Job:In Prison
F_6_2_S11	Check Status Before 2_th Job:In School
F_6_2_S12	Check Status Before 2_th Job:Other
F_6_2_S2	Check Status Before 2_th Job:Unemployed but N
F_6_2_S3	Check Status Before 2_th Job:Illness
F_6_2_S4	Check Status Before 2_th Job:Give Birth
F_6_2_S5	Check Status Before 2_th Job:Take Care of Chi
F_6_2_S6	Check Status Before 2_th Job:Take Care of Eld
F_6_2_S7	Check Status Before 2_th Job:Take Care of Oth
F_6_2_S8	Check Status Before 2_th Job:Leisure,Travelin
F_6_2_S9	Check Status Before 2_th Job:Training
F_6_3_S1	Check Status Before 3_th Job:Unemployed but L
F_6_3_S10	Check Status Before 3_th Job:In Prison
F_6_3_S11	Check Status Before 3_th Job:In School
F_6_3_S12	Check Status Before 3_th Job:Other
F_6_3_S2	Check Status Before 3_th Job:Unemployed but N
F_6_3_S3	Check Status Before 3_th Job:Illness
F_6_3_S4	Check Status Before 3_th Job:Give Birth
F_6_3_S5	Check Status Before 3_th Job:Take Care of Chi
F_6_3_S6	Check Status Before 3_th Job:Take Care of Eld
F_6_3_S7	Check Status Before 3_th Job:Take Care of Oth
F_6_3_S8	Check Status Before 3_th Job:Leisure,Travelin
F_6_3_S9	Check Status Before 3_th Job:Training
F_6_4_S1	Check Status Before 4_th Job:Unemployed but L
F_6_4_S10	Check Status Before 4_th Job:In Prison
F_6_4_S11	Check Status Before 4_th Job:In School
F_6_4_S12	Check Status Before 4_th Job:Other
F_6_4_S2	Check Status Before 4_th Job:Unemployed but N
F_6_4_S3	Check Status Before 4_th Job:Illness
F_6_4_S4	Check Status Before 4_th Job:Give Birth
F_6_4_S5	Check Status Before 4_th Job:Take Care of Chi
F_6_4_S6	Check Status Before 4_th Job:Take Care of Eld
F_6_4_S7	Check Status Before 4_th Job:Take Care of Oth
F_6_4_S8	Check Status Before 4_th Job:Leisure,Travelin
F_6_4_S9	Check Status Before 4_th Job:Training
F_6_5_S1	Check Status Before 5_th Job:Unemployed but L
F_6_5_S10	Check Status Before 5_th Job:In Prison
F_6_5_S11	Check Status Before 5_th Job:In School
F_6_5_S12	Check Status Before 5_th Job:Other
F_6_5_S2	Check Status Before 5_th Job:Unemployed but N
F_6_5_S3	Check Status Before 5_th Job:Illness
F_6_5_S4	Check Status Before 5_th Job:Give Birth
F_6_5_S5	Check Status Before 5_th Job:Take Care of Chi
F_6_5_S6	Check Status Before 5_th Job:Take Care of Eld
F_6_5_S7	Check Status Before 5_th Job:Take Care of Oth
F_6_5_S8	Check Status Before 5_th Job:Leisure,Travelin
F_6_5_S9	Check Status Before 5_th Job:Training
F_6_6_S1	Check Status Before 6_th Job:Unemployed but L
F_6_6_S10	Check Status Before 6_th Job:In Prison
F_6_6_S11	Check Status Before 6_th Job:In School
F_6_6_S12	Check Status Before 6_th Job:Other
F_6_6_S2	Check Status Before 6_th Job:Unemployed but N
F_6_6_S3	Check Status Before 6_th Job:Illness
F_6_6_S4	Check Status Before 6_th Job:Give Birth
F_6_6_S5	Check Status Before 6_th Job:Take Care of Chi
F_6_6_S6	Check Status Before 6_th Job:Take Care of Eld
F_6_6_S7	Check Status Before 6_th Job:Take Care of Oth
F_6_6_S8	Check Status Before 6_th Job:Leisure,Travelin
F_6_6_S9	Check Status Before 6_th Job:Training
F_6_7_S1	Check Status Before 7_th Job:Unemployed but L

