



SHARE WORKING PAPER SERIES

Combining the Retrospective Interviews of Wave 3 and Wave 7: The Third Release of the SHARE Job Episodes Panel

Agar Brugiavini, Cristina E. Orso, Mesfin G. Genie, Rinaldo Naci,
Giacomo Pasini

Working Paper Series 36-2019

SHARE-ERIC | Amalienstr. 33 | 80799 Munich | Germany | share-eric.eu

Combining the retrospective interviews of wave 3 and wave 7: the third release of the SHARE Job Episodes Panel¹

Agar Brugiavini, Cristina E. Orso, Mesfin G. Genie, Rinaldo Naci, Giacomo Pasini

Ca' Foscari University of Venice

Abstract

Data about working life histories, migration histories, fertility histories and marriage/cohabitation histories collected in the third wave of the SHARE survey were organised into a retrospective panel as described in Brugiavini, Cavapozzi, Pasini, and Trevisan (2013) and Antonova, Aranda, Pasini, and Trevisan (2014). In wave 7 respondents from the new countries were included in the survey and refresher sample respondents from the “old” countries were administered the retrospective questionnaire for the first time. We now combine all the life histories into a single panel dataset and add a new variable reporting the working/non-working condition each year.

Keywords: panel data, retrospective interview, dataset management

JEL Classification: C81, C83

¹ This paper uses data from SHARE Waves 1, 2, 3, 4, 5, 6 and 7 (DOIs: [10.6103/SHARE.w1.700](https://doi.org/10.6103/SHARE.w1.700), [10.6103/SHARE.w2.700](https://doi.org/10.6103/SHARE.w2.700), [10.6103/SHARE.w3.700](https://doi.org/10.6103/SHARE.w3.700), [10.6103/SHARE.w4.700](https://doi.org/10.6103/SHARE.w4.700), [10.6103/SHARE.w5.700](https://doi.org/10.6103/SHARE.w5.700), [10.6103/SHARE.w6.700](https://doi.org/10.6103/SHARE.w6.700), [10.6103/SHARE.w7.700](https://doi.org/10.6103/SHARE.w7.700)), see Börsch-Supan et al. (2013) for methodological details. The SHARE data collection has been primarily funded by the European Commission through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812) and FP7 (SHARE-PREP: N°211909, SHARE-LEAP: N°227822, SHARE M4: N°261982). Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C) and from various national funding sources is gratefully acknowledged (see www.share-project.org)

Introduction

The third wave of SHARE (Survey of Health, Ageing and Retirement in Europe) provided, for the first time, life-history information about a representative sample of about 27,000 respondents aged 50 or over and living in Europe. The domains of interest included family relationships, housing, working history, health and health care. The life history interview was then repeated in wave 7 for all respondents who did not participate in wave 3. Wave 7 took place in 2017 in 28 countries, reaching full coverage of the EU. Moreover, many countries included in wave 3 substantially enlarged their samples in waves 4 to 6. The result is that about 62,561 respondents took part in the retrospective interview of wave 7.

Both wave 3 and wave 7 data are released as an individual-level dataset organizing sequences of life events in a flat file format (Stuck, Zuber, Korbmacher, Hunkler, Kneip and Schröder, 2010; Malter, Schuller and Börsch-Supan, 2018). The first release of the Job Episodes Panel (JEP) organized the information about working life available in wave 3 into a retrospective panel, where each respondent contributes as many observations as there are years of age from birth to the age at which they are observed at the moment of the interview (Brugiavini, Cavapozzi, Pasini and Trevisan, 2013). The second release adds migration and household composition, as well as pension legislation respondents were exposed to still obtained from wave 3 (Antonova, Aranda, Pasini, and Trevisan, 2014). This paper describes the third release of the Job Episodes Panel (JEP), where information about wave 7 respondents is added. The third release of JEP (Job Episodes Panel release 7.0.0) also incorporates wave 3 with the addition of the new variables (i.e., situation and after last job situation) describing the activities carried out between the job spells.

Section 2 describes the variables included in the dataset derived from information collected in wave 3 and wave 7 and focuses on the differences with the previous releases of the JEP, and section 3 describes the variables obtained linking institutional information about pension legislation to the dataset. The last section concludes discussing how to incorporate further retrospective information in the dataset and how to merge with regular waves of SHARE.

