SHARE WAVE 7 METHODOLOGY:

Panel innovations and life histories

SHARE
SURVEY OF HEALTH, AGEING AND RETIREMENT IN EUROPE

SHARE
SURVEY OF HEALTH, AGEING AND RETIREMENT IN EUROPE

SHARE WAVE 7 METHODOLOGY:

Panel innovations and life histories

SHARE Wave 7 Methodology:
Panel innovations and life histories

Edited by:
Michael Bergmann
Annette Scherpenzeel
Axel Börsch-Supan

Authors:
Alin Marius Andrieș
Mircea Asandului
Diāna Baltmane
Michael Bergmann
Arne Bethmann
Radim Bohacek
Axel Börsch-Supan
Noam Damri
Giuseppe De Luca
Edwin de Vet
Marvin Formosa
Sabine Friedel
Stefan Gruber
Andrejs Ivanovs
Hendrik Jürges
Antanas Kairys
Ľubica Knošková
Manuel Kronenberg
Clemens Lechner
Michal Levinsky
Howard Litwin
Ekaterina Markova
Maurice Martens
Miika Mäki
Michal Myck
Giacomo Pasini
Yuri Pettinicchi
Jeny Tony Philip
Alexandros Polycarpou
Claudio Rossetti
Anna Rotkirch
Gregor Sand
Annette Scherpenzeel
Karin Schuller
Nikos Theodoropoulos
Signe Tomsone
Iggy van der Wielen
Dana Vokounová
Melanie Wagner
Guglielmo Weber
Gabriela Yordanova
Olga Zamalijeva

Published by:
Munich Center for the Economics of Aging (MEA)
at the Max Planck Institute for Social Law and Social Policy (MPISOC)
Amalienstrasse 33
80799 Munich
Tel: +49-89-38602-0
Fax: +49-621-38602-390
www.mea.mpisoc.mpg.de

Layout and printing by:
Valentum Kommunikation Gmbh
Bischof-von-Henle-Str. 2b
93051 Regensburg

www.valentum-kommunikation.de

Picture credits:
Adobe Stock: gonin

Suggested citation:
Bergmann, M., Scherpenzeel, A., & Börsch-Supan, A. (2019). SHARE Wave 7 Methodology: Panel innovations and life histories. Munich: MEA, Max Planck Institute for Social Law and Social Policy.

ISBN 978-3-00-062956-3

TABLE OF CONTENTS

1 SHARE Wave 7: Combining panel and retrospective data on life histories in eight new countries 15

1.1 Innovations and methodology in Wave 7 15

1.2 Acknowledgements 17

2 Questionnaire innovations 21

2.1 Questionnaire innovations in the seventh wave of SHARE 21

2.2 SHARELIFE 23

2.2.1 Introduction 23

2.2.2 Integrating SHARELIFE into SHARE 24

2.2.3 Implementation of the life history calendar 25

2.2.4 Technical challenges 27

2.3 Personality traits: The Ten-Item Big Five Inventory (BFI-10) 29

2.3.1 Introduction 29

2.3.2 Validation of the BFI-10 in SHARE: Looking for the Big Five dimensions 30

2.3.2.1 Dimensionality and reliability in the pooled sample 30

2.3.2.2 Country differences 31

2.3.3 Using the data: Generated variables in the SHARE release 33

2.3.4 Concluding remarks 34

2.4 New measures for interpersonal environment during childhood 35

2.4.1 Introduction 35

2.4.2 Questionnaire development 35

2.4.3 Preliminary explication of the data 35

2.4.4 Concluding remarks 37

2.5 Incidence and implications of dramatic life events: Extending the interview to cover experiences of respondents' parents 38

2.5.1 Items on experiences of persecution and dispossession 38

2.5.2 Items on dramatic life events in the lives of respondents' parents 40

2.5.3 Concluding remarks 41

2.6 Intergenerational cohabitation at older ages 43

2.6.1 Motivation 43

2.6.2 Parents and parents-in-law 43

2.6.3 Children 47

2.6.4 Concluding remarks 49

2.7 Palliative care 50
2.7.1 Introduction 50
2.7.2 Questionnaire development 50
2.7.3 Preliminary data analysis 53
2.7.4 Concluding remarks 54

3 Software innovations 57

4 Becoming a new SHARE country 63
4.1 Bulgaria 64
4.1.1 Introduction 64
4.1.2 Funding and assembling a national working group 64
4.1.3 Survey implementation 64
4.1.4 Summary 65
4.2 Cyprus 66
4.2.1 Introduction 66
4.2.2 Funding and assembling a national working group 66
4.2.3 Survey implementation 66
4.2.4 Summary 66
4.3 Finland 68
4.3.1 Introduction 68
4.3.2 Funding and assembling a national working group 68
4.3.3 Survey implementation 68
4.3.4 Summary 69
4.4 Latvia 70
4.4.1 Introduction 70
4.4.2 Funding and assembling a national working group 70
4.4.3 Survey implementation 70
4.4.4 Summary 70
4.5 Lithuania 72
4.5.1 Introduction 72
4.5.2 Funding and assembling a national working group 72
4.5.3 Survey implementation 72
4.5.4 Summary 72
4.6 Malta 74
4.6.1 Introduction 74
4.6.2 Funding and assembling a national working group 74
4.6.3 Survey implementation 74

4.6.4 Summary 74
4.7 Romania 76
4.7.1 Introduction 76
4.7.2 Funding and assembling a national working group 76
4.7.3 Survey implementation 76
4.7.4 Summary 76
4.8 Slovakia 78
4.8.1 Introduction 78
4.8.2 Funding and assembling a national working group 78
4.8.3 Survey implementation 78
4.8.4 Summary 78

5 Sampling design in SHARE Wave 7 81
5.1 Introduction 81
5.2 The SHARE target population 81
5.3 The SHARE sampling protocol 81
5.4 Sampling frames and population coverage 82
5.5 Sampling designs 83
5.5.1 Stratification 83
5.5.2 Clustering 83
5.5.3 Selection probabilities 84
5.5.4 Sample composition 84
5.6 Sampling variables in the released SHARE data 86

6 Interviewer training 89
6.1 Introduction 89
6.2 Train-the-trainer (TTT) sessions 90
6.2.1 Structure and logistics 90
6.2.1.1 Location of the training and timeline of data collections 90
6.2.1.2 Division by country groups 90
6.2.1.3 Agenda and duration of training 90
6.2.1.4 Attendees 92
6.2.1.5 Training resources/materials 92
6.2.2 Content and goals 93
6.2.2.1 Combination of theory and practice 93
6.2.2.2 Mock interview 93
6.2.2.3 General training of standardised interviews 93
6.2.2.4 Project-specific training 94

6.2.2.5 Group exercises 94
6.2.2.6 Question cards 94
6.2.2.7 Interviewer manual 94
6.3 National training sessions (NTS) 94
6.3.1 Content adaptation and conformity with the TTT programme 95
6.3.2 Logistics 95
6.3.3 Attendees 95
6.3.4 Excursion: Feedback from one instalment of NTS 97
6.3.4.1 Issues observed 97
6.3.4.2 Available solutions 97
6.4 Concluding remarks 97

7 The third round of the SHARE interviewer survey 101
7.1 The importance of the interviewer 101
7.2 Interviewer effects in surveys 101
7.3 The Wave 7 interviewer survey 102
7.4 First results: Interviewer characteristics overall, and comparisons within and between countries 103
7.5 Concluding remarks 106

8 Monitoring and managing SHARE fieldwork 109
8.1 Fieldwork monitoring and survey participation in the seventh wave of SHARE 109
8.1.1 Introduction 109
8.1.2 Fieldwork periods and survey agencies 109
8.1.3 Monitoring fieldwork 112
8.1.3.1 Classification of survey outcomes 112
8.1.3.2 Formulas to compute survey outcomes 114
8.1.3.3 Baseline/refreshment samples 115
8.1.3.3.1 Contacting households 116
8.1.3.3.2 Household cooperation and response rate 118
8.1.3.3.3 Individual participation of baseline/refreshment samples 120
8.1.3.3.4 Summary of baseline/refreshment samples 121
8.1.3.4 Panel samples 123
8.1.3.4.1 Contacting households 125
8.1.3.4.2 Household cooperation and response rate 127
8.1.3.4.3 Individual participation of panel samples 129
8.1.3.4.4 Summary of panel samples 132
8.1.4 Concluding remarks 135
8.2 Improving the efficiency of data quality back checks: A new procedure to prevent curbstoning 154

8.2.1 Development of a technical procedure to identify interview fabrication 154
8.2.2 Implementation of the new procedure in SHARE Wave 7 155
8.2.3 Outcome of the new back check procedure 156
8.2.4 Concluding remarks 158

9 Weights and imputations 167

9.1 Introduction 167
9.2 Composition of the Wave 7 sample: SHARELIFE and regular subsamples 167
9.3 Weighting strategies 168
9.3.1 The calibration procedure 168
9.3.2 Calibrated cross-sectional weights 169
9.3.3 Calibrated longitudinal weights 170
9.3.4 Structure of the SHARE weights in Release 7.0.0 170
9.3.5 Supplementary material and user guide on calibrated weights 171
9.4 Imputations of missing values due to item nonresponse 171
9.4.1 Imputations of variables collected in the regular interview 171
9.4.2 Imputations of variables collected in the SHARELIFE interview 174

10 The SHARE Data & Documentation Tool 191

10.1 Background 191
10.2 Technical implementation 191
10.3 Main features 192

CHAPTER 1

SHARE Wave 7: Combining panel and retrospective data on life histories in eight new countries

01

1 SHARE WAVE 7: COMBINING PANEL AND RETROSPECTIVE DATA ON LIFE HISTORIES IN EIGHT NEW COUNTRIES

Michael Bergmann, Annette Scherpenzeel, Axel Börsch-Supan – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC) and the Technical University of Munich (TUM)

This volume documents the most important questionnaire innovations, methodological advancements and new procedures introduced during the seventh wave of the Survey of Health, Ageing and Retirement in Europe (SHARE; Börsch-Supan et al., 2013). SHARE is a research infrastructure to better understand and cope with the challenges and chances of population ageing. The main objective of SHARE is to provide excellent data for aging research through a combination of (a) transdisciplinarity, studying the interactions between bio-medical and socio-economic factors; (b) longitudinality, combining a prospective panel structure and retrospective life histories; and (c) European coverage with strict cross-national comparability through the use of ex-ante harmonized survey tools and methodologies. All countries are on the same fieldwork schedule, use the same survey specifications given by a model contract, and administer the same questionnaire and interviewing software. In addition, data collection and response rates in all countries are centrally monitored.

In Wave 7, SHARE achieved full coverage of all 26 Continental EU Member States, in addition to Switzerland and Israel. After Croatia joined in Wave 6, SHARE already covered 18 countries. With a grant from DG Employment, Social Affairs and Inclusion, the remaining eight EU Member States were integrated in Wave 7: Bulgaria, Cyprus, Finland, Latvia, Lithuania, Malta, Romania, and Slovakia. In addition, Hungary was also recovered in Wave 7. Jointly with harmonized data from the English Longitudinal Study of Ageing (ELSA) and the Irish Longitudinal Study on Ageing (TILDA), pan-European research on effects of our ageing societies and their implications can be extended to all EU countries, providing an observatory of the ongoing reforms of pension and health care systems in Europe.

The second special feature of SHARE's Wave 7 was the collection of retrospective data on life histories in combination with the traditional panel approach of SHARE. This extended the life history data collected in Wave 3 by including all new countries and all refreshment samples which were added since them, thus giving a detailed picture of the current status of individuals in the complete EU with a view across their entire life courses. As described in Chapter 2, the retrospective questionnaire contained all important areas of the respondents' life histories, ranging from parental and childhood conditions, partners and children, housing and financial history and employment history to detailed questions on health and health care.

The data collection of SHARE Wave 7 started in March 2017 and ended on 31 October 2017. About 80,000 interviews (including end-of-life interviews) were collected in Wave 7, the vast majority conducted face-to-face by one of our approximately 2000 interviewers all across Europe and Israel.

1.1 Innovations and methodology in Wave 7

It was decided in an early design stage of the Wave 7 preparations, during the first Wave 7 Questionnaire Board meeting in February 2015 in Brussels, to repeat the retrospective collection of life histories of Wave 3. The reason for this decision was that since Wave 3, many new countries and new samples had been included for which no life history data was available. The design of Wave 7 was, however, much more complicated than the Wave 3 design, since it consisted of several different instruments for different subsamples. Firstly, a retrospective life history instrument was designed for all respondents who had not participated in Wave 3. Secondly, a regular panel interview was designed for all respondents who already had done a retrospective life history interview in Wave 3. In addition to that (and contrary to the Wave 3 interview), those respondents who were asked about their life histories in addition received a shorter, condensed set of questions from the regular panel questionnaire. This was done to not loose one wave of normal panel data about the current lives of the respondents. Thirdly, a baseline interview was combined with the retrospective life history instrument for the new samples in the eight new countries and other sample members who started participating for the first time in Wave 7 (e.g. new partners of existing panel members). In Chapter 2.1, Melanie Wagner, Jeny Tony Philip and Hendrik Jürges describe the design and content of these different instruments, as well as the different subsamples for whom

the instruments were used. In addition, in Chapter 3, Maurice Martens and Iggy van der Wielen discuss the technical design and programming of these instruments.

The Wave 7 life history instrument included some new questionnaire content which was not yet part of the instrument of Wave 3, such as new measures on the interpersonal environment during childhood, on intergenerational cohabitation, and on experiences of war, persecution, or discrimination experienced by the parents of the respondents. In addition to the innovations in the life history instrument, some other new questionnaire content was also included in Wave 7. For example: the 10-item Big Five Inventory to measure personality was implemented in Wave 7 after it had been postponed in Wave 6, and questions on palliative care were added to the end-of-life questionnaire. All new Wave 7 content is described in the subchapters of Chapter 2.

The decision to repeat the retrospective life histories and the instrument design for Wave 7 was presented at the SHARE meeting in May 2015 in Graz, Austria, to the SHARE country teams of the "old" countries (countries already participating in SHARE before Wave 7). Further fine-tuning of the design was performed at the following Questionnaire Board meetings in Bol, Croatia, September 2015, and in Munich, Germany, April 2016. The general kick-off for Wave 7 for all country teams also took place in Bol, Croatia, September 2015. Furthermore, two special meetings were organized for the eight "new" countries (countries joining SHARE for the first time in Wave 7) in February and June 2016, both in Munich. During these meetings, the new countries were introduced to the harmonized SHARE instruments and fieldwork operations, and trained in the country teams' organizational and preparatory tasks. In this respect, the SHARE manual ("MASHA"), a practical introduction and reference book for SHARE country teams that is internally available, was of great value. In Chapter 4, each of the eight new country teams has described how it managed to integrate in SHARE in a short period of time, before the start of the wave, and the challenges encountered in doing so.

One of the main challenges which the new countries as well as the central SHARE management in Munich faced, was the design and drawing of eight new baseline samples. Some of the new country teams and contracted survey agencies had no prior experience in obtaining probability based samples for surveys, and in addition not all countries had (access to) a population register to draw a person sample from. Figure 5.1 in Chapter 5 shows the type of sampling frame that was used in each of these countries. Furthermore, Michael Bergmann, Arne Bethmann, and Giuseppe De Luca describe all aspects of the sampling designs used in the new countries and provide an overview of the type of sample design used in all "old" countries that ever participated in SHARE until Wave 7.

An important tool to achieve SHARE's aim of strict harmonisation of fieldwork procedures across countries are the train-the-trainer (TTT) sessions, preceding each data collection. The TTT sessions serve as a template for the national interviewer trainings in each country, and are attended by representatives of each contracted survey agency as well as the country team operators. For Wave 7, the TTT sessions were especially important, because eight new country teams and survey agencies had to get fully acquainted with the SHARE procedures and instruments for the first time. Separate TTT sessions were therefore held for "old" and "new" countries, in order to keep the intensive teaching style manageable. In Chapter 6, Gregor Sand, Jeny Tony Philip and Yuri Pettinicchi describe the structure and content of all Wave 7 trainings (TTT sessions and national interviewer trainings). A total of 4 two-day training sessions were conducted in 2016 in Munich to prepare all countries for the two test runs (pretest & field rehearsal), and two additional training sessions took place in 2017 for the main data collection.

Another tool which contributes to SHARE's harmonisation of interviewer training and monitoring is the interviewer survey. Since Wave 5, SHARE invites all participating countries to let their interviewers participate in the SHARE interviewer survey, prior to the start of the main data collection. This survey gives information about the behaviour, attitudes and feelings of the interviewers with regard to their work for SHARE. The interviewer information can be linked to the survey data that each interviewer collected on his or her respondents. Since the new countries in Wave 7 also participated in the interviewer survey, we now have unique data about interviewers recruiting new panel members and interviewing them for the first time. Sabine Friedel, Arne Bethmann and Manuel Kronenberg explain, in Chapter 7, how the interviewer survey is done and what results were found.

While the main design decisions about questionnaire content and design in SHARE take place on the international level, the national teams of all countries translate the instrument, adapt it to national institutions (e.g. health care, pension system, asset classes), and test the questionnaire in a pretest and a field rehearsal data collection, preceding the main data collection of each wave. In the first months of 2016, the "old" SHARE countries conducted pretest interviews with the new Wave 7 instruments. The eight new countries had their pretests later, in summer 2016. For the second test round, the field rehearsal, the 28 old and new countries were together at the same schedule, and collected data in September and October 2016. The results of the pretest were evaluated at the SHARE meeting in March 2016 in Vilnius, Lithuania, and the results of the field rehearsal at the SHARE meeting in November 2016 in Starnberg, Germany. After another round of revisions to the various instruments, the main data collection of Wave 7 started in March 2017 and lasted until October 2017. In Chapter 8.1, Gregor Sand

gives all details of the fieldwork management and outcomes of the data collection across countries.

In Chapter 9 about weighting and imputations, which is closely connected to the descriptions of the sampling designs in Chapter 5 and the overview of response outcomes in Chapter 8.1, Giuseppe De Luca and Claudio Rossetti explain the SHARE weighting models. This chapter takes into account the joint probabilities of initial samples and refreshment samples at each wave and adjust for problems of unit nonresponse and sample attrition. They also explain how multiple imputation procedures are used in SHARE to deal with another nonresponse problem, namely item nonresponse on specific items of the SHARE instruments.

Two more important innovations are contained within this book. First, Chapter 8.2 by Michael Bergmann and Karin Schuller describes how a special monitoring programme was implemented in Wave 7 to identify individual interviewers producing deviant data quality. This program used a more complex approach than usual in survey research, incorporating indicators from CAPI (computer-assisted personal interviewing) data as well as paradata combined in a multivariate cluster analysis to distinguish deviant patterns of interviewing behaviour. The results were used to provide survey agencies with an informed sample for interviewer back checking during the ongoing fieldwork, thus improving the quality of the SHARE data at each stage of the survey. Second, Edwin de Vet, Maurice Martens, and Stefan Gruber introduce, in Chapter 10, the SHARE Data & Documentation Tool. This new tool is a fast, customisable, easy-to-use web interface for browsing and searching the SHARE (meta)data, and as such adds to the already existing comprehensive documentation material. SHARE's unique combination of a cross-national study with a panel design, including many waves, constitutes a complex data set. The new tool allows researchers to more easily track questions and variables across waves and countries.

1.2 Acknowledgements

As in previous waves, our greatest thanks belong first and foremost to the participants of this study. None of the work presented here and in the future would have been possible without their support, time, and patience. It is their answers that allow us to sketch solutions to some of the most daunting problems of ageing societies. The editors and researchers of this book are aware that the trust given by our respondents entails the responsibility to use the data with the utmost care and scrutiny.

The country teams are the backbone of SHARE and provided invaluable support: Rudolf Winter-Ebmer, Nicole Halmdienst and Michael Radhuber (Austria); Tim Goedeme, Koen Decancq, Daniela Skugor, Martine Vandervelden, Jorik Vergauwen (Belgium-NL), and Sergio Perelman, Stephanie Linchet, Jean-François Reynaerts, Laurent Nisen, Marine Maréchal, Xavier Flawinne and Jérôme Schoenmaeckers (Belgium- FR); Ekaterina Markova, Gabriela Yordanova, Diana Nenkova, Eugenia Markova and Vassil Kirov (Bulgaria); Šime Smolić, Ivan Čipin (Croatia); Nikolaos Theodoropoulos and Alexandros Polycarpou (Cyprus); Radim Bohacek, Jan Kroupa, Filip Pertold, Jiri Satava and Jaroslava Pospisilova (Czech Republic); Karen Andersen-Ranberg, Jørgen Lauridsen, Mette Lindholm Eriksen and Astrid Roll Vitved (Denmark); Luule Sakkeus, Kaia Laidra, Tiina Tambaum, Liili Abuladze, Orsolya Soosaar, Kati Karelson, Ardo Matsi, Maali Käbin, Lauri Leppik, Anne Tihaste, Lena Rõbakova and the whole team of Statistics Estonia, who carried out the fieldwork (Estonia); Anna Rotkirch and Miika Mäki (Finland); Marie-Eve Joël, Florence Jusot, Anne Laferrère, Nicolas Briant, Romain Sibille, Ludivine Gendre, Emily Bourgeat and Benjamin Levy (France); Annette Scherpenzeel, Felizia Hanemann, Michael Bergmann and Imke Herold (Germany); Antigone Lyberaki, Platon Tinios, Tassos Philalithis, Clive Richardson, George Papadoudis, Thomas Georgiadis (Greece); Anikó Bíró, Gábor Kézdi, Réka Branyiczki (Hungary); Howard Litwin, Lahav Karady, Noam Damri, Ella Schwartz and Maayan Levinson (Israel); Guglielmo Weber, Elisabetta Trevisan, Martina Celidoni, Andrea Bonfatti, Marco Bertoni, Angela Crema (Italy-Padua) and Agar Brugiavini, Giacomo Pasini, Michele Belloni, Elena Raluca Buia (Italy-Venice); Andrejs Ivanovs, Diana Baltmane and Signe Tomsone (Latvia); Antanas Kairys and Olga Zamalijeva (Lithuania); Maria Noel Pi Alperin, Gaetan de Lanchy, Jordane Segura (Luxembourg); Marvin Formosa and Katia Mifsud (Malta); Adriaan Kalwij and Marika de Bruijne (the Netherlands); Michał Myck, Monika Oczkowska, Mateusz Najsztub, Dominika Duda, Wojciech Pauksztelo and Michał Kundera (Poland); Alice Delerue Matos, Pedro Pita Barros, Katiusce Perufo, Fátima Barbosa, Cristina Barbosa, Patrícia Silva, Mara Silva, Cláudia Cunha and Gina Voss (Portugal); Alin Marius Andrieș, Mircea Asandului, Bogdan Căpraru, Iulian Ihnatov and Daniela Viorică (Romania); Ján Košta, Gabriela Dovaľová, Dana Vokounová, Ľubica Knošková (Slovakia); Pedro Mira and Yarine Fawaz (Spain); Josep Garre-Olmo, Laia Calvò-Perxas, Secundí López-Pousa and Joan Vilalta-Franch (Spain, Girona); Gunnar Malmberg, Mikael Stattin, Filip Fors Connolly and Jenny Olofsson (Sweden); Carmen Borrat-Besson (FORS), Sarah Vilpert (IUMSP), Jürgen Maurer (HEC), Alberto Holly (HEC) and Peter Farago, Nora Dasoki and Georg Lutz (FORS) (Switzerland); Boris Majcen, Saša Mašič, Andrej Srakar and Sonja Uršič (Slovenia).

The innovations of SHARE rest on many shoulders. The combination of an interdisciplinary focus and a longitudinal approach has made the English Longitudinal Survey on Ageing (ELSA) and the US Health and Retirement Study (HRS) our main role models. We are grateful to James Banks, Carli Lessof, Michael Marmot, James Nazroo and Andrew Steptoe from ELSA; to Mi-

chael Hurd, Mary Beth Ofstedal, Jim Smith, David Weir and Bob Willis from HRS; and to the members of the SHARE scientific monitoring board (Arie Kapteyn, chair, Orazio Attanasio, Lisa Berkman, Nicholas Christakis, Mick Couper, Finn Diedrichsen, Michael Hurd, Annamaria Lusardi, Daniel McFadden, Pierre Pestieau, Norbert Schwarz, Andrew Steptoe, Arthur Stone and Robert Willis) for their intellectual and practical advice, and their continuing encouragement and support.

We are very grateful to the contributions of the five area coordination teams involved in the design process. Agar Brugiavini (Ca' Foscari University of Venice) led the work and retirement area with Giacomo Pasini and Elena Raluca Buia. Guglielmo Weber (University of Padua) led the income and wealth area with Anne Laferrère, Andrea Bonfatti and Chiara Dal Bianco. The health area was led by Karen Andersen-Ranberg and assisted by Camilla Riis Nielsen and Jens Elmelund Rise (University of Southern Denmark). Health care and health services utilization fell into the realm of Hendrik Jürges in the first preparatory stage of Wave 7 (University of Wuppertal) and in the second stage of Florence Jusot (University of Paris-Dauphine). The fifth area, family and social networks, was led by Howard Litwin from Hebrew University with assistance from Ella Schwartz, Noam Damri and Maayan Levinson.

The coordination of SHARE entails a large amount of day-to-day work which is easily understated. We would like to thank Kathrin Axt, Corina Lica, Karl Riedemann and Andrea Oepen for managing SHARE ERIC and SHARE finances; Philipp Beck, Verena Coscia, Veronika Máté and Tobias Roeckl for their efforts in public relations; and Renate Eggenreich, Hannelore Henning, Stephanie Lasson and Susanne Schmid at MEA in Munich for their support throughout various phases of the project. Thorsten Kneip and Frederic Malter acted as assistant coordinators and were first accompanied, later followed by Annette Scherpenzeel as international coordinator. Preparing the data files for the fieldwork, monitoring the survey agencies, testing the data for errors and consistency are all tasks which are essential to this project. Many thanks therefore go to Michael Bergmann, Tim Birkenbach, Johanna Bristle, Theresa Fabel, Fabio Franzese, Sabine Friedel, Stefan Gruber, Felizia Hanemann, Imke Herold, Markus Kotte, Julie Korbmacher, Judith Kronschnabl, Yuri Pettinicchi, Gregor Sand, Daniel Schmidutz, Karin Schuller, Elena Sommer, Stephanie Stuck, Senta-Melissa Pflüger, Jeny Tony Philip, Melanie Wagner, Luzia Weiss and Sabrina Zuber for questionnaire development, sampling, training, data cleaning and monitoring services at MEA in Munich. We owe thanks to Giuseppe De Luca and Claudio Rossetti for weight calculations and imputations in Palermo and Rome.

Programming and software development for the SHARE survey was done by CentERdata in Tilburg. We want to thank Eric Balster, Marcel Das, Maurice Martens, Lennard Kuijten, Marije Oudejans, Iggy van der Wielen, Arnaud Wijnant, Edwin de Vet and Bart van Nieuwburg for their support, patience and dedication to the project.

The fieldwork of SHARE relied in most countries on professional survey agencies: IFES (Austria), CELLO (Belgium), GfK Bulgaria (Bulgaria), GfK (Croatia), RAI (Cyprus), SC&C (Czech Republic), SFI-Survey (Denmark), Statistics Estonia (Estonia), Taloustutkimus (Finland), KANTAR (formerly known as TNS Sofres) (France), KANTAR (formerly known as TNS Infratest) (Germany), KAPA research (Greece), TARKI (Hungary), Cohen Institute for Public Opinion Research (Israel), IPSOS (Italy), Institute of Sociological Research (Latvia), KANTAR (Lithuania), LISER (Luxembourg), EMCS Group (Malta), CentERdata (Netherlands), KANTAR (formerly known as TNS Polska) (Poland), DOMP (Portugal), GfK (Romania), GfK (Slovakia), IPSOS (Slovenia), KANTAR (formerly known as TNS Demoscopia) (Spain and sub-study in the Region of Girona), IPSOS (Sweden), and LINK (Switzerland). We thank their representatives for a productive and successful cooperation. We especially appreciate their constant feedback, the many suggestions, their patience in spite of a sometimes arduous road to funding, and their enthusiasm to embark innovative survey methods and contents. Much gratitude is owed to the nearly 2000 interviewers across all countries whose cooperation and dedication was, is and will be crucial to the success of SHARE.

Collecting these data has been possible through a sequence of contracts by the European Commission and the U.S. National Institute on Aging, and the support by the member states. The EU Commission's contribution to SHARE through the 6th, 7th and 8th framework program (SHARE, AMANDA, SHARE-LIFE, SHARE PREP, SHARE-LEAP, SHARE-I3, SHARE-M4) and H2020 (SHAREDEV3, No. 676536) is gratefully acknowledged. The SHARE-M4 and DEV 3 project financed the international coordination outside of Germany. We thank, in alphabetical order Peter Droell, Ales Fiala, Philippe Froissard, Jean-David Malo, Robert-Jan Smits, Dominik Sobczak, Maria Theofilatou and Harry Tuinder in DG Research for their continuing support of SHARE. We are also grateful for the support by DG Employment, Social Affairs, and Equal Opportunities through Georg Fischer, Ralf Jacob, Ettore Marchetti, Fritz von Nordheim, Ana Carla Pereira and Valdis Zagorskis.

Substantial co-funding for add-ons such as the physical performance measures, the train-the-trainer program for the SHARE interviewers, and the respondent incentives, among others, came from the U.S. National Institute on Aging (P30 AG12815, R03 AG041397, R21 AG025169, R21 AG32578, R21 AG040387, Y1 AG4553-01, IAG BSR06-11, OGHA 04-064, R01 AG052527, R01 AG052527, R01 AG056329, and HHSN271201300071C). We thank John Haaga, Georgeanne Patmios and John Phillips for their enduring support and in-

tellectual input. We are especially grateful to the late Richard Suzman who is sorely missed and who was instrumental for SHARE from its very beginning.

The German Ministry of Science and Education (BMBF) and the Max-Planck-Society (MPG) financed all coordination activities at MEA, the coordinating institution. Regarding the BMBF and its associated project administration we owe special thanks to Angelika Willms-Herget and Klaus Schindel who served as chairs of the SHARE-ERIC Council, and, in alphabetical order, Gisela Helbig, Hans Nerlich, Ranyana Sarkar, Brunhild Spannhake, Monika van Ooyen, Beatrix Vierkorn-Rudolph and Monika Wächter who helped us with determination and patience to set up SHARE as a research infrastructure in Germany. Representative for all those who contributed to the SHARE project within the MPG, we are grateful to Presidents Peter Gruß and Martin Stratmann, Vice President Angela Friederici, Secretary-Generals, Ludwig Kronthaler and Rüdiger Willems, as well as Corinna Heel, Sabine Mellinghoff and Maximilian Prugger.

The bulk of funding of Waves 6 and 7 came from national sources of the member states. We are grateful for the efforts it took to fund SHARE in each SHARE country, the perseverance of our ERIC delegates and ministry appointees in times in which funding social sciences and public health is all but trivial. The funding from all national funding sources is gratefully acknowledged (see www.share-project.org).

References

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., et al. (2013). Data resource profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). International Journal of Epidemiology, 42(4), 992-1001. https://doi.org/10.1093/ije/dyt088.

CHAPTER 2

Questionnaire innovations

02

2 QUESTIONNAIRE INNOVATIONS

2.1 Questionnaire innovations in the seventh wave of SHARE

Melanie Wagner, Jeny Tony Philip – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC), Hendrik Jürges – University of Wuppertal

The seventh wave of data collection was the most ambitious one till date in several respects. SHARE achieved full coverage of all continental European Union countries by including eight new countries (see Chapter 1). In the questionnaire design, new ground was explored by the combination of two questionnaires into one: the life history questionnaire and the regular SHARE panel questionnaire. The regular SHARE panel approach tracks the same people aged 50 years and older over time. In every SHARE wave (except for Wave 3), the respondents answer the SHARE panel questionnaire, which collects information that provides an overview of the current lives of the respondents. SHARE documents how the respondents react to the same questions and measurements across several waves. However, by focusing on people aged 50 and over, the life experiences of the first 50 years of each SHARE respondent are not available to a researcher working with SHARE data. This absence of these background circumstances makes it difficult for researchers to contextualise these measures. To overcome this shortcoming, in Wave 3, a so-called SHARELIFE questionnaire was fielded. It focuses on people's life histories and asks retrospectively about the initial conditions – the lives of respondents before they were first surveyed by SHARE. The SHARELIFE questionnaire includes all important areas of the respondents' lives (see Chapter 2.2). This perspective is especially useful for the analysis of long-term effects, for instance, the implications of childhood health for health in later life, of the socio-economic status of parents on their adult children's financial situation, or of employment history on pension income.

In Wave 7, the SHARELIFE interview was administered to respondents for whom information on their life histories was still missing. This interview concerned all respondents from countries that joined SHARE after Wave 3 and respondents from "old countries" who were not interviewed in Wave 3, namely, new spouses and respondents from so-called refreshment samples. In total, SHARELIFE data from over 60,000 respondents from 27 countries were collected. This massive data source will be useful for researchers around the world. Respondents whose life histories were already collected in Wave 3 were asked the regular SHARE panel questionnaire (approximately 13,000 respondents). This combination of two questionnaires into one resulted in a highly complicated hybrid questionnaire (see Figure 2.1 and Chapter 3 for the technical challenges). In addition, respondents who received the SHARELIFE interview also received a condensed set of questions from the regular SHARE panel questionnaire. This was done to maintain the panel dimension for key respondent characteristics. These questions included demographic information, basic questions on children and household consumption, details of present health status, cognitive functions and employment situation, retirement expectations, consumption, income and health care. Along with the grip strength measurement, which was also administered to SHARELIFE respondents, these questions were asked to obtain additional information on the current lives of the respondents.

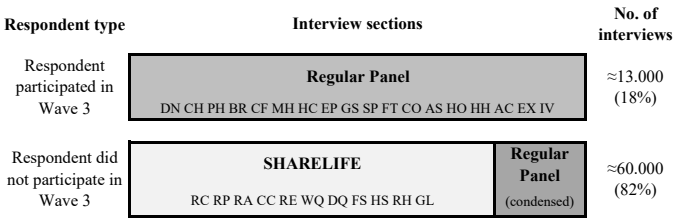


Figure 2.1: Structure of the Wave 7 interview

The Netherlands did not collect life history data as they repeated their multi-mode survey. In Wave 6, the Netherlands administered the SHARE survey in CAWI (computer-assisted web interviewing) and CATI (computer-assisted telephone interviewing) modes, which made it possible for them to stay in the SHARE survey given their restricted funding situation (see Das et al., 2017 for further details). In Wave 7, the regular panel questionnaire was used again online.

A procedural innovation was the obligatory inclusion of cognitive pretesting for newly developed items. While in former waves, the decision to use cognitive pretesting was the prerogative of the respective researcher who proposed the new content, it was now made mandatory to perform cognitive pretesting. For completely new items, cognitive pretesting was conducted by two external enterprises (GESIS & WESTAT) to explore whether the respondent understood

the questioning as intended by the originators of the question and was able to answer with confidence. A cognitive pretesting was performed, for example, with respect to a question on the intergenerational effects of parental exposure to extreme persecution or discrimination from dictatorial regimes (see also Chapter 2.5).

The seventh wave of SHARE also holds several novelties with regard to the questionnaire content. Detailed reports on such innovations are presented in the next subchapters. In Chapter 2.2, Jeny Tony Philip and Melanie Wagner explain the structure of the SHARELIFE interview in more detail. They provide insights on the theoretical background of the collection of life history data and on the technical implementation in this context; they also present the deletions and additions to the SHARELIFE questionnaire compared to Wave 3. In Chapter 2.3, Michael Levinsky, Howard Litwin, and Clemens Lechner introduce the 10-item Big Five Inventory, which was administered to the full Wave 7 SHARE sample to measure personality. Personality traits are increasingly included in the analyses of late-life outcomes or as predictors of other outcomes. The authors explain the development of the scale and present country differences in the Big Five Inventory. In Chapter 2.4, Noam Damri and Howard Litwin present new measures on the interpersonal environment during childhood as an addition to the SHARELIFE questionnaire. Retrospective indicators of respondents' relationships with parents, physical abuse by parents and others, friendships, neighbourhoods, and family religiosity were included to develop a broader understanding of how the interpersonal childhood circumstances of early life impact outcomes at older ages. In Chapter 2.5, Radim Bohacek and Michal Myck describe an extension to the module on "General life and persecution" to cover experiences of war, persecution, or discrimination experienced by the parents of the respondents. This information will increase knowledge of how parents' experiences of such dramatic life events affect life course developments. Chapter 2.6 by Giacomo Pasini and Guglielmo Weber describes an addition to the accommodation history section. The authors describe new questions on intergenerational cohabitation at older ages, which will add to awareness of how the family composition affects late-life outcomes. Finally, Chapter 2.7 by Hendrik Jürges introduces additional questions on palliative care to the end-of-life questionnaire that will allow for a comparative assessment of end-of-life care across countries. Aspects such as pain and pain treatment, feelings of anxiety and the possible need for help, or whether the respondent died in a special unit for palliative care are covered.