F_6_7_S10	Check Status Before 7_th Job:In Prison
F_6_7_S11	Check Status Before 7_th Job:In School
F_6_7_S12	Check Status Before 7_th Job:Other
F_6_7_S2	Check Status Before 7_th Job:Unemployed but N
F_6_7_S3	Check Status Before 7_th Job:Illness
F_6_7_S4	Check Status Before 7_th Job:Give Birth
F_6_7_S5	Check Status Before 7_th Job:Take Care of Chi
F_6_7_S6	Check Status Before 7_th Job:Take Care of Eld
F_6_7_S7	Check Status Before 7_th Job:Take Care of Oth
F_6_7_S8	Check Status Before 7_th Job:Leisure,Travelin
F_6_7_S9	Check Status Before 7_th Job:Training
F_6_8_S1	Check Status Before 8_th Job:Unemployed but L
F_6_8_S10	Check Status Before 8_th Job:In Prison
F_6_8_S11	Check Status Before 8_th Job:In School
F_6_8_S12	Check Status Before 8_th Job:Other
F_6_8_S2	Check Status Before 8_th Job:Unemployed but N
F_6_8_S3	Check Status Before 8_th Job:Illness
F_6_8_S4	Check Status Before 8_th Job:Give Birth
F_6_8_S5	Check Status Before 8_th Job:Take Care of Chi
F_6_8_S6	Check Status Before 8_th Job:Take Care of Eld
F_6_8_S7	Check Status Before 8_th Job:Take Care of Oth
F_6_8_S8	Check Status Before 8_th Job:Leisure,Travelin
F_6_8_S9	Check Status Before 8_th Job:Training
F_6_9_S1	Check Status Before 9_th Job:Unemployed but L
F_6_9_S10	Check Status Before 9_th Job:In Prison
F_6_9_S11	Check Status Before 9_th Job:In School
F_6_9_S12	Check Status Before 9_th Job:Other
F_6_9_S2	Check Status Before 9_th Job:Unemployed but N
F_6_9_S3	Check Status Before 9_th Job:Illness
F_6_9_S4	Check Status Before 9_th Job:Give Birth
F_6_9_S5	Check Status Before 9_th Job:Take Care of Chi
F_6_9_S6	Check Status Before 9_th Job:Take Care of Eld
F_6_9_S7	Check Status Before 9_th Job:Take Care of Oth
F_6_9_S8	Check Status Before 9_th Job:Leisure,Travelin
F_6_9_S9	Check Status Before 9_th Job:Training

Partnerships

Wave	Variable	Label	Type
LH	PRTNSTATE15-80	prtnstate:r partnership state at age 15-80	Categ

Descriptive Statistics

Variable	N	Mean	Std Dev	Minimum	Maximum
PRTNSTATE15	20307	1.02	0.14	1.00	2.00
PRTNSTATE16	20306	1.03	0.18	1.00	2.00
PRTNSTATE17	20306	1.06	0.24	1.00	2.00
PRTNSTATE18	20305	1.11	0.32	1.00	2.00
PRTNSTATE19	20303	1.18	0.39	1.00	2.00
PRTNSTATE20	20302	1.29	0.46	1.00	2.00
PRTNSTATE21	20300	1.40	0.49	1.00	2.00
PRTNSTATE22	20299	1.52	0.50	1.00	2.00
PRTNSTATE23	20299	1.64	0.48	1.00	2.00
PRTNSTATE24	20297	1.74	0.44	1.00	2.00
PRTNSTATE25	20297	1.82	0.39	1.00	2.00
PRTNSTATE26	20294	1.87	0.34	1.00	2.00
PRTNSTATE27	20294	1.90	0.29	1.00	2.00
PRTNSTATE28	20292	1.93	0.26	1.00	2.00
PRTNSTATE29	20291	1.94	0.24	1.00	2.00
PRTNSTATE30	20288	1.95	0.21	1.00	2.00
PRTNSTATE31	20287	1.96	0.20	1.00	2.00
PRTNSTATE32	20283	1.96	0.20	1.00	2.00
PRTNSTATE33	20280	1.96	0.19	1.00	2.00
PRTNSTATE34	20277	1.96	0.19	1.00	2.00
PRTNSTATE35	20272	1.97	0.18	1.00	2.00
PRTNSTATE36	20257	1.97	0.18	1.00	2.00
PRTNSTATE37	20244	1.97	0.18	1.00	2.00
PRTNSTATE38	20215	1.97	0.18	1.00	2.00
PRTNSTATE39	20184	1.96	0.18	1.00	2.00
PRTNSTATE40	20129	1.96	0.19	1.00	2.00