Variables derived from the retrospective interviews

The starting point are the 28,465 individuals interviewed in wave 3 and the 62,561 individuals who are administered the retrospective questionnaire in wave 7, i.e. all SHARE respondents taking part to wave 7 who did not take part to wave 3.² Using the year of birth and year of the interview from the cover screen section, we determine for each respondent the age at the time of the interview. The next step is to expand the dataset to have each individual contributing as many observations as the years of age from birth to the age at the time of interview. 27 retrospective respondents from wave 7 did not report their year of birth, and therefore are not used for the JEP. The base dataset obtained contains 6,116,307 person-year observations. Table 1 reports the distribution of individuals and person-year observations by country.

Table 1: Number of individuals and number of person-year observations by wave and country

Country	Number of individuals	N. of person-year observations
Austria	3,722	258,703
Germany	4,902	326,451
Sweden	4,091	284,913
Netherlands	2,258	148,848
Spain	5,702	395,682
Italy	5,529	369,415
France	4,686	315,451
Denmark	4,105	266,062
Greece	4,252	280,749
Switzerland	2,972	200,956
Belgium	6,200	411,025
Israel	2,131	150,830
Czech Republic	5,115	348,781
Poland	5,499	353,903
Ireland	855	57,009
Luxembourg	1,254	83,508
Hungary	1,538	106,180
Portugal	508	34,724
Slovenia	3,692	255,422
Estonia	5,117	356,127
Croatia	2,407	159,908
Lithuania	2,032	134,562
Bulgaria	2,002	132,991
Cyprus	1,233	85,081
Finland	2,007	132,795

² 457 respondents entered SHARE after wave 3 and were administered a panel since their first appearance in the sample. These respondents were administered a baseline interview in wave 7 and not a retrospective one.

Latvia	1,754	116,932
Malta	1,260	84,246
Romania	2,112	137,727
Slovakia	2,064	127,326
Total	90,999	6,116,307

Table 2 reports the full set of variables derived from wave 3 and wave 7 data, and for each of them, the variables in the regular SHARE waves as well as in the two retrospective interviews used to build them. The first set of variables allows merging the JEP with regular waves: *mergeid*, *hhid3*, *hhid7* are the person and household identifiers, *jep_w* takes value 3 if the corresponding individual took the retrospective interview in wave 3 and 7 if she/he took it in wave 7. *Year* is the calendar year and age is the age of the respondent in the year. *Country* is the country of interview of the respondent.

Table 2: Retrospective panel variables: description and corresponding questionnaire variables

Variables	Description	Questionnaire variables
<i>mergeid</i>	Person identifier fix across modules and waves	
<i>hhid7</i>	Household identifier wave 7	
<i>hhid3</i>	Household identifier wave 3	
<i>jep_w</i>	Number of Wave	
<i>gender</i>	Gender respondent	<i>gender</i> (cv_r module),
<i>yrbirth</i>	Year of birth respondent	<i>yrbirth</i> (cv_r module)
<i>Age</i>	Age respondent	<i>int_year</i> , <i>yrbirth</i> (cv_r module)
<i>Year</i>	Year	<i>age</i> , <i>yrbirth</i> (cv_r module)
<i>country</i>	Country of residence at the time of interview	<i>country</i> (cv_r module)
<i>situation</i>	Situation	<i>re003</i> , <i>re007</i> , <i>re0010_*</i> , <i>re033_*</i> (re module)
<i>working</i>	Working spell	<i>re011_*</i> , <i>re026_*</i> (re module)
<i>unemployed</i>	Unemployment spell	<i>re031_*</i> , <i>re033_*</i> , <i>re006_</i> , <i>re007_</i> , <i>re035_*</i> (re module)
<i>in_education</i>	In full time education	<i>re002_</i> (re module)
<i>retired</i>	Retirement spell	<i>re031_*</i> , <i>re033_*</i> , <i>re039a_*</i> (re module) Waves 1, 2,4,5,6: <i>ep329</i> (ep module)
<i>mainjob</i>	Main job spell	<i>re040_</i> , <i>re011_*</i> , <i>re026_*</i>
<i>ordjob</i>	Job spell numbering	<i>re011_*</i> (re module)
<i>industry</i>	Job industry	<i>re014_*</i> (re module)
<i>job_title</i>	Employee, civil servant or self-employed	<i>re015_*</i> (re module)
<i>working_hours</i>	Full time or part time	<i>re016_*</i> , <i>re018_*</i> , <i>re020_*</i> (re module)
<i>first_wage</i>	First wage for each job – nominal currency	<i>re021_*</i> (re module)
<i>currency_fw</i>	Currency coded first wage	<i>re022c_*</i> (re module)
<i>first_income</i>	First job income self-employment – nominal currency	<i>re023_*</i> (re module)
<i>currency_fi</i>	Currency coded first job income	<i>re024c_*</i> (re module)
<i>reason_endjob</i>	Reason left job	<i>re031_*</i> (re module)
<i>afterlast</i>	Situation after last job	<i>re035_*</i> (re module)
<i>lastwage</i>	Monthly wage at the end of main job	<i>re041_</i> (re module)
<i>currency_lw</i>	Currency of monthly wage at the end of main job	<i>re042_</i> , <i>re022c_</i> (re module)
<i>lastincome</i>	Monthly income at the end of main job	<i>re043_</i> (re module)
<i>currency_li</i>	Currency of monthly income at the end of main job	<i>re044_</i> , <i>re024c_</i> (re module)
<i>first_pension</i>	First monthly pension benefit when retired	<i>re036_*</i> (re module). Wave 1,2,4,5 and 6: <i>ep213_</i> (ep module)
<i>currency_fp</i>	Currency of first monthly pension benefit when retired	<i>re037c_*</i> (re module)