References

Das, M., de Bruijne, M., Janssen, J., & Kalwij, A. (2017). Experiment: Internet interviewing in the sixth wave of SHARE in the Netherlands. In F. Malter & A. Börsch-Supan (Eds.), SHARE Wave 6: Panel innovations and collecting dried blood spots (pp. 151-162). Munich: Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC).

2.2 SHARELIFE

Jeny Tony Philip, Melanie Wagner – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC)

2.2.1 Introduction

Figure 2.2: SHARELIFE: a retrospective measure of people's life histories

The SHARELIFE interview, which was introduced in Wave 3 and repeated in Wave 7, focuses on people's life histories and asks retrospectively about the "initial conditions" – the lives of respondents before they were first surveyed by SHARE. The life history questionnaire in Wave 7 was redesigned from its first implementation in Wave 3 (for details of Wave 3, see Börsch-Supan & Schröder, 2011). The original implementation was based on recent insights into how the human memory retrieves information. Our so-called autobiographical memory is relevant for retrieving past events and thus remembering. Conway and Rubin (1993) discuss three levels of specificity of autobiographical recollection: event-specific knowledge, related to specific events; general events, related to eventful periods; and lifetime periods, related to major periods such as childhood, adolescence, etc.

People do not reproduce events from the past without error (Rubin, 1996; Jürges, 2005). The environment in which data are collected, the characteristics of the individual, the type of data collected, the nature of the remembered event, survey practices and the period of recall all impact the reliability of recall (Rubin & Baddeley, 1989). In addition to whether events are remembered at all, the quality of retrospective data could be adversely influenced by poor accuracy of recall (Smith & Thomas, 2003). Gaskell et al. (2000) suggest that respondents can be prone to telescoping, a situation in which respondents report events as occurring more recently (forward telescoping) or further back (backward telescoping) than they actually did. As a result, studies collecting data retrospectively were not originally considered as credible as studies that interview people directly at different points in time. However, this attitude has evolved with the onset of validation studies that focus on salient key life events (Smith & Thomas, 2003) in comparison to external prevalence rates available in the past (Smith, 2009). This trend suggests that if the focus in retrospective studies is on key "pivotal" events instead of an elaborate tapestry of past events, retrospective data have great potential for causal analysis.

The most simplistic depiction of the life course would be unidimensional, recording events in the respondent's lifetime on a horizontal axis. Extending this representation along another dimension would give rise to the event history calendar (Freedman et al., 1988; Blane, 1996). The life course is not a unidimensional series of events unfolding and evolving over time but a simultaneous unfolding of many dimensions, all interwoven temporally and causally in complex ways (Freedman et al., 1988). According to Belli (1998), life history calendars enhance the respondent's ability to recall, as standard recall mechanisms, which are related to these memories, are triggered by this approach. These recall mechanisms are characterised by Belli (1998) as a hierarchical network that includes extended (inclusive of but not restricted to primary lifetime events) events, summarised events (a class of events of the same kind), and specific events. Belli suggests that these factors permit the retrieval of information along manifold pathways: first, a top-down recollection with general memories leading to specific incidents; second, a chronological or sequential recollection within life themes that unify extended events; and third, parallel recall for contemporaneous and synchronic sequential events.

An additional advantage of such calendars is that they enable cross-validation of responses by both the interviewer and the respondent, enabling them to crosscheck events and dates against each other, leading to more accurate verification of information (Börsch-Supan & Schröder, 2011). This helps to reduce the risk of misreporting. Written timing cues and visual representation of years before and after a particular event facilitate sequential recollection by respondents as they attempt to piece together the sequence of events in their lives (Axinn et al., 1999). Retrieving information in this manner is sometimes called "sequencing", since the event is recalled as part of an event sequence (Belli, 1998). Sequencing strategies help to contextualise events and allow them to be reported as a narrative, thereby reducing the risk of events being omitted (Glasner & van der Vaart, 2008). This approach also helps to curtail the problem of telescoping to

some extent. Another possible respondent behaviour that is held in check by the calendar implementation is the issue of time expansion, i.e., reporting an event as having been more or less frequent than it actually was. As certain experiences serve as temporal landmarks for personal histories, the landmarks may be actively and spontaneously used by individuals as bounding cues when performing recall and dating tasks (Shum, 1998; Belli, 1998). According to Shum (1998), temporal landmarks in an individual's memory are (1) events that are those in which the individual was personally involved, (2) events that are of great personal importance to the individual, and (3) events that serve as points of reference in the personal histories of the individual. Dating accuracy could thus be improved through the use of parallel event questioning (Auriat, 1993). The retrospective life history module can be designed as a "life grid" (Blane, 1996) or "event history calendar" (Belli, 1998), where a horizontal line represents time. Multiple lines can be used to represent different topics (dimensions of interest), including an external dimension representing years and landmark events. Respondents are free to use sequential (through time), cross-reference (across dimensions), or both methods of remembering, which increases the accuracy and detail of the recalled events.

The existing literature shows that events are more memorable when one salient life event that temporally coincides with another helps to trigger memories that are associated with the events (Dex, 1995). In the domain grid representation, a grid format is used to represent events in different areas in the calendar. A display with a calendar grid is used in which the lines represent different life areas and the columns represent different years. This display helps to provide an overview of the duration of various spells of employment, accommodation and relationships graphically, hence reducing the potential for erroneous statements and improving the quality of the retrospective data. The grid offers survey respondents sequential, parallel, and logic-based memory cues for dating autobiographical events (Glasner et al., 2012). The visual nature of the life history calendar, which allows respondents to see whether they have correctly reported the coincidence or ordering of various events, may also improve temporal recall and simultaneously help to structure the interviewers' questions (Axinn et al., 1999). The graphical representation of the information already recorded in the calendar renders the detection of gaps and inconsistencies very easy for the interviewers, who receive cues to probe accordingly (Brüderl et al., 2017). In calendar interviewing, the acquisition of valid retrospective events is assumed to depend on some flexibility to encourage respondents to remember the interrelationship of events that have occurred in their autobiographical past. Belli and Callegaro (2009) indicate that ordering questions is the most natural way for each respondent to increase his/her recall ability. Freedman et al. (1988) report instances of implementation of the life history calendar in which the interviewer has some flexibility in the ordering of the questions.

In 2007, the English Longitudinal Study of Ageing (ELSA) implemented the life history calendar to facilitate its face-to-face interviews (Scholes et al., 2009). These interviews served as the blueprint for the introduction of the calendar during the Wave 3 SHARELIFE history data collection, as SHARE is closely related to ELSA in terms of respondent characteristics. SHARE's sister surveys (HRS and CHARLES) also follow a similar structure to collect life histories in the US and China, respectively, which has the additional advantage of maintaining harmonisation across different surveys. The SHARELIFE data collection in Wave 3 included the retrospective life histories of 30,000 SHARE respondents from 13 countries (Austria, Belgium, Czech Republic, Denmark, France, Germany, Greece, Italy, Netherlands, Poland, Spain, Sweden and Switzerland; see Börsch-Supan et al., 2013). In Wave 7, approximately 60,000 SHARELIFE interviews were conducted.

2.2.2 Integrating SHARELIFE into SHARE

The SHARELIFE interviews (also known as life histories), which were conducted in Wave 3 (2008/2009), were repeated in Wave 7. Structured life histories were administered to all respondents who had been added to the SHARE panel since the third wave in 2008 or to those who did not participate in Wave 3. This group included respondents from refreshment samples in existing SHARE countries, new spouses of participating SHARE respondents, and all respondents in countries that joined SHARE after Wave 3, namely, Bulgaria, Croatia, Cyprus, Estonia, Finland, Latvia, Lithuania, Luxembourg, Malta, Portugal, Romania, Slovakia and Slovenia, and Israel, which did not collect life histories in Wave 3 either. Respondents who already performed the life history interview in Wave 3 received the regular panel interview that was conducted in all other waves.

The SHARELIFE interview in Wave 7 spanned various salient domains of a respondent's life course:

- The retrospective children questions collected retrospective information on children and deceased children, including information about pregnancies, births, adoptions, characteristics of children and maternity benefits and leave. The information obtained in these areas was contextualised by accompanying follow-up questions about employment and income at the instance of occurrence of certain salient events. For example, there were questions on income sources at the time of motherhood.

- The retrospective partner questions collected retrospective information on all relationships until the present, including information on living arrangements, cohabitation, marriages, separation, divorces, and death of partners.

- The retrospective accommodation questions collected information on past and current accommodations, including details of household establishment, residences (country, region), special accommodation events, moves, types of accommodation, cohabitation with parents/children, and ownership.

- The retrospective employment questions collected data on employment spells, including information about employment status, job characteristics, income, retirement benefits, and employment after retirement. Data are also collected on work quality, job satisfaction, and career breaks due to ill health or disability, disability allowances, insurance and computer skills.

- The retrospective health questions elicit information about health and healthcare history during both childhood and adulthood, including details about hospital stays, illnesses, injuries, diseases, vaccinations, doctor visits, preventive check-ups, health behaviours, reasons for not going to the doctor, forgone medication, and impact of financial situation on health care.

- Information was also collected on childhood circumstances (e.g., childhood health, academic performance, relationship with parents, features of accommodation, books read, companions).

- With respect to finances (e.g., insurance, housing, investments), information was collected on financial investments that the respondent may have made during his/her life, including investments in stocks or shares, mutual funds or managed investment accounts, life insurance uptake, business ownership, and overall household income (amount).

- Information was also collected on general life events (e.g., periods of hunger, periods of happiness, stress, financial hardship, discrimination at work, respondent and parental experiences of persecution, oppression, and dispossession).

There were also some key additions in Wave 7 to the life history portion. These additions were based on experiences with Wave 3 data, in which it became clear that the following spheres of people's life histories could be investigated in more detail: persecution of parents (Myck & Bohacek, 2011), cohabitation with parents and children at older ages, forgone retrospective health care use and computer use at work. In addition, questions were added to the childhood circumstances module to elicit information about the respondent's relationship with parents, physical abuse before age 17, loneliness, religion in the childhood family environment and friendships during childhood. This information was supplemented by questions about financial circumstances, number of books, and school performance at the age of 10, which was already part of Wave 3.

To make room for the new questions, it was decided to drop several questions that had been in the original SHARELIFE implementation in Wave 3. These questions included those regarding miscarriages, stillborn children, and the start and stop date of the menstrual periods, which were considered too personal to be asked in a face-to-face interview situation. The number of questions on health during childhood and health care use during adulthood was reduced, as they were considered too detailed and because other questions on health and health care were sufficient to predict later life outcomes. Questions on part-time work during the life span were also dropped to save interview time.

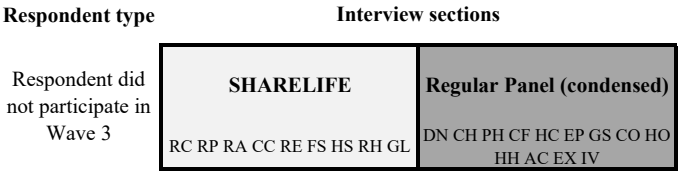


Figure 2.3: Integrating SHARELIFE into the SHARE questionnaire

The domains covered in SHARELIFE and the SHARE regular panel interview overlap largely. However, there is a key distinction in the mode of collecting information. The SHARELIFE interview is administered completely at the individual level with no transfer of information between participating partners. The regular panel interview, however, transfers information between interviews within a couple. For example, in the case of common children, the SHARELIFE interview asks for the children's complete history from both partners, while the SHARE regular panel interview has a family respondent answering all questions on children for both the respondent and the partner.

2.2.3 Implementation of the life history calendar

SHARELIFE uses a life history calendar as described in Section 2.2.1 above, in which the main events of a person's life are displayed visually to aid the recollection of events in juxtaposition to other key life experiences. The SHARELIFE history calendar records events into a grid, which spans across the respondent's life course in years, time being the horizontal dimension. Table 2.1 is a tabular itemisation of the SHARE Wave 7 life history calendar, which displays different dimensions of a respondent's life history:

- Information about the respondent's offspring
- Information about partners
- Accommodation history
- Employment history
- Health history

Table 2.1: Domains and landmarks on the SHARELIFE history calendar

Modules displayed	Milestones displayed
Retrospective children	Birth/death of biological child
	Year of adoption/death of adopted child
Retrospective partner	Start/end of living together with partners (marriages and live-in partners)
	Year of start of relationship (for long-distance relationships)
Retrospective accommodation	Start/end of living at each residence
	Year of first household establishment
Retrospective employment	Year of finishing school
	Start/end of employment/job
Health history section	Start/end of periods of ill health/disability

Figure 2.4 shows an example of the SHARELIFE calendar as it is implemented in Wave 7, with landmark events listed below the calendar. The calendar itself depicts various salient events or the progression of events as they unfold across the years.

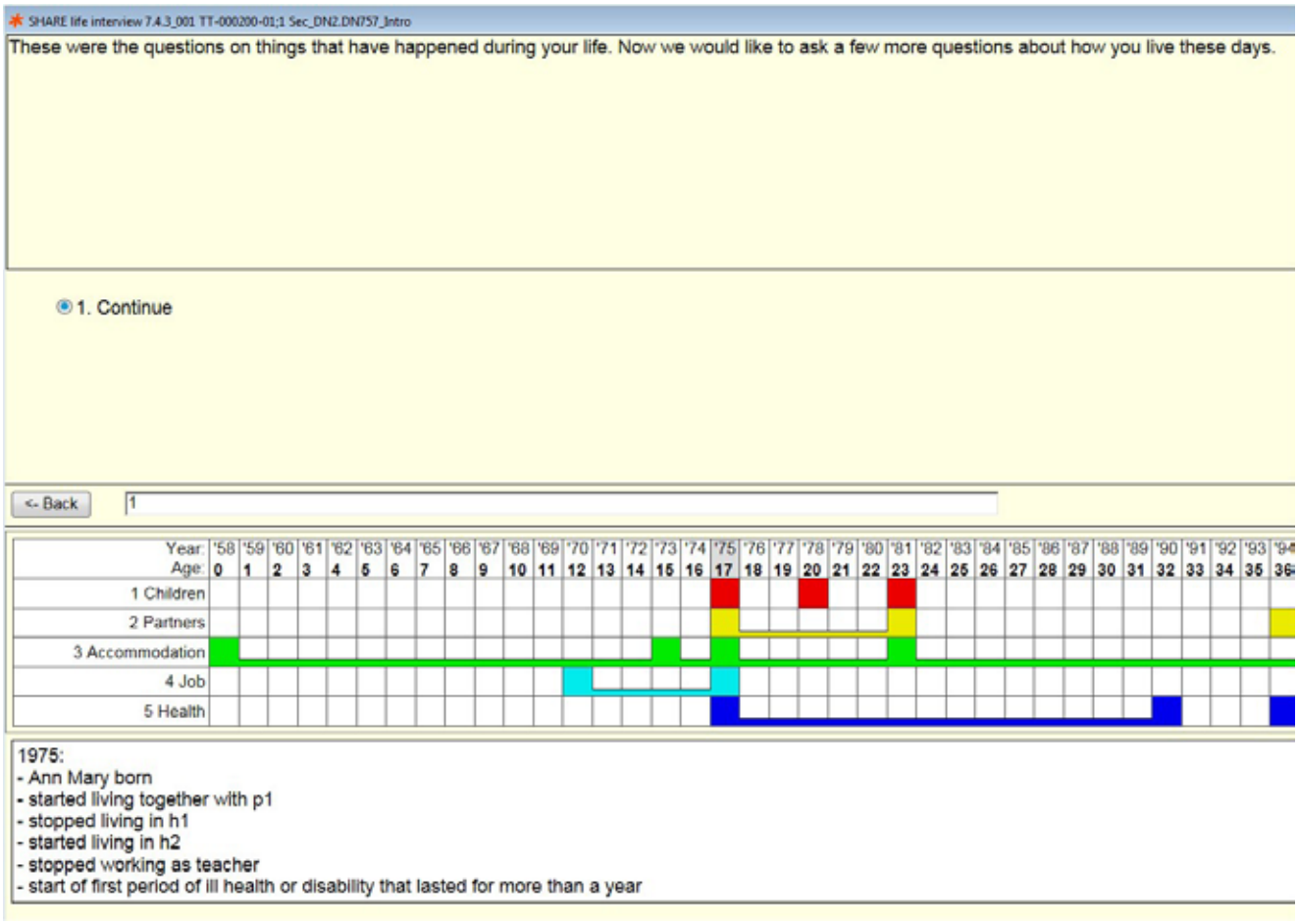


Figure 2.4: SHARELIFE calendar facilitating sequential and cross-dimensional recall

Figure 2.4 also shows how multiple life events that took place in the same year are not only displayed but also listed below the life history calendar, thus providing a multidimensional overview of all events that occur in a particular year. This overview would aid in the cross-dimensional recollection of events by the respondents, thus encouraging parallel and top-down retrieval, thereby facilitating the recollection of coincidental events and reducing the risk of underreporting.

The SHARE life history calendar thus uses the concept of "temporal landmarks", based on the presumption that once certain notable personal events in one's life, such as the birth of a child, marriage, death or divorce, are remembered, they can be used to anchor the respondent's memory and situate other personal events relative to the landmark. For example, the respondent might not know when exactly something happened, but if she knows it was the year after he/she got married, this information can be placed in context.

The life history calendar display in SHARELIFE Wave 7 is designed to facilitate ease of navigation for the interviewer across the calendar on various levels by clicking on a specific calendar module to change or complete recorded information in that domain, improving the completeness of responses. If the respondent is motivated to date an event as accurately as possible, he/she might retrieve dating cues even after providing a sufficient answer to double-check the accuracy of their statement (Glasner et al., 2012). By clicking on specific coloured cells corresponding to landmark events, the interviewer can recall them to the respondent or even change them if any subsequent questions elicit more accurate information of a previously recorded event. For example, in Figure 2.4, the interviewer can, if needed, click on the calendar line "2 Partners" to change the date when the respondent moved in together with his/her partner.

2.2.4 Technical challenges

There were several technical challenges to reimplementing the SHARELIFE calendar in Wave 7. In Wave 3, the implementation had focused purely on the collection of life history information, while in Wave 7, the focus broadened to the collection of both life history and panel information in one instrument. The initial plan to reuse the software used for Wave 3 was shelved after field testing in the pilot phase of the project showed the technology to have become outdated. Consequently, a completely new solution to support the visualisation of life histories had to be developed. This endeavour proved particularly challenging due to the multilingual nature of the SHARE study, wherein some calendar labels for languages with Cyrillic characters in the SHARELIFE calendar did not display properly but rather appeared as jumbled characters. This issue was resolved by using alternative encoding of problematic characters in HTML format. This approach involved integrating an HTML display and using the Blaise interface in the Sample Management System (SMS) to complete this task (for details, see Martens, 2016; Chapter 3). The new solution also made it possible to define the calendar concept within the Blaise questionnaire itself, allowing the questionnaire to be better linked to the translation environment.

References

Auriat, N. (1993). "My wife knows best": A comparison of event dating accuracy between the wife, the husband, the couple, and the Belgium population register. Public Opinion Quarterly, 57(2), 165-190. https://doi.org/10.1086/269364.

Axinn, W. G., Pearce, L. D., & Ghimire, D. (1999). Innovations in life history calendar applications. Social Science Research, 28(3), 243-264. https://doi.org/10.1006/ssre.1998.0641.

Belli, R. F. (1998). The structure of autobiographical memory and the event history calendar: Potential improvements in the quality of retrospective reports in surveys. Memory, 6(4), 383-406. https://doi.org/10.1080/741942610.

Belli, R. F., & Callegaro, M. (2009). The emergence of calendar interviewing: A theoretical and empirical rationale. In R. F. Belli, F. P. Stafford & D. F. Alwin (Eds.), Calendar and time diary methods in life course research (pp. 31-52), Thousand Oaks: Sage. https://dx.doi.org/10.4135/9781412990295.

Blane, D. B. (1996). Collecting retrospective data: Development of a reliable method and a pilot study of its use. Social Science & Medicine, 42(5), 751-757. https://doi.org/10.1016/0277-9536(95)00340-1.

Börsch-Supan, A., & Schröder M. (2011). Retrospective data collection in the Survey of Health, Ageing and Retirement in Europe. In M. Schröder (Ed.), Retrospective Data Collection in the Survey of Health, Ageing and Retirement in Europe: SHARELIFE Methodology. Mannheim: Mannheim Research Institute for the Economics of Aging.

Börsch-Supan, A., Brandt, M., & Schröder, M. (2013). SHARELIFE - One century of life histories in Europe. Advances in Life Course Research, 18(1), 1-4. http://dx.doi.org/10.1016/j.alcr.2012.10.009.

Brüderl, J., Castiglioni, L., Ludwig, V., Pforr, K., & Schmiedeberg, C. (2017). Collecting event history data with a panel survey: Combining an electronic event history calendar and dependent interviewing. methods, data, analyses, 11(1), 45-66. https://doi.org/10.12758/mda.2016.013.

Conway, M. A., & Rubin, D. C. (1993). The structure of

autobiographical memory. In A. Collins, S. Gathercole, M. A. Conway & P. E. Morris (Eds.), Theories of memory (pp. 103-138). Hove: Lawrence Erlbaum Associates. https://doi.org/10.1007/978-94-015-7967-4_10.

Dex, S. (1995). The reliability of recall data: A literature review. Bulletin of Sociological Methodology, 49(1), 58-89. https://doi.org/10.1177%2F075910639504900105.

Freedman, D., Thornton, A., Camburn, D., Alwin, D., & Young-DeMarco, L. (1988). The life history calendar: A technique for collecting retrospective data. Sociological Methodology, 18, 37-68. https://doi.org/10.2307/271044.

Gaskell, G. D., Wright, D. B., & O'Muircheartaigh, C. A. (2000). Telescoping of landmark events: Implications for survey research. Public Opinion Quarterly, 64(1), 77-89. https://doi.org/10.1086/316761.

Glasner, T. J., & van der Vaart, W. (2008): Cognitive processes in event history calendar interviews: A verbal report analysis. Proceedings of the Seventh International Conference on Survey Methodology, Naples. Retrieved from http://library.wur.nl/WebQuery/wurpubs/fulltext/28676.

Glasner, T. J., van der Vaart, W., & Belli, R. F. (2012). Calendar interviewing and the use of landmark events – Implications for cross-cultural surveys. Bulletin of Sociological Methodology, 115(1), 45-52. https://doi.org/10.1177/0759106312445701.

Jürges, H. (2005). Retrospective error in SOEP calendar data: The case of unemployment. Schmollers Jahrbuch: Journal of Applied Social Science, 125(1), 157-165.

Martens, M. (2016). A web compatible DEP. Proceedings of the 17th International Blaise Users Conference, The Hague (pp. 128-135). Retrieved from http://www.blaiseusers.org/2016/papers/3_5.pdf.

Myck, M., & Bohacek, R. (2011). Long shadows of history: Persecution in Central Europe and its labour market consequences. IZA Discussion Papers (6130). Retrieved from http://ftp.iza.org/dp6130.pdf.

Rubin, D. C. (1996). Remembering our past: Studies in autobiographical memory. Cambridge: Cambridge University Press.

Rubin, D. C., & Baddeley, A. D. (1989). Telescoping is not time compression: A model. Memory & Cognition, 17(6), 653-661. https://doi.org/10.3758/BF03202626.

Scholes, S., Medina, J., Cheshire, H., Cox, K., Hacker, E., & Lessof, C. (2009). Living in the 21st century: Older people in England. The 2006 English Longitudinal Study of Ageing. London: National Centre for Social Research.

Shum, M. S. (1998). The role of temporal landmarks in autobiographical memory processes. Psychological Bulletin, 124(3), 423-442. https://doi.org/10.1037//0033-2909.124.3.423.

Smith, J. P. (2009). Reconstructing childhood health histories. Demography, 46(2), 387-403. https://dx.doi.org/10.1353%2Fdem.0.0058.

Smith, J. P., & Thomas, D. (2003). Remembrances of things past: Test–retest reliability of retrospective migration histories. Journal of the Royal Statistical Society: Series A (Statistics in Society), 166(1), 23-49. https://doi.org/10.1111/1467-985X.00257.

2.3 Personality traits: The Ten-Item Big Five Inventory (BFI-10)

Michal Levinsky, Howard Litwin – The Hebrew University of Jerusalem, Clemens Lechner – GESIS

2.3.1 Introduction

"Personality" refers to characteristic differences in how people think, feel and act. Personality is associated with a broad range of life outcomes, including income, health, well-being, marital stability, and social participation (Roberts et al., 2007). During the first six waves of SHARE, the domain of personality was not yet assessed. In light of the growing body of evidence attesting to the relevance of personality traits for ageing research, Wave 7 introduced, for the first time, an established personality inventory. The 10-item Big Five Inventory (BFI-10), following Rammstedt and John (2007), measures the five "Big Five" personality dimensions with two items each. In this chapter, we will outline the rationale of this instrument's construction, analyse its psychometric properties in SHARE Wave 7, and give recommendations as to how SHARE users can use the instrument in their research.

The question of how personality should best be conceptualised and measured has been the focus of research for many decades. Several variable sets and taxonomies of personality structure have been suggested to reflect the complexity of personality. Since the 1940s, an effort has been made to create a parsimonious yet comprehensive taxonomy of personality traits that would allow for the study of personality in a systematic fashion, including across age groups and across cultures. By the late 1980s, the five-factor model of personality had emerged as the dominant schema. This five-factor paradigm is, to date, the most widely used and empirically best-validated framework of individual differences in personality (John & Srivastava, 1999). The five global factors, known simply as "the Big Five", are openness to experience (vs. closed-mindedness), conscientiousness (vs. lack of direction), extraversion (vs. introversion), agreeableness (vs. antagonism), and neuroticism (vs. emotional stability). They are sometimes referred to by the acronym OCEAN.

Initially, the classification into five dimensions was based on various personality questionnaires with hundreds of items. Factor analysis of such questionnaires revealed the same five personality dimensions on a consistent basis. Goldberg (1990) attempted to reduce the length of the questionnaires so that a smaller number of questions would still reflect the five factors. Since the 1990s, several versions of the Big Five questionnaire were developed with even fewer items. Rammstedt and John (2007) introduced the BFI-10 as an ultra-short measure of personality suitable especially for multi-theme surveys in which assessment time and questionnaire space are limited. The authors began their research initially using a version of the questionnaires that included 44 items. Their aim was to reduce the length of that questionnaire to only 10 items: two per personality dimension. They selected the best items[1] of each dimension and utilised a variety of samples and validation methods to increase the generalisability of their short scale. Within each pair of items reflecting the same dimension, one item was formulated in the positive direction and one in the negative direction. This procedure was performed to control for acquiescent responding ("yes-saying"), one of the most frequent content-independent response styles. Items were to be answered on a five-point Likert-type scale ranging from 1 ("fully disagree") to 5 ("fully agree"). On average, the part-whole correlation of the BFI-10 with the BFI-44 was r=.85, and the overall mean correlation of retest reliability across six weeks was r=.75.

Despite the sufficient reliability of the abbreviated questionnaire, some losses in reliability were found for the BFI-10 Agreeableness dimension. In the development of the questionnaire, therefore, it was recommended that researchers particularly interested in the Agreeableness dimension add a third Agreeableness item.[2] Consequently, towards the use of the BFI-10 in SHARE, the third item recommended by Rammstedt and John (2007) from the domain of Agreeableness was added to increase the reliability of these dimensions and to find the most fitting combination of items for use in SHARE. Following this recommendation, SHARE Wave 7 includes these BFI-10(+1) items. The exact wording is displayed in Table 2.2.

1 The two items for each dimension were selected according to several criteria, including the following: (1) having both a direct and a reverse scale item, (2) measuring the core aspects of the Big Five traits, (3) having identical English and German versions to enable cross-cultural use, and (4) measuring the empirical correlation between the items and the full BFI scale. For further reading on the selection of the items and the creation of the instrument, see Rammstedt and John (2007).

2 Namely, "I see myself as someone who is considerate and kind to almost everyone" (Table 2.2).

Table 2.2: BFI-10(+1) items

Trait	Name of variable	Question
Openness	ac705_	I see myself as someone who has few artistic interests
	ac710_	I see myself as someone who has an active imagination
Conscientiousness	ac703_	I see myself as someone who tends to be lazy
	ac708_	I see myself as someone who does a thorough job
Extraversion	ac701_	I see myself as someone who is reserved
	ac706_	I see myself as someone who is outgoing, sociable
Agreeableness	ac702_	I see myself as someone who is generally trusting
	ac707_	I see myself as someone who tends to find fault with others
	ac711_	I see myself as someone who is considerate and kind to almost everyone
Neuroticism	ac704_	I see myself as someone who is relaxed, handles stress well
	ac709_	I see myself as someone who gets nervous easily

2.3.2 Validation of the BFI-10 in SHARE: Looking for the Big Five dimensions

After the data collection in the main survey in Wave 7, we performed a series of dimensionality analyses to determine whether the five expected dimensions emerged in the SHARE sample and estimated the five dimensions' reliabilities. We performed these analyses for the pooled sample (i.e., all countries together) and for each country separately by using factor analysis.

2.3.2.1 Dimensionality and reliability in the pooled sample

As a first step, we ran a principal component analysis on the raw item responses in the pooled sample to test whether the five expected dimensions emerged. We repeated the analysis four times, each time with a different combination of the three Agreeableness items. Overall, the results for the loadings of the varimax-rotated factors did not reflect the Big Five structure. To be more precise, except for the Neuroticism dimension, at least one item from each dimension loaded onto more than one factor. Consequently, we investigated whether this lack of replicability of the Big Five structure might be due to individual differences in response styles, as has been frequently reported in previous studies (Costello & Roodenburg, 2015; Rammstedt & Farmer, 2013), including those conducted in older adults (Lechner & Rammstedt, 2015).

Psychological research investigates the potential biases that response styles may introduce in survey responses. Response styles refer to characteristic ways of answering survey items that are unrelated to the item content and, hence, that can distort respondents' answers away from their "true score" on the construct in question. An individual may answer inaccurately or falsely to questions in accordance with his/her enduring response tendency, regardless of their meaning. Examples of such response styles include extreme response, midpoint response, and acquiescence. Measures that use rating scales, such as the BFI-10, are particularly susceptible to acquiescence, which refers to individuals' general tendency to agree with items by selecting "yes" or "true" response options, regardless of the item content (Rammstedt & Farmer, 2013). Unless accounted for analytically, acquiescence can introduce substantial bias into individuals' responses to personality questionnaires, inflating any covariance-based statistic such as correlations or factor loadings. Acquiescent responding is typically found to be highest among individuals with lower cognitive ability and education; moreover, there are cultural differences in respondents' acquiescent response tendencies. For example, studies clearly show that adults with lower cognitive ability (Lechner & Rammstedt, 2015) and lower education exhibit a higher tendency for acquiescent response bias than do adults with higher cognitive ability and education (Costello & Roodenburg, 2015).

To test whether acquiescent responding may have caused the inequivalence of the Big Five factor structure in the SHARE data as well, an acquiescence index was computed for each respondent. We controlled for individual differences in acquiescence response tendencies by computing each individual's mean score across all items, subtracting the mean from each item response separately, and then analysing these mean-corrected scores (Rammstedt et al., 2010). Table 2.3 shows the factor loadings of the BFI-10(+1) items for the SHARE sample after the acquiescence correction. The fact that the hypothesised Big Five structure emerged with much greater clarity after correcting for acquiescence sug-

gests that the tendency for acquiescent responding did, as expected, blur the factor structure of the Big Five scales in the SHARE sample. In all the combinations of the 10 and 10(+1) items, the resultant factor structures for the acquiescence-corrected indices demonstrated greater correspondence to the idealised factor matrix than did the factors derived from the raw items. As seen in Table 2.3, for all 10(+1) of the personality probes, the results show that the items in the acquiescence-corrected analysis loaded primarily on the factor to which they conceptually belonged, with one exception: the third Agreeableness item loaded on two factors. In the combination of 10 items excluding the "find fault with others" Agreeableness item, two of 10 items did not load on their corresponding factor. In the combination of 10 items excluding the "generally trusting" Agreeableness item, four factors were formed instead of five (in this case, Conscientiousness and Agreeableness loaded onto the same factor). However, analysis of the 10-item combination that included the two original Agreeableness items ("generally trusting" and "find fault with others") demonstrated a clear and consistent Big Five pattern as expected (see the highlighted column in Table 2.3). Thus, the original BFI-10 combination demonstrated the best results in the SHARE sample as a whole and is recommended for use.

The reliability for the two items within each factor varied slightly (as calculated by the Spearman-Brown coefficient): Openness – 0.45, Conscientiousness – 0.50, Extraversion – 0.53, Agreeableness – 0.39 (for the original two items; the three items did not lead to a higher reliability in this sample), and Neuroticism – 0.67. Given the brevity of the instrument, these values can be seen as acceptable, with the exception of agreeableness.

Table 2.3: Varimax-rotated factor structures of the BFI-10(+1) items: Loadings based on items corrected for acquiescence

	BFI-10(+1) Agreeableness item					**BFI-10 with "kind" item ("trusting" dropped)**					**Original BFI-10 item set**				
Big Five Inventory items	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Openness – few interests	0,02	0,04	0,03	-0,23	**-0,87**	-0,08	0,01	0,08	**-0,87**	-	0,05	0,06	0,02	**-0,89**	-0,18
Openness – imagination	0,13	0,10	0,13	-0,25	**0,69**	0,02	0,06	0,16	**0,70**	-	0,11	0,18	0,12	**0,68**	-0,21
Conscientiousness – lazy	**-0,78**	0,04	-0,15	-0,03	0,10	**-0,69**	0,16	-0,14	0,05	-	0,06	**-0,83**	-0,11	0,09	-0,02
Conscientiousness – thorough	**0,68**	0,15	0,01	0,00	0,19	**0,60**	0,09	0,07	0,21	-	0,14	**0,69**	0,02	0,20	0,07
Extroversion – reserved	0,03	-0,12	**-0,89**	0,00	-0,05	0,04	-0,11	**-0,90**	-0,06	-	-0,12	0,05	**-0,90**	-0,04	0,04
Extroversion – outgoing	0,31	0,06	**0,66**	0,18	0,05	0,39	0,11	**0,64**	0,05	-	0,08	0,36	**0,62**	0,07	0,21
Agreeableness – trusting	-0,18	0,07	0,28	**0,77**	0,03						0,04	-0,08	0,31	0,05	**0,82**
Agreeableness – find fault	-0,32	-0,21	0,17	**-0,61**	0,09	**-0,58**	-0,31	0,23	0,13	-	-0,21	-0,30	0,21	0,08	**-0,66**
Agreeableness – kind	**0,50**	0,01	0,11	**0,51**	0,09	**0,68**	0,12	0,13	0,07	-					
Neuroticism – relaxed	0,03	**0,86**	0,05	0,06	0,00	0,06	**0,84**	0,09	0,01	-	**0,84**	0,06	0,06	0,01	0,10
Neuroticism – nervous	-0,08	**-0,83**	-0,13	-0,12	-0,06	-0,09	**-0,83**	-0,11	-0,07	-	**-0,85**	-0,03	-0,12	-0,05	-0,07

Note: SHARE Wave 7 data (Release 0). Loadings higher than 0.4 are bolded.

2.3.2.2 Country differences

SHARE facilitates research at the European level, but it also allows for cross-country comparisons as well as within-country studies. To examine the Big Five instrument for use in cross-country and within-country inquiries, we repeated the analyses described above separately for each country. Beyond the issue of how closely the instrument's structure conformed to the theoretical Big Five structure in each setting, we were interested in understanding how similar the instrument's structures are to each other in the respective countries in terms of their measurement invariance. For cross-national comparisons to be valid and unbiased, measurement variance needs to be established. Different levels of measurement invariance are required for different purposes. For example, research questions involving a comparison across countries of associations between the Big Five and a variable of interest (patterning effects) require identical Big Five factor loadings across countries (metric invariance). Research questions that involve mean-level comparisons of

the Big Five across countries (positioning effects) additionally require invariant item intercepts (scalar or strong invariance) (Vandenberg & Lance, 2000).