PRTNSTATE41	20041	1.96	0.19	1.00	2.00
PRTNSTATE42	19881	1.96	0.19	1.00	2.00
PRTNSTATE43	19539	1.96	0.20	1.00	2.00
PRTNSTATE44	19021	1.96	0.20	1.00	2.00
PRTNSTATE45	18493	1.96	0.20	1.00	2.00
PRTNSTATE46	17859	1.96	0.21	1.00	2.00
PRTNSTATE47	17285	1.95	0.21	1.00	2.00
PRTNSTATE48	16671	1.95	0.21	1.00	2.00
PRTNSTATE49	15950	1.95	0.22	1.00	2.00
PRTNSTATE50	15229	1.95	0.23	1.00	2.00
PRTNSTATE51	14422	1.94	0.23	1.00	2.00
PRTNSTATE52	13549	1.94	0.24	1.00	2.00
PRTNSTATE53	12943	1.94	0.24	1.00	2.00
PRTNSTATE54	12569	1.93	0.25	1.00	2.00
PRTNSTATE55	12050	1.93	0.26	1.00	2.00
PRTNSTATE56	11583	1.92	0.27	1.00	2.00
PRTNSTATE57	10963	1.92	0.28	1.00	2.00
PRTNSTATE58	10321	1.91	0.29	1.00	2.00
PRTNSTATE59	9677	1.90	0.30	1.00	2.00
PRTNSTATE60	8921	1.89	0.31	1.00	2.00
PRTNSTATE61	8188	1.89	0.32	1.00	2.00
PRTNSTATE62	7534	1.88	0.33	1.00	2.00
PRTNSTATE63	6835	1.87	0.34	1.00	2.00
PRTNSTATE64	6227	1.86	0.35	1.00	2.00
PRTNSTATE65	5631	1.85	0.36	1.00	2.00
PRTNSTATE66	5080	1.83	0.37	1.00	2.00
PRTNSTATE67	4585	1.82	0.38	1.00	2.00
PRTNSTATE68	4121	1.81	0.39	1.00	2.00
PRTNSTATE69	3726	1.79	0.40	1.00	2.00
PRTNSTATE70	3354	1.77	0.42	1.00	2.00
PRTNSTATE71	3003	1.76	0.43	1.00	2.00

PRTNSTATE72	2688	1.73	0.44	1.00	2.00
PRTNSTATE73	2348	1.72	0.45	1.00	2.00
PRTNSTATE74	2042	1.70	0.46	1.00	2.00
PRTNSTATE75	1781	1.68	0.47	1.00	2.00
PRTNSTATE76	1563	1.66	0.47	1.00	2.00
PRTNSTATE77	1338	1.64	0.48	1.00	2.00
PRTNSTATE78	1131	1.62	0.49	1.00	2.00
PRTNSTATE79	936	1.60	0.49	1.00	2.00
PRTNSTATE80	796	1.57	0.50	1.00	2.00

Categorical Variable Codes

Value
.a:age not applicable
.m:missing
1.not married/living alone
2.married/living with partner

How Constructed

Partnership history variables include PRTNSTATE15-80. These variables indicate whether the respondent was respondents were in a partnership and lived together, or were married, at each age between 15 and 80 years old.

PRTNSTATE15-80 are coded as: 1. Not married/living alone and 2. Married/living with partner. Special missing .a is assigned when the respondent has not yet reached the age indicated by the particular variable at the time of the life history survey. For example, if the respondent completed the life history survey at age 60, PRTNSTATE61 to PRTNSTATE80 would be assigned .a.

Partnership histories were constructed using information on each partnership reported in the interview. Specifically, CHARLS used information on when respondents started living with a partner or spouse (beginning of spell) and the age when they stopped living with them or got divorced (end of spell). To identify the annual partnership state, CHARLS checked for each age from age 15 to age 80 if a partnership spell covered the respective ages.

Cross Wave Differences in CHARLS Life History

There is currently only one wave of CHARLS Life History.

Differences with the Harmonized SHARE/ELSA Life History

CHARLS ask for married and unmarried partnerships separately (different loops according to marital status), while ELSA asks for partnerships in general (same loops regardless of marital status) in addition to information on marital status as part of each partnership-loop.

CHARLS Variables Used

Life History:

P001A	Have You Ever Been Married?
P002_1_	The Year of Marriage
P002_2_	The Year of Marriage
P002_3_	The Year of Marriage
P002_4_	The Year of Marriage

P002_5_	The Year of Marriage
P007_1_	Still Living with Spouse
P007_2_	Still Living with Spouse
P007_3_	Still Living with Spouse
P007_4_	Still Living with Spouse
P007_5_	Still Living with Spouse
P010_1_	The Year of Stopping Living with Spouse
P010_2_	The Year of Stopping Living with Spouse
P010_3_	The Year of Stopping Living with Spouse
P010_4_	The Year of Stopping Living with Spouse
P013A_1_	The Year of Starting Living with Partner
P013A_2_	The Year of Starting Living with Partner
P013A_3_	The Year of Starting Living with Partner
P013A_4_	The Year of Starting Living with Partner
P017_1_	Still Living with Partner
P017_2_	Still Living with Partner
P017_3_	Still Living with Partner
P017_4_	Still Living with Partner
P020_1_	The Year of Stopping Living with Partner
P020_2_	The Year of Stopping Living with Partner
P020_3_	The Year of Stopping Living with Partner

Children

Wave	Variable	Label	Type
LH	CHLDSTATE15-80	chldstate:r total number of living children at age 15-80	Cont
LH	YCHDSTATE15-80	ychdstate:r total number of children under the age of 18 at	Cont