country_res	Country of residence in a given year	Wave 7: ra006_*, ra007_*, ra013_*, ra014c_*, ra015c_*, ra021_* (ra module). Wave 3 (ac module). Waves 1, 2,4,5,6: dn009 (dn module)
nchildren_nat	Number of natural children alive in a given year	rc023 rc024_* rc027_* rc028_* (rc module)
nchildren	Number of children alive (including adopted children) in a given year	rc023,rc024_*, rc027_*, rc028_*, rc038_,rc039_, rc041*, rc043*, rc044*, rc045*(rc module)
age_youngest_nat	Age of the youngest natural child alive in a given year	rc023, rc024_*, rc027_*, rc028_* (rc module)
age_youngest	Age of the youngest child alive, including adopted children, in a given year	rc023, rc024_*, rc027_*, rc028_*, rc038_, rc039_, sl_rc041*, sl_rc043*, sl_rc044*, rc045*(rc module)
withpartner	Dummy, takes value 1 if Resp is cohabiting with a partner	rp003, rp004b_*, rp011_*, rp012_* (rc module)
married	Dummy, takes value 1 if Resp is married	rp008_*, rp014_* (rc module)

Note: an asterisk on a questionnaire variable name means we used all variables relating to a given questionnaire item. In wave 3, the variables are named with an additional prefix: *sl_*. Moreover, the module used for the construction of *country_res* was *ac module*.

Employment (and non-employment) histories

The dummy variable “*working*” was generated using the start and end date of each job spell, and the assumptions made to fill the missing values and correct typing mistakes extend what we did in previous waves:

- If the **starting year of the first job spell was missing**, we used information about the age at which the respondent finished full-time education and on the gap between the end of full-time education and the entry in the labour market;
- If the **starting year of any other job spell was missing**, we used information on the gap between jobs. In other words, if the respondent declared she started the new job right after the old one, we assumed that the starting year of the new job was equal to the year in which the previous job finished. We did not make any assumption about respondents who reported that they started the new job more than 6 months after or before the end of the previous one, thus coding as missing the starting year in the retrospective panel too.
- If the **new job started before the end of the previous one** and the respondent declared to have started the new job right after the old one, we changed the starting year and set it equal to the ending year of the previous job.
- If the **ending year of any intermediate job spell was missing**, we used information on the gap between jobs: if the respondent moved from job to job without more than 6 months gap,

we assumed that the ending year of the previous job was equal to the starting year of the new one.

- If the **ending year of the last job spell was missing** and the respondent retired right after, we used the information about the year of retirement to fill the missing value.
- If the **beginning year of a job spell was after the end year of it**, we inverted them using the declared gap between the previous and subsequent job (ra010* and ra033*). We did not change the dates if the reported gap was inconsistent with the two dates.
- If a **beginning year or an ending year** was inconsistent and there was a gap before/after the job spell, we checked consistency with the reported beginning and end of gap between jobs to decide whether to set them to missing or to amend them.

In the previous release, we also added a series of case-by-case corrections to amend those cases where the beginning and end dates of a job spell seemed to be inverted, or where we thought there was clearly a typing mistake. An example is the beginning year of job n being 1964, end year 1959, and beginning year of job $n + 1$ being 1970. In this release, we decided to limit data manipulation to the minimum in order to leave to the researcher decisions that may involve a certain degree of arbitrariness. Therefore, we did not implement any case-by-case correction.

Table 3 summarises the number of corrections we discussed above for each set of variables, we report separately the change made in the set of variables:

- From wave 3 (“sl_re006_” start of first paid job, “sl_re011_*”, beginning years for the job spells, “sl_re026_*”, ending years for the job spells).
- From wave 7 (“re006_” start of first paid job, “re011_*”, beginning years for the job spells, “re026_*”, ending years for the job spells).