The country analyses showed that most of the countries demonstrated the same principal component structure as the entire SHARE sample that was pooled across countries. However, some of the countries deviated from the Big Five model, even after correcting for acquiescence. A formal test of measurement invariance using confirmatory factor analyses (CFA) and exploratory structural equation modelling (ESEM) with a latent acquiescence factor (Aichholzer, 2014) confirmed that, as the PCA models suggested, the BFI-10 is not measurement-invariant across all the SHARE countries. CFA and even the much less restrictive ESEM models showed a poorer fit to the data when metric invariance was imposed across countries. That is, the pattern of loadings differed somewhat between at least a subset of countries. The lack of metric invariance implies that the meaning of the five principal components or factors is not identical across all countries and therefore does not always correspond perfectly to the theoretical Big Five structure.

To test how well the data in each country fit the Big Five model, we tested the extent to which each loading pattern of the BFI-10 resembled an idealised Big Five structure in each country. This was achieved by rotating the empirical PCA loading matrix towards an "ideal" target matrix containing only –1, 1, and 0 loadings and then computing the congruency coefficient c (Lechner & Rammstedt, 2015). Consistent with a previous study in this field (Lorenzo-Seva & Ten Berge, 2006), we judged a congruence coefficient of c >.85 as indicating fair similarity or better and values below 0.85 as signifying insufficient similarity. Table 2.4 presents the list of the congruency coefficients of the principal component solution to the Big Five model by country. It is important to emphasise that our results for the pooled sample showed a strong congruency (c =.94) between the idealised Big Five structure and the actual scores. This finding underscores that the SHARE data for all countries combined is, indeed, a good measure of factor similarity. Moreover, in 17 out of 27 countries, the congruency coefficient suggested good or fair correspondence of the empirical and ideal Big Five loading patterns. In these countries, therefore, the Big Five structure holds. However, in ten countries, the Big Five structure did not emerge with sufficient clarity.

Table 2.4: Congruency coefficients per country

Congruency coefficients – similarity to the theoretical Big Five structure		
Fair or good similarity	Sweden	0.97
	Estonia	0.97
	Luxembourg	0.97
	Denmark	0.96
	Poland	0.96
	Austria	0.96
	Finland	0.96
	Belgium	0.96
	Germany	0.96
	Switzerland	0.96
	Czech Republic	0.95
	France	0.94
	Italy	0.93
	Portugal	0.93
	Spain	0.92
	Lithuania	0.92
	Croatia	0.92
insufficient similarity	Hungary	0.84
	Cyprus	0.83
	Israel	0.81
	Slovenia	0.78
	Latvia	0.77
	Malta	0.77
	Greece	0.74
	Slovakia	0.74
	Romania	0.73
	Bulgaria	0.70

We note that even in the countries in which the full five-factor structure did not form, several of the dimensions nevertheless emerged. In fact, the Neuroticism dimensions emerged in all of the countries. Table 2.5 shows the countries in which the respective dimensions emerged, based upon factor loadings of 0.4 or greater.

Table 2.5: Big Five factor analyses per country – the countries in which the dimensions emerged

	Openness	Conscien-tiousness	Extra-version	Agree-ableness	Neuro-ticism
Sweden	X	X	X	X	X
Estonia	X	X	X	X	X
Luxembourg	X	X	X	X	X
Denmark	X	X	X	X	X
Poland	X	X	X	X	X
Austria	X	X	X	X	X
Finland	X	X	X	X	X
Belgium	X	X	X	X	X
Germany	X	X	X	X	X
Switzerland	X	X	X	X	X
Czech Republic	X	X	X	X	X
France	X	X	X	X	X
Italy	X	X	X	X	X
Portugal	X		X	X	X
Spain	X	X	X		X
Lithuania	X	X		X	X
Croatia	X	X	X	X	X
Hungary	X	X			X
Cyprus		X		X	X
Israel		X			X
Slovenia	X	X	X	X	X
Latvia	X	X	X	X	X
Malta					X
Greece			X		X
Slovakia	X	X	X		X
Romania		X			X
Bulgaria	X	X		X	X

Note: The countries are presented in the same order as in Table 2.4

In sum, the lack of measurement invariance for some of the Big Five dimensions in some of the countries implies that caution is needed when engaging in specific cross-national comparisons involving the Big Five personality dimensions. The non-equivalence of the structure of the BFI-10 (or at least some of its dimensions) in at least some countries can complicate the interpretation of potential cross-national differences and introduce potential bias. We therefore advise researchers engaging in specific personality-focused cross-national comparisons to test the measurement invariance of the BFI-10 in the subsamples they use. Advanced analytical techniques, such as approximate measurement invariance using the alignment method (Muthén & Asparouhov, 2012), may prove helpful in this regard. Moreover, researchers should test the sensitivity of their substantive conclusions to different assumptions about measurement invariance.

2.3.3 Using the data: Generated variables in the SHARE release

In accordance with the analyses presented earlier in this chapter and as shown in Table 2.3, the BFI-10 items emerged as the preferred five factorial solutions in the pooled SHARE Wave 7 sample after corrections were made for acquiescence. To facilitate the use of these data, five generated personality trait variables were derived from the raw data. Each variable was aggregated from the set of original ten items (based on the two original Agreeableness items – see the highlighted column in Table 2.3). Each such variable consists of the mean of the pair of respective items (one reverse-coded).

The generated variables are stored in a separate generated variables module under the name of the Big Five dimensions (gv_big5). The five generated variables are bfi10_open (for openness), bfi10_consc (for conscientiousness), bfi10_extra (for extroversion), bfi10_agree (for agreeableness), and bfi10_neuro (for neuroticism). In addition, the raw data of the original 10(+1) items are accessible in the AC module (under the item names, which appear in Table 2.2 above, e.g., ac705_). We should note that in the calculation of the generated variables, we used the original raw data, as it is the accepted practice in empirical analysis and is advised by the experts in this area at GESIS. The acquiescence correction was executed solely to evaluate the impact of bias in factor analyses. However, in the calculation of generated variables, it is generally accepted that acquiescence is implicitly controlled (even if the raw item scores are used). Acquiescence can be conceived of as an additive constant to a person's true score in the "agree" direction of the scale. In "balanced item sets" (i.e., scales with an equal number of positively and negatively worded items per dimension), the negatively worded items have an additive constant to the negative direction, and the positively worded items have an additive constant to the positive direction. Taking the mean across all items, the additive constant is repealed, and acquiescence bias is automatically corrected. This results in a score that is unbiased by acquiescence (Schriesheim & Hill, 1981). Researchers may use all the generated scale scores of the full five-factor model or focus on a specific personality dimension of interest. The personality trait variables

can function as predictors, outcomes or control variables. As noted earlier, personality is associated with a variety of outcomes in late life.

2.3.4 Concluding remarks

The inclusion of the BFI-10 is an important innovation within the SHARE survey. It allows researchers to address personality traits as a key component of late-life outcomes, such as health, depression and well-being. SHARE now provides users with two sources of personality data, one containing the eleven raw Big Five items (these items may be found in the AC module data) and one containing the generated variables of the five personality dimensions (found in the gv_big5 data file). On the whole, the BFI-10 that was employed in SHARE yielded results conforming to the five-factor model of personality across the entire SHARE sample and in most countries individually. The BFI-10 should be used cautiously in specific personality-focused cross-national comparisons, as the instrument is not measurement-invariant across all the SHARE countries. Notwithstanding this limitation, the BFI-10 adds a new dimension to SHARE-based inquiries.

References

Aichholzer, J. (2014). Random intercept EFA of personality scales. Journal of Research in Personality, 53, 1-4. https://doi.org/10.1016/j.jrp.2014.07.001.

Costello, S., & Roodenburg, J. (2015). Acquiescence response bias – Yeasaying and higher education. The Australian Educational and Developmental Psychologist, 32(2), 105-119. https://doi.org/10.1017/edp.2015.11.

Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. Journal of Personality and Social Psychology, 59(6), 1216-1229. https://doi.org/10.1037/0022-3514.59.6.1216.

John, O. P., & Srivastava, S. (1999). The Big-Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), Handbook of personality: Theory and research (2nd ed.) (pp. 102-138). New York: The Guilford Press.

Lechner, C. M., & Rammstedt, B. (2015). Cognitive ability, acquiescence, and the structure of personality in a sample of older adults. Psychological Assessment, 27(4), 1301-1311. https://doi.org/10.1037/pas0000151.supp.

Lorenzo-Seva, U., & ten Berge, J. M. F. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. European Journal of Research Methods for the Behavioral and Social Sciences, 2(2), 57-64. https://doi.org/10.1027/1614-2241.2.2.57.

Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. Psychological Methods, 17(3), 313-335. https://doi.org/10.1037/a0026802.

Rammstedt, B., & Farmer, R. F. (2013). The impact of acquiescence on the evaluation of personality structure. Psychological Assessment, 25(4), 1137-1145. https://doi.org/10.1037/a0033323.

Rammstedt, B., Goldberg, L. R., & Borg, I. (2010). The measurement equivalence of Big-Five factor markers for persons with different levels of education. Journal of Research in Personality, 44(1), 53-61. https://doi.org/10.1016/j.jrp.2009.10.005.

Rammstedt, B., & John, O. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. Journal of Research in Personality, 41(1), 203-212. https://doi.org/10.1016/j.jrp.2006.02.001.

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. Perspectives on Psychological Science, 2(4), 313-345. https://doi.org/10.1111/j.1745-6916.2007.00047.x.

Schriesheim, C., & Hill, K. D. (1981). Controlling acquiescence response bias by item reversals: The effect on questionnaire validity. Educational and Psychological Measurement, 41(4), 1101-1114. https://doi.org/10.1177/001316448104100420.

Vandenberg, R. J., & Lance, C. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. Organizational Research Methods, 3(1), 4-70. https://doi.org/10.1177/109442810031002.

2.4 New measures for interpersonal environment during childhood

Noam Damri, Howard Litwin – The Hebrew University of Jerusalem

2.4.1 Introduction

The third wave of SHARE (SHARELIFE) showed very clearly that early-life experiences impact late-life outcomes (Börsch-Supan et al., 2011). However, the variables that were examined in SHARELIFE in 2008 did not take into account the respondents' interpersonal environment while they were growing up. One's relationship with his or her parents during childhood is a key factor in shaping his or her personality and traits and is the context in which future relationships and psychological developments take place (Bowlby, 1988). Studies have documented the long-lasting effect of such relationships on domains such as health and well-being (Shaw et al., 2004).

2.4.2 Questionnaire development

To broaden the understanding of the complex relationship between early-life relations and different outcomes at later stages in life, new retrospective indicators that reflect the nature of the interpersonal environment (family and social network) in early and midlife were formulated for SHARE. These additional measures provide information on the extent of social embeddedness at key points in the life course and allow for the examination of the relationship between the family and social network at younger ages on the one hand and that between health status and other relevant late-life outcomes on the other hand.

The new retrospective social indicators in SHARE were derived from a number of prominent surveys. Among them were the Panel Study of Income Dynamics Childhood (PSID) Retrospective Circumstances Study questionnaire from 2014 (PSID, 2015), the National Survey of Midlife Development in the United States (MIDUS) Self-Administered Questionnaire from 1995-1996 (Brim et al., 1999), and the English Longitudinal Study of Ageing (ELSA) Life History Self-Completion Questionnaire from 2009 (Ward et al., 2009). A preliminary set of questions was drafted based on these questionnaires. The respective measures probed issues during childhood such as relationships with parents, physical abuse by parents or other people, friendships and social connections, quality of the neighbourhood in which one grew up and family religiosity. Due to the time constraints of the SHARE interview, not all of these questions could be included in the final version of the questionnaire. The preliminary set of questions was tested in the pretest that was conducted in February 2016, in which data from approximately 1,900 respondents from 17 countries were collected. Based on the results from the pretest, a subsequent abbreviated version of the question set was formulated and tested in the field rehearsal. The final set of questions queried in the main data collection, along with their response options, is presented in Table 2.6.

2.4.3 Preliminary explication of the data

The retrospective questions were given in Wave 7 to all respondents who had not participated in the previous life history questionnaire that was administered in Wave 3 (SHARELIFE). Hence, respondents who had participated in Wave 3 were not asked the childhood interpersonal environment probes. The number of respondents receiving the retrospective interpersonal environment questions, therefore, was about 60,000, and they constituted approximately 82 percent of all respondents aged 50+ in Wave 7 (based on Release 0 data).

Table 2.6: Overview of the new questions on interpersonal environment during childhood

Question code	Question text (interviewer instruction)	Response options	Other studies using similar item
CC721_understand*	How much did your mother/your father (or the woman/man who raised you) understand your problems and worries?	1. A lot 2. Some 3. A little 4. Not at all	PSID Childhood Retrospective Circumstances Study (2014)

Question code	Question text (interviewer instruction)	Response options	Other studies using similar item
CC722_relationship*	How would you rate the relationship with your mother/ your father (or the woman/man who raised you)?	1. Excellent 2. Very good 3. Good 4. Fair 5. Poor	PSID Childhood Retrospective Circumstances Study (2014)
CC725_Harm*	How often did your mother/your father push, grab, shove, throw something at you, slap, or hit you?	1. Often 2. Sometimes 3. Rarely 4. Never	PSID Childhood Retrospective Circumstances Study (2014)
CC727_HarmElse	How often did anybody else physically harm you in any way?	1. Often 2. Sometimes 3. Rarely 4. Never	PSID Childhood Retrospective Circumstances Study (2014)
CC728_Religion	How important was religion in your home when you were growing up?	1. Very important 2. Somewhat important 3. Not very important 4. Not at all important	The National Survey of Midlife Development in the United States (MIDUS) Self-Administered Questionnaire (1995-1996)
CC729_Lonely	Please look at card 16. Now I would like you to think back to your childhood, how often were you lonely for friends? *(Interviewer instruction: Childhood is during school years, between ages 616)*	1. Often 2. Sometimes 3. Rarely 4. Never	PSID Childhood Retrospective Circumstances Study (2014)
CC730_Comfortable	And how often did you have a group of friends that you felt comfortable spending time with? *(Interviewer instruction: Childhood is during school years, between ages 6-16)*	1. Often 2. Sometimes 3. Rarely	PSID Childhood Retrospective Circumstances Study (2014)

Note: *Question asked twice, once for each parent.

The results from an exploratory factor analysis (see Table 2.7) highlight the way in which the questions tap different aspects of the interpersonal environment during childhood. Three separate factors emerged from the analysis: (1) relationship with parents during childhood, (2) any physical harm respondent suffered (from parents or anyone else), and (3) relationships with friends during childhood. One additional question, the importance of religion at home when growing up, did not load on any of the resultant factors.

Table 2.7: Factor analysis of the new questions on interpersonal environment during childhood

	Factor 1	Factor 2	Factor 3	Uniqueness
Mother understands problems and worries	0,72			0,41
Father understands problems and worries	0,74			0,34
Relationship with mother	0,73			0,40
Relationship with father	0,74			0,35
Mother physical harm		0,58		0,42
Father physical harm		0,58		0,39
Anybody else physical harm		0,58		0,52
Importance of religion at home when growing up				0,91
Lonely for friends in childhood			0,70	0,32
Group of friends felt comfortable spending time with			-0,74	0,31
Eigenvalue	3,04	1,36	1,23	
% of total variance	25,5%	16,5%	14,2%	

Note: SHARE Wave 7 data (Release 0). Results from principal-component factor analysis with varimax (orthogonal) rotation. N= 54,625. Rotated factor loadings and unique variances (blanks represent abs(loading)<.5).

2.4.4 Concluding remarks

The new questions address important social aspects of childhood circumstances that were not solicited in the previous SHARELIFE questionnaire. The new items can improve the understanding of the associations among the interpersonal environment, relationships with parents and the extent of social embeddedness during childhood on the one hand and among several middle- and late-life outcomes on the other hand. Thus, for example, one can examine whether positive and/or negative childhood relations with parents are related to subsequent educational achievements over the life course, whether having had friends during one's youth is related to the extent of social engagement in later life, and whether factors such as personality, financial status or disability moderate the effects of the childhood interpersonal environment on one's mental health in old age.

References

Börsch-Supan, A., Brandt, M., Hank, K., & Schröder, M. (2011). The individual and the welfare state: Life histories in Europe. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-17472-8.

Bowlby, J. (1988). A secure base: Parent-child attachment and healthy human development. New York: Basic Books.

Brim, O. G., Baltes, P. B., Bumpass, L. L., Cleary, P. D., Featherman, D. L., Hazzard, W. R., Kessler, R. C., Lachman, M. E., Markus, H. R., Marmot, M. G., Rossi, A. S., Ryff, C. D., & Shweder, R. A. (1999). National Survey of Midlife Development in the United States (MIDUS 1), 1995-1996. Ann Arbor: Inter-university Consortium for Political and Social Research. https://doi.org/10.3886/ICPSR02760.v15.

McGonagle, K., & Freedman, V. A. (2015). The Panel Study of Income Dynamics, Childhood Retrospective Circumstances Study (PSID-CRCS) user guide: Final release 1. Retrieved from https://psidonline.isr.umich.edu/CRCS/2014UserGuide.pdf.

Shaw, B. A., Krause, N., Chatters, L. M., Connell, C.M., & Ingersoll-Dayton, B. (2004). Emotional support from parents early in life, aging, and health. Psychology and Aging, 19(1), 4-12. https://doi.org/10.1037/0882-7974.19.1.4.

Ward, K., Medina, J., Mo, M., & Cox, K. (2009). ELSA Wave 3: Life history interview. A user guide to the data. Version 1. Retrieved from http://doc.ukdataservice.ac.uk/doc/5050/mrdoc/pdf/5050_Wave_3_Life_History_Documentation.pdf.

2.5 Incidence and implications of dramatic life events: Extending the interview to cover experiences of respondents' parents

Radim Bohacek – The Czech Academy of Sciences, Michal Myck – Centre for Economic Analysis (CenEA)

There is growing evidence of the importance of major life events in individual life-course development and of their role in determining lifetime trajectories in outcomes related to health, labour market status, family structure or income and wealth (e.g., Acemoglu et al., 2011; Kesternich et al., 2014; Bohacek & Myck, 2017). Such analyses, among others, have been made possible through the availability of retrospective information, such as the data collected in the original SHARELIFE survey in Wave 3 of SHARE. With this in mind, this chapter describes the reasoning behind the extension of the set of questions focused on major life events in the SHARELIFE questionnaire, which was implemented in Wave 7.

The SHARELIFE battery of questions related to the experience of persecution, which was administered in the initial life history interview in Wave 3, had a specific focus on the relationship between the individual and the state (Bohacek & Myck, 2011). From the perspectives of the individuals comprising the cohorts participating in the survey, this relationship was of particular importance in the countries of Central and Southern Europe. It proved valuable, for example, in the analysis of the implications of political persecution in the former Communist countries on long-term outcomes, such as life satisfaction, material conditions and health (Myck & Bohacek, 2011), or in the identification of the role of persecution for lifetime earnings and pensions (Bohacek & Myck, 2017). However, given the socio-political developments in Europe in the 20th century and the extent of the possible experiences of tragic events among the analysed cohorts, it was decided to extend the original battery of questions to include some information on the implications of these developments on the lives of the respondents' parents.

This chapter briefly describes the set of questions focused on the experience of persecution, which were implemented in the same way as in Wave 3, and subsequently provides the rationale for extending this section to cover the specific experiences of respondents' parents. It is worth noting that SHARE is the first international survey that provides micro data on persecution, discrimination and dispossession from a representative sample.

2.5.1 Items on experiences of persecution and dispossession

The SHARE Wave 7 questionnaire included the same items focused on the experience of persecution as the original SHARELIFE interview. Individuals were asked whether they ever experienced persecution or discrimination, given the following introductory definition:

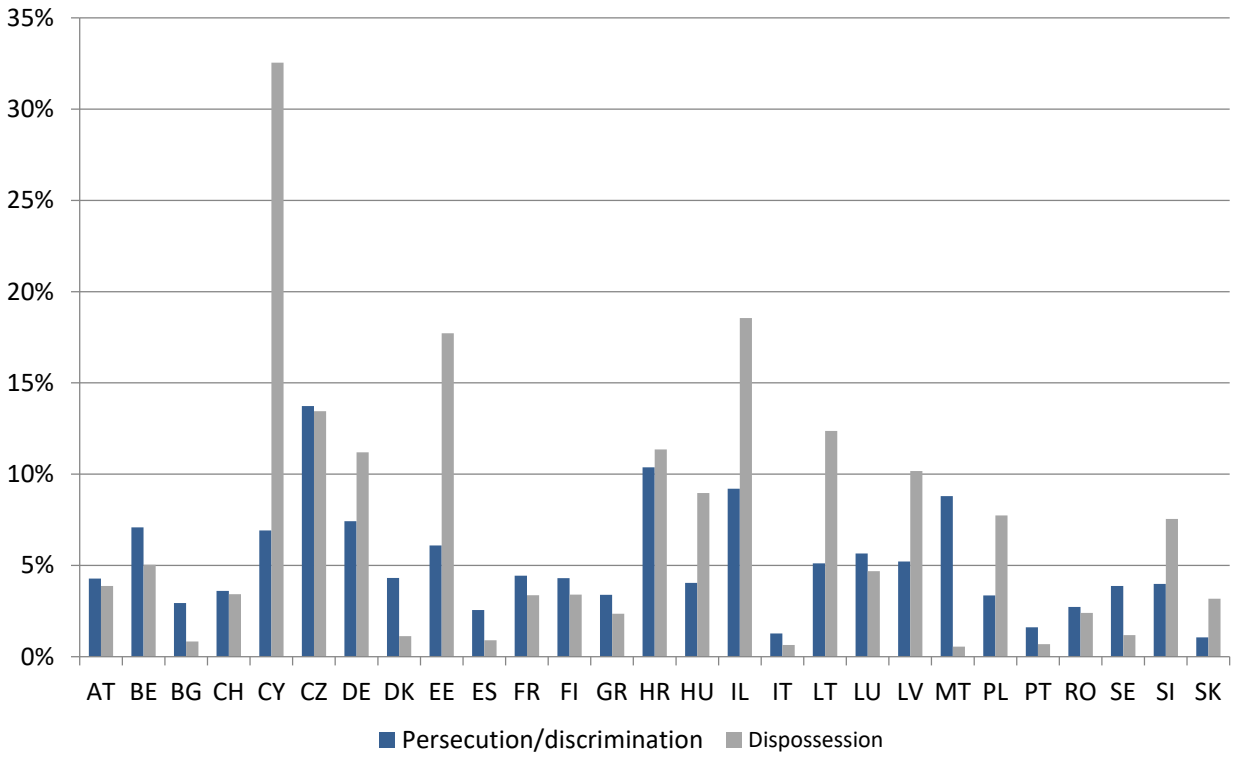
[GL022_EverVictPers] "There are times in which people are persecuted or discriminated against, for example, because of their political beliefs, religion, nationality, ethnicity, sexual orientation or their background. People may also be persecuted or discriminated against because of the political beliefs or the religion of their close relatives."

"Have you ever been the victim of such persecution or discrimination?"

If such experience were reported in the survey, the questionnaire asked for the main reason for such persecution (GL023_ReasPersec) and its principal consequences, with respect to the following:

- job loss (GL024_PersecStopWork)
- other negative consequences at work, e.g., denied promotions, harassment, pay cuts (GL026_ConsPersec)
- difficulties finding a new job (GL028_DiffFindJob)
- imprisonment, labour camp, exile, etc. (GL030_CampPersec)

The persecution items were asked at the end of the life history interview. This meant that, in cases when respondents reported job losses or on-the-job consequences of persecution, they could be asked to identify these jobs from a job list created earlier in the interview.



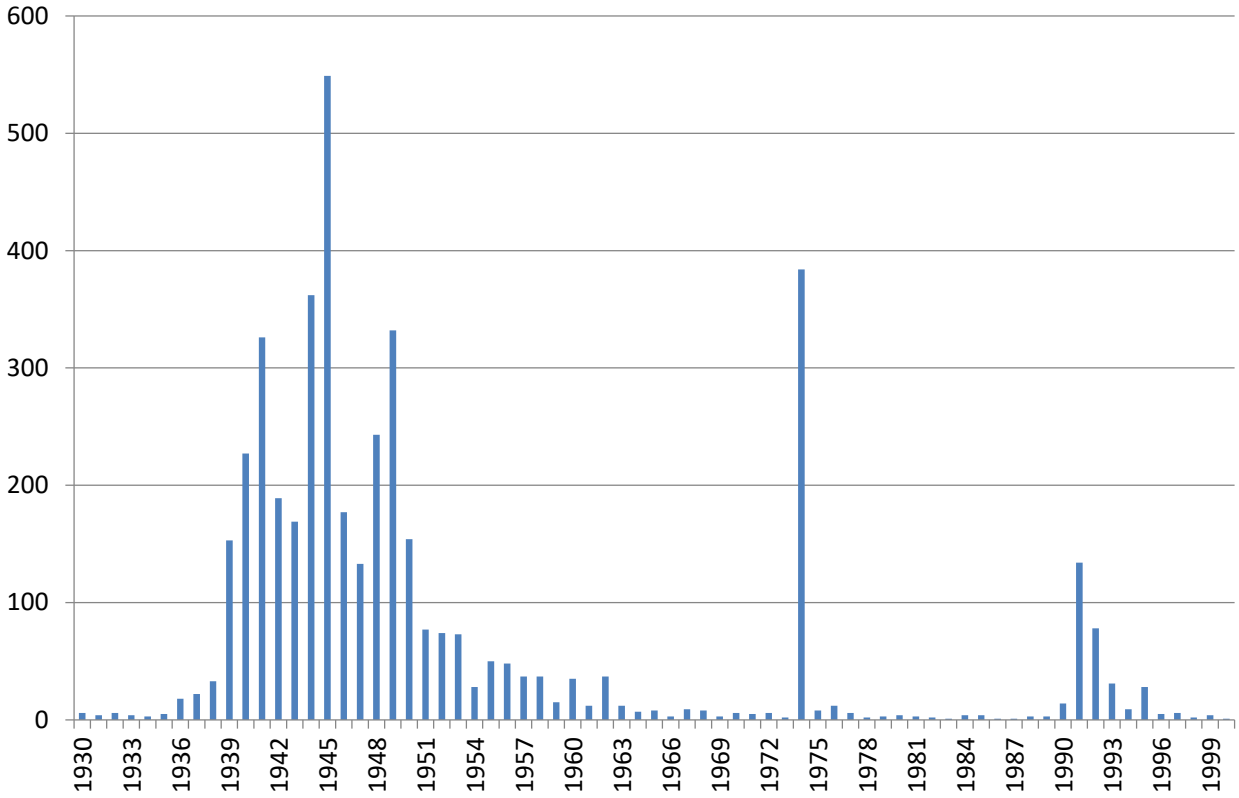
Note: Own calculations using SHARE Wave 7 data (Release 0), unweighted.
Figure 2.5: Experience of persecution/discrimination and dispossession in SHARE

The second set of items related to the implications of persecution focuses on the experiences of dispossession as a result of war or persecution and opens with the following question:

> *[GL031_PropDissp] "There may be cases when individuals and their families are dispossessed of their property as a result of war or persecution. Were you or your family ever dispossessed of any property as a result of war or persecution?"*

This item is followed by detailed questions on the timing of dispossession (GL033_WhenPropAway), the type of property that had been lost (GL032_TypePropDissp), and whether respondents' families had ever been compensated for their loss (GL034_TypePropDissp).

In Figure 2.5, we show the (unweighted) percentages of respondents who declared in the Wave 7 interview that they had experienced persecution or discrimination and who reported that they or their families were ever dispossessed. Figure 2.6, on the other hand, shows the timing of dispossession as reported in the interviews – by year of dispossession. The figure reflects a number of historic developments in different European countries over the years, including World War II, the early years of Communism, the consequences of the Turkish invasion of Cyprus in 1974, and the Balkan wars of the early 1990s.



Note: Own calculations using SHARE Wave 7 data (Release 0), unweighted.

Figure 2.6: Incidents of property dispossession among SHARE respondents and their families by year of dispossession

2.5.2 Items on dramatic life events in the lives of respondents' parents

The primary goal of SHARE is to better understand the interplay between the state and different dimensions of individual life courses. Thus, while it would be extremely interesting to collect detailed information on the social surroundings of the respondents (e.g., close relatives, extended families, members of broader social networks), the survey focuses on developments that could translate into substantial implications for the respondents' life courses. To keep the questionnaire concise, the choice of questions therefore focuses only on the life events of the parents of the respondents, as parents are, in most cases, the closest relatives and clearly – in one way or another – shape the respondents' lives very strongly. Asking about life events as experienced by the parents attempts to identify the implications of major historical developments to which individuals and their families were exposed. A further selection was made to include events that potentially identify differences in the life course and that especially might have long-lasting consequences for the later life outcomes of the respondents. The selection of life events formed the answer categories for GL738 and GL739 (see wording below).

The following items were asked in Wave 7 as extensions of the persecution section described above. The selected experiences are likely to have occurred in all SHARE countries but particularly in those affected by World War II and Communism. Since these questions were not included in the Wave 3 SHARELIFE questionnaire, they were asked to all respondents in Wave 7 (those who did and those who did not participate in Wave 3). The questions were asked separately about the experiences of the respondent's mother and father. Respondents were asked to "select all that apply" from a list with which they were presented. For those who were not asked the persecution questions, the new items were preceded by the following introduction:

"There are times in which people are persecuted or discriminated against, for example, because of their political beliefs, religion, nationality, ethnicity, sexual orientation or background."

[GL737_intro_discrM/GL739_intro_discrF] "Now we would like to ask whether your parents were affected by war, persecution or discrimination. Please think about your [mother/father] now."

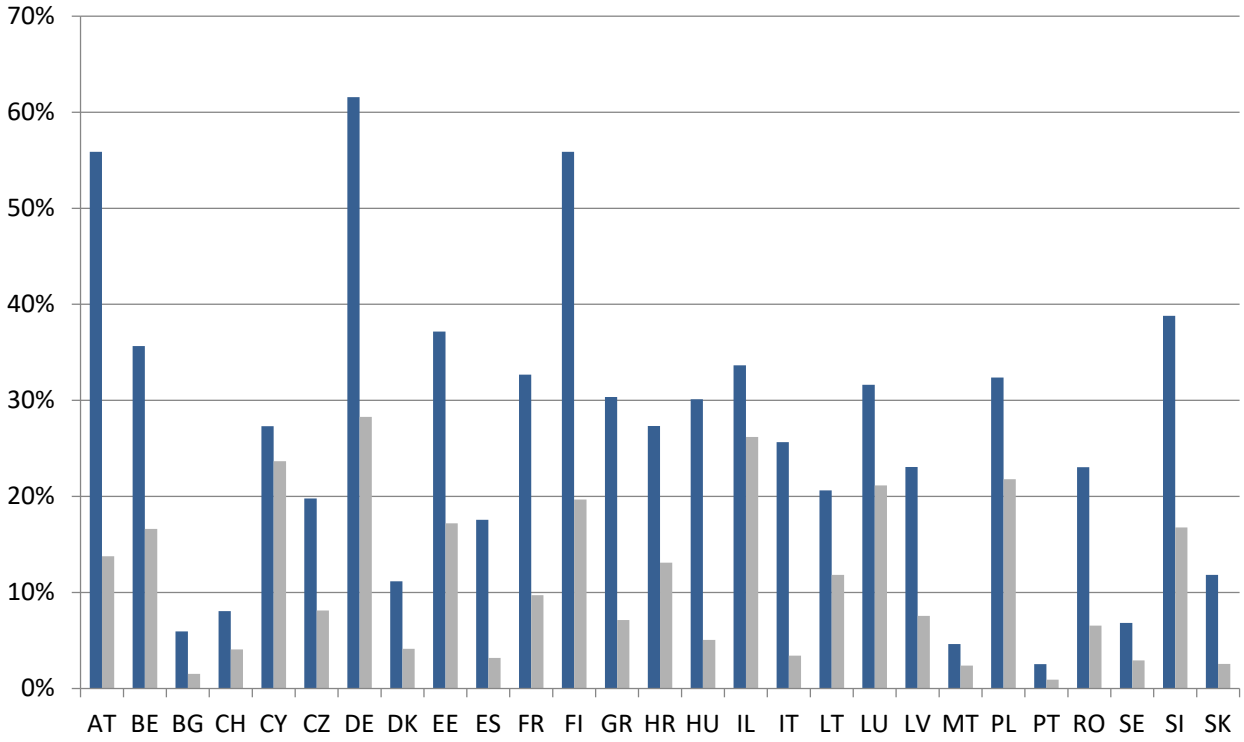
[GL738_discrmother/GL739_discrfather] "Please look at show card {SHOWCARD_ID}. Has your [mother/father] experienced the following situations and consequences because of war, persecution or discrimination?"

1. Imprisonment
2. Labour camp
3. Concentration camp
4. Deportation, forced displacement or flight
5. Engaged in combat operations/fighting
6. Serious damage to health or injury – includes damage to physical or mental health
7. Death

96. None of these

Based on discussions and feedback at the pretest and field rehearsal stages, the following interviewer instructions were added: "By 'war', we also mean the time of occupation during the World War II. Death only includes death of the mother/father of the respondent as a direct consequence of war, persecution or discrimination. For example, if the mother died in a labour camp, mark both labour camp and death."

In Figure 2.7, we show (unweighted) percentages of respondents who declared that their mothers and fathers experienced any of the above situations and consequences of war, persecution or discrimination. As we can see, the highest frequencies of respondents whose fathers experienced any of the listed situations or consequences were reported in Germany (59.1 percent), Finland (55.8 percent), Austria (55.0 percent), Slovenia (38.8 percent) and Estonia (37.2 percent), while the highest frequencies among mothers were reported in Germany (27.4 percent), Israel (26.2 percent), Cyprus (23.7 percent), Luxemburg (21.1 percent) and Poland (20.9 percent).



Note: Own calculations using SHARE Wave 7 data (Release 0), unweighted.

Figure 2.7: Dramatic life events in the life of respondents' parents

2.5.3 Concluding remarks

Given the dramatic developments in the 20th century in most countries that are part of SHARE and the significant role of the family and household context in determining individual life courses, the record of major life events as experienced by the respondents' parents should be helpful in better understanding life course developments and later life outcomes among Europeans aged 50 and older. The questions regarding life events of respondents' parents, which are new to SHARE in Wave 7, complement the information collected in Wave 3 and Wave 7 concerning respondents' experiences of persecution and discrimination. First, the combined data will complement the existing information by including life events to which the respondents were indirectly exposed, and second, the data will shed light on the importance of major socio-political events that respondents and their families experienced in their lives.

References

Acemoglu, D., Hassan, T. A., & Robinson, J. A. (2011). Social structure and development: A legacy of the holocaust in Russia. Quarterly Journal of Economics, 126(2), 895-946. https://doi.org/10.1093/qje/qjr018.

Bohacek, R., & Myck, M. (2011). Persecution in Central Europe and its consequences on the lives of SHARE respondents. In A. Börsch-Supan, M. Brandt, K. Hank & M. Schröder, (Eds.), The individual and the welfare state: Life histories in Europe (pp. 271-285). Berlin, Heidelberg: Springer. http://doi.org/10.1007/978-3-642-17472-8_24.

Bohacek, R., & Myck M. (2017). Economic consequences of political persecution. IZA Discussion Papers (11136). Retrieved from http://hdl.handle.net/10419/174046.

Kesternich, I., Siflinger, B., Smith, J. P., & Winter, J. K. (2014). The effects of World War II on economic and health outcomes across Europe. The Review of Economics and Statistics 96(1), 103-118. https://doi.org/10.1162/REST_a_00353.

Myck, M., & Bohacek, R. (2011). Long shadows of history: Persecution in Central Europe and its labor market consequences. IZA Discussion Papers (6130). Retrieved from http://hdl.handle.net/10419/58834.

2.6 Intergenerational cohabitation at older ages

Giacomo Pasini – Università Ca' Foscari Venezia, Guglielmo Weber – University of Padova

2.6.1 Motivation

In the retrospective questionnaire of Wave 7, respondents answered questions about episodes of intergenerational co-residence. In particular, they reported whether and when they lived together with their parents or parents-in-law after they left the parental home to set up their own household. Further, if their children lived with them at the time of the interview, respondents recorded whether and when they came back to the parental nest after they had left it to set up their own household. Therefore, the intergenerational perspective in SHARE is covered in two directions by including the younger generation (living with children) and the older generation (living with parents) of the respondent. The research issues behind this set of questions are as follows:

a) To what extent do particular trajectories in family composition affect outcomes at older ages, such as socio-economic status, wealth, health, social engagement and social support?

b) What is the mechanism by which such effects operate? Does a larger family provide economic support and/or help with family chores that allow its members to be more involved in the labour market?

c) What kind of welfare changes in the twentieth century shaped the size and composition of households? Did the expansion of the welfare state (social housing, childcare, maternity leave and health care) crowd out the need for insurance within the family? Or, was it the transformation of household size and composition that led to the need for such policies?