Descriptive Statistics

Variable	N	Mean	Std Dev	Minimum	Maximum
CHLDSTATE15	20655	0.03	0.31	0.00	7.00
CHLDSTATE16	20654	0.04	0.32	0.00	7.00
CHLDSTATE17	20653	0.05	0.36	0.00	7.00
CHLDSTATE18	20652	0.08	0.40	0.00	7.00
CHLDSTATE19	20650	0.13	0.47	0.00	7.00
CHLDSTATE20	20649	0.21	0.58	0.00	7.00
CHLDSTATE21	20647	0.33	0.70	0.00	8.00
CHLDSTATE22	20646	0.50	0.82	0.00	8.00
CHLDSTATE23	20646	0.70	0.93	0.00	8.00
CHLDSTATE24	20644	0.92	1.02	0.00	9.00
CHLDSTATE25	20644	1.15	1.09	0.00	10.00
CHLDSTATE26	20641	1.36	1.14	0.00	10.00
CHLDSTATE27	20641	1.56	1.19	0.00	10.00
CHLDSTATE28	20639	1.73	1.23	0.00	10.00
CHLDSTATE29	20638	1.88	1.27	0.00	10.00
CHLDSTATE30	20635	2.01	1.30	0.00	10.00
CHLDSTATE31	20634	2.13	1.34	0.00	11.00
CHLDSTATE32	20630	2.22	1.37	0.00	11.00
CHLDSTATE33	20627	2.30	1.40	0.00	11.00
CHLDSTATE34	20624	2.37	1.42	0.00	11.00
CHLDSTATE35	20618	2.42	1.45	0.00	11.00
CHLDSTATE36	20603	2.47	1.48	0.00	11.00
CHLDSTATE37	20590	2.51	1.50	0.00	12.00
CHLDSTATE38	20561	2.55	1.53	0.00	12.00
CHLDSTATE39	20529	2.57	1.55	0.00	13.00

CHLDSTATE40	20473	2.60	1.56	0.00	13.00
CHLDSTATE41	20385	2.62	1.58	0.00	14.00
CHLDSTATE42	20223	2.64	1.59	0.00	15.00
CHLDSTATE43	19877	2.67	1.60	0.00	15.00
CHLDSTATE44	19352	2.71	1.61	0.00	15.00
CHLDSTATE45	18818	2.75	1.61	0.00	15.00
CHLDSTATE46	18172	2.79	1.62	0.00	17.00
CHLDSTATE47	17583	2.83	1.63	0.00	17.00
CHLDSTATE48	16960	2.87	1.63	0.00	17.00
CHLDSTATE49	16229	2.91	1.65	0.00	17.00
CHLDSTATE50	15497	2.96	1.66	0.00	17.00
CHLDSTATE51	14683	3.01	1.68	0.00	17.00
CHLDSTATE52	13799	3.06	1.69	0.00	17.00
CHLDSTATE53	13185	3.10	1.70	0.00	17.00
CHLDSTATE54	12806	3.13	1.71	0.00	17.00
CHLDSTATE55	12282	3.17	1.72	0.00	17.00
CHLDSTATE56	11806	3.21	1.73	0.00	17.00
CHLDSTATE57	11182	3.25	1.74	0.00	17.00
CHLDSTATE58	10527	3.32	1.76	0.00	17.00
CHLDSTATE59	9875	3.38	1.78	0.00	17.00
CHLDSTATE60	9110	3.45	1.79	0.00	17.00
CHLDSTATE61	8363	3.53	1.81	0.00	17.00
CHLDSTATE62	7698	3.60	1.82	0.00	17.00
CHLDSTATE63	6990	3.68	1.85	0.00	17.00
CHLDSTATE64	6374	3.76	1.87	0.00	17.00
CHLDSTATE65	5772	3.84	1.90	0.00	17.00
CHLDSTATE66	5213	3.92	1.93	0.00	17.00
CHLDSTATE67	4705	3.99	1.95	0.00	17.00
CHLDSTATE68	4232	4.06	1.98	0.00	17.00
CHLDSTATE69	3833	4.12	2.01	0.00	17.00
CHLDSTATE70	3455	4.17	2.02	0.00	17.00

CHLDSTATE71	3097	4.22	2.07	0.00	17.00
CHLDSTATE72	2778	4.26	2.09	0.00	17.00
CHLDSTATE73	2432	4.31	2.12	0.00	17.00
CHLDSTATE74	2121	4.36	2.16	0.00	17.00
CHLDSTATE75	1853	4.39	2.19	0.00	17.00
CHLDSTATE76	1631	4.43	2.23	0.00	17.00
CHLDSTATE77	1402	4.45	2.28	0.00	17.00
CHLDSTATE78	1187	4.49	2.32	0.00	17.00
CHLDSTATE79	991	4.48	2.38	0.00	17.00
CHLDSTATE80	846	4.41	2.40	0.00	17.00
YCHDSTATE15	20655	0.03	0.27	0.00	6.00
YCHDSTATE16	20654	0.03	0.28	0.00	6.00
YCHDSTATE17	20653	0.04	0.32	0.00	6.00
YCHDSTATE18	20652	0.07	0.37	0.00	6.00
YCHDSTATE19	20650	0.12	0.44	0.00	7.00
YCHDSTATE20	20649	0.20	0.55	0.00	7.00
YCHDSTATE21	20647	0.33	0.67	0.00	8.00
YCHDSTATE22	20646	0.49	0.80	0.00	8.00
YCHDSTATE23	20646	0.69	0.90	0.00	8.00
YCHDSTATE24	20644	0.91	0.99	0.00	8.00
YCHDSTATE25	20644	1.14	1.07	0.00	9.00
YCHDSTATE26	20641	1.35	1.12	0.00	9.00
YCHDSTATE27	20641	1.55	1.16	0.00	9.00
YCHDSTATE28	20639	1.72	1.21	0.00	9.00
YCHDSTATE29	20638	1.86	1.24	0.00	9.00
YCHDSTATE30	20635	1.99	1.28	0.00	10.00
YCHDSTATE31	20634	2.11	1.31	0.00	10.00
YCHDSTATE32	20630	2.20	1.34	0.00	10.00
YCHDSTATE33	20627	2.27	1.37	0.00	10.00
YCHDSTATE34	20624	2.33	1.39	0.00	10.00
YCHDSTATE35	20618	2.37	1.41	0.00	10.00
YCHDSTATE36	20603	2.40	1.42	0.00	11.00