Table 3: Summary of the corrections

Variables	Non-missing values		Missing values (Don't know and refusals)	
	N. of original non-missing values	N. of value changes	N. of original missing values	N. of missing values set to real values
re006_	57957	2192	182	0
re011_*	150354	2432	1768	106
re026_*	150441	21131	1663	212
sl_re006_	25346	646	37	0
sl_011_*	73772	1517	330	27
sl_026_*	73851	9377	249	42

The variables “*retired*” and “*in_education*” are based on retrospective and standard waves and follows exactly the same assumptions as in Brugiavini, Cavapozzi, Pasini, Trevisan (2013).

The new variable “*situat*” defines the job market status or the self-defined status if not in the labour force for each year. Respondents are asked to report what they did in gaps between jobs (re033) if two job spells are not adjacent, moreover they are asked to report their activity on the same item list if they never worked. Those who never worked can report up to 8 non-working spells. This information is combined with those about working, retired and in full time education spells. If a given year is reported to be both a working and a non-working year, *situat* reports it as working. The same applies to a year reported to be both a non-working and a retirement year, or a working and a retirement year: in both cases *situat* report the individual to be retired. Note that those overlaps are not necessary an indicator of recall bias. Most of them occur in the year of transition between a working and a non-working spell or in the year of retirement.

The order of job episodes, “*ordjob*”, counts all job episodes, this means that it can have “jumps” if some information is not available. Being precise, if we know an individual had a job spell in a given period but we do not have the exact begin and end date, “*situat*” and the derived dummies are set to missing. If for individual i “*ordjob*” is equal to n in years corresponding to the last job before the period with missing information, it will take value $n+2$ for the first job spell after the period with missing information. In general, “*ordjob*” may jump from n to $n+j$ if there is a period with $j-1$ job spells within it for which we do not have the exact beginning and end year. Note this is a new feature of the third release of the JEP: until JEP release 2, as soon as a start or end date of a job spell was missing, all the variables that refer to a job episode were set to missing for each year until the interview year.

“*ordjob*” allows to attach to each job spell its characteristics, all built as in previous releases. “*Industry*”, “*job title*” (i.e. employee, self-employed or civil servant), “*reason left job*”, “*working_hours*”, “*mainjob*” are attached to each year of the job spell, first monthly wage and its currency, first monthly income and its currency for self-employed in the first year of the spell for employees. The last wage/income from self-employment earned (and the currency in which it is expressed) are attached at the end of the “*mainjob*” spell. The first pension benefit (i.e. “*first_pension*”) and its currency are attached to the proper person-year observation using the

information on the year in which respondents received the first pension benefit obtained in regular waves. Table 2 describes the variables contained in the retrospective panel and lists:

Country of residence

Respondents are asked to report all the changes in accommodation they had throughout their lives since they established their own household after being living in their parental home. If they were living in a different country in some period, they were asked to report the country of residence. Respondents in wave 3 can answer with the name of a country out of a list that includes all SHARE countries, UK, USA, Russia, Finland, Norway, Slovakia, Russia and two broader categories: “other European country” or “other non-European country”. SHARE wave 7 incorporates a “country-coder”, i.e. a built-in program that recognizes text strings and code them into a country name. This means respondents could answer virtually any country, including those that do not exist anymore as e.g. West Germany or Czechoslovakia.

As in JEP release 2, we checked the consistency of the starting and ending date of each residence spell following the same rules we use for job spells. Country codes are those of wave 7, therefore, they are not consistent with the release 2, nor with the public release of wave 3.

The novelty introduced with the “country-coder” is important: in table 3 we report the number of individual-year observations by country of residence. 36,18% individual-year observations from wave 7 refer to countries that in w3 would have been classified as “other European country” or “other non-European country”. In these categories fall also the new SHARE country, but 138,587 are from non-SHARE countries that are coded thanks to the country-coder.