We stress that evidence on current co-residence is available from standard SHARE waves, but standard waves are silent on intergenerational cohabitation (co-residence) back in time and the reasons for it. Therefore, the retrospective questionnaire of Wave 7 includes a set of questions on household structure that should overcome this limitation; these questions ask for the first and last period of cohabitation with parents/parents-in-law and the motivation for cohabitation ("to help them, to receive help, both, none of them"). Similar questions were asked about children currently living with the respondents, that is, if they had previously left the parental nest.

In this chapter, we present some descriptive statistics based on SHARE Wave 7 data (Release 0). We start with questions about the co-residence of the respondents with their parents/parents-in-law. Section 2.6.2 looks at cohabitation of respondents with their parents and parents-in-law; Section 2.6.3 examines cohabitation with their children. We draw conclusions in the last section.

2.6.2 Parents and parents-in-law

Table 2.8 reports the number of respondents who report having co-resided with their mother, father, mother-in-law or father-in-law. We see that over three-quarters report never having lived with any of these individuals, while almost 12 percent have co-resided with their mother, 10 percent with their mother-in-law, and smaller percentages with their father (8 percent) or father-in-law (7 percent), which is consistent with the tendency for women to survive their husbands. The most interesting feature that emerges from the table is the very low number of item nonresponses (items marked "don't know", refusals and missing values account for just 0.1 percent of all responses).

Table 2.8: Who lived in the household (parents of the respondent)

	Mother		Father		Mother-in-law		Father-in-law		None of these	
	n	%	n	%	n	%	n	%	n	%
Not selected	52010	88,1	54217	91,8	53069	89,9	54767	92,8	13359	22,6
Selected	6955	11,8	4748	8,0	5896	10,0	4198	7,1	45606	77,3
Missing	42	0,1	42	0,1	42	0,1	42	0,1	42	0,1
Don't know	24	0,0	24	0,0	24	0,0	24	0,0	24	0,0
Refusal	8	0,0	8	0,0	8	0,0	8	0,0	8	0,0
Total	59039	100,0	59039	100,0	59039	100,0	59039	100,0	59039	100,0

Note: SHARE Wave 7 data (Release 0).

In Table 2.9, we display the most common types of co-residence. We see that living with both parents is the most common type (reported by almost 7 percent), followed by living with both in-laws (6 percent) and living with the mother (4 percent) or mother-in-law (3 percent) alone. Other types of co-residence are less common. The table also shows (absolute) frequencies by gender and current age.

Table 2.9: Who lived with the respondent, by gender and age

	Overall		Gender		Age	
			Males	Females	70+	50-69
	n	%	n		n	
Mother alone	2522	4,3	951	1571	1162	1360
Father alone	476	0,8	225	251	204	271
Mother and father	4021	6,8	1967	2054	1595	2426
Mother-in-law alone	1949	3,3	762	1187	886	1063
Father-in-law-alone	413	0,7	126	287	166	246
Mother- and father-in-law	3554	6,0	1226	2328	1508	2045
Two parents, not couple	160	0,3	65	95	87	73
3 or 4 parents	264	0,5	102	162	104	160
None of these	45606	77,3	19754	25851	17666	27919
Missing	74	0,1	34	40	37	36
Total	59039	100,0	25212	33826	23415	35599

Note: SHARE Wave 7 data (Release 0).

A possible concern with retrospective information is that respondents may provide inaccurate information on the exact timing of the episodes they report. In the case of co-residence, for instance, some individuals may fail to report dates or state that a co-residence period ended before it started.

Table 2.10 addresses this issue by reporting the number of cases in which both the start and end years are reported and for which the end year is strictly after the start year. It does so separately for the first co-residence period and then for the last (if different). The consistency variable takes a value of 1 if the condition is met and a value of 0 otherwise.

Table 2.10: Consistency of year information

First co-residence period								
	Mother		Father		Mother-in-law		Father-in-law	
	n	%	n	%	n	%	n	%
1	6729	95,6	4528	94,3	5798	97,9	4112	97,6
0	307	4,4	272	5,7	124	2,1	101	2,4
Total	7036	100,0	4800	100,0	5922	100,0	4213	100,0
Last co-residence period, if more than one								
	Mother		Father		Mother-in-law		Father-in-law	
	n	%	n	%	n	%	n	%
1	607	91,4	306	93,3	401	89,3	189	90,0
0	57	8,6	22	6,7	48	10,7	21	10,0
Total	664	100,0	328	100,0	449	100,0	210	100,0

Note: SHARE Wave 7 data (Release 0).
1 = consistent information (end year is strictly greater than start year); 0 = inconsistent information on start and end year of co-residence.

We see that in over 94 percent of all cases, the consistency check is passed for the first co-residence period; the percentage falls by a few points for the last co-residence period, probably reflecting respondent fatigue.

In Table 2.11, we provide evidence on the number of cases in which parents co-resided with the respondent at least twice and in which there is a gap of more than one year between the end of the first and the beginning of the last cohabitation period. In this case, there might be further episodes of co-residence that were not reported because the respondent was only asked about first and last co-residence periods. The variable takes a value of 1 if there is a period in which the respondent might have been co-residing with a given parent. We see that this could have happened in 15 percent of the cases for co-residence with the mother, 12 percent for co-residence with the mother-in-law, and in less than 10 percent of cases for co-residence with the father or father-in-law.

Table 2.11: Gaps in information about co-residing parents

	Mother		**Father**		**Mother-in-law**		**Father-in-law**	
	n	%	n	%	n	%	n	%
No gap years	5884	84,6	4300	90,6	5209	88,4	3920	93,4
Gap years	1071	15,4	448	9,4	687	11,7	278	6,6
Total	6955	100,0		100,0	5896	100,0	4198	100,0

Note: SHARE Wave 7 data (Release 0).

The questionnaire also elicited information on the reasons why such co-residence took place. A respondent had to choose among four mutually exclusive possibilities. Co-residence could take place to help the respondent, to help the other named person (mother, father, mother-in-law, father-in-law), to help both or to help neither. Table 2.12 lists the absolute and relative frequencies for each type of person and for the current, first and last episodes.

Table 2.12: Reasons for cohabitation with parents/parents-in-law

Mother	Current help		First help		Last help	
	n	%	n	%	n	%
Help mother	226	35,4	1526	23,7	224	36,1
Help respondent	15	2,4	705	11,0	38	6,1
Help both	326	51,0	3129	48,6	286	46,1
Help neither	69	10,8	1054	16,4	69	11,1
Don't know	2	0,3	22	0,3	4	0,6
Refusal	1	0,2	4	0,1		
Total	639	100,0	6440	100,0	621	100,0

Father	Current help		First help		Last help	
	n	%	n	%	n	%
Help father	45	19,9	752	16,4	98	30,8
Help respondent	5	2,2	544	11,9	20	6,3
Help both	133	58,9	2355	51,4	158	49,7
Help neither	40	17,7	895	19,5	40	12,6
Missing			1	0,0		
Don't know	2	0,9	30	0,7	2	0,6
Refusal	1	0,4	7	0,2		
Total	226	100,0	4584	100,0	318	100,0

Mother-in-law	Current help		First help		Last help	
	n	%	n	%	n	%
Help mother-in-law	109	28,6	1092	19,6	163	37,8
Help respondent	7	1,8	557	10,0	27	6,3
Help both	226	59,3	2904	52,2	188	43,6
Help neither	39	10,2	999	18,0	50	11,6
Don't know			7	0,1	3	0,7
Total	381	100,0	5559	100,0	431	100,0

Father-in-law	Current help		First help		Last help	
	n	%	n	%	n	%
Help father-in-law	25	17,86	592	14,51	63	31,03
Help respondent	4	2,86	435	10,66	11	5,42
Help both	94	67,14	2272	55,69	99	48,77
Help neither	17	12,14	771	18,90	28	13,79
Don't know			9	0,22	2	0,99
Refusal			1	0,02		
Total	140	100,0	4080	100,0	203	100,0

Note: SHARE Wave 7 data (Release 0).

Three general features emerge from the table. First, in all cases, mutual help is the modal answer. This finding is not surprising, given that co-residence affords major savings, as there are economies of scale in the production of household services (shelter, heating, utilities, food, etc.). However, we also see that if the respondent states that co-residence helps only one person, this person is someone else. This finding is in line with expectations, given that respondents are aged 50 years and older, and their parents and parents-in-law will tend to be some 25-35 years older. However, there may also be some reporting bias due to the natural reluctance to admit one's own dependence on others. Finally, there is a time (or age) gradient: the last episodes of co-residence are more often intended to help the named person than are the first episodes.

2.6.3 Children

Respondents who lived with children at the time of the interview were asked since when the co-residence was taking place, and when (if at all) each cohabiting child left the parental home for the first time to establish his/her own household. Moreover, respondents reported the motive for co-residence in exactly the same way as for parents and parents-in-law.

In Table 2.13, we check for consistency of dates. The variable takes a value of 0 if dates are correctly reported, that is, if the co-residing child never left the parents' household, or if the date on which he/she established his/her own household preceded the date on which the child started the current cohabitation period. We show the consistency variable separately for each child (up to the fifth natural child) and collectively for the remaining natural children on the one hand and for adopted and foster children on the other. The results show that over 92 percent of the information about cohabiting children is correct.

Table 2.13: Consistency of years

	1st child		2nd child		3rd child		4th child		5th child	
	n	%	n	%	n	%	n	%	n	%
1	5113	93,1	4833	94,5	2325	94,7	830	93,8	287	92,9
0	379	6,9	280	5,5	129	5,3	55	6,2	22	7,1
Total	5492	100,0	5113	100,0	2454	100,0	885	100,0	309	100,0

	Other natural children		Adopted or foster children	
	n	%	n	%
1	253	96,9	127	96,9
0	8	3,1	4	3,1
Total	261	100,0	131	100,0

Note: SHARE Wave 7 data (Release 0).
1 = consistent information (end year is strictly greater than start year);
0 = inconsistent information on start and end year of co-residence

Table 2.14 further elaborates on these data. We now consider only the "valid responses", i.e., those for which the child is currently cohabiting and information on dates is usable (value "1" in previous table). We check for observations for which we do not have enough information to cover the entire life span. There are no gap years if the child either always lived in the household or left the household and returned the subsequent year. There are gap years if the child established his/her own household at least two years prior to the year in which the current cohabitation with parents started. In this case, we do not know whether there were other cohabiting periods besides the current one or whether the child lived continuously on his/her own. We see from Table 2.14 that for the vast majority of cohabiting children, there are no gap years.

Table 2.14: Gaps in information about co-residing children

	1st child		2nd child		3rd child		4th child		5th child	
	n	%	n	%	n	%	n	%	n	%
No gap years	4796	93,5	4581	94,7	2205	94,7	795	95,2	275	95,2
Gap years	331	6,5	258	5,3	124	5,3	40	4,8	14	4,8
Total	5127	100,0	4839	100,0	2329	100,0	835	100,0	289	100,0

	Other natural children		Adopted or foster children	
	n	%	n	%
No gap years	232	91,7	118	92,9
Gap years	21	8,3	9	7,1
Total	253	100,0	127	100,0

Note: SHARE Wave 7 data (Release 0).

Finally, Table 2.15 reports the motive for cohabiting with a given child. As we already saw in Table 2.12, mutual help is the modal answer. However, the fraction of respondents who report that they are cohabiting to help the named child is substantial. This finding is in line with the evidence about the most recent cohabitation of respondents with their parents and parents-in-law reported in Table 2.12. In that case, the respondents more often claim they are co-residing in order to help their parents rather than to receive help. This situation is perfectly possible, given that we are focusing on the middle generation (which is expected to provide help to both the older and the younger generations). However, another explanation is that the perception of the respondents suffers from a "warm glow" bias: respondents over-report playing the active role in the helping relationship.

Table 2.15: Reasons for cohabitation with child/children

	1st child		2nd child		3rd child		4th child		5th child	
	n	%	n	%	n	%	n	%	n	%
Help child	1555	27,7	1524	29,3	766	49,2	279	30,9	86	27,4
Help respondent	154	2,7	131	2,5	62	4,0	33	3,7	17	5,4
Help both	2544	45,3	2324	44,6	105	6,7	350	38,7	116	36,9
Help neither	1359	24,2	1231	23,6	624	40,1	242	26,8	95	30,3
Total	5612	100,0	5210	100,0	1557	100,0	904	100,0	314	100,0

	Other natural children		Adopted and foster children	
	n	%	n	%
Help child	70	25,5	48	35,8
Help respondent	7	2,5	2	1,5
Help both	124	45,1	47	35,1
Help neither	74	26,9	37	27,6
Total	275	100,0	134	100,0

Note: SHARE Wave 7 data (Release 0).

2.6.4 Concluding remarks

In this chapter, we evaluated the quality of the data collected with the new questions in Wave 7 designed to reconstruct the cohabitation history of respondents with their parents and children to evaluate their potential for future research. Data quality is remarkably high: item nonresponse is negligible, and respondents report dates correctly in the vast majority of cases. Data quality is a prerequisite for usefulness in research, but even from this brief analysis, some further indications emerge. First, for most respondents, we are in the position of reconstructing the exact composition of their households for their entire lives, meaning that various indicators of family composition, such as number of members, average age of members and number of cohabiting generations, can be constructed, added to SHARELIFE in its retrospective panel format and used as a determinant of outcomes at older ages, as explained in Section 2.6.1. Second, the prevalence of "boomerang children" (Mitchell & Gee, 1996), i.e., adult children returning to the parental home, is lower than one might expect. This evidence deserves further investigation, for example, to answer the question of whether there are differences across cohorts and/or countries. Finally, the evidence for the direction of assistance among generations is consistent with the notion of a "sandwich generation" that provides help to members of both elderly and younger generations (Miller, 1981). However, it may also point to a "warm glow" response bias, as noticed in the literature on inter vivos gifts (Alessie et al., 2014).

References

Alessie, R., Angelini, V., & Pasini, G. (2014). Is it true love? Altruism versus exchange in time and money transfers. De Economist, 162(2), 193-213. http://doi.org/10.1007/s10645-013-9221-5.

Miller, D. A. (1981). The 'sandwich' generation: Adult children of the aging. Social Work, 26(5), 419-423. https://www.jstor.org/stable/23712207.

Mitchell, B. A., & Gee, E. M. (1996). "Boomerang kids" and midlife parental marital satisfaction. Family Relations, 45(4), 442-448. https://doi.org/10.2307/585174.

2.7 Palliative care

Hendrik Jürges – University of Wuppertal

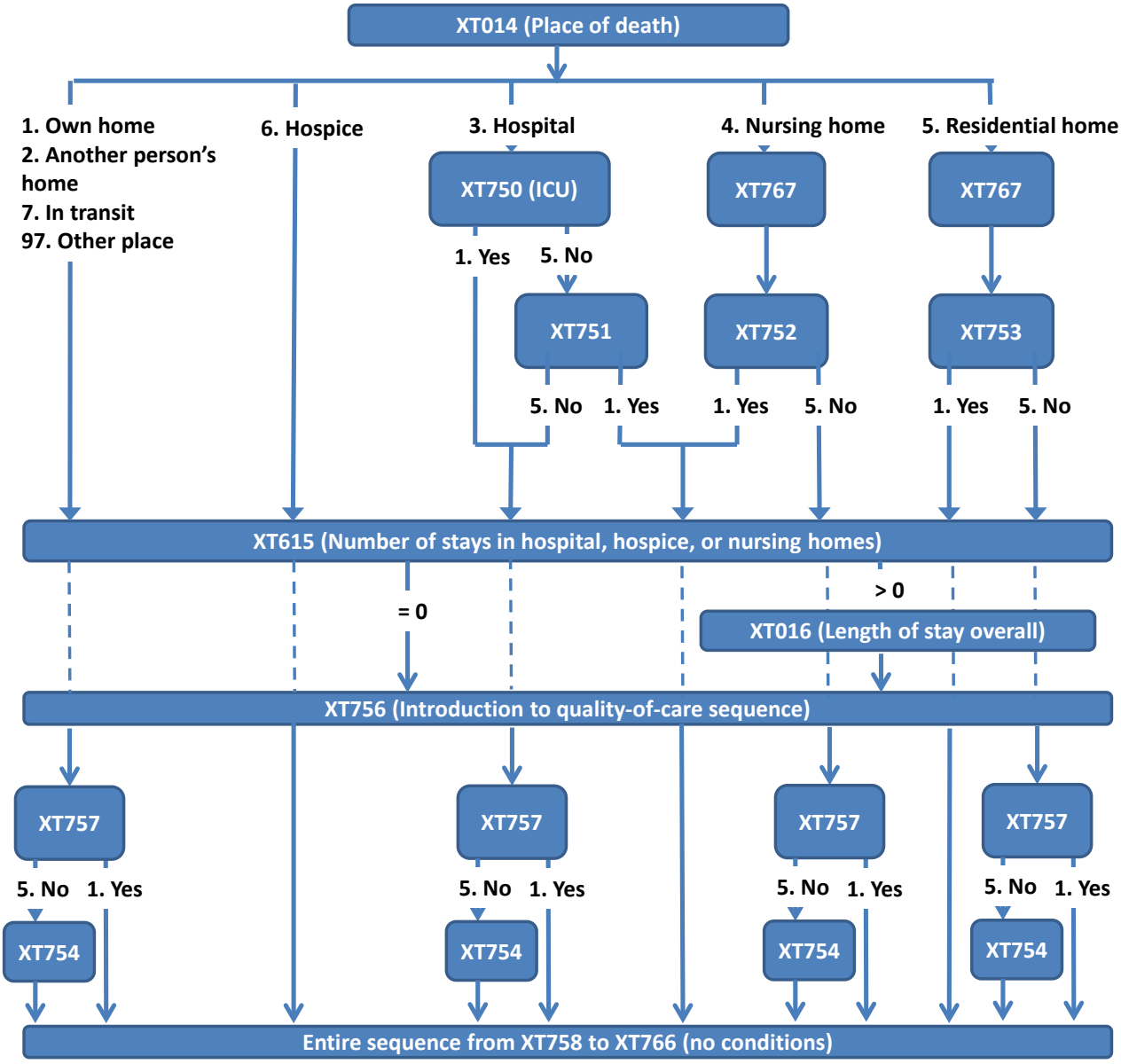
2.7.1 Introduction

For many people, the last year of life is associated with pain and anxiety. Caregivers, especially family members, can also bear a considerable burden. For both the terminally ill and their caregivers, palliative care aims to improve quality of life "through the prevention and relief of suffering by means of early identification and impeccable assessment and treatment of pain and other problems, physical, psychosocial and spiritual" (WHO Definition of Palliative care). The SHARE end-of-life questionnaire in Wave 7 includes, for the first time, a series of questions on places and quality of end-of-life care as well as satisfaction with it. These questions have two aims: first, to measure how many respondents have been cared for in institutions that are specialised in palliative care or have had palliative care at their own home; second, to measure the experience of end-of-life care among those with and without palliative care and to identify possible unmet need. Under the assumption that the need for palliative care is similar in each SHARE country, international comparison based on the availability of palliative care will aid our understanding of the benefits of palliative care in institutionalised and non-institutionalised settings.

2.7.2 Questionnaire development

In selecting suitable questions for the SHARE end-of-life questionnaire, we were looking for question formulations that can be asked of caregivers (proxy respondents) both face-to-face and over the telephone. Questions should also be appropriate to be asked of everyone independent of place of death (hospital, palliative care unit or hospice, at home, etc.). Due to limitations in survey length, we needed to choose broad and salient dimensions of end-of-life care. Moreover, we decided to focus our questions on the deceased, i.e., our primary SHARE respondent. Thus, we were leaving out almost all the concerns the family may have had. Finally, the new set of questions should blend seamlessly into the existing end-of-life questionnaire.

The first part of questionnaire development was concerned with the factual assessment of whether any form of palliative care was received by the deceased SHARE respondent. A key challenge here was to ascertain that heterogeneous types of palliative care were covered in a single set of questions for all countries avoiding country-specific routing. A flowchart is shown in Figure 2.8. Starting from the original question on place of death, we asked for each type of institution if the care at the place of death was given by a specialised palliative care or hospice unit (XT751 to XT753). For those who did not die at a location with such specialised care, e.g., at home, in an intensive care unit (ICU), or in a nursing home but not in a palliative care unit, we asked whether they had any type of palliative care in the last four weeks of their lives (XT757). If they did not, we asked why among the following reasons: because it was not needed or wanted, because it was needed or wanted but not available, or because it was needed or wanted but too expensive. This question was included to measure unmet need.



Note:
XT751: Was that in a palliative care or inpatient hospice unit?
XT752: Was that an inpatient hospice unit?
XT753: Was the residential housing provided by hospice?
XT757: In the last four weeks of his/her life, did [name of the deceased] have any hospice or palliative care?
XT754: What was the reason that he/she did not have hospice or palliative care?

Figure 2.8: Flowchart of palliative care questions

The second part of the questionnaire development was concerned with assessing, independently of the end-of-life care setting, the quality of care in the last month of life. A review of existing multidimensional palliative care assessment tools (available in English) was conducted. Our shortlist included the questionnaire of the "After-Death Bereaved Family Member Interview"[3], the "Quality of Dying and Death Questionnaire for Family Members" (QODD)[4], and the "Family Evaluation of Palliative Care" (FEPC)[5]. Eventually, we chose a subset of questions from the FEPC, which was until recently used by the US National Hospice and Palliative Care Organization (NHPCO)[6] for continuous quality assessment of palliative care. Table 2.16 shows the set of questions chosen for the SHARE end-of-life interview together with the response options. Items cover the following dimensions: pain, trouble breathing, anxiety and sadness, personal care needs, and being treated with respect, as well as an overall assessment of quality of care.

3 http://www.chcr.brown.edu/pcoc/linkstoinstrumhtm.htm.
4 http://depts.washington.edu/eolcare/products/instruments/
5 https://www.nhpco.org/performance-measures/family-evaluation-palliative-care-fepc.
6 www.nhpco.org

Table 2.16: Overview of the new questions on quality of life and palliative care

Question code	Routing	Question text	Response options
XT756	All	The next couple of questions are about the care [Name of the deceased] received in the last month of his/her life. Please answer these questions based on your experience and the deceased's experience while he/she was receiving care. Some of the questions ask about the staff. By staff, we mean doctors, nurses, social workers, chaplains, nursing assistants, therapists, and other personnel.	1. Continue
XT758	All	In his/her last month of life, did [Name of the deceased] have pain or take medicine for pain?	1. Yes 5. No
XT759	If XT758 == 1	Did the deceased receive too much, too little, or just the right amount of medication for his/her pain?	1. Too much 2. Too little 3. Right amount
XT760	All	In his/her last month of life, did [Name of the deceased] have trouble breathing?	1. Yes 5. No
XT761	If XT760 == 1	How much help in dealing with his/her breathing did the deceased receive - too little or just the right amount?	1. Too little 2. Right amount
XT762	All	In his/her last month of life, did [Name of the deceased] have any feelings of anxiety or sadness??	1. Yes 5. No
XT763	If XT762 == 1	How much help in dealing with these feelings did the deceased receive - too little or just the right amount?	1. Too little 2. Right amount
XT764	All	How often were the deceased's personal care needs - such as bathing, dressing, and changing bedding - taken care of as well as they should have been? (READ OUT)	1. Always 2. Usually 3. Sometimes 4. Never 5. Help was not needed or wanted for personal care
XT765	All	During his/her last month of life, how often overall was the staff who took care of him/her kind, caring, and respectful?	1. Always 2. Usually 3. Sometimes 4. Never

Question code	Routing	Question text	Response options
XT766	All	Overall, how would you rate the care the deceased received in his/her last month of life?	1. Excellent
			2. Very good
			3. Good
			4. Fair
			5. Poor

2.7.3 Preliminary data analysis

The palliative care questions were answered by 3,342 proxy respondents (SHARE Wave 7 data, Release 0) from 19 countries that participated in SHARE before Wave 7: Austria, Germany, Sweden, Spain, Italy, France, Denmark, Greece, Switzerland, Belgium, Israel, Czech Republic, Poland, Luxembourg, Hungary, Portugal, Slovenia, Estonia, and Croatia. Among those proxy respondents, 40 percent were spouses of the deceased, 26 percent were children of the deceased, 14 percent were other relatives (e.g., children-in-law or grandchildren), and the remainder of 20 percent were non-relatives. Two-thirds of proxy respondents had daily contact with the deceased in the last year of his/her life. Only 11 percent of proxy respondents had less than weekly contact with the deceased.

As detailed in Figure 2.8, we ascertain whether the deceased died in a facility that provided palliative care (13 percent) and if not, whether the deceased had any form of palliative care in the last four weeks of their lives (24 percent). Thus, proxy respondents reported some palliative care for more than one-third of the respondents. Of those who neither received palliative care at their place of death or other setting, only 5 percent did not receive such care because it was not available or not needed, according to the proxy respondents' assessment. This low number raises concerns regarding whether the need for palliative care was correctly assessed; this topic will be subject to further scrutiny.

Table 2.17: Quality of end-of-life care by type of care; raw percentages (N = 3,342)

	No palliative care (%)	Any type of palliative care (%)
Pain or took medicine for pain: Yes	49	75
Right amount of medication for pain	87	87
Too much	6	8
Too little	7	6
Deceased had trouble breathing: Yes	42	53
Right amount of help	87	91
Too little help with breathing	13	9
Deceased had any feelings of anxiety or sadness: Yes	44	59
Right amount of help	80	82
Too little help	20	18
Personal care needs taken care of as well as they should have been*		
Always	72	77
Usually	16	18
Sometimes	6	4
Never	5	1

	No palliative care (%)	Any type of palliative care (%)
Staff was caring and respectful		
Always	72	75
Usually	17	20
Sometimes	4	4
Never	7	2
Overall rating of end-of-life care		
Excellent	38	34
Very good	33	37
Good	21	22
Fair	6	5
Poor	3	2
N	2128	1214

Note: SHARE Wave 7 data (Release 0), unweighted.
* Excluding cases in which care was "not wanted or needed".

Table 2.17 shows the raw percentages for quality of care in the last four weeks of life. No attempt has been made to correct for selectivity, e.g., differential availability of proxy respondents across countries or types of care. The overall conclusion from Table 2.17 is that patients who had received palliative care more often suffered from pain, trouble breathing, and anxiety or sadness. This finding likely reflects selection into palliative care, especially with respect to pain. Conditional on suffering from any of those problems, however, quality of care seems to be only slightly better with specialised palliative care. Additionally, with respect to personal care needs, treatment by staff or the overall assessment of end-of-life care, proxy respondents reported no major differences between palliative and non-palliative care.

2.7.4 Concluding remarks

The SHARE end-of-life questionnaire as administered in Wave 7 allows, for the first time, for a comparative assessment of end-of-life care across countries or within countries between different socio-economic groups. Before any definitive conclusions can be drawn from the data, however, careful evaluation of answer patterns is warranted. One important limitation of our data is that they allow between-subject comparisons only. For instance, the need for palliative care may only be realised once such care is available or when subjective assessments may be influenced by expectations. If relatives are cared for in palliative care units, one may expect higher standards of care than those associated with care at home by relatives or professional staff. These are exciting new topics of research into the quality of life at the end of life and its measurement in social surveys.

CHAPTER 3

Software innovations

03

3 SOFTWARE INNOVATIONS

Maurice Martens – CentERdata; Tilburg University, Iggy van der Wielen – CentERdata

Harmonised software tools are essential for the management of fieldwork for a project such as SHARE. Although most of the software can be reused in the various waves of SHARE, every wave has new challenges that call for changes and updates to the software. In Wave 7, it was decided that part of the sample would receive a questionnaire based on the Wave 3 questionnaire, which was a history calendar. We developed this history calendar in 2006 using Visual Basic 6 (VB6) and the Blaise Application Programming Interface (API) (see Das et al., 2010). The Wave 7 respondents, who already completed the history calendar in Wave 3, should receive the regular (or rather, updated) Wave 6 questionnaire.

Since the questionnaire was used satisfactorily in Wave 3, we initially hoped we could reuse it. However, we were not sure to what extent this setup would still run on newer versions of Windows and how easy it would be to introduce right-to-left support for Hebrew and Arabic versions. The integration of other tools we have developed since then could cause problems as well, for example, support for the display of movies and support for look-up tables, which were used to classify countries and occupations during the interview. To examine the feasibility of reusing this version of the VB6 software, we started further developing this software and tested it in a pretest in 25 of the SHARE countries. To a certain extent, this approach actually seemed to work. Although the tool was very slow on some laptops and it was not easy to find the correct libraries to install, it operated properly, and we received data from all countries. However, the burden on support and the risk of unexpected unsolvable issues were considerable. If problems occurred during fieldwork, we were not able to guarantee a time frame in which these problems could be solved. Furthermore, it was reported that the tool was very much outdated. After another review, we decided that it was no longer feasible to support the tool. Although the concept worked in general (building a layer around the survey software for displaying the calendar), VB6 was the wrong software for doing so.

This decision left us with a challenge: how should we display this calendar if we could not use the old tool anymore? Could we think of a way to do it in the 4.8 Blaise Data Entry Program (DEP)? Would Blaise 5 present a solution? Should we develop something like the VB6 tool in another Integrated Development Environment (IDE)? Should we perhaps simplify the functionality of the calendar to displaying images? We had some experience with an online version of the history calendar (see Martens, 2013), but the backend was programmed in PHP and used BlaiseIS. If we were to pursue this approach, we would need to install a server on every laptop, which seemed irrational. Should we perhaps leave Blaise altogether, develop a tool ourselves, or look for another software vendor that might support the features we need? It was clearly time to make an overview and properly discuss what path we should take to solve the challenge within the given time (see Table 3.1). In the following table, we present our thoughts and findings.

Table 3.1: Review of potential solutions

Solution	Advantage	Disadvantage
Reuse VB6 build on Blaise 4.8 API	• Proven • Works with current questionnaire	• Outdated • Limited recent experience • Challenging to support • No future solution
Blaise 4.8 DEP	• Stable	• Cannot develop a rich interface
Blaise 5 DEP	• New Blaise features seem to fit the regular questionnaire very well	• Cannot develop a rich interface • Limited experience • Complete redesign of all tools involved

Solution	Advantage	Disadvantage
Build something on Blaise 4 API	• Experience	• What IDE?
Build something on Blaise 5 API	• Future proof • Build up experience	• Complete redesign of all tools involved • What IDE?
Use the earlier developed web solution	• Proven • Easily adapted to work with current questionnaire	• Questionnaire in browser • Install Blaise IS on every laptop
Other vendor	• Could try to find something that fully supports all our requests	• No experience • Complete redesign of all tools involved • Cannot reuse already-developed code
Develop own survey system	• Full control	• Cannot reuse already-developed code • Would take very long

Since we only had a few months to find a solution, the paths that would probably cost too much time were ignored. For both Blaise native solutions, we decided that it would not be possible to create the tools we needed. We decided to either develop a tool in addition to Blaise 4.8 API again or find a way to enable the web tool to operate on laptops.

The code of the web tool we developed earlier was an HTML frontend, with a style sheet for the layout, which was obtained from a Blaise IS questionnaire. The calendar was generated by a PHP script, driven by JavaScript calls defined in the Blaise question texts. We wanted to avoid installing server software on laptops, so other software would have to take over the role of PHP. Ideally, the questionnaire would not be displayed in a browser. Deactivating the browsers' standard menus, shortcuts, and other behaviour would not be possible. Since our Sample Management System (SMS) already launched the DEP to start an interview, a logical step was to use a standard browser class to launch the browser window from within the SMS, creating our own browser. This standard browser enables full control; it would not require too many structural changes, and the overall architecture remained largely intact.

Once this idea was born, everything fell in place. The Java environment using a library called com4j was needed to call the Blaise API, much like our VB6 solution had done before. The browser window could display the HTML solution. This feature allows for the (re)use of (web-) techniques such as cascading style sheets, jQuery, JSON objects, and HTML5 in a CAPI (computer-assisted personal interviewing) environment. It provides the freedom to develop an interface and questionnaire behaviour according to one's own preferences. In addition, with this web-compatible DEP, we could reuse many of the tools we already developed for our websites. In our design, we would need to show only one web page, and all submitted actions performed by this website would be detected by SMS and translated into Blaise actions. The SMS will call the Blaise API and transfer the information that needed to be displayed on the website. A set of high-level operations was defined as follows:

- NextQuestion
- PreviousQuestion
- GotoLastQuestion
- GotoFirstQuestion
- GotoQuestion

When one of these actions needed to be performed by the webpage, a JSON structure would be formulated as follows:

```
PreviousQuestion
    Key1
    Key2
    Answer
    Status
    RemarkText
    Suppress
```

This minimum information is transferred from the webpage to the browser and tunneled through to the Blaise API, ultimately generating a new interview state. This new state is detected by the Java environment, and a JSON object is returned to the website:

BlaiseObj
- Key1
- Key2
- Name
- Text
- questionText
- Remarked
- RemarkText
- Interface
- ErrorMessage
- HardErrorMessage
- HardErrorInvolved
- SoftErrorMessage
- SoftErrorInvolved
- DontKnowAllowed
- RefusalAllowed
- Required
- Value
- Status
- FieldDef
  - MinValue
  - MaxValue
  - DataType
  - Categories
  - IsSet
  - TextAsSetString

This set allows us to display the questionnaire and interact with the questionnaire. JavaScript reads out the new values and adapts the webpage accordingly (see Figure 3.1).

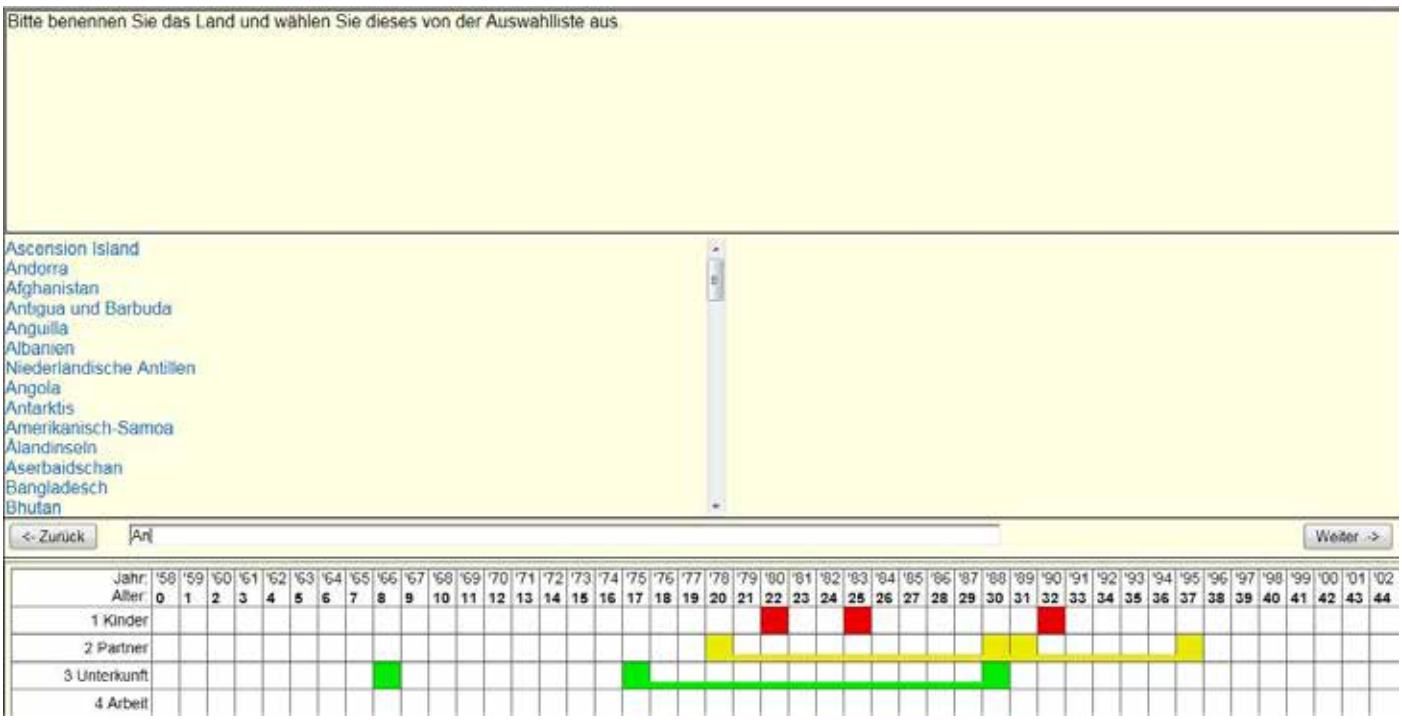


Figure 3.1: SHARE Wave 7 instrument

After the basic properties of the question were defined, in the object named BlaiseObj, the object was further extended to support some extra sub-objects:

- BlaiseObj['Movie']; a list of words with their duration is defined in the Blaise questionnaire. Using the setTimeout function, we can trigger when the words are shown. Before, we used a movie to show this list, which we needed to generate for all translations by hand.
- BlaiseObj['Coder']; several lookup tables were defined as JavaScript arrays, countries, job titles, languages, currencies. Searching in these tables can now be performed in JavaScript, giving us full control of the behaviour of these tools.
- BlaiseObj['Calendar']; the information that needs to be displayed in a calendar is translated into an HTML structure that displays the calendar.