YCHDSTATE37	20590	2.39	1.42	0.00	12.00
YCHDSTATE38	20561	2.34	1.42	0.00	12.00
YCHDSTATE39	20529	2.24	1.42	0.00	12.00
YCHDSTATE40	20473	2.10	1.42	0.00	12.00
YCHDSTATE41	20385	1.92	1.43	0.00	13.00
YCHDSTATE42	20223	1.72	1.42	0.00	12.00
YCHDSTATE43	19877	1.51	1.39	0.00	12.00
YCHDSTATE44	19352	1.32	1.35	0.00	11.00
YCHDSTATE45	18818	1.14	1.30	0.00	11.00
YCHDSTATE46	18172	0.98	1.25	0.00	11.00
YCHDSTATE47	17583	0.84	1.18	0.00	11.00
YCHDSTATE48	16960	0.72	1.11	0.00	11.00
YCHDSTATE49	16229	0.61	1.04	0.00	9.00
YCHDSTATE50	15497	0.52	0.96	0.00	9.00
YCHDSTATE51	14683	0.44	0.89	0.00	8.00
YCHDSTATE52	13799	0.38	0.82	0.00	8.00
YCHDSTATE53	13185	0.32	0.76	0.00	8.00
YCHDSTATE54	12806	0.26	0.68	0.00	7.00
YCHDSTATE55	12282	0.22	0.62	0.00	7.00
YCHDSTATE56	11806	0.18	0.55	0.00	7.00
YCHDSTATE57	11182	0.14	0.50	0.00	7.00
YCHDSTATE58	10527	0.12	0.45	0.00	6.00
YCHDSTATE59	9875	0.10	0.41	0.00	5.00
YCHDSTATE60	9110	0.08	0.37	0.00	5.00
YCHDSTATE61	8363	0.07	0.35	0.00	5.00
YCHDSTATE62	7698	0.06	0.32	0.00	5.00
YCHDSTATE63	6990	0.05	0.30	0.00	5.00
YCHDSTATE64	6374	0.05	0.27	0.00	6.00
YCHDSTATE65	5772	0.04	0.26	0.00	6.00
YCHDSTATE66	5213	0.04	0.25	0.00	6.00
YCHDSTATE67	4705	0.03	0.23	0.00	5.00

YCHDSTATE68	4232	0.03	0.24	0.00	5.00
YCHDSTATE69	3833	0.03	0.22	0.00	6.00
YCHDSTATE70	3455	0.02	0.19	0.00	5.00
YCHDSTATE71	3097	0.03	0.21	0.00	6.00
YCHDSTATE72	2778	0.02	0.20	0.00	6.00
YCHDSTATE73	2432	0.02	0.18	0.00	5.00
YCHDSTATE74	2121	0.02	0.22	0.00	5.00
YCHDSTATE75	1853	0.02	0.22	0.00	5.00
YCHDSTATE76	1631	0.02	0.21	0.00	5.00
YCHDSTATE77	1402	0.02	0.24	0.00	5.00
YCHDSTATE78	1187	0.02	0.19	0.00	4.00
YCHDSTATE79	991	0.03	0.24	0.00	3.00
YCHDSTATE80	846	0.03	0.26	0.00	4.00

How Constructed

Children histories provide information regarding whether the respondent had children and how many children they had at each age between age 15 and age 80. CHLDNSTATE15-80 counts the total number of children a person had at a specific age. YCHDSTATE15-80 also considers the age of children and counts the number of children below age 18 for each age. In both cases, all children (biological, adopted) are considered. With each variable indicating the number of children at a particular age, if the respondent has not yet reached the variable age, special missing .a is assigned, as it is not applicable. For example, if the respondent's age were 60, CHLDNSTATE61 to CHLDNSTATE80 and YCHDSTATE61 to YCHDSTATE80 would be assigned .a.