Table 4: Individual-year observation by country

Country of residence	Freq.	Country of residence	Freq.
Africa	27	Liberia	4
Afghanistan	86	Libyan Arab Jamahiriya	239
South America	43	Liechtenstein	21
Albania	146	Lithuania	132,046
Antarctica	22	Luxembourg	71,749
Algeria	855	Madagascar	69
American Samoa	27	Malawi	18
Andorra	10	Malaysia	12
Angola	417	Maldives	8
Azerbaijan	624	Mali	2
Argentina	1,457	Malta	82,34
Australia	895	Martinique	49
Austria	251,001	Mauritania	49

Bahrain	10	Mauritius	131
Armenia	172	Mexico	219
Belgium	397,334	Monaco	4
Bermuda	3	Mongolia	10
Bolivia	80	Moldova, Republic of	964
Bosnia and Herzegovina	12,262	Montenegro	330
Brazil	658	Morocco	3,701
British Indian Ocean Territory	22	Mozambique	135
Virgin Islands (british)	33	Namibia	26
Bulgaria	132,315	Nepal	45
Burundi	96	Netherlands	148,62
Belarus	4,813	Netherlands Antilles	171
Cambodia	91	New Caledonia	7
Cameroon	73	New Zealand	41
Canada	941	Nicaragua	16
Cape Verde	280	Niger	5
Cayman Islands	3	Nigeria	21
Central African Republic	185	Norway	1,325
Sri Lanka	109	Pakistan	241
Chad	18	Panama	14
Chile	248	Papua New Guinea	3
China	308	Paraguay	67
Taiwan	9	Peru	199
Colombia	380	Philippines	227
Congo, Republic of	249	Poland	356,752
Congo, Democratic Republic of (was zair)	896	Portugal	36,878
Costa Rica	4	Guinea-Bissau	11
Croatia	149,619	Qatar	5
Cuba	152	Reunion	69
Cyprus	82,462	Romania	139,322
Czechoslovakia	238,016	Russian Federation	34,667
Czech Republic	107,272	Rwanda	41
Benin	24	Saudi Arabia	63
Denmark	259,113	Senegal	132
Dominican Republic	114	Serbia	3,173
Ecuador	279	Singapore	62
El Salvador	24	Slovakia	128,905
Ethiopia	144	Viet Nam	344
Estonia	330,75	Slovenia	246,619
Fiji	3	Somalia	33
Finland	132,651	South Africa	398
Aaland Islands	40	Zimbabwe	50
France	304,008	Spain	384,562
French Guiana	17	South Sudan	5
French Polynesia	7	Suriname	92
Djibouti	21	Swaziland	21
Gabon	9	Sweden	273,741
Georgia	742	Switzerland	188,168
Gambia	26,158	Syrian Arab Republic	308
Palestinian Territory, occupied	304	Tajikistan	52
Germany	136,618	Thailand	154
German Democratic Republic (East German)	33,118	Togo	46

Federal Republic of Germany (West German)	129,823	United Arab Emirates	52
Ghana	50	Tunisia	840
Gibraltar	12	Turkey	1,338
Greece	276,922	Turkmenistan	153
Greenland	11	Uganda	6
Guadeloupe	11	Ukraine	10,825
Guatemala	6	Macedonia, The former Yugoslav Republic	538
Guinea	79	U.S.S.R.	71
Haiti	83	Egypt	483
Honduras	88	United Kingdom	7,875
Hong Kong	55	Tanzania, United Republic of	54
Hungary	105,685	United States of America	3,59
Iceland	36	U.s. Miscellaneous Pacific Islands	3
India	435	Virgin islands (U.S.)	12
Indonesia	68	Burkina Faso	8
Iran (Islamic Republic of)	701	Uruguay	180
Iraq	753	Uzbekistan	2,13
Ireland	52,095	Venezuela	317
Israel	119,152	Samoa	1
Italy	369,476	Yemen	439
Cote d'Ivoire	189	Socialist Federal Republic of Yugoslavia	34
Japan	133	Zambia	25
Kazakhstan	2,311	Congo (both)	6
Jordan	54	Stateless	14
Kenya	35	EU-Citizenship	65
Korea, Republic of	8	Former Eastern Terr. of German Reich	17
Kuwait	49	Kosovo	566
Kyrgyzstan	567	Kurdistan (region)	5
Lao People's Democratic Republic	43	Chechnya	45
Lebanon	415	Korea (North or South)	3
Lesotho	11	W3_Other European country (incl. Turkey)	4,663
Latvia	109,332	W3_ Non-European country	14,961
TOTAL			6,065,770

Another advantage of the country coder is a more precise treatment of changes of residence involving Czechoslovakia, Czech Republic and Slovakia, as well as Germany before and after reunification. As regards Czechoslovakia, in the previous release a (limited) number of respondents reported a period of residence in Czechoslovakia after 1994. This mistake was probably due to the fact the respondent had to choose from a drop-down list that included Czechoslovakia even after 1994. There are no observations with this kind of mistake (nor the opposite, reporting Czech Republic or Slovakia before the velvet revolution). The country-coder requires the respondent to spell out the country of residence relative to each change of accommodation without choosing from a pre-determined list. This way retrieving information is more demanding, but less prone to mistakes. The same advantage applies to respondents who lived in east or west Germany before 1989 but were living in a different country at the time of the interview. The country coder induces them to report FRG or DDR, while the dropdown

list of wave 3 included only Germany. As regards individuals living in Germany at the moment of the interview, as for the second release of the JEP we exploit the question DN009, asked in waves 1,2,4,5 and 6.³

Fertility and relationship histories

SHARE retrospective interviews contain information on the date of birth and death of each child of respondents, as well as on the year of adoption for the adopted children and a number of questions about marriages and relationships. We use this information to generate *nchildren_nat*, *nchildren*, *age_youngest_nat*, *age_youngest*, *married* and *withpartner* exactly as in the previous release.