The HTML page where we load the questionnaire is very basic; it is defined in a set of div-tags, as seen in the code below.

```
<div id="completedep">
	<div id="questiontext-outer">
		<div id="questiontext">
		</div>
	</div>
	<div id="moviediv"></div>
	<div id="errordiv"></div>
	<div id="helptext"></div>
	<div id="answeroptions" class="answeroptions"></div>
	<div id="navigation" class="navigation">
		<div class="navigation-inner">
			<input name="previousbtn" value="&#8592; Back" />
			<div style="display: inline-block;">
				<div id="remarked" class="remarked"> </div>
				<div id="answerdiv" class="answerdiv">
					<input type="text" name="answerfield" />
				</div>
			</div>
			<input name="nextbtn" value="Next &#8594;" />
		</div>
	</div>
	<div id="CalendarDiv"></div>
	<div id="remark">
		<span id="remarklbl"></span>
		<textarea id="remarkText"></textarea><br>
		<input type="button" name="savebtn" />
		<input type="button" name="cancelbtn" />
	</div>
	<div id="console"></div>
</div>
```

When we finally managed to complete this setup, two problems were detected. One problem was speed. A long period of time passed, before we got a response from the API. When tracking it down, we found that this issue was due to the moment we stored the responses. To solve this issue, we changed the order and now save the answers after the response is returned to the browser, that is, when the interviewer is already looking at the next question.

A larger problem was the use of open answers. We wanted the instrument to be completely UTF-8 encoded. However, Blaise 4.8 stores strings in ANSI internally, but depending on the Windows installation, it will use different code pages for this instrument. When we want to use open-ended responses to complete a subsequent question, the open answers that were coded with this different code page would appear to be nonsensical. To solve this problem, we needed to recode the open-answer strings to something that could be formulated in the first 128 characters of the code pages because these strings are always identical for every code page. Since we now presented the questions in a browser, a logical candidate for this would be to encode special characters

into htmlchars (e.g., Ӓ). However, 7 characters are needed to encode each non-Latin character. Since we sometimes need to integrate these open answers into fills, we now encounter that these fills can become too long. Since these composed fills break off when they exceed a length of 256, we could still encounter nonsensical texts when open answers are too long. To avoid this, we adapted the Blaise source questionnaire and created extra fills for long answers, so that we could divide them over multiple fills. However, this process was not fully automated. As soon as we encountered difficulty, we had to create new fills manually.

A key change in Wave 7 concerned the development of a single multilingual instrument for countries with more than one language. In Wave 7, interviewers were asked to choose which language should be used, once before the coverscreen and again before the CAPI (see Figure 3.2). This procedure was preferred over having CAPIs in different languages, such that the interviewer had to know beforehand which language would be spoken in the household.

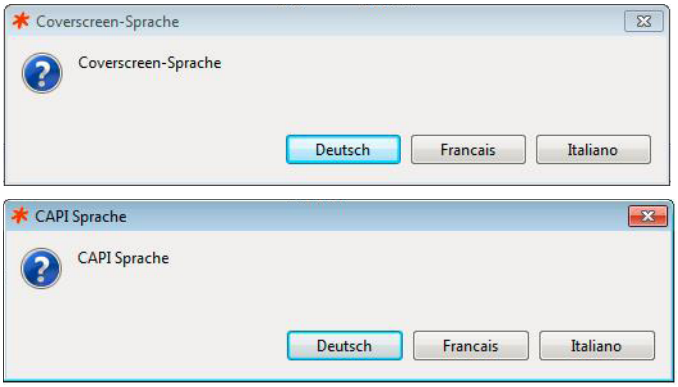


Figure 3.2: SHARE Wave7 instrument for countries with more than one language

The implications of this process were as follows:

- No language switching occurred within the interview
- The (consent to preload) pop-up window between the coverscreen and CAPI needed to show all languages

In Wave 7, we once again found a creative solution to allow the SHARE questionnaire to run. We did not encounter major problems, although perhaps some speed issues were encountered in certain countries; however, overall, the wave was a success. Nevertheless, we do think that the underlying Blaise software is completely outdated; we find ourselves taking too much time to find creative solutions for issues that should be supported natively, such as Unicode. Furthermore, the code for SMS and Sample Distributor (SD) is also nearly out of date. For future waves, we urge the investigation of a more sustainable solution for SHARE software to avoid risks, to lower the burden on support and to improve fieldwork.

References

Das, M., Martens, M., Wijnant, A. (2010). Survey instruments in SHARELIFE. In M. Schröder (Ed.), Retrospective Data Collection in the Survey of Health, Ageing and Retirement in Europe. SHARELIFE Methodology (pp. 20-28). Mannheim: Mannheim Research Institute for the Economics of Aging.

Martens, M. (2013). Jumping around in Blaise IS. Proceedings of the 15th International Blaise Users Conference, Washington, DC (pp. 39-48). Retrieved from http://www.blaiseusers.org/2013/papers/2b.pdf.

CHAPTER 4

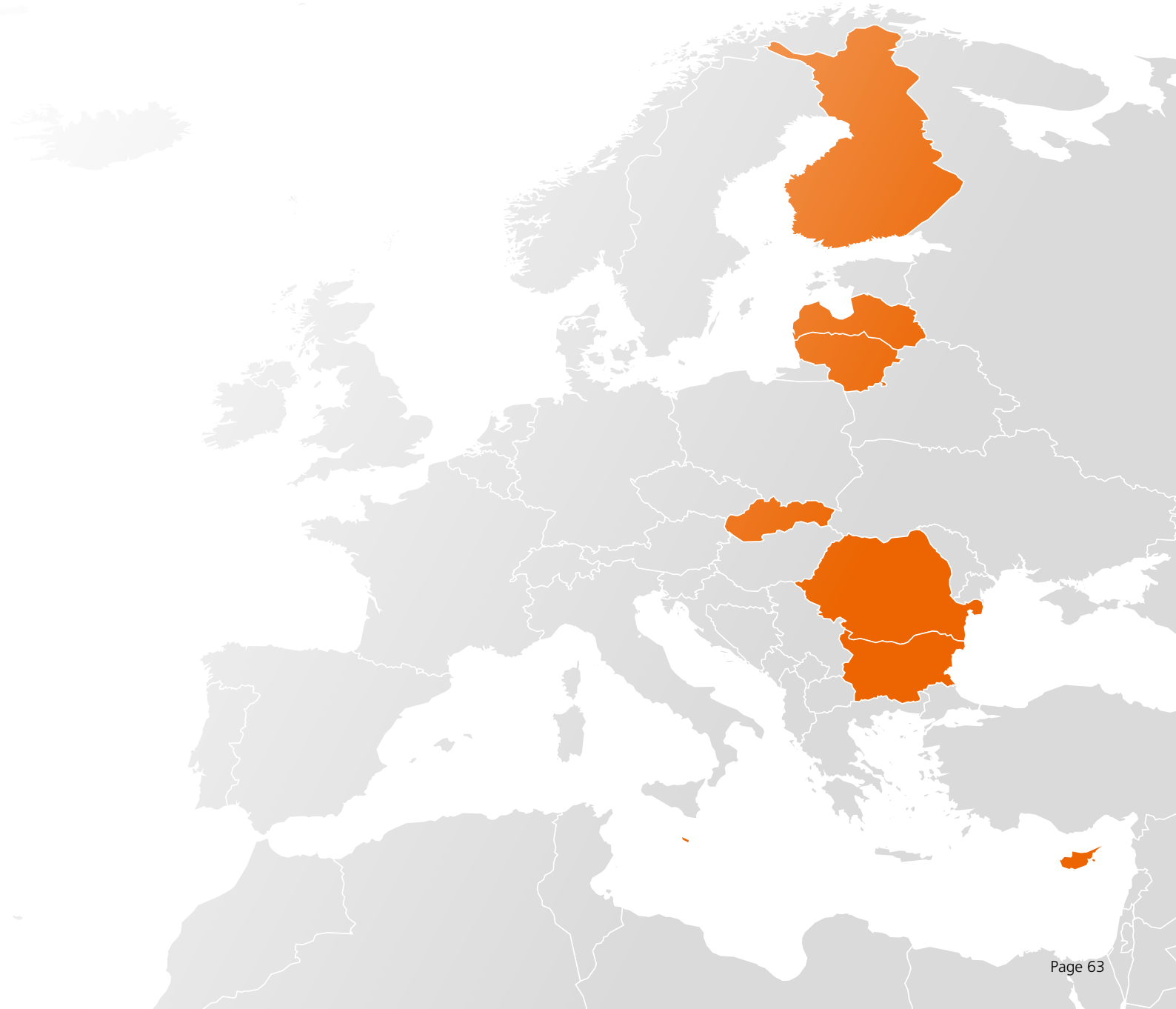
Becoming a new SHARE country

04

4 BECOMING A NEW SHARE COUNTRY

This chapter reflects the experience of eight countries that joined SHARE as new countries in Wave 7. We asked the authors to give an account of what motivated them to become part of SHARE, to describe the obstacles and challenges they encountered and how they were overcome, and to provide a brief overview of future directions.[7] The countries appear in alphabetical order.

7 We greatly appreciate the help received from Yuri Pettinicchi in collecting the countries' responses.



4.1 Bulgaria

Ekaterina Markova, Gabriela Yordanova – Institute for the Study of the Societies and Knowledge at the Bulgarian Academy of Sciences (ISSK-BAS)

4.1.1 Introduction

Bulgaria has one of the most imbalanced demographic compositions of countries throughout the world (United Nations, 2017), and it is expected to lose more than 15 percent of its population by 2050. Except for Eurostat's provided data, important targeted statistics for ageing Bulgarians are missing in various EU reports, such as the latest report entitled "Challenges in long-term care in Europe" (Spasova et al., 2018). Bulgaria is not part of many comparative surveys, so its participation in SHARE is a major advantage, filling such a data gap and supporting researchers and policy makers. The Bulgarian team from the ISSK-BAS is honoured to be part of SHARE and contribute to its expertise in survey methodology and policy evaluation.

SHARE provides unique data about physical and mental health status, quality of life, labour and retirement, and social networks for ageing Bulgarians. SHARE will be used to convince policy makers to adjust the current demographic strategy, health system and long-term care; national and comparative analyses will also be performed from SHARE life histories of Wave 7.

4.1.2 Funding and assembling a national working group

Bulgaria is joining SHARE Wave 7 thanks to the financial support of the European Commission with its grant VS/2016/0135. SHARE-Bulgaria, coordinated by ISSK-BAS, is part of the National Roadmap for Research Infrastructures (2017-2023). For 2018, SHARE-ERIC Bulgaria received partial state funding for creating the national SHARE centre and university courses on SHARE. The roadmap regulations require the avoidance of double funding of project activities, so the National Roadmap cannot fund SHARE fieldwork and coordination.

Dr. Ekaterina Markova (Bulgarian SHARE Country Team Leader, CTL) has been a senior research fellow at ISSK-BAS since 2004, specialising in survey research methodology, policy impact measurement and evaluation, and demographic change. Dr. Gabriela Yordanova (Country Team Operator, CTO) has combined applied survey research practice, university teaching and scientific research at ISSK-BAS since 2006. Currently, she is a senior research associate at ISSK-BAS, specialising in survey methodology and labour sociology. The Bulgarian SHARE team is supported by Dr. Vassil Kirov (expert), Associate Professor at ISSK-BAS and Associate Researcher at the Centre Pierre Naville, University of Evry and at the European Trade Union Institute (ETUI), with long-term experience in demographic change, labour sociology and policy evaluation. Diana Nenkova, a highly experienced sociologist, supports the team as an administrative assistant.

In general, the ISSK-BAS, as well as the entire BAS, is funded mainly by international organisations. However, the ISSK-BAS is facing permanent issues with EU funding that requires advanced personal payment and in-kind contributions. To provide its own funding in advance, the research team was forced to ask for a loan at BAS on difficult conditions, taking not institutional but rather personal responsibility. The ISSK-BAS team is maintaining its deadlines, contributing the highest possible quality of work, and working under pressure and on a tight schedule. The ISSK-BAS team did not face any operational obstacles in coordinating Wave 7 of SHARE. However, the efforts of the team members remained underestimated by EU funding.

4.1.3 Survey implementation

Translating and testing the SHARE questionnaire of Wave 7 was the greatest challenge for the Bulgarian team. The workload was doubled and included checking previous translations made by a translation company, including missing labels, and correcting "copy/paste" issues and meaningless automatic translation. The translating management tool used also caused problems; corrected labels were not saved in the system, so we repeatedly entered the same information. Despite the bugs in the system and the translation issues, we were supported by SHARE Central, with which we communicated truly actively and solved all the problems.

The Bulgarian SHARE team coordinated the work with the survey agency without any difficulties, both in person and electronically. The previous experience of SHARE team members in survey research facilitated this coordination; in fact, Bulgaria was among the countries that first completed SHARE Wave 7 data collection. Our experience from SHARE shows that continuous monitoring of all stages of the survey implementation, with special attention to the interviewers' network, is of great importance. Communication with

SHARE Central and the survey agency is also crucial because the lack of information about any element in the preparation and execution of the fieldwork could lead to serious negative effects.

4.1.4 Summary

The Bulgarian SHARE team organised several meetings with stakeholders and media interviews to raise awareness of the possible use of SHARE data for research and policies. In the future, we will continue such activities, also making efforts to establish a national SHARE centre with the support of state funding within the National Roadmap for Research Infrastructures. In this respect, it was an important step that Bulgaria joined SHARE-ERIC. Furthermore, the university courses on SHARE that are planned to be developed will also increase the usage of SHARE data among wider audience. The challenge for SHARE in general is mainly the way it will be funded in the future. The Bulgarian SHARE team will continue to search for national funding opportunities. Completing a team with an experienced survey methodologist with wide practical knowledge, in conjunction with finding the best survey agency, could ensure successful work on SHARE. In addition, one must put significant efforts into planning the workload, following deadlines and avoiding compromises with data quality.

References

United Nations, Department of Economic and Social Affairs, Population Division (2017). World Population Prospects: The 2017 Revision, Volume I: Comprehensive Tables. Retrieved from https://population.un.org/wpp/Publications/Files/WPP2017_Volume-I_Comprehensive-Tables.pdf.

Spasova, S., Baeten, R., Coster, S., Ghailani, D., Peña-Casas, R., & Vanhercke, B. (2018). Challenges in long-term care in Europe: A study of national policies. Brussels: European Commission. https://doi.org/10.2767/84573.

4.2 Cyprus

Nikos Theodoropoulos, Alexandros Polycarpou – University of Cyprus

4.2.1 Introduction

Among the challenges Cyprus will face in the next decades is population ageing due to the low fertility rate and a large increase in life expectancy. According to Eurostat, the fertility rate was 1.33 children per female in 2017 and will be 1.56 children per female in 2060. The share of individuals aged 65 and above is, however, expected to increase from 15.5 percent in 2017 to 31.5 percent by 2060. Furthermore, the old dependency rate is expected to increase from 22.5 percent to 55.2 percent by 2060. These changes will have significant impacts on the sustainability of the pension system and on health care. Actions should be taken to offset the adverse effects of population ageing, and SHARE data can be utilised to provide evidence-based policy reforms. SHARE data will help Cyprus to gather data, monitor the process of population ageing, and study how economic, health and social factors shape the living conditions of older people. With respect to research, the SHARE data offer a unique opportunity to researchers to study many interesting socio-economic phenomena. Learning about the data from "within" helps to clarify their structure and ways to approach one's research question. Another motivation for joining SHARE is participation in a wider research network where many positive spillover effects are generated from research presentations and networking with top academics in their respective fields.

4.2.2 Funding and assembling a national working group

Funding for Wave 7 in Cyprus was provided by DG Employment grant VS/2016/0135, which extended the coverage of the SHARE survey to all member states. Co-funding was provided in-kind by the University of Cyprus. To receive co-funding for future waves from national sources as well as to become a SHARE-ERIC member, we initially approached interested ministries and gave a number of presentations on the usefulness of SHARE data for policy making and for academic research.

Forming the Cyprus SHARE country team was accomplished easily. The team is comprised of Nikos Theodoropoulos, Assistant Professor at the Department of Economics, University of Cyprus, who is the CTL, and Alexandros Polycarpou, researcher at the Economics Research Centre, University of Cyprus, who is the CTO.

4.2.3 Survey implementation

SHARE Wave 7 was the first wave in Cyprus, and the SHARELIFE questionnaire was implemented. The translation process of the questionnaire ran smoothly, since a first version of the translated questionnaire was available from the Greek national team. Significant changes had to be made, however, due to the different institutional characteristics between Greece and Cyprus (i.e., differences in the pension and social security systems). Expert knowledge of the "standard" terminology was needed for some questions. In those instances, we asked for help from colleagues from other academic departments. Testing and reviewing the translation was also a demanding process, especially after the field rehearsal, where interviewers and survey respondents made comments and remarks.

Coordination with the survey agency was achieved with regular phone calls, various meetings and in national training sessions, where we reviewed the process of fieldwork and discussed problems faced during the interview. In general, we had a very good collaboration with the survey agency, and all problems we encountered were adequately addressed. The main challenge we faced was that we could not have access to the restricted micro-level census data. Although we exerted great efforts to access it, our efforts were not fruitful. Thus, other sources of information had to be used for the sample design. Fortunately, the survey agency has a complete and updated telephone registry for Cyprus. To this end, the telephone registry was used for drawing the sample.

4.2.4 Summary

SHARE data are a great source of information. Both primary and secondary analysis of SHARE data can provide policy makers with facts on which to base their decisions. The Cyprus team is planning to use the SHARE data to study the factors that affect the retirement decisions of individuals and to determine what drives them to early retirement. In addition, we will use the data to study the demographic changes in Cyprus due to population ageing, low fertility and migration and what consequences they may have on the strength of the health and social systems. The SHARE data can also be used to answer whether the newly adopted National Health Insurance System of Cyprus is well designed to face the challenges of population ageing.

The first SHARE wave in Cyprus has taught us important lessons. SHARE is a well-managed survey; however, due to its nature, the workflow is not equally divided in time, and time constraints can sometimes be tight. Thus, team members should be well organised with flexible agendas and be able to work under pressure to fulfil the requirements of the project. Further, since funding for future waves is not secured, it is important to initiate the procedure for obtaining funding early by preparing research proposals and giving presentations to government organisations and to other interested (and potentially funding) parties about the existence and the strength of the SHARE data in shaping policy and its potential impact in the research community. In this respect, it was an important step that Cyprus joined SHARE-ERIC in the meantime.

4.3 Finland

Anna Rotkirch, Miika Mäki – Population Research Institute, Väestöliitto

4.3.1 Introduction

The Family Federation of Finland is a large family welfare organisation with its own demographic research institute, which conducts research and provides policy briefings on family relations and population change. This Population Research Institute (PRI) represents leading Finnish expertise in the demographic and sociological study of grandparenting and parenting, with extensive expertise in collecting and using high-quality and representative survey data in Finland, as well as in combining survey and register data.

Finnish scholars interested in intergenerational exchange had lamented the absence of Finland from SHARE for several years. In 2007, we launched a Finnish longitudinal survey on generational transmissions, which included parts of the central SHARE modules. In 2016, the PRI was contacted by the SHARE Central team in Munich and was told that the EU, through DG Employment, had provided financing for Finland for Wave 7 and asked whether we would be interested in acting as the scientific partner institution of SHARE in Finland. The PRI, being acquainted with SHARE and having used SHARE data in our own research on grandparenting, was highly motivated to do so. Eija Koivuranta, CEO of The Finnish Family Federation, to which PRI belongs, was also enthusiastic about this possibility to contribute to a truly European project with real relevance for ageing people and their family members.

The Finnish team will use the SHARE data in its four-year research project LoveAge. LoveAge investigates couple and social relations among older adults (aged 50+) in Finland and the European Union. We apply a life course and life history perspective and study the elderly as both receivers and providers of different kinds of support. We are interested in how partnership dynamics affect family relations and social networks.

4.3.2 Funding and assembling a national working group

In Wave 7, the European Commission DG Employment funded a minimum sample for Finland (VS/2016/0135). The PRI contributed own funding and in-kind expenses for country team salaries and media outreach. The original Finnish research team included Anna Rotkirch (research director, CTL), Anneli Miettinen (researcher, CTO) and Tiina Helamaa (information specialist). We also engaged journalist Barita Rosenström to work with public relations (PR) and journalist Baba Lybeck to act as the figurehead for SHARE Finland. The main problem for the PRI was the relatively low contribution to the country team, which meant the Family Federation had to support SHARE Finland extensively. Since the scientific team had extensive experiences with survey questionnaire design and register data use, we were qualified for the tasks. However, we joined SHARE quickly and well into preparations for Wave 7. The number of specific work tasks required was also larger than expected, which, in combination with the timetable, created some pressures.

The current Finnish research team includes the following: Miika Mäki (coordinator and researcher, CTO); Tiina Helamaa (information specialist); post-doc researchers and register specialists Mirkka Danielsbacka, Antti Tanskanen, and Andreas Weiland; and Milja von Lerber and Nina Östman (research assistants).

From the beginning, we wanted to ensure maximal awareness and use of Finland in SHARE Wave 7, as well as in the upcoming waves. In 2017, we appointed a SHARE Finland steering group, consisting of 16 members representing experts from ministries, universities and pension institutions. The task of the steering group is to follow and guide the Finnish country team. The academics of the steering group use SHARE data as part of their teaching, while the experts in ministries and pension institutions will stress the significance of the research infrastructure within their fields and contribute to fundraising efforts. The current team consists of 18 members and is listed here.[8] The steering group meets yearly, and group members provide advice and support outside meetings.

4.3.3 Survey implementation

The national CAPI (computer-assisted personal interviewing) was handled by the CTO with assistance from the CTL and with the help of two research assistants: one checked the Finnish translations and another checked the translations into Swedish, Finland's second national language. This was a

8 https://www.vaestoliitto.fi/in_english/population_research_institute/?x5184217=7402482.

labour-intensive stage, but we received very good guidance from SHARE Central in Munich and help both during the meetings and through e-mails.

Taloustutkimus was selected as the Finnish survey agency. The main challenge in selecting the agency and agreeing on the contract was related to estimating the interview costs. Finland is a sparsely populated country with comparatively high wages, which means that personal interviews are very expensive. Furthermore, the special nature of SHARE data collection means that the survey agency had to budget costs for training and travels, which did not resemble their previous experiences and were difficult to estimate. Ultimately, all things were agreed upon, and the data collection went very well. The fieldwork was successful, and we did not have any major issues, although the length of the interviews turned out to be comparatively long. The interviewers were clearly highly motivated and did their best to ensure a good response rate. A couple of interviewers had to be contacted because of the feedback we received; these cases were handled professionally by the survey agency.

As a whole, the collaboration with the Finnish survey agency has been smooth. The agency also appreciates the important place the interviewers and the agency are given in the SHARE community. Currently, the Finnish country team meets with the representatives of Taloustutkimus on a regular basis while phone calls and e-mails are exchanged on a weekly, and sometimes even on a daily, basis.

4.3.4 Summary

The Finnish team is working hard to raise national funding for Wave 8 and Wave 9. A main goal is to get SHARE introduced as part of the Finnish national research infrastructures and to raise external national funding for each wave. Before and after the data release of Wave 7, we have recommended, initiated and facilitated SHARE-related research projects in Finland. Although many Finnish scholars already use SHARE, it is crucial to obtain vast use of the Finnish data within different research disciplines. We see the organisation of the SHARE spring meeting in 2020 as a great chance to promote SHARE among scholars, funders and laymen alike. In addition to the conference, we will arrange SHARE-related satellite lectures and meetings in Helsinki. The steering group has proven to be very useful for obtaining advice, finding the right connections and promoting SHARE nationally. The Finnish country team can highly recommend other countries to consider forming similar bodies. Additionally, the hiring of PR services was very useful in ensuring national TV and other media coverage around SHARE and can be recommended especially during the establishment of SHARE in a country and as new data are released.

4.4 Latvia

Signe Tomsone, Andrejs Ivanovs, Diāna Baltmane – Riga Stradins University

4.4.1 Introduction

Latvia had no well-established traditions of multidisciplinary and longitudinal studies, especially involving people older than working age. Before the SHARE project, there have only been a few international and local projects focusing on the research of people in retirement age, for instance, the Eurobarometer or Statistical reports by Central Statistical Bureau of Latvia. However, it was not possible to compare the obtained results from these studies with those of other EU countries, or these studies did not occur over years that would allow for the comparison of the dynamics of life situations for people in retirement age. As one of the scientific research priorities of Riga Stradins University (RSU) is society's ageing problems, it has developed a foundation within this field. Therefore, after receiving an invitation from the SHARE project management group to join the project, it was assessed as a great opportunity to participate in one of the most recognised projects among academics and policy makers in Latvia, as SHARE is a multidisciplinary and longitudinal research involving people over retirement age that allows for the analysis of the interplay of different factors in the course of ageing.

4.4.2 Funding and assembling a national working group

Latvia was invited to participate in SHARE Wave 7 in February 2016, based on approved EU grant (VS/2016/0135). There was a possibility of coordinating the SHARE project in Latvia, and researchers at RSU were approached. After a period of internal discussions, the RSU researchers team was formed. Associate Professor Signe Tomsone (Faculty of Rehabilitation, SHARE CTL) and Andrejs Ivanovs (Statistics Unit, SHARE CTO) took coordination roles; later, the team was joined by Mirdza Kursīte (Statistics Unit) as a CTO assistant, followed by Diāna Baltmane (Statistics Unit), also as a CTO assistant. The EU grant included funding for data collection; however, no funding was allocated for project coordination activities. In addition, local co-funding of the project was needed. Attempts to obtain immediate financial support from the state agencies (e.g., Ministry of Education and Science, Ministry of Welfare) were not successful; opportunities would involve later periods. Therefore, the RSU administration made a crucial decision to implement internal financial resources to support the activities of SHARE project coordination.

4.4.3 Survey implementation

Latvia implemented SHARE Wave 7 as its first round. The collaboration with the survey agency responsible for data collection, the Institute of Sociological Research (Socioloģisko pētījumu institūts), was beneficial, and the sample design and testing of the national CAPI was performed in close contact. Our main challenges in implementing SHARE were mostly related to a gradual learning of all the specific procedures related to data collection and data transfer, which have been developed within a project's lifetime. These challenges put us into a very tight economic situation. Our main challenge, as experienced by other SHARE countries such as Estonia, was the translation of the generic CAPI questionnaire not only in Latvian but also in Russian. Moreover, the Latvian team received tremendous help and advice from its neighbouring countries: Estonia (a more experienced SHARE member) and Lithuania (also a newcomer to the SHARE project).

4.4.4 Summary

The main challenge for SHARE in Latvia appears to be securing sustainable funding for future longitudinal waves. The national funders in countries where no country-specific results from SHARE are yet available seem to be hesitant to include SHARE on the national ESFRI (European Strategy Forum on Research Infrastructures) roadmap; however, we are currently communicating with the Ministry of Education and Science to include SHARE on the national ESFRI roadmap, which would allow us to obtain necessary co-funding. Although this co-funding arrangement has not yet been made, RSU provides co-funding for the SHARE project. A promising turn for the SHARE Latvia team was the establishment of the Institute of Public Health at RSU, which took over the coordination role for this project in Latvia. The Institute assists in the communication process with local ministries and promotes SHARE in Latvia.

4.5 Lithuania

Antanas Kairys, Olga Zamalijeva – Vilnius University

4.5.1 Introduction

Similarly to other EU countries, Lithuania is facing the issue of major demographic change. Population ageing has an impact on the functioning of public services, requires data-based adjustments to social policy and affects the lives of individuals, especially those from vulnerable groups. Apart from common challenges, there are country-specific circumstances that make the process of addressing these issues even more demanding. Since joining the EU in 2004, Lithuania has been experiencing a negative impact related to the emigration of working-age adults to other member states, which only reinforced the change in demographic structure. Moreover, Lithuania has the highest number of deaths due to ischaemic heart disease, the highest suicide rate (which is nearly three times the average rate in the EU), and the highest levels of alcohol consumption in Europe. The need for a better understanding of the underlying processes of these problems has been the main motivation for collecting comparative interdisciplinary micro-level data in Lithuania.

4.5.2 Funding and assembling a national working group

Wave 7 of SHARE in Lithuania was funded by the Directorate-General for Employment, Social Affairs and Inclusion of the European Commission (grant no. VS/2016/0135) and co-funded by the Research Council of Lithuania (grant no. MIP-010/2015), with Vilnius University providing in-kind contribution. The grant from the Research Council was an invaluable contribution, which allowed the Lithuanian team to start the preparation for the survey even before the other new countries joined SHARE for Wave 7. On the other hand, the conditions of this grant involved additional tasks that had to be carried out within a strictly limited period. These tasks included conducting a very complicated public procurement process that had to be coordinated not only with SHARE-ERIC but also with the National Public Procurement Office and producing scientific publications based on other data sources.

The Lithuanian team was assembled at Vilnius University Faculty of Philosophy. The team consisted of two principal coordinators (Antanas Kairys as the CTL and Olga Zamalijeva as the CTO). Faculty administration had to carry the burden of various organisational tasks, and several faculty members helped immensely with the translation of difficult questionnaire items.

4.5.3 Survey implementation

Preparation for SHARE Wave 7 included two major tasks – translation of the questionnaire and deciding on the best possible sampling design. All translations were conducted by the country team members. Both principal members of the country team work in the Laboratory of Applied Psychology of Vilnius University and have vast experience with the translation and adaptation of various psychometric assessment tools. The most difficult part of the translation was the adaptation of education, pension system and accommodation questions. As Wave 7 focused on the personal history of the respondents, many questions had to be translated in a way that is suitable for different historical periods of Lithuania (e.g., independent state, part of the Soviet Union) without deviating from the generic questionnaire version. In addition, several sampling options were considered. Although there is a population register, access to the data is restricted. Moreover, for a long period of time, there has been no obligation to declare one's place of residence, and the data in the population register are partially outdated or incomplete. Consequently, an address register was used for the sampling.

The fact that data collection was highly successful may be related to not only the experience, work ethic and quality standards of the survey agency but also the extensive preparation of and close cooperation between the country team and the survey agency. For instance, before starting the fieldwork, the team leader and the survey agency representative attended a popular morning show on national television, where they encouraged potential respondents to participate in the survey as well as informed them on how to identify a SHARE interviewer, with the dynamometer becoming an easily recognisable attribute. Information about the start of SHARE Wave 7 was also published in regional media.

4.5.4 Summary

The main issue with the SHARE project in Lithuania has been funding, and without the European Commission grant, it would be close to impossible to carry out the Wave 7 survey. The most important future objectives for SHARE Lithuania are to promote use of the data and recognition of the project. After the public release of the data, the country team is planning a series of public events for policy makers and the scientific community in Lithuania. Expanding the

pool of data users by increasing awareness of the scope, quality, cross-country comparability, potential for research and benefits for data-based policy decisions might increase the chances for more stable funding of the SHARE project in Lithuania.

4.6 Malta

Marvin Formosa – University of Malta

4.6.1 Introduction

SHARE is of unprecedented importance to researchers' understanding of the social and health determinants of active and successful ageing. The key reason that motivated our decision to join the SHARE project is that such data are warranted to shed more light on the quandary that whilst the life expectancy has increased gradually in the past 20 years, European rates of healthy life expectancies have not followed similar rates of increase. This finding shows that longevity is much more complex than geneticists and social scientists assume and that the key to understanding better well-being in late adulthood lies in multidisciplinary research. The Malta country team looks forward to the emergent data as to better understand the dynamics of physical, psychological and social capital on one hand and the quality of life of older persons on the other. It is the first time that Malta will have such an extensive pool of data on gerontological and geriatric issues.

4.6.2 Funding and assembling a national working group

Malta was funded through a European Union grant (grant no. VS/2016/0135), the Junior Ministry for Persons with Disability and Active Ageing, and some contributions from the Department of Gerontology within the University of Malta. The team is composed of Professor Marvin Formosa, head of the Department of Gerontology and Dementia Studies at the University of Malta, and Katia Mifsud, a Department of Gerontology and Dementia Studies alumnus. The Malta team experienced major challenges in finding the right people for the team, as the country has full employment now. Thus, trainee gerontologists and social scientists are rare, as many prefer to leave Malta to take up bureaucratic posts in Brussels and Luxembourg. The fact that SHARE teams must work under pressure due to a tight schedule also hindered the situation, as many young social scientists have young families. However, all difficult situations generate learning possibilities, and we have learned much from this experience, mainly that much psychological and social support is necessary for a successful country team. I mention SHARE frequently in my lectures, and this renders students curious and more ready to listen to and read SHARE-induced reports. Malta joined SHARE in Wave 7, and we hope that the country will be able to take part in more future waves so that we can build longitudinal models of data.

4.6.3 Survey implementation

The procedures to finalise the national CAPI, such as translating and testing, were challenging. It took time initially to build a working relationship with the survey agency so that we could understand each other fully and hence limit misunderstandings. We therefore set many meetings, sometimes at the University of Malta, sometimes at the survey agency's office, and other times in neutral places. Common meetings in Munich for both the country team and the survey agency helped us to build a sense of fellowship but always with different roles and responsibilities. One key challenge for the Malta country team and the survey agency was that we were both completing these tasks for the first time, and in some instances, neither of us knew what to do. However, luckily for us, the SHARE Central team was very helpful and always there to guide us and assist us in any way possible. This is one of the key reasons why SHARE is an all-around sustaining experience: participants not only learn a great deal of information but they also engage in satisfying human relations with all the European Union member states.

4.6.4 Summary

The next step for the Malta country team is to present the data to the Maltese government and state policy makers so that policy making is informed by such data. We also need to locate suitable researchers to engage with SHARE data in their degree theses so that released data are analysed and eventually published. Such data will be extremely useful in policies relating to pensions, as the government will be able to decipher what is a fair lowest possible pension. Taking part in Wave 7 was a very fruitful and absorbing experience, and hopefully, we can provide important advice to other national teams.

4.7 Romania

Alin Marius Andrieș, Mircea Asandului – Alexandru Ioan Cuza University of Iași

4.7.1 Introduction

Romania was a country dominated for the better part of the last century by a strong and strict Communist regime, in which virtually no one was on welfare or social benefits. Indeed, unemployment rates before 1989 were negligible, staying employed came with special perks (free housing, holidays, etc.), and begging was illegal. Since the Iron Curtain fell, Romania has become a welfare paradox, with pensions and child benefits that rival wages but also with the lowest unemployment benefits in the EU. Currently, nearly one in three Romanians (totalling approximately 6.5 million people) receives some form of social benefit (National Institute of Statistics, 2018). Moreover, like other European Union member states, Romania is also facing a serious demographic ageing crisis. The crisis is associated with an increasing number of young people who immigrate to more developed countries. Between 2007 and 2015, approximately 3.4 million Romanians have emigrated, placing the country in second place globally with respect to the emigration growth rate between 2007 and 2015, after Syria (United Nations, 2017). These aspects of the Romanian society have generated significant transformations in the structure of the population, leading to an increased proportion of elderly individuals compared to young people. This imbalance can create major challenges, such as the sustainability of the public pension system or excess capacity of care centres for elderly persons.