To define the beginning of individual children histories, CHARLS used the year of birth of children. In addition, CHARLS used information on year of death in case that a child died. CHARLSE first checked for each age from age 15 to age 80 how many children the respondent had, specifically, how many children were born at or before the respective age and not died before. This leads to the first set of variables counting the total number of living children. Secondly, CHARLS considered the age of the child (based on year of birth) and counted how many children respondents had that were below 18 for each age of the respondent from age 15 to age 80.

In sum, the harmonized data allows to count whether and how many children a respondent had for each year of age between 15 and 80, either irrespective the age of children or for those who have not reached the legal age yet (below 18). In case users are interested in counting the number of children at other ages, users may modify the provided do-files of the Gateway.

Cross Wave Differences in CHARLS Life History

There is currently only one wave of CHARLS Life History.

Differences with the Harmonized SHARE/ELSA Life History

SHARE asks for biological and adopted children separately (different loops according to relationship status), while CHARLS asks for children in general (same loops regardless of relationship status) in addition to information on the relationship status towards each child.

CHARLS Variables Used

Life History:

C001A_1_10_	The Year of Death
C001A_1_11_	The Year of Death
C001A_1_1_	The Year of Death
C001A_1_2_	The Year of Death
C001A_1_3_	The Year of Death
C001A_1_4_	The Year of Death
C001A_1_5_	The Year of Death
C001A_1_6_	The Year of Death
C001A_1_7_	The Year of Death
C001A_1_8_	The Year of Death
C001A_1_9_	The Year of Death
CHILDYEAR_10_	The Birth Year of Child
CHILDYEAR_11_	The Birth Year of Child
CHILDYEAR_12_	The Birth Year of Child
CHILDYEAR_13_	The Birth Year of Child
CHILDYEAR_14_	The Birth Year of Child
CHILDYEAR_15_	The Birth Year of Child
CHILDYEAR_16_	The Birth Year of Child
CHILDYEAR_17_	The Birth Year of Child
CHILDYEAR_18_	The Birth Year of Child
CHILDYEAR_1_	The Birth Year of Child
CHILDYEAR_2_	The Birth Year of Child
CHILDYEAR_3_	The Birth Year of Child
CHILDYEAR_4_	The Birth Year of Child
CHILDYEAR_5_	The Birth Year of Child
CHILDYEAR_6_	The Birth Year of Child
CHILDYEAR_7_	The Birth Year of Child
CHILDYEAR_8_	The Birth Year of Child
CHILDYEAR_9_	The Birth Year of Child

Accommodations

Wave	Variable	Label	Type
LH	HKSTATE15-80	hkstate:r state of hukou type at age 15-80	Categ

Descriptive Statistics

Variable	N	Mean	Std Dev	Minimum	Maximum
HKSTATE15	20118	1.11	0.34	1.00	4.00
HKSTATE16	20117	1.11	0.34	1.00	4.00
HKSTATE17	20117	1.11	0.34	1.00	4.00
HKSTATE18	20117	1.11	0.34	1.00	4.00
HKSTATE19	20115	1.11	0.34	1.00	4.00
HKSTATE20	20116	1.11	0.34	1.00	4.00
HKSTATE21	20114	1.11	0.34	1.00	4.00
HKSTATE22	20113	1.11	0.34	1.00	4.00
HKSTATE23	20113	1.11	0.35	1.00	4.00
HKSTATE24	20111	1.11	0.35	1.00	4.00
HKSTATE25	20111	1.11	0.35	1.00	4.00
HKSTATE26	20108	1.11	0.35	1.00	4.00
HKSTATE27	20108	1.11	0.35	1.00	4.00
HKSTATE28	20106	1.11	0.35	1.00	4.00
HKSTATE29	20105	1.11	0.35	1.00	4.00
HKSTATE30	20102	1.11	0.35	1.00	4.00
HKSTATE31	20101	1.11	0.35	1.00	4.00
HKSTATE32	20097	1.11	0.35	1.00	4.00
HKSTATE33	20094	1.12	0.35	1.00	4.00
HKSTATE34	20091	1.12	0.35	1.00	4.00
HKSTATE35	20085	1.12	0.35	1.00	4.00
HKSTATE36	20070	1.12	0.35	1.00	4.00
HKSTATE37	20059	1.12	0.35	1.00	4.00
HKSTATE38	20032	1.12	0.35	1.00	4.00
HKSTATE39	20001	1.12	0.35	1.00	4.00
HKSTATE40	19949	1.12	0.35	1.00	4.00