Pension legislation variables

Table 5 reports the set of variables describing the pension legislation the respondent faced throughout their lives, which we linked to individual information in the job episodes panel.

Table 5 : Contextual variables

Variable	Description
contrib_employee	Pension contribution rate by the employee
contrib_employer	Pension contribution rate by the employer
ret_age	Statutory retirement age
early_age	Early retirement age
early_ret_reduction	Early retirement reduction rate
currency_min_pension	Currency of minimum pension
currency_max_pension	Currency of maximum pension
min_pension	Minimum pension benefits
max_pension	Maximum pension benefits

Information is collected from the Mutual Information System on Social Protection (MISSOC) website⁴, and from the Social Security Administration (SSA) website⁵ for Israel and to fill a few information not available from the MISSOC tables for specific countries. Both sources provide information starting from 2004⁶, therefore all the variables in table 5 are set to missing before that date. This is a difference from the previous release. In that case, the SHARE country teams collected information about pension legislation in place until 2004. At the time of the second release, obtaining

³ DN009 question asks “Where have you lived on November 1st 1989, that is before the Berlin wall came down - in the GDR, in the FRG, or elsewhere? “

⁴The MISSOC data are publicly available online at: <http://www.missoc.org/>.

⁵ The SSA data are publicly available online at: <https://www.ssa.gov/policy/docs/progdsc/ssptw/index.html>

⁶ For countries that joined the EU after 2004, information is available since the time they joined the Union. For instance, information for Croatia is available since 2013, for Bulgaria and Romania since 2007.

these data was difficult, and the consistency between country-specific data sources and the MISSOC varied across countries. Moreover, after the second release of the JEP, a few SHARE users noted that alternative country-specific databases presented different information. We, therefore, decided to rely only on cross-country information validated by the EU for European Countries, and on the SSA database for Israel because the information regarding countries included both in the MISSOC tables and the SSA database are consistent. This more conservative choice regarding data sources forced us to limit the period covered to 2004-2018.

As already highlighted in the JEP release 2, producing comparable measures across countries inevitably requires a number of simplifying assumptions. This means the proposed variables should be used to account for macro differences in the pension incentives faced by SHARE respondents in different countries and in different points in time, or to construct aggregate pension generosity measures as e.g. the Social Security Wealth measure included in wave 4⁷. On the other end, a detailed analysis on the incentives embedded in the pension system specific of each country may require more information than what is available in the proposed dataset. The key assumptions made in the previous release are maintained and integrated with those needed in case of a change in pension legislation in countries participating both in wave 3 and 7, and those needed to deal with the new countries participating in wave 7. For the sake of readability, we list all of them here rather than referring to Antonova, Aranda, Pasini, Trevisan (2014) for those assumptions that did not change.

The first key assumption is that for countries with several types of pension benefits (like in France or Denmark), the values provided are in accordance with the main or core part of the pension plan, i.e., the one arguably affecting the large majority of the population. Moreover, legislation regarding hazardous, heavy or unhealthy employment is not considered.

Observations from the Job Episodes Panel are merged by year and country of residence at the time of interview (*country*) with relevant contextual variables regarding pension programs. This is the first necessary assumption, it is coherent with previous waves, but it is worth noting that it is not innocuous. The alternative would have been to merge by year and country of residence in the given year (*country_res*). We chose *country* and not *country_res* as linking variable because we collected consistent information about pension systems only for EU countries since 2004. Merging by *country_res*, would create so many missing to make the new variables impossible to use.

⁷ See Belloni et al. (2019) for further details and an application of such a measure.

Whenever a variable depends not only on country and year but also on demographic characteristics (e.g., gender), this is used as an additional merging criterion. The assumption behind this approach is the following: not being aware of future changes in pension legislation, individuals base their choices on the existing legislation at each moment in time. Nevertheless, a different principle is used when merging information regarding early retirement age for countries where the number of contribution or insurance years to the pension system is relevant in determining such age (Czech Republic, Germany, Italy, Bulgaria, Croatia, and Estonia). The early retirement age for respondents from these countries is given as missing until each respondent reaches the required number of years of contributions/insurance. After the threshold is met, the missing value is replaced by the early retirement age allowed by each country's legislation at that point in time.