4.7.2 Funding and assembling a national working group

The country-representing scientific partner institution of SHARE in Romania is Alexandru Ioan Cuza University of Iaşi[9], which has been carrying on a tradition of excellence and innovation in the fields of education and research. The Romanian research team includes Alin Marius Andrieș as CTL, Mircea Asandului as CTO and Bogdan Căpraru, Iulian Ihnatov and Daniela Viorică as experts. The team's main challenge was related to the novelty of the survey in Romania and the way it was perceived by the respondents in the sample. Another major source of pressure was associated with the fact that the entire national schedule needed to be coordinated with the international framework of the survey. The main funding of Wave 7 in Romania was made through European Commission DG Employment grant no. VS/2016/0135, and the project was co-financed by Alexandru Ioan Cuza University of Iași.

4.7.3 Survey implementation

In March 2016, the SHARE-ERIC Management Board invited several survey agencies with a proven track record of running large-scale CAPI surveys in Romania. For SHARE, a probabilistic sample was largely used, based on individual addresses randomly selected from the Direction for Population Registration and Administration of Databases, which is part of the police department. For rural areas without information on inhabitants, a listing procedure was applied. When we started working with the agency, we found a very enthusiastic team in the fieldwork department that was accustomed to dealing with long questionnaires.

One of the challenges that appeared was the length of the questionnaire. Given that this questionnaire included elements from the respondents' whole lifespan, a large amount of time was needed for translating and adapting it into Romanian. Another challenge that occurred was related to the households in large cities, where the response rate was quite low. This low rate was related to the fact that during summer, many people were on holiday. These issues were solved by placing greater stress on the importance of this survey and fielding more questionnaires in September. The collaboration with the agency went well during all the phases of the project, and together with their team, we solved all the issues that appeared in the field.

4.7.4 Summary

Implementing SHARE Wave 7 in Romania was very challenging because the necessary resources for translating, adapting and finalising the questionnaire were much higher than initially considered. We are convinced that once the data are public, the SHARE project will become well known in our country and that policy makers will take into account the results of the analyses that will be carried out based on these data. We hope that the key actors in the field will understand the utility of the information gathered by the survey

9 www.uaic.ro

and that we manage to include the SHARE project on the National Roadmap for Research Infrastructures.

References

National Institute of Statistics (2018). Regional economic and social benchmarks: Territorial statistics 2018. Retrieved from http://www.insse.ro/cms/en/content/regional-economic-and-social-benchmarks-territorial-statistics.

United Nations, Department of Economic and Social Affairs, Population Division (2017). World Population Prospects: The 2017 Revision, Volume I: Comprehensive Tables. Retrieved from https://population.un.org/wpp/Publications/Files/WPP2017_Volume-I_Comprehensive-Tables.pdf.

4.8 Slovakia

Ľubica Knošková, Dana Vokounová – University of Economics in Bratislava

4.8.1 Introduction

The development of population processes in Slovakia clearly confirms the intensive ageing of the population, which is manifested by the increase of the number of people in post-productive age and a reduction in the share of the population in childhood. Population ageing is one of the biggest demographic challenges in society. It also means a threat to the current pension and healthcare schemes. For the third consecutive year, the average age is over 40, with half of the population older than 39.8 years (based on data from the Statistic Office of the Slovak Republic, demography and social statistics). The reason for this trend is seen in reproductive behaviour, as fertility is approaching a critical level. As a consequence, the Slovakian population is ageing at an increasingly faster pace. The Slovakian population increase is very small, mainly due to migration reasons. Challenges arising from population ageing and opportunities to provide country-specific findings in the area of health, ageing and retirement of the population aged 50+ led Slovakia to join the SHARE project. Wave 7 results, which are the first SHARE results for Slovakia, would give fact-based inputs for policy makers in the areas of pension systems and health care. Our team will focus on investigating the lifestyle of the ageing population with a link to health by applying comparative analysis at the European level.

4.8.2 Funding and assembling a national working group

SHARE Slovakia was funded by EC grant no. VS/2016/0135. No other funding was obtained for Wave 7. Securing financing on the local level is a complicated and demanding process and requires full attention. Two institutions formed the first country team: the Slovak Academy of Sciences, which assumed leadership responsibilities, and the University of Economics in Bratislava, which assumed responsibilities mainly in fieldwork coordination. The situation changed at the end of Wave 7, when University of Economics in Bratislava overtook the country leadership. The country team consists of economists; due to changes, it is a new team and needs completion and funding.

4.8.3 Survey implementation

The fieldwork in the first weeks was far behind schedule. Only a few interviews were successfully completed in the first month of fieldwork. The SHARE country team initiated a status meeting with the survey agency. We discussed the reasons for this slow progress. It turned out that there are no serious issues and no blockers at all. However, interviewers were given a 2.5-month time limit to visit all assigned addresses and make corresponding interviews. Many interviewers decided to postpone their work to later weeks because of this generous time limit. We agreed with the survey agency to introduce an additional bonus scheme for interviewers that would motivate them to make household visits on a regular basis. We monitored the number of visits and completed interviews on a weekly basis. Those interviewers who made enough progress and reached their target were granted special financial bonuses on the top of original bonus agreement. This approach helped to significantly increase the number of completions and facilitated the successful completion of the fieldwork within the original time frame.

4.8.4 Summary

Given the lack of consistent and reliable longitudinal data on population ageing in Slovakia, we would like to introduce the results of SHARE Wave 7 in the areas of health status and health behaviour, socio-economic matters, cognitive functioning, social networks and health care to both the Ministry of Labour, Social Affairs and Family of the Slovak Republic, which is responsible for employment support, social care and the functioning of the pension scheme, and the Ministry of Health of the Slovak Republic, which participates in the creation of state health policy and is responsible for providing health care. We learned that the sample size of 2,000 respondents that we obtained in Wave 7 needs to be enlarged in future waves to improve statistical reliability in smaller segments to develop evidence-based policies. Thus, we need to guarantee local co-funding for the SHARE project, which is a very challenging and lengthy process in Slovakia, as projects in the social sciences are underfinanced.

CHAPTER 5

Sampling design in SHARE Wave 7

05

5 SAMPLING DESIGN IN SHARE WAVE 7

Michael Bergmann, Arne Bethmann – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC) and the Technical University of Munich (TUM), Giuseppe De Luca – University of Palermo

5.1 Introduction

The aim of the SHARE survey design is to draw inferences about the population of people aged 50 years and older across countries by using probability-based sampling. This is a complex process for all cross-national surveys since the samples in each country must do justice to national specificity but at the same time be internationally comparable. This chapter documents the sampling design adopted in SHARE. Starting with a definition of the SHARE target population (Section 5.2), we describe the protocol that is followed to harmonise and document the sampling procedure (Section 5.3) and present the sampling frames used by the countries that recruited a baseline or refreshment sample in Wave 7 (Section 5.4). We then discuss some important aspects of the SHARE sampling design, such as stratification, clustering, variation in selection probabilities and sample composition (Section 5.5). Finally, we provide additional information about the sampling variables included in the released SHARE dataset (Section 5.6).

5.2 The SHARE target population

The target population of SHARE consists of all persons aged 50 years and older at the time of sampling (i.e., 2016 in Wave 7) who have their regular domicile in the respective SHARE country. Persons are excluded if they are incarcerated, hospitalised or out of the country during the entire survey period, unable to speak the country's languages[10], cannot be located due to errors in the sampling frame (e.g., non-existent address, vacant house), or have moved to an unknown address. Spouses/partners of people aged 50 and older are included in the target population, regardless of their own age, because the household level is important for many of the variables collected in SHARE. Therefore, the target population of SHARE could also be defined in terms of households, i.e., all households with at least one member belonging to the target population of individuals. In contrast to many other studies, SHARE includes persons living in nursing homes and residential care whenever they are covered in the sampling frame from which the baseline/refreshment samples are drawn (whether this is the case differs between countries; see section 5.4 and Schanze, 2017) as well as in the longitudinal part of the sample. Further information on eligibility for the study can be found in the SHARE release guide that is publicly available on the SHARE website.[11]

5.3 The SHARE sampling protocol

The SHARE sampling protocol follows a four-stage process. First, each country that draws a baseline or refreshment sample in a wave of the study is initially required to provide a filled sample design form (SDF), containing a complete description of both the chosen sampling frame and the associated sampling design. In the second stage, the sampling proposal is carefully evaluated by the central coordination (SHARE Central) in Munich. Open questions and uncertainties are clarified on a bilateral basis with the country team and the survey agency before the sampling design is ultimately approved. The third stage consists of drawing the sample according to the approved sampling design process and is carried out by the country team or the survey agency. Finally, the country team provides a gross sample via the filled gross sample template (GST). This template contains all selected persons or households, the associated sampling frame information needed for the computation of the selection probabilities (e.g., household-level and population-level information about stratification and clustering), the household-level information about NUTS and LAU codes and (if any) additional auxiliary variables that could be used for ex-post compensation of non-sampling errors (see Section 5.6). After another round of checks by SHARE Central to detect inconsistencies with the proposed sampling design, the GST forms the baseline/refreshment sample component (in addition to a possible panel sample) of the Sample Distributor (SD) software, which is used to assign new respondents to the interviewers' laptops.

10 If a language is spoken by more than ten percent of the population in a certain country, the questionnaire is translated also into that language to include the language group in SHARE and to avoid under-coverage of important migrant groups (e.g., Russian in Estonia).

11 www.share-project.org

5.4 Sampling frames and population coverage

In the ideal case, all countries included in SHARE would have a probability-based sample from an official population register covering the population of interest and providing individual information. However, the availability of population registers that can be used as sampling frames varies greatly across countries. Furthermore, the regulations about who can or cannot access the registers and what information can be obtained from them are often country-specific. In addition, under- and over-coverage errors in the available registers may introduce non-sampling errors, which may jeopardise the standard properties of sample-based inference.

All countries in SHARE that draw a baseline/refreshment sample are requested to use the best sampling frame available at each wave, implying that sampling frames can, in principle, differ between waves and/or countries. For the target population of SHARE, a key feature any frame must fulfil is the availability of reliable information on age. If this information is not available from a given sampling frame, a preliminary screening procedure using the SD software must be applied before starting fieldwork to identify sample members aged 50 years and older.

Figure 5.1 shows which countries recruited a baseline or refreshment sample and on what type of sampling frame these samples were based. Baseline/refreshment samples from central or local population registers with individual information on age were drawn in Finland, Israel, Malta and Poland. In Croatia, a national health service/insurance register was used, while in Malta, the baseline sample was based on an election register.[12] Bulgaria and Latvia used a register that allowed for the selection of households with at least one person aged 50 years and older. Thus, although a screening procedure had to be applied, the size of the gross sample drawn could at least be minimised here. In contrast, the available address register in Lithuania and the telephone directory used in Cyprus did not include any information on the age of the selected units. Finally, Romania and Slovakia recruited a baseline sample using a combination of listing households and a subsequent screening for age eligibility. All other countries participating in Wave 7 interviewed only their longitudinal samples and did not draw a refreshment sample (see also Section 5.5.4).

12 Due to uncertainties in the available data in Malta, the information on age was verfied again at the beginning of the interview.

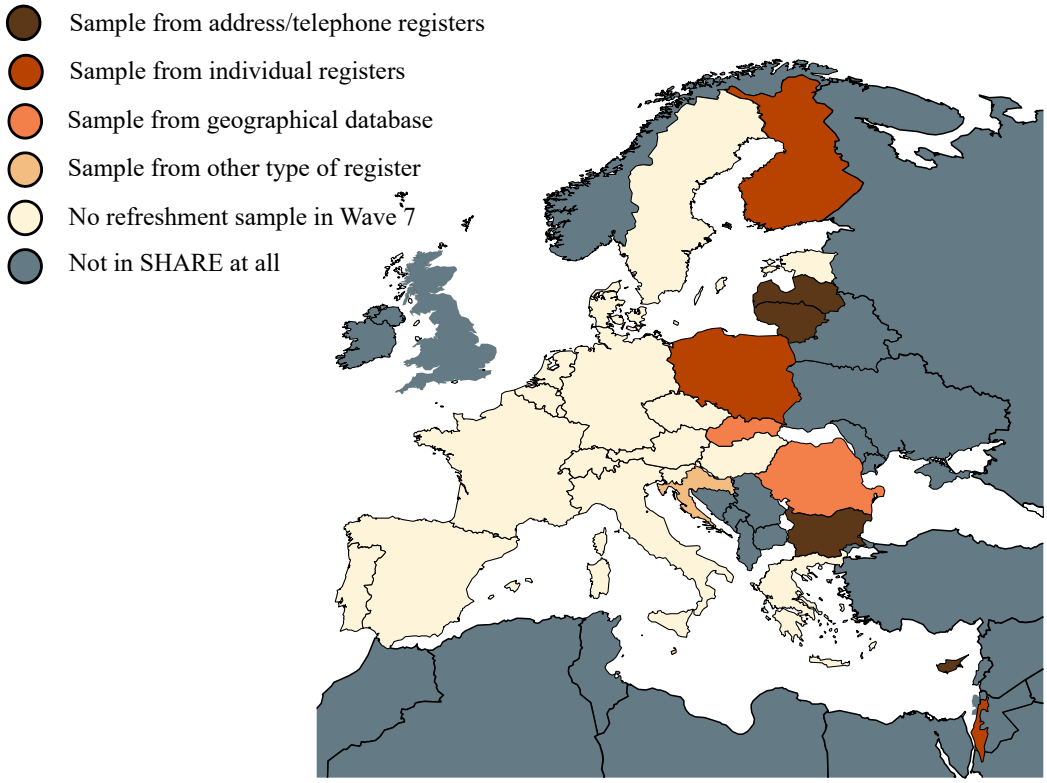


Figure 5.1: Types of sampling frames used in SHARE countries for baseline and refreshment samples in Wave 7

In Europe, 2.7 million people are aged 65 or older and live in a retirement home, nursing home or health care institution (Laferrère et al., 2013). Within the group of the Europeans aged 85 or older, about 13 percent live in an institution. As described in Section 5.2, persons aged 50 years and older who are living in a nursing home or another institution for the elderly are part of the SHARE target population. SHARE countries do not use specific sampling methods for these groups but include them as part of the general population sample. Differences in sampling frames used across countries, however, can lead to country-specific under-coverage of the nursing home population. Other sources of error might be due to either inaccuracies in the sampling frame (persons incorrectly registered as living in a private household) or interviewer mistakes (interviewer entering the code for "private household" instead of "nursing home"). In the longitudinal samples, respondents who lived in a private household before but moved to a nursing home or another institution for the elderly between two waves remain in the sample and are interviewed in the institution. In this respect, SHARE has developed special target measures to help interviewers gain access to nursing home respondents.

5.5 Sampling designs

After choosing the best sampling frame available in each country, the next step is the selection of a particular design for the national sampling schemes (i.e., a concrete procedure to draw a national sample from the national sampling frame). The rationale of the SHARE sampling design in Wave 7, as in all foregoing waves, is the same that all advanced population-based survey programmes apply at present. Kish (1994, p. 173) provided the underlying idea:

> *"Sample designs may be chosen flexibly and there is no need for similarity of sample designs. Flexibility of choice is particularly advisable for multinational comparisons, because the sampling resources differ greatly between countries. All this flexibility assumes probability selection methods: known probabilities of selection for all population elements."*

Thus, the sampling design is not restricted to be the same in all SHARE countries, but the basic principles of probability sampling with minimal coverage errors guide the choice of the national sampling designs. However, several features of the sampling design may still affect the precision of the estimates. For this reason, general advice on stratification, clustering, variation in selection probabilities and sample composition are provided in each wave to all participating countries by means of a sampling guide (see Bethmann et al., 2019) as well as bilateral discussions with the SHARE Central coordination team. We summarise these important aspects of the sampling design in the following subsections.

5.5.1 Stratification

The most frequently used sampling design in SHARE is a multistage stratified sampling design. Regional stratification schemes are recommended to ensure a good representation of different geographical areas within each country, improve the efficiency of the survey estimates and reduce the costs of the interview process. If other relevant characteristics are available from the sampling frame – such as age and gender in the case of population registers – then countries are advised to also use these characteristics for stratification.

5.5.2 Clustering

SHARE aims to use sampling schemes with a minimum variation of selection probabilities and a minimum amount of clustering. However, the design of sampling schemes with such characteristics is not always possible due to the lack of suitable sampling frames. Such a scenario applies, for example, if a country only has access to a list of households without individual information on age and an eligible person must then be selected from all eligible target persons of a sampled household (i.e., screening). In this case, variation in the selection probabilities cannot be avoided, and the national sampling scheme necessarily introduces a so-called "design effect" due to unequal selection probabilities:

$$\mathrm{Deff_p} = n\frac{\sum_{i=1}^{n} w_i^2}{(\sum_{i=1}^{n} w_i)^2},$$

where n is the sample size and w_i is the design weight, defined as the inverse of the selection probability.

Other studies (e.g., the European Social Survey; see Lynn et al., 2018) have shown that $\mathrm{Deff_p}$ usually ranges between 1.1 and 1.3 for designs that involve the random selection of one adult per household, depending on the variation of household sizes in a country. For SHARE, $\mathrm{Deff_p}$ is smaller than these values, as it depends only on the number of age-eligible units per household rather than the total number of adults per household, where an age-eligible unit is defined as either a single person aged 50 years and older or a couple comprising at least one age-eligible person. In most countries in SHARE, the majority of households do not contain more than one age-eligible unit, and very few have more than two.

In Wave 7, most countries (except Romania and Slovakia) had access to some form of official register, and sampling schemes yielding equal selection probabilities for all elements of the sample could therefore be implemented. In most of them, however, some sort of geographical clustering of the sample was used for cost efficiency reasons. This is especially true in countries with a large regional spread, where the cost efficiency of cluster sampling is relatively high

due to the reduction in interviewers' travel costs. The most common cluster design in SHARE was two-stage sampling with geographical areas (usually municipalities) as primary sampling units (PSUs) and households or individuals as secondary sampling units (SSUs). The main drawback of cluster sampling is concerned with statistical efficiency. For any estimator $\hat{\theta}$ of a parameter θ, the design effect due to clustering can be measured by the following:

$$\text{Deff} = \frac{Var(\hat{\theta})_c}{Var(\hat{\theta})_s},$$

where $Var(\hat{\theta})_c$ and $Var(\hat{\theta})_s$ are, respectively, the variances of $\hat{\theta}$ under the actual cluster sampling and a hypothetical simple random sampling. In principle, this indicator can be either smaller or greater than 1, indicating that cluster sampling can yield better or worse results (in terms of precision) than simple random sampling. In practice, however, clusters tend to be internally homogeneous. This intra-cluster homogeneity increases standard errors and thus decreases the statistical precision of our estimators. The stratification of the population of clusters can help to counteract this efficiency loss and was hence strongly advised. Furthermore, the countries were instructed to choose a mean cluster size that was as small as possible and to select as many PSUs as possible (see Section 5.6 for an overview of the sampling design variables included in the released SHARE dataset).

5.5.3 Selection probabilities

The calculation of selection probabilities in SHARE is subject to three difficulties. First, these probabilities must take into account the mentioned country-specific features of the various national sampling schemes and possible differences over waves. Second, the national sampling frames frequently do not contain any information about marital status, partnership and age of the spouse/partner, which is required to compute the selection probabilities of couples with two age-eligible persons. Third, as the panel proceeds, many countries attempt to maintain the representation of the younger cohorts that were not age-eligible in the previous waves by combining the refreshment subsample drawn in the current wave with the longitudinal subsample drawn in previous waves. The main problem here is that since these two subsamples are drawn from a partly overlapping target population, the elements of the longitudinal subsample may have a nonzero probability of being selected in the refreshment subsample, and the elements of the refreshment subsample may have, in turn, a nonzero probability of being selected in the longitudinal sample. Furthermore, the sampling frame information needed to compute these nonzero "cross-selection probabilities" is available only in a few countries where sampling is based on a simple design (e.g., Denmark and Sweden). Of course, these issues do not reflect specific limitations in the design of SHARE as such, but rather general problems faced in the implementation of longitudinal and cross-national sample surveys involving interviews with multiple household members at each wave (see, e.g., Lynn, 2009; Smith et al., 2009).

To ensure that the sampling strategy adopted to cope with these issues is harmonised as much as possible across countries and waves, the computation of selection probabilities is carried out by the SHARE Central coordination team in Munich together with the SHARE weighting experts in Palermo and Rome. More precisely, we address the lack of sampling frame information about the spouse/partner of each sample member by using the household composition data collected through the coverscreen module, which is administered at the beginning of the SHARE interview. The main problem is that these data are only available for respondents and not for the whole sample. Thus, we cannot compute selection probabilities for the subsample of non-respondents. Moreover, we account for the contribution of nonzero cross-selection probabilities by applying the "pooling method" proposed by Watson (2014). For countries using a complex sampling design involving stratification and clustering, this approach requires estimating the unknown cross-selection probabilities with available sampling frame information such as strata, age, gender and regional indicators. Although this stage introduces some randomness in the computation of selection probabilities, Monte Carlo simulations performed by Watson (2014) suggest that the pooling method outperforms many other ad hoc solutions to the problem of unknown cross-selection probabilities and is hence also applied in SHARE.

5.5.4 Sample composition

Sample composition, including the size of the national sample, is an additional feature of the sampling design affecting the efficiency of cross-sectional and longitudinal analyses. Table 5.1 gives an overview of all countries that ever participated in SHARE until Wave 7 and the composition of their samples in the respective wave(s).

Table 5.1: Sample type by wave and country

Country	Wave 1	Wave 2		Wave 3	Wave 4		Wave 5		Wave 6		Wave 7	
	Baseline	Panel	Refreshment /Baseline	Panel	Panel	Refreshment /Baseline	Panel	Refreshment /Baseline	Panel	Refreshment /Baseline	Panel	Refreshment /Baseline
AT	≤1954	✓		✓	✓	≤1960	✓		✓		✓	
BE_FR	≤1954	✓	≤1956	✓	✓	≤1960	✓	≤1962	✓	≤1964	✓	
BE_NL	≤1954	✓		✓	✓	≤1960	✓	≤1962	✓	≤1964	✓	
BG												≤1966
CH	≤1954	✓	≤1956	✓	✓	≤1960	✓		✓		✓	
CY												≤1966
CZ				✓	✓	≤1960	✓	≤1962	✓		✓	
DE	≤1954	✓	≤1956	✓	✓		✓	≤1962	✓		✓	
DK	≤1954	✓	≤1956	✓	✓	[1957-1960]	✓	≤1962	✓	[1963-1964]	✓	
EE						≤1960	✓		✓	[1963-1964]	✓	
EG								≤1962	✓		✓	
ES	≤1954	✓	≤1956	✓	✓	≤1960	✓		✓		✓	
FI												≤1966
FR	≤1954	✓	≤1956	✓	✓	≤1960	✓		✓	≤1964	✓	
GR	≤1954	✓	≤1956	✓					✓	≤1964	✓	
HR										≤1964	✓	≤1966
HU						≤1960					✓	
IE			≤1956	✓								
IL	≤1954	✓	≤1956				✓	[1953-1962]	✓		✓	≤1966
IT	≤1954	✓	≤1956	✓	✓	≤1960	✓	≤1962	✓	≤1964	✓	
LT												≤1966
LU								≤1962	✓	≤1964	✓	
LV												≤1966
MT												≤1966
NL	≤1954	✓	≤1956	✓	✓	≤1960	✓	≤1962	✓		✓	
PL			≤1956	✓	✓				✓	[1963-1964]	✓	≤1966
PT						≤1960			✓		✓	
RO												≤1966
SE	≤1954	✓	≤1956	✓	✓		✓	≤1962	✓		✓	
SI						≤1960	✓	≤1962	✓	≤1964	✓	
SK												≤1966

Note:

≤1966	Baseline sample
≤1966	Full range refreshment sample
[1963-1964]	Refreshment sample of youngest cohorts only

All SHARE respondents who were interviewed in any previous wave are part of the longitudinal sample. In addition, refreshment samples are drawn regularly for the following reasons: 1) to maintain the representation of the younger cohorts of the target population in Wave 7 (i.e., people born in 1965 and 1966) that were not age-eligible in the previous waves and 2) to compensate for the reduction in the size of the panel sample due to attrition. In Wave 7, refreshment samples were recruited in Croatia, Israel and Poland. In addition, baseline samples were drawn in the eight "new" countries of Bulgaria, Cyprus, Finland, Lithuania, Latvia, Malta, Romania and Slovakia, which joined SHARE for the first time. Only longitudinal samples, without the addition of refreshment samples, were interviewed in Austria, Belgium, Switzerland, the Czech Republic, Germany, Denmark, Estonia, Spain (including Girona), France, Greece, Hungary (participating again for the first time since Wave 4), Italy, Luxembourg, the Netherlands (although with another mode; see Das et al., 2017), Portugal, Sweden and Slovenia. For all countries, no panel rotation method was used to maximise the sample size available for longitudinal analyses. In other words, all units in the panel sample were considered eligible for an interview in the seventh wave, including nonresponding partners of panel members who were interviewed in some previous wave.

The choice of conducting a refreshment sample is mainly made by the countries because they must apply to their national funding agencies for their own funding. Since funding and sampling resources vary across participating countries, SHARE does not define a minimum net sample size. Instead, SHARE advises countries to maximise their net sample size with the available budget. In Wave 7, all countries that drew a refreshment sample deemed it necessary to select the full age range of people born in 1966 or earlier to compensate for the effect of panel attrition on all age cohorts. Where possible, these full-range refreshment samples included an over-sampling of the youngest cohorts that were not age-eligible in the previous refreshment samples to maintain the representation of younger cohorts.

5.6 Sampling variables in the released SHARE data

The SHARE Release 7.0.0 includes a generated module (gv_weights) with variables providing information about the sampling design in each country. Hence, the variable *subsample* identifies the various subsamples (e.g., different age groups when some form of oversampling has been applied) drawn in a specific country and wave of the SHARE panel sample, while the indicators *psu*, *ssu*, and *stratum* provide information about stratification and clustering in each subsample. In addition, the gv_housing module contains regional information (so-called NUTS areas; *nuts1-nuts3*) about the interviewed respondents who are also part of the GST (see Bethmann et al., 2019 for further information). Table 5.2 provides an overview of these variables, which are necessary to construct appropriate weights addressing problems of unit nonresponse and attrition (see Chapter 9 on weighting and imputation strategies).

Table 5.2: Sampling design variables

Variable	Description	Unit of analysis
subsample	Subsamples within country	Household & individual
psu	Primary sampling unit	Household & individual
ssu	Secondary sampling unit	Household & individual
stratum1	First stratum	Household & individual
stratum2	Second stratum	Household & individual
nuts1-nuts3	Regional classification of unit	Household & individual

References

Bethmann, A., Bergmann, M., & Scherpenzeel, A. (2019). SHARE Sampling Guide – Wave 8. SHARE Working Paper Series (33-2019). Munich: Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC).

Das, M., de Bruijne, M., Janssen, J., & Kalwij, A. (2017). Experiment: Internet interviewing in the sixth wave of SHARE in the Netherlands. In F. Malter & A. Börsch-Supan (Eds.), SHARE Wave 6: Panel innovations and collecting Dried Blood Spots (pp. 151-162). Munich: Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC).

Kish, L. (1994). Multipopulation survey designs: Five types with seven shared aspects. International Statistical Review, 62(2), 167-186. https://doi.org/10.2307/1403507.

Laferrère, A., Van den Heede, A., Van den Bosch, K., & Geerts, J. (2013). Entry into institutional care: Predictors and alternatives. In A. Börsch-Supan, M. Brandt, H. Litwin & G. Weber (Eds.), Active ageing and solidarity between generations in Europe (pp. 253-264). Berlin: De Gruyter.

Lynn, P. (2009). Methods for longitudinal surveys. In P. Lynn (Ed.), Methodology of longitudinal surveys (pp. 1-19). Chichester: Wiley. https://doi.org/10.1002/9780470743874.ch1.

Lynn, P., Scherpenzeel, A., Liberts, M., Kaminska, O., & Goedeme, T. (2018). European Social Survey Round 9 sampling guidelines: Principles and implementation. London: ESS ERIC Headquarters.

Schanze, J.-L. (2017). Report on sampling practices for the institutionalised population in social surveys. Deliverable 2.16 of the SERISS project funded under the European Union's Horizon 2020 research and innovation programme GA No: 654221. Retrieved from www.seriss.eu/resources/deliverables.

Smith, P. A., Lynn, P., & Elliot, D. (2009). Sample design for longitudinal surveys. In P. Lynn (Ed.), Methodology of longitudinal surveys (pp. 21-34). Chichester: Wiley. https://doi.org/10.1002/9780470743874.ch2.

Watson, N. (2014). Evaluation of weighting methods to integrate a top-up sample with an ongoing longitudinal sample. Survey Research Methods, 8(3), 195-208. http://dx.doi.org/10.18148/srm/2014.v8i3.5818.

CHAPTER 6

Interviewer training

06

6 INTERVIEWER TRAINING

Gregor Sand, Jeny Tony Philip, Yuri Pettinicchi – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC)

6.1 Introduction

Compared to self-completion online or paper-and-pencil questionnaires, face-to-face surveys such as SHARE that are conducted by interviewers always require the recruitment and training of interviewers. Interviewers play a crucial role in establishing contact and gaining the cooperation of respondents to collect survey data. Personal characteristics, training and its practical application, and acquired job experience play an important role in achieving cooperation and in conducting high-quality interviews. To avoid or reduce interviewer effects and optimise data quality, it is important to focus on the recruitment of suitable and competent interviewers and their preparation and motivation for data collection through intensive general and project-specific training (Billiet & Loosveldt, 1988).

In smaller studies, recruitment and training can be performed in-house; however, in large-scale studies, they are usually carried out by commercial and professional survey agencies that are subcontracted. All interviewer trainings should comprise general and project-specific interviewing techniques, theoretical and practical sessions, last a sufficient amount of time to cover the main aspects of the study, and take place shortly before the data collection (Alcser et al., 2016; Schnell et al., 2011; Stiegler & Biedinger, 2015).

The SHARE training programme was established to ensure cross-national comparability within SHARE and the HRS (Health and Retirement Study) (Alcser & Benson, 2005). In SHARE, interviewers are recruited by national survey agencies or statistical institutes (from now on subsumed under survey agencies) that have their own procedures and standards of contracting interviewers. While recruitment is outsourced, interviewer training in SHARE follows a multiplier principle in which trainers of the survey agencies are trained by SHARE Central; this approach is called a train-the-trainer (TTT) programme. This means that trainers are trained as if they were interviewers. The TTT sessions are hands-on trainings to prepare the survey agencies before each data collection on all aspects necessary for successful data collection. The TTT programme provides the model agenda for the so-called national training sessions (NTS), which take place after all survey agency staff returned to their respective countries to teach their interviewers what they learned at the centralised training.

A large proportion of the existing literature focuses on improving data quality by reducing interviewer effects and nonresponse. For instance, Blom et al. (2010) analyse interviewer and country effects on nonresponse. They find that there are systematic country differences in contact strategies and respondent cooperation, which can be partly explained by differences in interviewer characteristics. Groves and McGonagle (2001) examine the effectiveness of interviewer trainings. They use an experimental design to test whether interviewer workshops improve the cooperation rates of interviewers. The authors compare interviewers' performance before and after a training workshop and use a control group that did not receive any training. Their findings suggest that, at least in face-to-face surveys, interviewer trainings increase respondent cooperation far beyond interviewers' on-the-job experience does, especially among low-performing interviewers. In an experimental study, Billiet and Loosveldt (1988) also analyse the improvement of data quality through interviewer training. The authors show significant effects of interviewer training; in other words, trained interviewers collect more complete information than untrained interviewers do. However, these effects depend on the question structure: larger effects are observed for questions that require more interviewer activity because interviewers can actually apply what they have learned. The authors also find some evidence of better data quality if interviewers know they are being checked.

Apart from guidelines about interviewing techniques (e.g., Prüfer & Stiegler, 2002), literature on the person-related factors that improve data quality – such as the recruitment and training of interviewers – is scarce. Two noteworthy studies focus on the structure and content of interviewer trainings to minimise interviewer effects and optimise interviewer efficiency. Alcser et al. (2016) provide an outline of the qualification and training of a small number of interviewers in a self-administered study. According to them, data quality can be improved by targeted interviewer selection, training, and quality control or back checks. Stiegler and Biedinger (2015) stress the importance of general interviewer training (GIT) for successful interviewing before study-specific training is conducted. Depending on the complexity of the study, the latter should provide the opportunity for practice and role playing, an elaborated interviewer project manual, software training, and the opportunity for trainees to give feedback. In very large and cross-country studies, a TTT model can

help ensure the implementation of standardised study-specific protocols. We contribute to the existing literature by focusing on the planning and implementation of interviewer trainings based on hands-on experience. This chapter gives an overview of its centralised multiplier trainings, the general structure of the trainings, logistical challenges, content, and goals of the sessions. Apart from that, SHARE Central provides first-hand experiences of a NTS.

6.2 Train-the-trainer (TTT) sessions

SHARE interviewers are hired by independent survey agencies in all participating countries. It therefore requires additional effort to ascertain that the interviewers across all participating countries follow standardised procedures in contacting respondents, interviewing them and performing, for example, the physical measurements. A centralised training approach for about 2,000 interviewers is impossible due to the many languages being used in SHARE. Multinational, multicultural, or multiregional surveys present a particular challenge, as the recruitment, selection and training of interviewers can vary greatly among different countries due to differences in the cultural environment, existing infrastructure, and resources available (Alcser, et al., 2016; Smith, 2007). Therefore, for each wave of data collection, SHARE utilises the TTT programme to facilitate decentralised training in member countries. This training was scripted by Michigan's Survey Research Centre (SRC), which conducts the HRS (for details see Alcser & Benson, 2005). The initial centralised trainings are designed to enable replication in a standardised manner across member countries, thus achieving the goal of ex-ante harmonisation.

6.2.1 Structure and logistics

6.2.1.1 Location of the training and timeline of data collections

The SHARE TTT sessions are usually held in Munich. They are hands-on trainings to prepare the survey agencies before each data collection on all aspects necessary for a successful data collection and can take up to two full days. During each wave, TTT sessions are needed at three stages, namely, preceding the pretest, the field rehearsal, and the main survey. The primary aim of the TTT programme at the pretest stage is to familiarise the agency with the SHARE survey and its procedural aspects, its questionnaire content, the technical aspects of the software installation, and the performance of an interview using the SHARE instrument. The field rehearsal is the dress rehearsal for the main data collection and therefore simulates an environment that is as close as possible to that of the fieldwork of the main survey. The TTT sessions preceding the field rehearsal focus on changes since the pretest phase. The revisions highlighted at this stage include modifications to the instrument and study protocols, as well as basic fieldwork monitoring procedures. Details of the questionnaire are also covered in these sessions. The TTT sessions at the third and final stage – the main survey – continue where the field rehearsal training left off and focus on gaining respondents' cooperation. The main difference between these final training sessions and the sessions that were conducted for the test runs is an even stronger focus on hands-on training in the use of the software and the interview scenario – from contacting a household and handling refusals to conducting the actual interview. Apart from that, the main TTT programme includes a session on fieldwork monitoring.

6.2.1.2 Division by country groups

SHARE uses separate TTT sessions targeted at experienced and less-experienced participants among the country teams and the contracted survey agencies to render the intensive teaching style more efficient and to avoid overly crowded training sessions of 100+ attendees. This approach evolved based on the experience of Wave 7, in which eight new countries were integrated into the survey. As indicated earlier, TTT sessions are organised not only preceding the main data collection but also before each pretest and field rehearsal. Adaptations were made in Wave 7 to create a separate schedule for the "new" countries, which effectively doubled all TTT sessions: one training session for "old" countries and one session for "new" countries, multiplied by three (pretest, field rehearsal, and main survey). As a result, a total of six TTT sessions were conducted for Wave 7.

6.2.1.3 Agenda and duration of training

For each wave, the prototype agenda for a two-day training covering all essential topics of a SHARE training programme is adapted to the needs of that specific wave. The training plan requires a total of 16 to 18 hours. On some occasions, the training is stretched over three days.

The agenda of a typical SHARE TTT programme includes the following:

- An overview of the SHARE project
- An overview of the SHARE questionnaire, various modules, and new content
- Presentation on the sampling procedures
- Overview of the software

- Overview of the functionality and technical requirements of the CAPI (computer-assisted personal interviewing) instrument
- Special interview situations, such as nursing home and end-of-life interviews
- Protocols/ethics for handling self-administered questionnaires
- Biomarkers/physical measurements
- Mock interview

The total required training times for each sample type[13] are as follows:

- Baseline (interviewers without SHARE experience): ~13 hours
- Panel (interviewers without SHARE experience): ~16 hours
- Panel (interviewers with SHARE experience): ~9 hours

Table 6.1 shows a prototype agenda for the TTT programme.