HKSTATE41	19865	1.12	0.35	1.00	4.00
HKSTATE42	19711	1.12	0.35	1.00	4.00
HKSTATE43	19397	1.12	0.36	1.00	4.00
HKSTATE44	18908	1.12	0.36	1.00	4.00
HKSTATE45	18415	1.12	0.36	1.00	4.00
HKSTATE46	17814	1.12	0.36	1.00	4.00
HKSTATE47	17257	1.12	0.36	1.00	4.00
HKSTATE48	16654	1.12	0.36	1.00	4.00
HKSTATE49	15938	1.12	0.37	1.00	4.00
HKSTATE50	15224	1.13	0.37	1.00	4.00
HKSTATE51	14429	1.13	0.37	1.00	4.00
HKSTATE52	13561	1.13	0.37	1.00	4.00
HKSTATE53	12957	1.13	0.37	1.00	4.00
HKSTATE54	12584	1.13	0.37	1.00	4.00
HKSTATE55	12064	1.13	0.37	1.00	4.00
HKSTATE56	11594	1.13	0.37	1.00	4.00
HKSTATE57	10977	1.12	0.37	1.00	4.00
HKSTATE58	10335	1.13	0.37	1.00	4.00
HKSTATE59	9693	1.12	0.37	1.00	4.00
HKSTATE60	8938	1.13	0.37	1.00	4.00
HKSTATE61	8204	1.13	0.37	1.00	4.00
HKSTATE62	7548	1.13	0.37	1.00	4.00
HKSTATE63	6845	1.13	0.37	1.00	4.00
HKSTATE64	6239	1.13	0.37	1.00	4.00
HKSTATE65	5646	1.13	0.37	1.00	4.00
HKSTATE66	5091	1.12	0.37	1.00	4.00
HKSTATE67	4594	1.12	0.37	1.00	4.00
HKSTATE68	4125	1.12	0.37	1.00	4.00
HKSTATE69	3735	1.12	0.37	1.00	4.00
HKSTATE70	3360	1.12	0.37	1.00	4.00
HKSTATE71	3009	1.12	0.37	1.00	4.00

HKSTATE72	2694	1.12	0.37	1.00	4.00
HKSTATE73	2352	1.13	0.37	1.00	4.00
HKSTATE74	2050	1.13	0.37	1.00	4.00
HKSTATE75	1785	1.13	0.38	1.00	4.00
HKSTATE76	1565	1.13	0.38	1.00	4.00
HKSTATE77	1342	1.14	0.38	1.00	4.00
HKSTATE78	1133	1.13	0.38	1.00	4.00
HKSTATE79	938	1.13	0.37	1.00	4.00
HKSTATE80	797	1.13	0.38	1.00	4.00

Categorical Variable Codes

Value
.a:age not applicable
.m:missing
1.Agricultural Hukou
2.Non-agricultural Hukou
3.Unified Residence Hukou
4.None

How Constructed

Hukou histories combine details on each reported hukou spell. Each respondent were categorized with their initial hukou type (agricultural hukou, non-agricultural hukou, unified residence hukou, none). In cases respondents changed their hukou type, the new hukou type was assigned to them for the respective age.

Cross Wave Differences in CHARLS Life History

There is currently only one wave of CHARLS Life History.

Differences with the Harmonized SHARE/ELSA Life History

There is no equivalent information in SHARE or ELSA as Hukou is a unique Chinese concept. China established the Hukou system in the 1950s. This system assigned each child to an agricultural or non-agricultural hukou and registered the child to a community, according to the status of their mother. Hukous are of special importance, and they offer access to certain public benefits with large differences between agricultural and urban hukous. Changes of the hukou status are rare but they do occur. Example of the sociological research on hukous and social inequality can be found in Zhang & Treiman 2013, an example of the relationship between hukous and health can be found in Zhang, Nazroo, & Vanhoutte 2021.

CHARLS Variables Used

Life History:

HK002_11_	Change The Hukou Type
HK002_1_10_	The Year of Hukou Type or Location Changed
HK002_1_11_	The Year of Hukou Type or Location Changed
HK002_1_1_	The Year of Hukou Type or Location Changed
HK002_1_2_	The Year of Hukou Type or Location Changed
HK002_1_3_	The Year of Hukou Type or Location Changed
HK002_1_4_	The Year of Hukou Type or Location Changed
HK002_1_5_	The Year of Hukou Type or Location Changed
HK002_1_6_	The Year of Hukou Type or Location Changed
HK002_1_7_	The Year of Hukou Type or Location Changed

HK002_1_8_	The Year of Hukou Type or Location Changed
HK002_1_9_	The Year of Hukou Type or Location Changed
HK004_1_	The New Hukou Type
HK004_2_	The New Hukou Type
HK004_3_	The New Hukou Type
HK004_4_	The New Hukou Type
HK004_5_	The New Hukou Type
HK004_6_	The New Hukou Type

Childhood

Wave	Variable	Label	Type
LH	RAMOMOCCUP_C	ramomoccup: r occupation of female guardian before 17	Categ
LH	RADADOCCUP_C	radadoccup: r occupation of male guardian before 17	Categ
LH	RACHCHLT	rachchlt: r health status relative to other children before	Categ
LH	RAMISCHLTH	ramischlth: r miss school for a month because of a health co	Categ
LH	RACHBEDHLTH	rachbedhlth: r confined to bed for a month because of a heal	Categ
LH	RACHHOSPITAL	rachhospital: r in hospital for a month because of a health	Categ
LH	RACHHOSPITAL3	rachhospital3: r in hospital 3 times because of a health con	Categ
LH	RACHVACCINE	rachvaccine: r vaccination during childhood	Categ