Statutory retirement age

Statutory retirement age in a country is the age at which a person is expected or required to cease working and at which she may be entitled to receive pension benefits. In many countries, statutory retirement age is different for males and females, in some countries other conditions may apply (e.g., qualifying minimum periods of contributions/insurance, the degree of disability, years of military service, etc.). For instance, in Malta the statutory retirement age varies according to the year of birth of respondents, while in Slovenia explicitly depends on the insurance years of individuals. We take account of these specific conditions in order to calculate statutory retirement age. A particular case is Czech Republic, where the statutory retirement age for females depends on the number of children raised. To account for such requirement, information on the number of natural and adopted children alive at the time of interview is included for the calculation of the retirement age of female respondents from the Czech Republic.

Early retirement age

Early retirement age is the minimum permitted age at which a person can start receiving pension benefits before she reaches the statutory retirement age, provided certain conditions are met. These conditions for early retirement vary substantially across countries.

For countries in which a given number of years of contributions is required (Germany, Italy, Croatia, Bulgaria and Estonia) the early retirement age is estimated by means of working years of each respondent. We assumed that the number of working years is equivalent to the number of years of contributions. For countries in which the number of years of insurance is the requirement (in our

specific case, only Czech Republic), the early retirement age is calculated using the information on working years and on eventual periods of unemployment.

Some remarks are at hand regarding more specific aspects of early retirement age at the individual country level. In Italy, the possibility of early retirement differs for employees and self-employed individuals, so we exploit the corresponding variable in SHARE to attach the institutional information to the early retirement age of Italian respondents. In France, eligibility for early retirement is related to the length of the working career of individuals. More precisely, early retirement age ranges from 56 to 60 depending on the age of commencement of activity and the duration of insurance period. Since information provided in the MISSOC tables is not sufficient to clearly identify eligible respondents, early retirement age in France is given as missing.

In Hungary, eligibility for early retirement depends on specific job characteristics. Indeed, a specific early retirement scheme is provided to individuals who experience hazardous working conditions. Moreover, early retirement is available for women regardless of age, who have at least 40 years of eligibility and cease gainful activity. Eligibility period includes periods gained with gainful activity or pregnancy-confinement benefit, child-care fee, child home care allowance, and child raising support or nursing fee. Due to the impossibility to control for all these conditions, early retirement age is given as missing.

Early retirement reduction

Early retirement reductions correspond to the amount of pension benefits given up by a respondent who decides to retire at an earlier age than what otherwise established by the legislation of his or her country. In the dataset, these rates are calculated in terms of percentage reductions for each year of anticipation of retirement with respect to the statutory retirement age. In Belgium from 2004 to 2008 –as well as in Switzerland from until 2010– early retirement reductions differ by gender. In Austria starting from 2008, the total reduction cannot exceed a limit of 15%, a limitation that is not reflected in the values as reported in the present dataset. Reduction rates in the Czech Republic are discontinuous: they are equal to 0.9% for each 90 days of anticipation (with a limit of 720 days) and 1.5% when this limit is surpassed. We consider only the first number and assume a consequent yearly reduction of 3.6% for each year of anticipation of Czech respondents.

Contribution of employee/employer

The dataset includes a yearly percentage contribution rate estimated separately for employees and employers by country. Yet, there are two countries which display an atypical system of contributions, and which require further clarification. In the Czech Republic, the figures provided are valid only for regular employees, while for self-employed they are equal to 28% of declared net earnings (total income subtracted by the total expenses) bounded by a ceiling that may vary across years. Moreover, and in contrast to all other countries in the dataset, survival and invalidity pensions are comprised in the given contribution rates provided for the Czech Republic. This results in inflated contribution rates; in other words, the reported rates are higher than what they should be were they to include only and exclusively contributions to old age pension benefits. Finally, no information is provided for Denmark, since its pension system is not based on contributions of employers and employees, but entirely financed through general taxation.

Minimum pension benefits

Considerable cross-country variation is observed in the data concerning the existence and amount of minimum and maximum pension benefits. For instance, minimum pension schemes display diverging eligibility criteria: while minimum pensions are guaranteed whenever sufficiency in years of contributions is attained in Belgium (corresponding to two-thirds of the amount of years required for the full pension), in Denmark three years of residency suffice to be eligible for a minimum pension. In Italy, the minimum pension applies only to people insured before 1/01/1996 while in Latvia it relies on insurance years.