13 Baseline interviews are conducted for respondents who participate in a regular interview for the first time. Baseline respondents are completely new to SHARE or have participated in a SHARELIFE interview in the previous wave. Panel interviews are conducted for respondents who have already participated in a baseline interview.

Table 6.1: SHARE training model agenda

Topic	Purpose (provided materials)	Time in minutes		
		(1)	(2)	(3)
Introduction, Welcome, and Logistics	Setting the stage for this intense training	10	10	10
SHARE Project and Questionnaire Overview	Explaining the goals of the project; importance of baseline and longitudinal sample	30	15	30
Sample Overview	Explaining how the sample was selected, sample eligibility, and response rate requirements	30	15	30
GIT Requirements	Covering minimal GIT requirements, including when and how to contact sample, probes, feedback, etc.	60	30	60
Overview of the Software and Practice Session	Operating the Sample Management System (SMS); assigning result codes, entering call notes; introducing non-contact scenarios and test results	60	30	60
Special Scenarios	Including split households, deceased individuals, new eligible respondents, additional result codes	30	15	--
Nursing Homes	Contacting respondents in nursing homes; working with gatekeepers/ potential proxy respondents	30	15	--
Proxy Interviews	Explaining how to identify proxy respondents; interviewing proxy respondents	30	15	30
End-of-Life (EOL) Interviews	Covering both concepts of the EOL interview, approaching respondents, and administering the interview (CAPI/CATI)	30	15	--
Overview of the Blaise Program	Handling Blaise, different types of questions, question text, data entry, interviewer instructions, etc.	45	15	45

Topic	Purpose (provided materials)	Time in minutes		
SHARE Questionnaire Walk-Through	Including a description of SHARE modules, a scripted review of the questionnaire ("mock") and mock materials; addressing main questions and issues that arise with different sections.	180	100	150
	Longitudinal: describing longitudinal differences and explaining preload; addressing different questions arising from re-interviews			
Drop-Off	Describing drop-offs and the procedure for identifying and labelling drop-offs appropriately; explaining the procedure for administering drop-offs; recording info in SMS	30	15	30
Physical Measurements	Hands-on exercises for each interviewer to demonstrate and prove their ability to conduct all physical measures	45	30	45
Response Rates and Contact Efforts	Importance of response rates and reiteration of required contact effort per line.	60	30	60
	Longitudinal: panel care and effort requirements; tracking effort			
Gaining Respondent Cooperation	Reviewing concerns that interviewers are likely to encounter; practising quick answers to several concerns; highlighting that a longitudinal sample is more likely to encounter different types of resistance	60	30	60
Practising Door-Step Situation	Interviewers team up in groups of 8-12 persons and practise introducing the study	60	30	60
Pair-Wise Questionnaire Walk-Through	Opportunity for interviewers to go through the questionnaire with a fellow interviewer using an abbreviated script, switching at the half-way point to complete the interview	90	90	90
Pair-Wise EOL Interview	Practise administering the EOL interview	30	30	--
Administrative Wrap-Up	Answer open questions	15	15	15

6.2.1.4 Attendees

Each participating country sends two to three representatives from the contracted survey agency to each TTT session. Agencies usually send the same representatives to the trainings of the pretest, field rehearsal, and main survey so that national trainings can focus on updates or changes in existing or new protocols. These are usually the same persons conducting the NTS in their respective countries. All appropriate computer skills required for conducting an interview are mandatory for these individuals. Parallel sessions are also conducted by the software development team CentERdata to instruct the survey agency IT managers on how to operate the software tools, in particular the Sample Distributor (SD; i.e., the server software that contains the national household sample). In addition to representation from the survey agency, attendance of at least one country team representative, preferably both the country team leader and his/her country team operator (usually a post-doc with survey experience), is mandatory for the TTT sessions. The country team is required, not only to provide a supporting function to motivate and guide the survey agency but also to ensure that the country team is conversant with the requirements for interviewers and the procedures and standards for the SHARE fieldwork at that stage. In Wave 7, these trainings were held at two separate dates and at two separate locations in Munich and involved approximately 80 survey agency representatives who were accompanied by representatives of the scientific partner institutions of their countries.

6.2.1.5 Training resources/materials

Integral to the training process at the TTT sessions are key training resources, namely, the installation software for the CAPI instrument, the SD software, the mock interview script (see Chapter 6.2.2.2), PowerPoint presentations, card exercises, interviewer manuals, recording booklets,

and showcards. In addition, if any biomarker measures are measured in a specific wave, relevant equipment is made available to ensure proper hands-on training. All materials are provided in English, and the survey agencies are responsible for translating them with guidance from the country teams if necessary.

6.2.2 Content and goals

6.2.2.1 Combination of theory and practice

A key aim of the TTT programme is to achieve cross-national comparability of data collected across all SHARE countries by ensuring standard interviewer behaviour across all countries. Each TTT session aims at providing specific guidelines on procedures that interviewers must follow in each country to ensure cross-national comparability of the results. The content of the TTT sessions address all technical, logistical, and managerial aspects of successful fieldwork. These are demonstrated in accompaniment with hands-on training exercises. In addition, great emphasis is placed on multi-modal teaching methods to ensure that standardised best practices of interviewing (such as active listening and being prepared to address respondent reluctance) are trained at the national level. These guidelines present strategies to optimise interviewer efficiency and minimise the effect of interviewer attributes on the data through appropriate recruitment, selection, and case assignment; they also present strategies to minimise the effect of interviewer behaviours on sampling error, nonresponse error, measurement error, and processing error through interviewer training (Alcser, et al., 2016). The format of the TTT sessions combines lectures with slide presentations and group exercises. This helps keeping the participants engaged and acknowledges that different people learn in different ways (James & Galbraith, 1985). It is recommended that the agencies translate and use these slides for the national sessions. Feedback from attendees on the content, structure, and mode of the training is used to improve subsequent TTT sessions.

6.2.2.2 Mock interview

The practical sessions help attendees to proceed from acquiring knowledge to acquiring actual skills (O'Brien et al., 2002). The most effective mode of training is the so-called "mock interview". The mock interview is the simulation of an actual SHARE interview with TTT attendees taking on different roles: all TTT attendees are encouraged to volunteer as interviewers and read out questions to a "mock" respondent. The mock interview is structured to include as many potentially complicated scenarios as possible, thus exposing the interviewers to difficult parts of the questionnaire. The mock interviews at different stages (pretest, field rehearsal, main survey) are also sufficiently diverse to ensure that the participants are exposed to different variants of a typical SHARE interview, for example, a panel interview, a baseline interview, a life history interview, an EOL interview, or a nursing home interview.

A concerted effort is made to minimise teaching with slides and maximise the use of interactive formats, such as the mock interview, group exercises, and mini-presentations, by the attendees. The mock interviews are not only useful for the trainees in their training in the reading and administration of the questionnaire but are also useful to the trainers in their assessments of attendees' skills, understanding, and skills development. It is emphasised that it is important for the agencies to re-enact the complete mock interview created by SHARE. Standardised question-asking is to be the norm with interviewers being trained to read each question exactly as written and to read the questions slowly and audibly (Fowler & Mangione, 1990). Interviewers are reminded to ask all questions exactly in the order in which they are presented in the questionnaire (Doyle, 2004). Emphasis is placed on following interviewer conventions, such as stressing words that are highlighted and not reading interviewer instructions aloud. The mock interview at the TTT session for the main data collection is also scripted to provide some practical examples of GITs other than the aforementioned.

6.2.2.3 General training of standardised interviews

To establish an effective blueprint for all subsequent trainings that will occur, a GIT component is essential (Alcser & Benson, 2005). Standard GITs implemented at the SHARE TTT sessions include standardised interviewing, probing and feedback conventions, and the collection and processing of data, including when a contact was attempted and what the result of each contact attempt was. Research indicates that GIT helps to improve survey data quality by (1) reducing item nonresponse (Billiet & Loosveldt, 1988), (2) increasing the amount and accuracy of information obtained (Billiet & Loosveldt, 1988), and (3) increasing survey participation by teaching interviewers how to identify and respond to respondents' concerns (O'Brien, et al., 2002). If a respondent's answer is inadequate, interviewers should be trained to employ unbiased techniques to encourage answers that are more complete, appropriate, and thoughtful (Cannell et al., 1977; Groves et al., 2009). Probes must be neutral; that is, they must avoid "sending a message" about what is a good or a bad response (Alcser et al., 2016). The response to the approach suggested above varies across survey agencies, with some agencies not finding this approach useful, as they already have their own standard in-house training, while others find it useful to have a standardised blueprint for training and implementing the survey.

6.2.2.4 Project-specific training

Detailed specifications are given to all attendees – namely, future NTS trainers – of participating survey agencies on how to relay information imparted at the TTT sessions to their interviewers at home. TTT sessions, which are conducted for new countries, are geared towards the specific needs of the newcomers. For example, several sessions on obtaining probability samples were organised when eight new countries were included in Wave 7, as it was anticipated that some countries would face challenges due to the absence of accessible population registers. In addition, the TTT sessions also included long bilateral consultations with many survey agency representatives to explain SHARE standards of sampling and address their specific questions. In special cases, such as when the SHARELIFE interview is conducted, participants are trained on using the calendar functionality of the questionnaire.

The session incorporating the training for the SMS includes several steps, starting with details of the functionality and handling of household information entered in the so-called coverscreen, including respondent selection/identification. It also includes launching an interview, interviewer observations, and documenting the call (call record). Additionally, part of this training is instructions on transferring a sample record if instructed to do so by a supervisor and on sending sample and interview information to the survey organisation. This training also entails standard field procedures, including the following: contacting sample households; maintaining and accounting for each selected sample element in the gross sample; writing "call records" for each contact attempt; interviewing population groups with limitations; and handling gatekeepers, refusals, and number of call attempts (Alcser et al., 2008). Interviewers are trained to record all contacts and contact attempts and to take notes in order to tailor approaches for maximising contact. This information allows researchers to observe considerable variation by country in how contact strategies are implemented (Alcser et al., 2011). Apart from this training, other training in all software innovations is provided, for example, new protocols to update information on deceased members of the household and difficult situations for the interviewer, such as nursing home interviews or refusals. Details on technical specifications of hardware, namely, laptops and tablets to be used for the interviews, are also provided.

An integral part of the training is a special slot to train the interviewer to perform special tasks, such as the collection of biomarkers. Examples of such areas that need specific attention and training are the grip strength exercise in Wave 7 and the dried blood spot collection in Wave 6 (for details see Börsch-Supan & Andersen-Ranberg, 2017). In addition, the conformity of the implementation in the SHARE countries with the standardised format provided through the central TTT sessions are assessed with the help of onsite visits during NTS (see Chapter 6.3) and interviewer debriefing reports (Alcser et al., 2011).

6.2.2.5 Group exercises

Another successful format that is very popular at the training sessions is group exercises in which attendees are asked to discuss in small groups their "best practices" to gain respondent cooperation. The results of these group exercises are then collated onto a flip chart and presented at a plenary session. An example of this is a group exercise conducted in the TTT sessions for the main survey of Wave 7 that split the attendees into different groups to discuss how to maximise response rates and manage difficult interview situations. This exercise resulted in many tips on how to handle difficult situations and led to the useful sharing of strategies to improve communication with interviewers and keep both interviewers and respondents motivated.

6.2.2.6 Question cards

A popular tool used in the TTT programme and recommended for individual agency training sessions are the question cards. These cards are made available throughout training for trainees to write questions that could not be accommodated on the spot due to time constraints. During breaks and at the end of the day, these questions are shared with the entire group, and answers are provided so that everyone receives the same information.

6.2.2.7 Interviewer manual

A key reference source is the SHARE interviewer manual that was completely re-designed for Wave 7 and now is structered according to a much shorter question-and-answer (Q&A) format. This new manual received a lot of positive feedback from the interviewers and supplements the TTT sessions by providing a reference to all of the SHARE protocols. Among other information, the manual provides interviewers with background information about SHARE, interview criteria, how to address respondent concerns, how to prepare for SHARE interviews, how to handle specific situations, and information about complicated technical terms used within the questionnaire.

6.3 National training sessions (NTS)

The NTS mimic the format of the TTT programme (see previous section). The outcome of the training depends on several features, ranging from the quality of the content to the

participation of the interviewers. The ultimate goal of the NTS is to prepare the interviewer to go into the field, adequately qualified to face the complexity of the survey. For the SHARE project, the NTS are performed by the national survey agencies. The content and the structure of the training are designed by the investigators at SHARE, but the survey agency staff must adapt it to the national context, jointly with the respective SHARE country team. The survey agency staff must also handle the logistics of the trainings, such as the location of the training, the number of interviewers in the training room, or their average experience. All of these aspects play an important role in conveying the project-specific information to the interviewers and are discussed in the following paragraphs, focusing on the experiences during SHARE Wave 7.

6.3.1 Content adaptation and conformity with the TTT programme

The NTS replicate the TTT sessions, both in structure and in content. While the structure is followed rigorously, the content needs to be adapted to the national context. The adaptation process is not only about the translation of the material from the source language (English) to the target language (national language), but it also involves some adjustments to place the content and the instructions properly into the national context. The following examples better convey the necessity of these adaptations. Pension systems or health care systems are very complicated, and a one-size-fits-all type of question does not help the interviewers handle special situations that might be quite common in a country. Therefore, the survey agency staff must provide country-specific instructions to capture special situations correctly and consistently and to make them comparable with the observations collected in other countries. As an example, some countries provide special pension benefits for having worked at the Chernobyl site in the late 1980s. In some regions, the high density of people who are not imprisoned but rather are confined to their homes implies the need to provide more detailed instructions to interviewers on how to code the situation of the respondent (i.e., private home, nursing home, how to address questions that assume freedom of movement).

6.3.2 Logistics

The NTS are mainly located in large cities, easily reached and can be spread out among several cities. Multiple NTS are not necessarily held in the same city. The choice should factor in the size of the country. Usually, multiple NTS should be held within a limited period of time before the fieldwork. Having the NTS close to each other helps the trainers to remain focused on the training, with continued emphasis on the crucial components. The number of trainings depends on the number of interviewers to be trained. This, in turn, depends on the sample size of the interviews to be conducted. Approximately 20 interviewers per training allow each of them to follow the training easily.

6.3.3 Attendees

There are two types of attendees in the NTS: the trainers and the interviewers. The trainers are survey agency staff who attend the TTT sessions (see previous section) and are responsible for the outcome of the training. They are in control of the workflow and rely on their expertise to make it effective. One member from the country team should also be present at the NTS to provide support during the training in case of difficult questions posed by the audience or if country-specific situations need to be handled. In such a case, the trainer can focus on the big picture and rely on the help of the country team members for very specific issues. IT staff from the survey agency provide support on the spot in case of software issues. This approach should reduce possible interruptions during the training due to technical malfunctioning, which could lower the level of the attention of the audience.

The grouping of interviewers in the training should be heterogeneous according to not only their age and gender but also their SHARE experience. The ideal group of interviewers comprises men and women with an average age close to that of the target population, that is, 50 years and older for SHARE. Being of a similar age improves the cooperation between the interviewer and the interviewee (Vercruyssen et al., 2017). The interviewers of a given training should also differ with respect to their job experience and, more importantly, SHARE-specific expertise. The right mix in the group of interviewers should help transfer knowledge horizontally and keep the audience more engaged and active (i.e., the new interviewers can benefit from the solutions implemented by the experienced ones when they faced difficult situations in previous waves).

Table 6.2 provides an overview of the NTS in the SHARE countries for Wave 7. The average number of NTS per country is approximately five, and they took place in the month before the fieldwork (31 days between the first and the last NTS, on average). On average, 14 interviewers, mainly women (70 percent), attended a training session; the average age of these interviewers was 53 years.

Table 6.2: Overview of NTS in Wave 7 (per country)

Country team	Number of NTS	Time span (days)	Total number of interviewers	Female (%)	Age (years)
Austria	4	44	73	57.5	53.7
Belgium (FR)	4	12	57	62.5	54.2
Belgium (NL)	5	35	59	65.4	57.9
Bulgaria	5	13	114	83.3	49.9
Croatia	7	36	81	86.4	43.5
Cyprus	4	12	27	89.5	43.7
Denmark	4	21	61	49.2	69.9
Estonia	5	32	81	89.2	53.6
Finland	2	5	41	78.0	59.6
France	4	16	139	58.9	55.8
Germany	5	15	141	49.3	64.2
Greece	10	31	150	71.0	39.8
Hungary	4	38	74	69.0	55.9
Israel	3	98	18	39.0	48.7
Italy	10	24	138	81.1	51.5
Latvia	3	49	24	94.6	55.2
Lithuania	3	4	58	93.2	53.6
Luxembourg	7	28	30	83.9	57.0
Malta	3	54	28	80.0	39.7
Poland	4	36	65	63.2	51.2
Portugal	4	157	11	45.4	41.8
Romania	8	9	87	79.3	43.8
Slovakia	3	6	31	76.4	n.a.
Slovenia	5	14	52	80.0	45.5
Spain	3	12	85	77.6	49.2
Sweden	3	30	64	54.5	58.1
Switzerland	10	17	71	53.3	58.9
Total	132	--	1860	70	53

6.3.4 Excursion: Feedback from one instalment of NTS

A member of SHARE Operations had the chance to participate in an instalment of NTS during the preparation for Wave 7. The aim was to collect first-hand impressions from the interviewers while they were trained and to determine which aspects of the training format could be improved. The desired output was to develop operational improvements that would make the training easier to attend and, therefore, more effective.

6.3.4.1 Issues observed

During that training, the interviewers had to process a large amount of information, ranging from the concept and meaning of questions to the correct coding of all possible scenarios. For SHARE Wave 7, there were more than 1,500 questions in the questionnaire plus all the auxiliary information collected before and after the interview. Such a vast amount of information can be challenging for attendees and requires a trade-off between available resources (i.e., individual attention, training time, and amount of information to be understood). The right equilibrium is always difficult to achieve by survey agency staff; however, in the present case, the content was clearly presented, with enough emphasis on the crucial steps as well as enough time for practical examples.

Another issue that affected the outcome of the training was connected with the questionnaire software. Interviewers, for example, struggled when several acronyms or short-cut keys that are specific to the project were used during the training. Interviewers usually handle several complicated surveys at the same time with overlapping topics. This approach could create confusion and make them prone to mistakes.

6.3.4.2 Available solutions

A set of solutions could be implemented to help interviewers address the information they need to process. The instructions for interviewers on how to code special situations should be streamlined. The instructions that cover common situations should be visible directly on the screen when they read the questions. These instructions should be short, clear and to the point (e.g., "if this is the case, then code 1"). The SHARE project provides a slim and clear manual for interviewers that is structured according to a question-and-answer (Q&A) format. It addresses common situations that could occur at the interview place and a short description on how to handle them. Moreover, it reports questions that the interviewee could ask with a possible answer to use.

Another related solution would be the provision of a priority list of the most crucial information to be trained. Having a clear idea about which information is more decisive will help the survey agency staff to tailor the training and to allocate more time to the more relevant parts. The SHARE project provides a glossary of the most important concepts used in the questionnaire (i.e., definition of household, nursing home, or eligible respondents). Moreover, country teams are allowed to add country-specific instructions to the interviewer manual for the questions that need to address special situations (e.g., Chernobyl pension benefit).

Moreover, the use of shortcuts can be beneficial once interviewers become familiar with the software, but in practice, they are only used by some interviewers. Therefore, it would be better to keep the shortcuts as a second option and allow interviewers to use the mouse to navigate through the questionnaire.

6.4 Concluding remarks

While extant literature has shown that interviewer trainings per se increase respondent cooperation, data completeness, and data quality, this chapter notes and discusses the main contributing factors: the structure, logistics, content, and goals of these interviewer trainings.

The SHARE TTT programme is modelled on the SRC training model, but its structure has been adapted to meet the requirements of a cross-national study by ensuring active engagement in the form of participation and multiple feedback loops from trainers at the national level. Centralised trainings allow for replication in a standardised manner across member countries. To ensure the goal of proper harmonisation across member countries, SHARE adheres to specific norms and processes essential for cross-national comparison, thus standardising results across participating countries. Over the course of time, the SHARE protocols and procedures have evolved and may have influenced other surveys (e.g., European Social Survey). In a similar vein, lessons learned at SHARE Central, from implementing a decentralised approach to training while simultaneously ensuring cross-national harmonisation, can be useful for all international studies.

Several factors contribute to the improvement of data quality and data harmonisation in SHARE: three data collections that require three standardised and decentralised trainings and corresponding material; the combination of general and project-specific content with practice sessions and a mock interview; and the attendance of country team members and survey agency representatives, including IT staff. While the recruitment of skilled interviewers and the determination of accessible NTS locations lies completely in the hands of the

survey agencies, they must replicate or adapt the content and structure of the centralised TTT sessions to the national context. NTS visits and evaluations by SHARE Central staff do help develop best practices that make the interviewer training easier and more effective.

In sum, the key aims and contributions of the SHARE TTT programme are as follows:

- Implementing standardised training practices for all interviewers across all member countries while simultaneously including country-level involvement
- Obtaining high-quality data that can be harmonised
- Obtaining results that are comparable with those of sister surveys (e.g., HRS)

References

Alcser, K. H., & Benson, G. D. (2005). The SHARE train-the-trainer program. In A. Börsch-Supan & H. Jürges (Eds.), The Survey of Health, Ageing And Retirement in Europe – Methodology (pp. 70-73). Mannheim: Mannheim Research Institute for the Economics of Aging.

Alcser, K. H., Benson, G. D., & Guyer, H. (2011). Pilot testing the SHARE quality profile. In M. Schröder (Ed.), Retrospective data collection in the Survey of Health, Ageing and Retirement in Europe. SHARELIFE Methodology (pp. 37-43). Mannheim: Mannheim Research Institute for the Economics of Aging.

Alcser, K. H., Clemens, J., Holland, L., Guyer, H., & Hu, M. (2016). Interviewer recruitment, selection, and training. Cross-cultural survey guidelines (pp. 419-468). Retrieved from http://ccsg.isr.umich.edu/images/PDFs/CCSG_Interviewer_Recruitment.pdf.

Alcser, K. H., Kirgis, N., & Guyer, H. (2008). Adapting U.S. based general interviewing and study-specific training protocols for cross-national survey data collection. Paper presented at the International Conference on Survey Methods in Multinational, Multiregional and Multicultural Contexts (3MC), Berlin.

Billiet, J., & Loosveldt, G. (1988). Improvement of the quality of responses to factual survey questions by interviewer training. Public Opinion Quarterly, 52(2), 190-211. https://doi.org/10.1086/269094.

Blom, A. G., Leeuw, E. D., & de Hox, J. J. (2010). Interviewer effects on nonresponse in the European social survey. MEA Discussion Papers. Mannheim: Mannheim Research Institute for the Economics of Aging. Retrieved from https://ub-madoc.bib.uni-mannheim.de/2966/1/meadp_202_10.pdf.

Cannell, C. F., Oksenberg, L., & Converse, J. M. (1977). Striving for response accuracy: Experiments in new interviewing techniques. Journal of Marketing Research, 14(3), 306-315. https://doi.org/10.2307/3150768.

Doyle, J. K. (2004). Introduction to interviewing techniques. In D. W. Wood (Ed.), Handbook for IQP advisors and students. Worcester: Worcester Polytechnic Institute.

Fowler, F. J., & Mangione, T. W. (1990). Standardized survey interviewing. In F. J. Fowler & T. W. Mangione (Eds.), Applied social research methods: Standardized survey interviewing (pp. 33-54). Thousand Oaks: SAGE Publications. https://doi.org/10.4135/9781412985925.

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). Survey methodology (2nd ed.). Hoboken: John Wiley & Sons.

Groves, R. M., & McGonagle, K. A. (2001). A theory-guided interviewer training protocol regarding survey participation. Journal of Official Statistics, 17(2), 249-265.

James, W. B., & Galbraith, M. W. (1985). Perceptual learning styles: Implications and techniques for the practitioner. Lifelong Learning, 8(4), 20-23.

Börsch-Supan, M., & Andersen-Ranberg, K. (2017). Collection of dried blood spot samples in SHARE Wave 6. In F. Malter & A. Börsch-Supan (Eds.), SHARE Wave 6: Panel innovations and collecting dried blood spots. Munich: Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC).

O'Brien, E. M., Mayer, T. S., Groves, R. M., & O'Neill, G. E. (2002). Interviewer training to increase survey participation. Proceedings of the Survey Research Methods Section, American Statistical Association, Alexandria (pp. 2502-2507). Retrieved from http://www.asasrms.org/Proceedings/y2002/Files/JSM2002-000530.pdf.

Prüfer, P., & Stiegler, A. (2002). Die Durchführung standardisierter Interviews: Ein Leitfaden. Mannheim: Zentrum für Umfragen, Methoden und Analysen (ZUMA). Retrieved from https://nbn-resolving.org/urn:nbn:de:0168-ssoar-201435.

Schnell, R., Hill, P. B., & Esser, E. (2011). Methoden der empirischen Sozialforschung (9th ed.). Munich: Oldenbourg Verlag.

Smith, T. W. (2007). Survey non-response procedures in

cross-national perspective: The 2005 ISSP Non-Response Survey. Survey Research Methods, 1(1), 45-54. https://doi.org/10.18148/srm/2007.v1i1.50.

Stiegler, A., & Biedinger, N. (2015). Interviewer Qualifikation und Training. GESIS Survey Guidelines. Mannheim: GESIS – Leibniz-Institut für Sozialwissenschaften. https://doi.org/10.15465/sdm-sg_013.

Vercruyssen, A., Wuyts, C., & Loosveldt, G. (2017). The effect of sociodemographic (mis)match between interviewers and respondents on unit and item nonresponse in Belgium. Social Science Research, 67, 229-238. https://doi.org/10.1016/j.ssresearch.2017.02.007.

CHAPTER 7

The third round of the SHARE interviewer survey

07

7 THE THIRD ROUND OF THE SHARE INTERVIEWER SURVEY

Sabine Friedel, Arne Bethmann, Manuel Kronenberg – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC) and the Technical University of Munich (TUM)

7.1 The importance of the interviewer

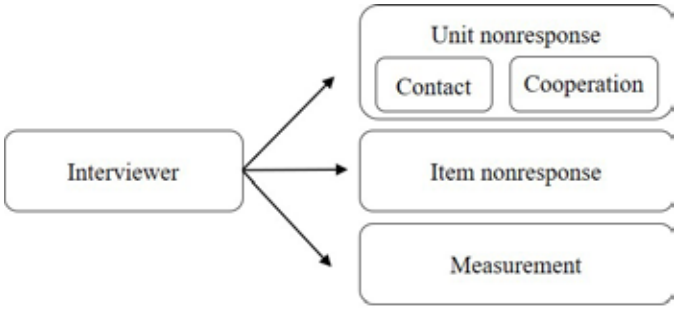
Interviewers play a crucial role in face-to-face surveys because they are the link between the researchers who developed and conducted a survey and the data that result from the survey. Interviewers must establish contact with the sampled persons, convince them to participate in the survey, administer the survey precisely, and answer questions that arise during the interview. Furthermore, interviewers may conduct specific measurements or tests and lay the foundation for successful future contacts in a panel survey (Groves & Couper, 1998; Schaeffer et al., 2010). However, interviewers do not perform similarly in these different tasks, and different performances can result in various interviewer effects. The term "interviewer effect" describes the bias of survey outcomes that are due to the interviewer. Such effects can be observed when respondents who are interviewed by the same interviewer answer questions more similarly than respondents interviewed by different interviewers (Blom & Korbmacher, 2013). Such differences in interviewer performance can affect the data quality.

Despite the essential role of interviewers, researchers know very little about them. A wide body of literature exists about identifying interviewer effects, but researchers still have relatively little knowledge about their process of formation. To shed more light on interviewer effects and to explain rather than describe them, detailed information on interviewers is required. Therefore, SHARE launched an interviewer survey to overcome the lack of information on interviewers and to offer researchers the opportunity to investigate interviewer effects in SHARE profoundly.

Since Wave 5, SHARE has invited all participating countries to additionally participate in the SHARE interviewer survey. The goal of this project is to make interviewer information available through the collection of data on the SHARE interviewers prior to fieldwork. Researchers can link the information gathered in this separate online survey to the SHARE survey data that each interviewer collected on his or her respondents.

7.2 Interviewer effects in surveys

Interviewer effects can occur in different steps of a survey. Figure 7.1 gives an overview of the three main aspects that are prone to interviewer effects.



Adapted from Blom and Korbmacher (2013)

Figure 7.1: Types of interviewer effects in surveys

A large body of literature discusses interviewer effects on contact and cooperation rates (e.g., Blom et al., 2011; Durrant et al., 2010; Groves & Couper, 1998; Hox & De Leeuw, 2002; Jäckle et al., 2013; Lipps & Pollien, 2011; Pickery et al., 2001). Thus, interviewers are differentially successful in recruiting survey participants, consequently affecting unit nonresponse. Research in this area has focused on interviewer attributes, such as experience, interviewer skills or interviewer-respondent interaction, and survey management characteristics, such as interviewer payment or interviewer burden (for an overview see West & Blom, 2017).

Interviewers can also affect the respondents' willingness to provide answers to certain questions (e.g., Pickery & Loosveldt, 2001; Singer et al., 1983). For instance, questions on income, drug use, and sexual behaviour are prone to item nonresponse. Respondents may not be willing to provide information on those kinds of questions because they are too sensitive and intimate (Tourangeau et al., 2000; Tourangeau & Yan, 2007). The way interviewers handle such situations can influence the respective item nonresponse rates.

Moreover, interviewers can affect the substantive answer a respondent gives during the interview or the result of a test. This topic is very complex, and interviewer effects vary for different measurements (Schaeffer, et al., 2010). Even the presence of an interviewer, interviewers' observable characteristics, and their actions during the interview can influence

the answers that respondents provide in a survey (Groves et al., 2009).

7.3 The Wave 7 interviewer survey

The SHARE interviewer survey was implemented as an online survey and was based on the conceptual framework developed by Blom and Korbmacher (2013) that distinguishes four dimensions of interviewer characteristics as possible sources of interviewer effects: interviewer attitudes (1), interviewers' own behaviour (2), interviewers' experience (3), and interviewers' expectations (4). In addition to basic demographics, information about attitudes towards surveys in general and about expectations and experiences with specific SHARE modules were collected. Additionally, hypothetical questions of how interviewers would behave as a SHARE respondent were posed.

The interviewer survey was coordinated centrally at the Munich Center for the Economics of Aging (MEA). In sum, 22 SHARE countries participated in the SHARE Wave 7 interviewer survey: Austria (AT), Belgium (BE), Bulgaria (BG), Croatia (HR), Cyprus (CY), Denmark (DK), Germany (DE), Estonia (EE), Finland (FI), Greece (GR), Hungary (HU), Italy (IT), Latvia (LV), Lithuania (LT), Luxembourg (LU), Poland (PL), Portugal (PT), Romania (RO), Spain (ES), Slovenia (SI), Slovakia (SK), and Sweden (SE). The participation of the countries in the interviewer survey and the participation of the interviewers were voluntary and confidential. In most countries, participation was compensated by a small conditional incentive. The amount of the incentive varied between 5 and 10 euros.

We invited interviewers to participate at the end of the national interviewer training sessions. The invitation letters were distributed randomly to interviewers and included the web link to the survey and a unique login code. The interviewers were asked to answer the survey before the beginning of the Wave 7 fieldwork to ensure that their answers were unaffected by their first experiences in the field. To link the interviewer survey data with the SHARE survey data, we asked the interviewers to provide their SHARE interviewer ID at the end of the interviewer survey. The number of interviewers working in each country and participation in the interviewer survey differed among countries (see Table 7.1). In most countries, not all interviewers who participated in the training session also conducted SHARE interviews. For practical reasons, we only refer to interviewers who participated in the national training session and worked later as a SHARE interviewer. Column 2 of Table 7.1 refers to that number and summarises how many interviewers per country worked for the seventh wave of SHARE. The participation rate, which is calculated by the number of interviewers who participated in the SHARE interviewer survey divided by the number of interviewers who conducted interviews for SHARE, varied greatly among countries, ranging from 100 percent in Italy, Poland, Portugal and Slovenia to 33 percent in Slovakia.

Table 7.1: Participation in the interviewer survey

Country	Interviewers in SHARE	Participation interviewer survey	Participation rate (in %)
Austria	73	62	85
Belgium	108	78	72
Bulgaria	108	99	92
Cyprus	19	14	74
Germany	144	124	86
Denmark	61	59	97
Estonia	92	41	45
Spain	123	114	93
Finnland	41	28	68
Greece	118	41	35
Croatia	85	66	78
Hungary	71	35	49
Italy	138	138	100
Lithuania	58	46	79

Country	Interviewers in SHARE	Participation interviewer survey	Participation rate (in %)
Luxembourg	31	22	71
Latvia	37	20	54
Poland	68	68	100
Portugal	10	10	100
Romania	87	71	82
Sweden	64	56	88
Slovenia	52	52	100
Slovakia	72	24	33

Note: Two different survey agencies conducted interviews in Belgium (Flemish- and French-speaking) and Spain; the results shown in this chapter combine the results of both surveys.

7.4 First results: Interviewer characteristics overall, and comparisons within and between countries

Overall, approximately half of the interviewers are aged between 40 and 59, and 17 percent are younger, i.e., between 19 and 39 years, and one-third is 60 years or older. However, the age distributions across countries differ (Figure 7.2). Denmark has the oldest interviewer population, with 95 percent of the interviewers being 60 years or older. In contrast, only one of Portugal's ten interviewers who provided age information in the survey is in the oldest age group. Half of the Portuguese interviewers are between 19 and 39 years old.

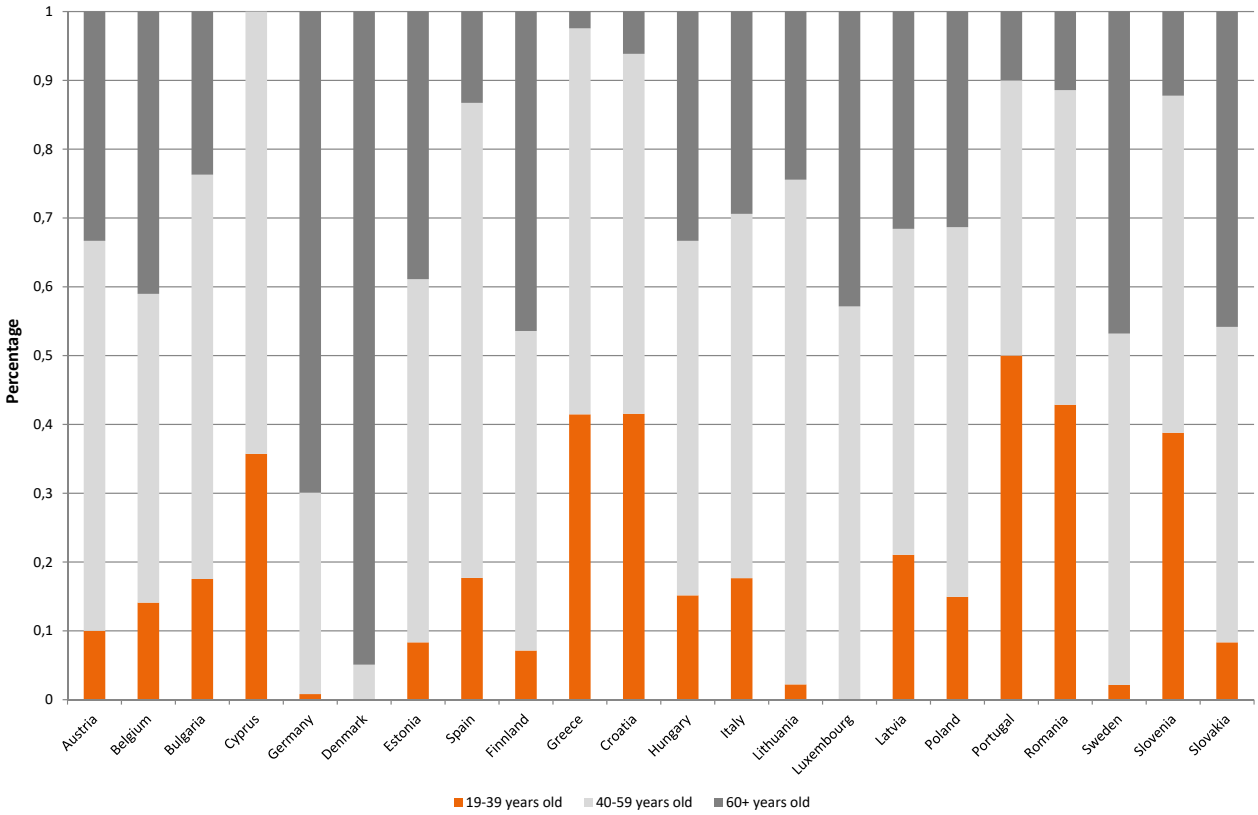


Figure 7.2: Age distribution of interviewers by country

The SHARE interviewers are predominantly female (72 percent). The only countries that have a larger share of males than females in their interviewer population are Denmark and Germany (both approximately 52 percent male). Upon closer examination of the distribution of interviewer ages for both genders, we see that among women, the largest group (54 percent) is in the age range from 40 to 59 years, while among men, the largest group (43 percent) is 60+. These patterns across age, gender, and countries could lead to differential interviewer effects (tables not shown here), since respondents may interact differently with interviewers of different gender and age groups. Even more important for the interaction between interviewer and respondent, and hence for the quality of survey data, are attitudes of the interviewers towards their job in general and towards the interview process. In the following, we will therefore provide an overview of the responses to two item batteries from the interviewer survey, namely, "reasons for being an interviewer" and "statements about interviewer behaviour".