Descriptive Statistics

Variable	N	Mean	Std Dev	Minimum	Maximum
RAMOMOCCUP_C	18067	1.07	0.26	1.00	2.00
RADADOCCUP_C	19168	1.20	0.40	1.00	2.00
RACHCHLT	20261	2.66	1.01	1.00	5.00
RAMISCHLTH	20299	0.04	0.20	0.00	1.00
RACHBEDHLTH	20208	0.06	0.23	0.00	1.00
RACHHOSPITAL	20214	0.02	0.15	0.00	1.00
RACHHOSPITAL3	20219	0.01	0.11	0.00	1.00
RACHVACCINE	20122	0.85	0.35	0.00	1.00

Categorical Variable Codes

Value	RAMOMOCCUP_C
.m:missing	412
.n:no female guardian	463
.w:didn't work during R's childhood	1714
1.Farming	16767
2.Non-agricultural	1300

Value	RADADOCCUP_C
.m:missing	445
.n:no male guardian	884
.w:didn't work during R's childhood	159
1.Farming	15265
2.Non-agricultural	3903

Value	RACHCHLT
.m:missing	395
1.Much healthier	3379
2.Somewhat healthier	3753
3.About average	10483
4.Somewhat less healthy	1618

```
5.Much less healthy                    |          1028

Value----------------------------------|    RAMISCHLTH
.m:missing                             |           357
0.no                                   |         19491
1.yes                                  |           808

Value----------------------------------|   RACHBEDHLTH
.m:missing                             |           448
0.no                                   |         19058
1.yes                                  |          1150

Value----------------------------------|  RACHHOSPITAL
.m:missing                             |           442
0.no                                   |         19745
1.yes                                  |           469

Value----------------------------------| RACHHOSPITAL3
.m:missing                             |           437
0.no                                   |         19991
1.yes                                  |           228

Value----------------------------------|   RACHVACCINE
.m:missing                             |           534
0.no                                   |          2970
1.yes                                  |         17152
```

How Constructed

RAMOMOCCUP_C and RADADOCCUP_C indicate the occupation of the male and female guardian before the age of 17. They are coded as 1. Farming and 2.Non-farming. A special missing .n is assigned to the respondent who indicates they did not have a female or male guardian. A special missing .w is assigned to respondent who indicates that their guardians did not work during their childhood. Other missing responses are coded as .m.

Childhood health was captured by several questions.

RACHCHLT contains a self-evaluation of respondents' health before the age of 15 compared with others. It is coded as follow, 1. Much healthier, 2. Somewhat healthier, 3. About average, 4. Somewhat less healthy, and 5. Much less healthy.

RAMISCHLTH, RACHBEDHLTH, and RACHHOSPITAL measure whether respondents ever missed school, were confined to bed, or were in a hospital for a month or more due to an illness, respectively. They are coded as 0. No and 1. Yes. In the case of "missed school", respondents with no education were coded as never missed school due to an illness. Other missing responses are coded as .m.

RACHHOSPITAL3 indicates whether respondents were ever hospitalized three times or more in a given year before the age of 15. It is coded as 0. No and 1. Yes. Other missing responses are coded as .m.

RACHVACCINE indicates whether respondents had been vaccinated in their childhood. It is coded as 0. No and 1. Yes. Other missing responses are coded as .m.

Cross Wave Differences in CHARLS Life History

There is currently only one wave of CHARLS Life History.

Differences with the Harmonized SHARE/ELSA Life History

SHARE and ELSA measure childhood conditions differently and are therefore not comparable with CHARLS. Childhood health in SHARE and ELSA was measured as a self-report and not in comparison to others.

CHARLS Variables Used

```
Life History:
    C1_A2                    Occupation of Female Dependents
```

C1_B2	Occupation of Male Dependents
HC001	Vaccination in Childhood
HS002	Childhood Health Status
HS003	Childhood Missing School
HS004	Confined to Bed in Childhood
HS005	Hospital in Childhood
HS006	Three Times in Hospital in Childhood

References

Gabadinho A, Ritschard G, Müller NS, & Studer M. 2011. Analyzing and Visualizing State Sequences in R with TraMineR. *Journal of Statistical Software, 40*(4), 1-37.

Halpin B. 2017. SADI: Sequence Analysis Tools for Stata. *Stata Journal, 17*(3), 546-572.

Zhao Y, Hu Y, Smith JP, Strauss J, & Yang G. 2014. Cohort Profile: The China Health and Retirement Longitudinal Study (CHARLS). *International Journal of Epidemiology,* 43, 61-38.

Zhang N, Nazroo J, & Vanhoutte B. 2021. The relationship between rural to urban migration in China and risk of depression in later life: An investigation of life course effects. *Social Science & Medicine,* 270.

Zhang Z & Treiman DJ. 2013. Social Origin, Hukou Conversion, and the Wellbeing of Urban Residents in Contemporary China. *Social Science Research*, 42(1), 71-89.