Marital status is an essential element in the estimation of minimum pensions in some European countries. In particular, in Austria, Belgium, Malta, Spain—and, since 2008, also Sweden—minimum pension benefits differ depending on whether the individual is single or married. Until 2008, having a dependent spouse resulted in a higher level of minimum pension benefits in Spain. From 2008 onwards, the minimum pension benefit for married respondents varies according to whether or not the spouse is dependent. For the sake of simplicity, we considered spouses as independent: thus, until 2016 equal amounts of minimum pensions for single and married respondents are reported for Spain. In the Netherlands, however, the defining criteria are not marriage, but cohabitation, starting from 2009 the minimum pension in the Netherlands is higher for single parents with at least one underage child (aged 17 or less). Whenever relevant, marriage and cohabitation requirements are incorporated when calculating minimum pension benefits in the different countries.

The amounts provided for minimum and maximum pension benefits are translated into yearly terms and provided in the local currency valid in the country at each particular year. Furthermore, minimum and maximum pension benefits are not applicable in the Netherlands, where a flat-rate pension system is in place, and in Finland.

Maximum pension benefits

Most country legislations also provide a maximum amount of public pension benefits. In Belgium, Denmark, and the Netherlands this amount differs by marital status; on the other hand, in Belgium (from 2004 to 2008), maximum pension benefits varied by gender. We take this heterogeneity into account. Finally, no maximum pension is provided for Poland in the dataset, since the amount of maximum benefits equates to the totality of the reference wage, and thus varies on an individual basis. Similarly, maximum pension is given as missing for Cyprus, where it depends on individual insurable earnings.

Conclusions

This paper describes the new release of the SHARE Job Episodes Panel, i.e. the dataset that reorganise the retrospective interviews administered in the third and seventh wave of SHARE into a panel dataset where each contributes with as many observations as years of age at the moment of the interview. This dataset is meant to simplify the data management phase of researchers willing to work with retrospective data in SHARE. The JEP is by no means exhaustive. Many other sections in the retrospective interviews, e.g. health conditions, health care use or general life collect the date or the relevant period of a specific event or situation. Examples are GL003 and GL004 that record the beginning and end of a period of particular happiness, or HS029 where it is reported the year in which the respondent started to have regular gynaecological visits. This information can be easily merged with the JEP using the person identifier and those year variables.

Standard waves of SHARE can be linked to the JEP in two ways. First, panel respondents that took part in wave 3 were also interviewed in the following waves. Data from these waves can be used to add data points to the JEP information for those respondents: as an example, values for the working, unemployed and retired JEP variables can be easily retrieved from EP005 in waves 4 to 6. In general, the Employment and Pensions section of the standard waves includes all the necessary information

to add data points to the variables related to the Job episodes. The second possibility is to use retrospective questions included in standard waves. As an example, PH009 asks the age at which a particular health problem was diagnosed. This information can be merged by ‘*mergeid*’ and age to the JEP.

We hope the JEP will help current SHARE users to save time in their data preparation, and at the same time will attract new researcher interested in life course analysis. As for the previous releases of the JEP and any new wave of SHARE data, we work to shed light on the fascinating process of ageing Europe is going through.

References

Antonova, L., Aranda, L., Pasini, G., & Trevisan, E. (2014). Migration, family history and pension: the second release of the SHARE Job Episodes Panel. *Working Paper Series*, 18.

Belloni, M., Brugiavini, A., Buia, R. E., Carrino, L., Cavapozzi, D., Orso, C.E., and G. Pasini (2019). "What do we learn about redistribution effects of pension systems from internationally comparable measures of Social Security Wealth?" *Journal of Pensions, Economics & Finance*, forthcoming.

Brugiavini, A., Cavapozzi, D., Pasini, G., and Trevisan, E. (2013). "Working life histories from SHARELIFE: a retrospective panel". *SHARE Working Paper Series* 11-2013.

Malter, F.; Schuller, K , Börsch-Supan, A. (2018). *SHARE Compliance Profiles – Wave 7*. Munich: MEA, Max Planck Institute for Social Law and Social Policy.

Stuck, S., Zuber, S., Korbmacher, J., Hunkler, C., Kneip, T., Schröder, M. (2010), "SHARELIFE release guide 1", Mannheim: Mannheim Research Institute for the Economics of Aging.

Orso, C.E., A. Brugiavini, D. Cavapozzi, G. Pasini (2017): *SHARE Job Episodes Panel*. Release version: 6.0.0 SHARE-ERIC. Dataset. DOI: 10.6103/SHARE.jep.600