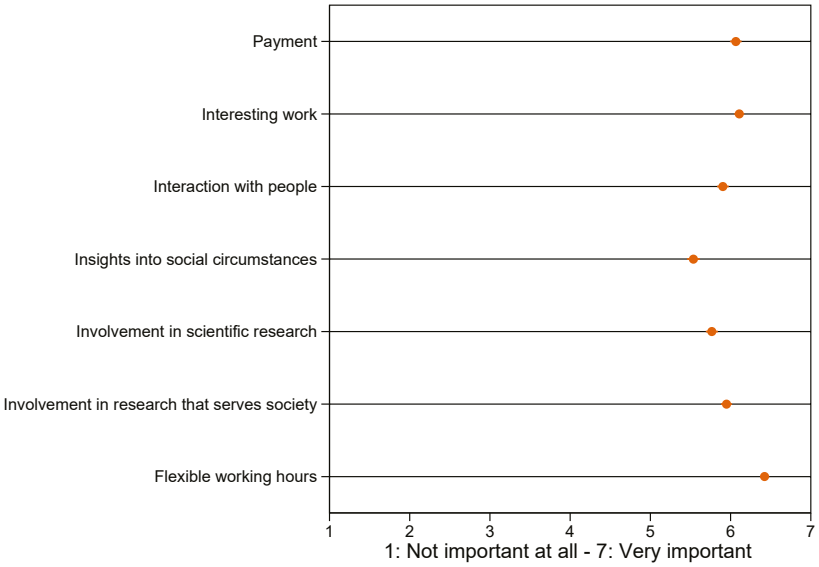


Figure 7.3: Means of the importance of reasons for being an interviewer

Interviewers do not vary much in the reported importance of the different reasons for being an interviewer (Figure 7.3). For all items, the averages across interviewers are above five, i.e., all reasons are rather important to them. "Flexible working hours", "interesting work", and "payment" are rated highest, with averages above six on the 7-point scale. In contrast, "insight into social circumstances" seems to be the least important reason, albeit still with an average score of 5.5.

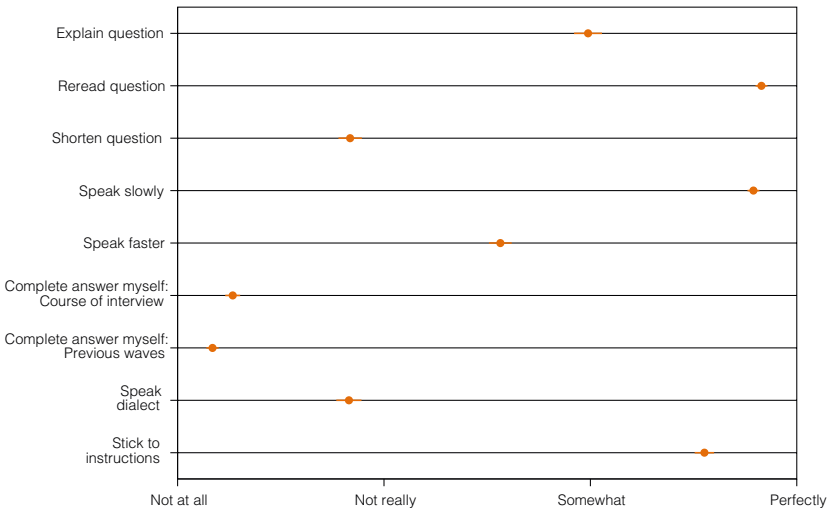


Figure 7.4: Mean of agreements with self-reported interviewer behaviour

More variation was observed within the self-reported interviewer behaviour in comparison to the reasons for being an interviewer. The average agreement with different statements about interviewer behaviour on a 4-point scale is plotted in Figure 7.4. The agreement is particularly high with statements that aim at behaviour addressed in interviewer training: "reread question", "speak slowly", and "stick to instructions". Agreement is lower with regard to behaviour that is possibly detrimental to the respondents' understanding of the questions, i.e., "shorten question", "speak faster", "speak dialect". Behaviour that would constitute a violation of appropriate scientific practice and possibly the labour contract is not condoned. Interestingly, "explain the question" received an average agreement of "somewhat". Interviewers explaining questions to respondents who have problems understanding them is a viable alternative to strictly standardised interviewing approaches in the literature on conversational interviewing (see, e.g., Schober & Conrad, 1997). The interviewers agreeing to this kind of behaviour, at least to a certain extent, suggests that they might be explaining questions during the interview, although it conflicts with their standardised interviewing instructions.

The main reason to examine these interviewer characteristics and attitudes is to gain insights into how and to what extent they may produce interviewer effects. West and Blom (2017) provide an overview of research on interviewer effects. They conclude that interviewers can influence the data process at various steps. They claim that the background characteristics of interviewers, such as age, gender, and experience, can affect the data and that psychometric characteristics, attitudes and expectations of interviewers may mediate these relationships. However, the results in the literature are mixed because the survey design or questions and the respondent moderate the relations.

We provide preliminary ideas of how researchers can use the SHARE data to contribute to the existing literature on interviewer effects, and we investigated the interviewer variance in three SHARE variables:

1. HH017 – household income as a dummy variable indicating the respondents' nonresponse
   - Approximately fifteen percent of respondents did not report their exact income (i.e. without accounting for estimations in the corresponding bracket question) in the analysis sample
   - Sample size: 28,658

2. PH003 – self-reported health
   - Wording: "Would you say your health is …"
   - Response categories: 1 Excellent; 2 Very good; 3 Good; 4 Fair; 5 Poor
   - Mean: 3.3; standard deviation: 1.1
   - Sample size: 42,332

3. HS003 – retrospective health in childhood
   - Wording: "Would you say that your health during your childhood was in general…"
   - Response categories: 1 Excellent; 2 Very good; 3 Good; 4 Fair; 5 Poor
   - Mean: 2.2; standard deviation: 1.0
   - Sample size: 34,596

For each item, we first estimated a random effects multilevel model – with individual respondents at the lower level and interviewers at the higher level – as a baseline model to assess the amount of variance in the responses that can be attributed to the interviewers. In a second step, we included a set of variables at the interviewer level in the model to account for the interviewers' influence and thus reduce the amount of interviewer variance. The following interviewer characteristics were included in our full models: gender, age, education, agreement with interviewer behaviour, and reasons for working as an interviewer. In addition, we included the Big Five personality traits to control for psychometric characteristics at the interviewer level.

In the income item nonresponse model, the percentage of variance attributable to the interviewer was estimated at 46.4 percent in the baseline model. Including all interviewer characteristics, this percentage declined slightly to 43.3 percent. For the subjective health item, the percentage of variance attributable to the interviewer could be reduced from 14.2 percent in the baseline model to 13.1 percent in the model with interviewer-level predictors. An even smaller reduction is found for the items on health during childhood. The percentage of variance attributable to the interviewer is 17.0 percent in the baseline model and could be reduced to 16.3 percent in the full model. Overall, the interviewer characteristics exhibit an influence on the responses (results not shown here), but they only reduce the interviewer variance marginally. Hence, other interviewer characteristics seem to be more important in explaining and reducing the random error introduced by the interviewer.

These results are only very preliminary, but we think that they provide a good first impression of potential analyses of interviewer effects in SHARE using the SHARE interviewer survey data. The proper investigation of interviewer effects in SHARE would definitely require more thorough analyses. Researchers would have to provide far more detail on understanding and modelling the specific processes with regard to types of questions, interviewer characteristics, respondent characteristics, interactions, etc. (see West & Blom, 2017 for a recent overview on the subject).

7.5 Concluding remarks

Interviewers are an important actor within the process of conducting a survey, and they can influence the data quality of a survey. Our SHARE interviewer survey offers the opportunity to analyse and understand interviewer effects. The descriptive comparison of interviewer characteristics shows the variation among interviewers within one country and among countries. This finding is an important prerequisite to the identification of interviewer characteristics that can explain interviewer effects. The data of the interviewer survey will be released soon. Information on how to obtain access to the data will then be available on the SHARE homepage.

References

Blom, A. G., De Leeuw, E., & Hox, J. J. (2011). Interviewer Effects on Nonresponse in the European Social Survey. Journal of Official Statistics, 27(2), 359-377.

Blom, A. G., & Korbmacher, J. M. (2013). Measuring interviewer characteristics pertinent to social surveys: A conceptual framework. Survey Methods: Insights from the Field, 1-16. https://doi.org/10.13094/SMIF-2013-00001.

Durrant, G. B., Groves, R. M., Staetsky, L., & Steele, F. (2010). Effects of interviewer attitudes and behaviors on refusal in household surveys. Public Opinion Quarterly, 74(1), 1-36. http://dx.doi.org/10.1093/poq/nfp098.

Groves, R. M., & Couper, M. P. (1998). Nonresponse in household interview surveys. New York: John Wiley & Sons. Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). Survey methodology (2nd ed.). Hoboken: John Wiley & Sons.

Hox, J., & De Leeuw, E. (2002). The influence of interviewers' attitude and behavior on household survey nonresponse: an international comparison. In R. M. Groves, D. A. Dillman, J. L. Eltinge & R. J. A. Little (Eds.), Survey nonresponse (pp. 103-120). New York: John Wiley & Sons.

Jäckle, A., Lynn, P., Sinibaldi, J., & Tipping, S. (2013). The effect of interviewer experience, attitudes, personality and skills and attitudes on respondent co-operation with face-to-face surveys. Survey Research Methods, 7(1), 1-15. http://dx.doi.org/10.18148/srm/2013.v7i1.4736.

Lipps, O., & Pollien, A. (2011). Effects of interviewer experience on components of nonresponse in the European Social Survey. Field Methods, 23(2), 156-172. https://doi.org/10.1177/1525822X10387770.

Pickery, J., & Loosveldt, G. (2001). An exploration of question characteristics that mediate interviewer effects on item nonresponse. Journal of Official Statistics, 17(3), 337-350.

Pickery, J., Loosveldt, G., & Carton, A. (2001). The effects of interviewer and respondent characteristics on response behavior in panel surveys. Sociological Methods & Research, 29(4), 509-523. https://doi.org/10.1177/0049124101029004004.

Schaeffer, N. C., Dykema, J., & Maynard, D. W. (2010). Interviewers and interviewing. In P. V. Mardsen & J. D. Wright (Eds.), Handbook of survey research (2nd ed.) (pp. 437-470). Bingley: Emerald Group Publishing Limited.

Schober, M. F., & Conrad, F. G. (1997). Does conversational interviewing reduce survey measurement error? Public Opinion Quarterly, 61(4), 576-602. https://www.jstor.org/stable/2749568.

Singer, E., Frankel, M. R., & Glassman, M. B. (1983). The effect of interviewer characteristics and expectations on response. Public Opinion Quarterly, 47(1), 68-83. https://doi.org/10.1086/268767.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). The psychology of survey response. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511819322.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. Psychological Bulletin, 133(5), 859-883. https://doi.org/10.1037/0033-2909.133.5.859.859.

West, B. T., & Blom, A. G. (2017). Explaining interviewer effects: A research synthesis. Journal of Survey Statistics and Methodology, 5(2), 175-211. https://doi.org/10.1093/jssam/smw024.

CHAPTER 8

Monitoring and managing SHARE fieldwork

08

8 MONITORING AND MANAGING SHARE FIELDWORK

8.1 Fieldwork monitoring and survey participation in the seventh wave of SHARE

Gregor Sand – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC)

8.1.1 Introduction

SHARE Wave 7 was the first wave to include 28 countries. SHARE Central undertook the ambitious effort to integrate the following eight new countries: Bulgaria, Cyprus, Finland, Lithuania, Latvia, Malta, Romania, and Slovakia (see also Chapter 4). Apart from this effort, after two waves of absence, Hungary obtained sufficient funding to join fieldwork again. The Netherlands could participate in SHARE again by conducting the survey via the Internet (see Das et al., 2017 for a description of this experiment in Wave 6). In addition to these changes, the following chapter is a continuation of the chapter about fieldwork monitoring in SHARE Wave 6 (see Malter & Sand, 2017), with all numbers and statistics adapted to the countries of Wave 7. Its conceptual basis was developed in the run-up to Wave 5 and is outlined in Kneip et al. (2015). Again, all indicators were conceptualised strictly in accordance with the 9th edition of standards set by the American Association for Public Opinion Research (AAPOR, 2016). Through this approach, we could report at any point in time what the response and retention rates[14] would be if fieldwork was terminated at that given moment. We are convinced that ensuring data quality must be a key concern of any population-level survey study while emphasising all the major components of the total sampling error, as described in Kneip et al. (2015).

8.1.2 Fieldwork periods and survey agencies

The integration of eight new countries was especially challenging in Wave 7. The countries were divided into two groups regarding their start of fieldwork to avoid delays caused by peak workload due to releasing 28 national questionnaire instruments by SHARE Central within a short time. Most countries of Wave 7 were able to put the planned schedule into action. Figure 8.1 shows that Wave 7 happened largely synchronously across countries. It can be seen that by the middle of March 2017, most countries in Group 1, consisting of all countries that had already participated in Wave 6, had completed their interviewer trainings (green squares) and had already conducted their first interviews. Notable exceptions were Austria and Greece, which show substantial delay (one month and more) between obtaining their sample software and delivering the first interview. One other notable exception was Portugal, which experienced issues with securing funding, thus precluding them from starting on time with the other countries in their group. Delays of this kind make the proper execution of fieldwork quite difficult. Most of the countries in Group 2, consisting of all countries that took part in SHARE during Wave 7 for the first time, started their fieldwork by the middle of April 2017.

14 In the following text, we differentiate between the terms "response" and "retention". We refer to response rates whenever we look at the first response of a unit (household or individual) in a baseline or refreshment sample, while we refer to retention rates when we analyse response behaviour in the longitudinal sample.

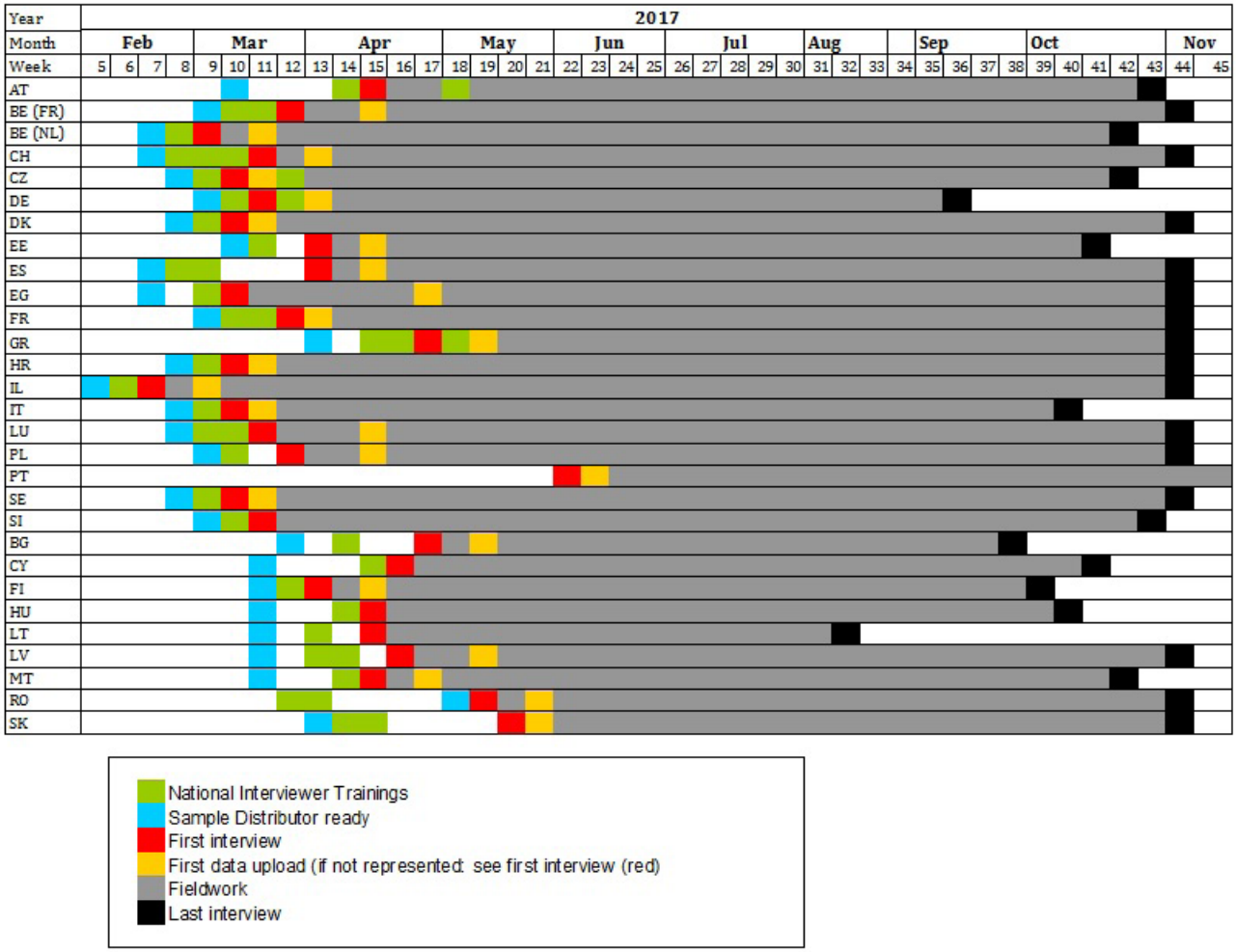


Figure 8.1: Fieldwork periods of SHARE Wave 7

The data collection of SHARE Wave 7 was terminated on 31 October 2017. The last export of data from the survey agencies to SHARE Central (via CentERdata) occurred on 2 November 2017. No new interview data could be accepted after that date. Some countries had completed fieldwork much earlier. The organisations in Table 8.1 below conducted the fieldwork in each wave. There has been high stability of contracted survey agencies over time in most countries.

Table 8.1: Survey agencies from Wave 1 to Wave 7 of countries participating in Wave 7

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7
AT	IMAS	same	IFES	same	same	same	same
BE-FR	PSBH, Liège Univ.	same	same	same	CELLO - Antwerp Univ.	same	same
BE-NL	PSBH, Antwerp Univ.	same	CELLO - Antwerp Univ.	same	same	same	same
BG	-	-	-	-	-	-	GfK
CH	MIS Trend	LINK	same	same	same	same	same
CY	-	-	-	-	-	-	RAI
CZ	-	SC&C	same	same	same	same	same
DE	infas GmbH	same	same	same	TNS Infratest	same	same (Kantar)
DK	SFI-Survey	same	same	same	same	same	same
EE	-	-	-	Statistics Estonia	GfK	Statistics Estonia	same
ES	TNS Demoscopia	same	same	same	same (incl. region of Girona)	same	same (Kantar)
FI	-	-	-	-	-	-	Taloustutkimus
FR	INSEE	same	same	INSEE (panel)/ GFK-ISL (refresh.)	GFK-ISL	GFK-ISL	Kantar
GR	Kapa Research	same	same	-	-	Kapa Research	same
HR	-	-	-	-	-	GfK	same
HU	-	-	-	Tarki	-	-	Tarki
IL	Cohen Institute, Tel Aviv Univ.	same	-	-	Cohen Institute, Tel Aviv Univ.	same	same
IT	DOXA S.p.A.	same	same	same	IPSOS	same	same
LT	-	-	-	-	-	-	Kantar
LU	-	-	-	-	CEPS/INSTEAD	LISER	same
LV	-	-	-	-	-	-	ISR
MT	-	-	-	-	-	-	EMCS
NL	TNS	same	same	same	same	CentERdata	same
PL	-	TNS-OBOP	same	same	TNS Polska	same	same (Kantar)
PT	-	-	-	GfK Metris	CECS, University of Minho	same	DOMP
RO	-	-	-	-	-	-	GfK
SE	Intervjubolaget IMRI	same	same	same	same	IPSOS Observer Sweden	same
SI	-	-	-	CJMMK	same	IPSOS	same
SK	-	-	-	-	-	-	GfK

8.1.3 Monitoring fieldwork

This section includes information about the classification and computation of survey outcomes and all final rates and figures of Wave 7 based on the last data export on 2 November 2017. All numbers and figures reported during fieldwork are based on information from the SHARE Sample Management System (SMS), which is the interviewer software used to document contact attempts and conduct the interview. As of Wave 7, all SMS data have been routinely cross-checked against interview data already during fieldwork. The separation between baseline/refreshment samples and panel samples known from the monitoring reports are applied to this chapter as well. All indicators are graphed over calendar weeks to visualise each country's progress of fieldwork over time. Final rates and interview numbers are then provided again in a final summary graph without trajectories to allow for easier comparison among countries.

8.1.3.1 Classification of survey outcomes

Identically to Waves 5 and 6, most representational indicators (i.e., those on unit nonresponse) were set out as quality targets in the specifications of the model contract of SHARE Wave 7. As in previous waves, we follow the newest edition of AAPOR guidelines and use data from the SHARE SMS to classify the baseline/refreshment and longitudinal gross samples[15] of each country into exhaustive and mutually exclusive categories reflecting the survey outcomes for each sample type. All contact information entered by interviewers into the SMS is continuously converted into a so-called "household state". The algorithm, which creates the household state, divides the sample into three mutually exclusive categories: (1) ineligible households, (2) eligible households, and (3) households of unknown eligibility.[16] This algorithm is performed in a hierarchical way. Once the eligibility status is determined, a new contact code cannot revert the eligibility status into "unknown" anymore. For the sake of completeness, we repeat the same basic concepts presented in Kneip et al. (2015): if a household is classified as ineligible, this is a "final state" that permanently closes a case (i.e., no more actions can be performed by interviewers). The same concept applies to sorting households into subcategories of the household state. A new contact results in a change of the household state only if it involves new information that conceptually trumps the previous information. For example, a household formerly classified as "non-contact" (NC) would be switched to "refusal" (R) if the interviewer establishes a successful contact, but the respondent refuses to participate. However, if the interviewer does not reach anyone ("non-contact") in an attempt to convert a previous refusal, the household state remains "R". The hierarchical order of the nexus contact code-household state is shown in Table 8.2.

15 Baseline/refreshment samples consist of respondents who participate in a regular SHARE interview for the first time. They are completely new to SHARE or have participated in a SHARELIFE interview for the first time. Panel or longitudinal samples comprise respondents who have already participated in a baseline or refreshment interview.
16 For details on SHARE's target population and eligibility criteria, see Kneip (2013) and Bergmann et al. (2017).

Table 8.2: Detailed list of SMS entries and fieldwork outcomes at the household level

SMS Contact Protocol Entry	Household State
Ineligible	NE
Deceased[3]	
In hospital[3]	
In nursing home[4]	
In prison	
Moved abroad	
Language barriers	
Moved, new address unknown[3]	
Address non-existent, house vacant[3]	
No eligible persons after CV	
Household screened as ineligible[5]	

SMS Contact Protocol Entry	Household State
Eligible	E
Completed interview (incl. end-of-life interview)	CI
Partial interview	PI
Interrupted interview	II
Refusal[1]	R
Too busy, no time	
Too old, poor health conditions	
No interest, against surveys	
Other reasons	
Other non-interview	O
Contact, no appointment	
Contact, appointment for another contact	
Contact, appointment for interview	
Deceased[3]	
In hospital[3]	
In nursing home[4]	
Moved, new address known	
Moved, new address unknown[3]	
Address non-existent, house vacant[3]	
Household screened as eligible	
Non-contact2	NC
Unknown Eligibility	UE
Screening refusal	UE_R
Other screening non-cooperation	UE_O
Screening non-contact	UE_{NC}
No contact attempted	UE_{NCA}

Notes:
1 For each category, interviewers could distinguish between a "soft" and a "hard" refusal, with the latter type requiring intervention from the agency. Neither of the refusal codes set by the interviewer closed a case.
2 Non-contact for the eligible proportion of the sample does not apply to the baseline/refreshment sample in the Czech Republic.
3 This circumstance led to ineligibility only in the baseline/refreshment sample, but not in the longitudinal sample.
4 Whether this circumstance led to ineligibility in the baseline/refreshment sample depended on a country's sampling frame. In the longitudinal sample, institutionalised cases were always considered eligible.
5 Subcategories are age, ineligible household, problems with phone, address non-existent, and language barriers.

8.1.3.2 Formulas to compute survey outcomes

Apart from eligibility, the household state variable provides information about a household's contact and cooperation status. Table 8.3 reports which fieldwork indicators are used and how they are computed based on the household state. Because the current state can be determined by the SMS for every household at any given point in time, we are able to report the state of fieldwork at any time as if it were completed. In terms of household cooperation, households are considered to be participating if at least one eligible household member is successfully interviewed. With respect to individual cooperation, several definitions of individual response rates are possible depending on how households with unknown eligibility are treated and how the number of eligible households with unknown composition is determined. These households may or may not contain eligible individuals. Different assumptions about their number directly affect the denominator of the response rate. In general, we assume that only a fraction p of the households with unknown eligibility are in fact eligible; we estimate this fraction by "E" /"E+NE" . Over the course of fieldwork, this estimate improves in precision as the non-attempted part of the sample declines.

The number of eligible persons is only known for households with a completed coverscreen interview (CV). Based on the assumption that, in each country, the average number of eligible persons in households without a CV does not systematically differ from that in households with a CV, we take the latter as an estimate for the baseline or refreshment samples. For households in the longitudinal sample without a CV, we can use preload information on the household composition to assess the number of eligible respondents. Here, the assumption is that this number has not changed since the last interview. By estimating the average number of eligible respondents $\bar{n}$ in a specific sample, the total number of eligible respondents – and thus the denominator of the individual response rate – is $\bar{n}(E+pUE)$.

Table 8.3: Outcome rate formulas

Estimated proportion of eligible households	$p=\frac{E}{E+NE}$
Percentage of households attempted	$\frac{(CI+PI+R+II+O+NC)+(UE_R+UE_O+UE_{NC})+NE}{GS}$
Household contact rate (AAPOR CON2)	$\frac{(CI+PI+R+II+O)+p(UE_R+UE_O)}{E+p\cdot UE}$
Household cooperation rate (cf. AAPOR COOP2)[1]	$\frac{(CI+PI)}{(CI+PI+R+II+O)+p(UE_R+UE_O)}$
Household response rate (AAPOR RR4)	$\frac{(CI+PI)}{E+p\cdot UE}$
Household refusal rate (AAPOR REF2)	$\frac{R+II+p(UE_R)}{E+p\cdot UE}$
Household other non-interview rate (AAPOR ONI2)	$\frac{O+p(UE_O)}{E+p\cdot UE}$
Individual response rate[2]	$\frac{(CI_r+PI_r)}{\bar{n}(E+p\cdot UE)}$
Individual response rate in subsample i[3]	$\frac{(CI_i+PI_i)}{\bar{n}_i(E+p\cdot UE)}$

1 $p(UE_R+UE_O)$ is not part of the denominator in AAPOR COOP2. The calculation method was adapted for equation RR=CON×COOP to hold.

2 $\bar{n}$ is the average number of eligible persons per household. For baseline/refreshment sample, $\bar{n}$ is estimated based on households with completed coverscreen. For the longitudinal sample, information on house hold composition is available for all households from the previous wave. CI_r and PI_r refer to the number of completed and partially completed interviews, respectively.

3 $\bar{n}_i$ is the average number of eligible persons from subsample i per household, where i = {A,B,C,D}.

8.1.3.3 Baseline/refreshment samples

While the eight new SHARE countries began with a baseline sample in Wave 7, Croatia, Israel, and Poland drew a refreshment sample. Figure 8.2 shows the size of the baseline and refreshment samples. While in Croatia and Israel the refreshment sample was rather small, it was much larger in Poland. Apart from the reasons leading to ineligibility in the longitudinal sample (i.e., incarceration, moving abroad, language barriers), baseline households are also considered ineligible in the following cases: death of the drawn respondent, inpatient treatment during the entire field time, unknown or invalid addresses, and the CV interview yielding no eligible persons in the household. In Bulgaria, Cyprus, Lithuania, Latvia, Romania, and Slovakia, the sample had to be screened for age eligibility first. Hence, ineligibility could also be an outcome of a screening contact. The fraction of ineligible households is the highest in Lithuania, which reflects the lack of the availability of sample frame information on which sampling was based. Households are classified as having "unknown eligibility" after any form of screening nonresponse (non-contact, refusal, other nonresponse). This fraction is also largest in Lithuania because the available sampling frame of addresses does not contain any information on the age of the individuals living at a certain address. However, in most other countries that require screening, it is guaranteed that all selected units (e.g., addresses) contain at least one person over 50 years.

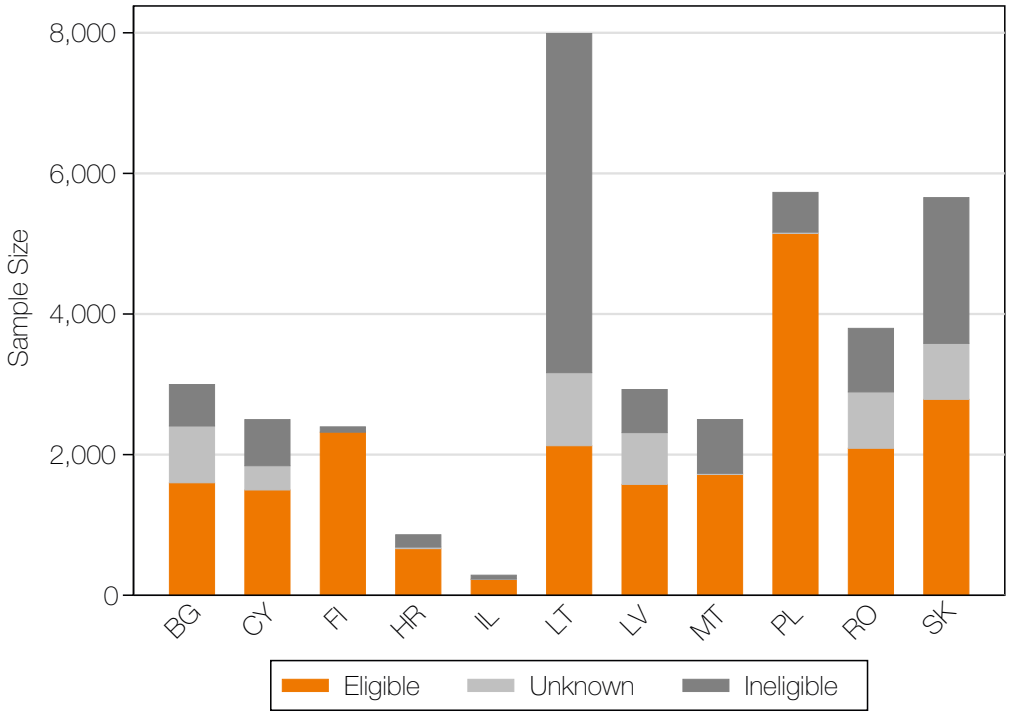


Figure 8.2: Baseline and refreshment samples by classification of sample units

8.1.3.3.1 Contacting households

Figure 8.3 shows the fraction of households of the refreshment/baseline samples in which a contact was attempted (i.e., all households in which either an interviewer reports a contact attempt but was unable to actually contact anybody or a contact is successful). By definition, these criteria include households with one or more conducted interviews.

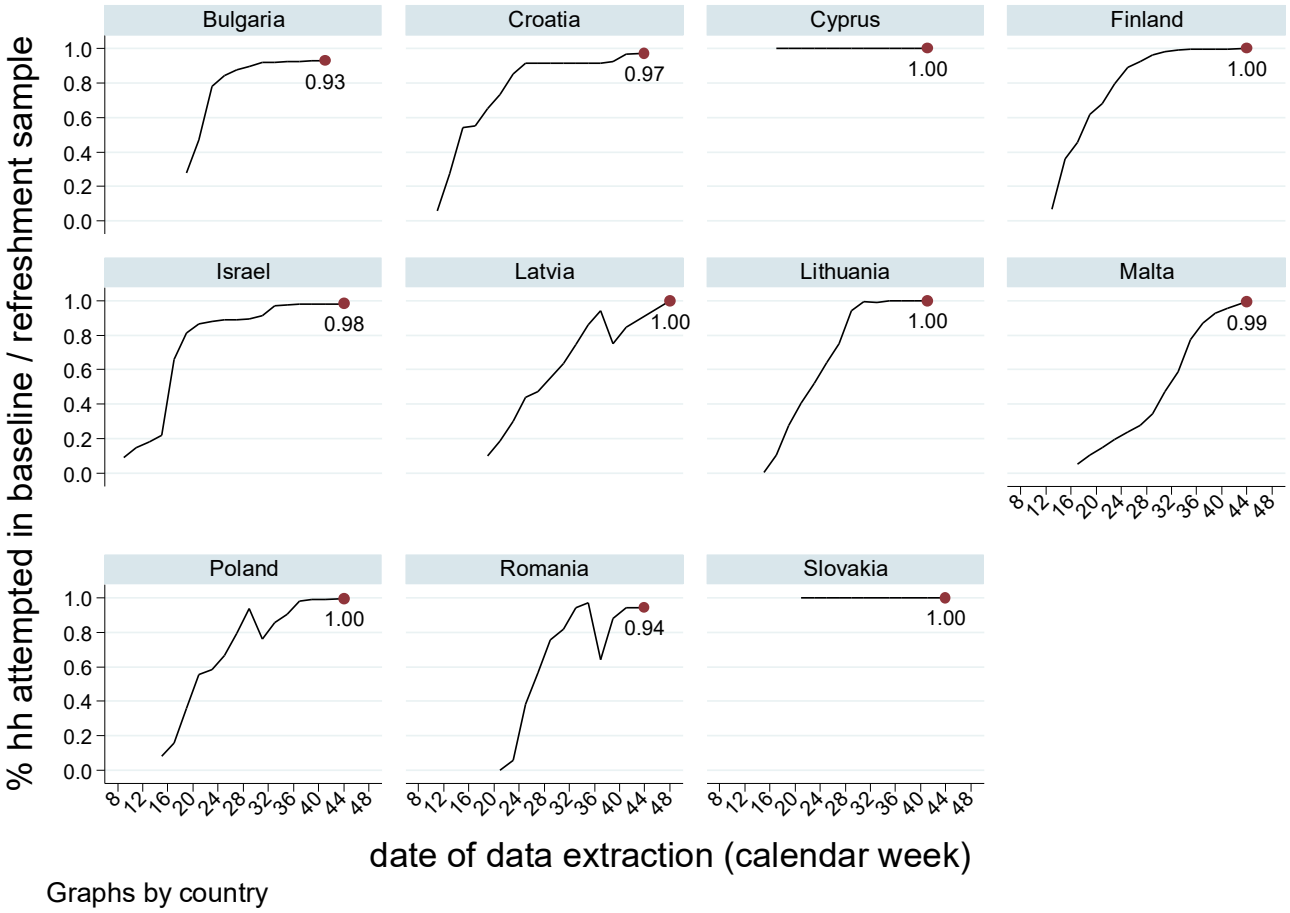


Figure 8.3: Fraction of baseline/refreshment sample households with contact attempts by country over time

In Wave 7, Croatia, Israel, and Poland had a refreshment sample. Among the new countries, only Romania did not completely exhaust its baseline sample in terms of contact attempts. Similar to the panel samples, most countries have a steep increase that levels out over time. Latvia, Poland, and Romania added additional batches in the course of fieldwork. Cyprus and Slovakia screened households for eligibility before starting fieldwork, which is why their contact attempt rates were 100 percent from the start.

Figure 8.4 shows household contact rates broken down by country. This figure contains contact attempts that resulted in an actual contact. By definition, this criterion may also include households with at least one completed interview.

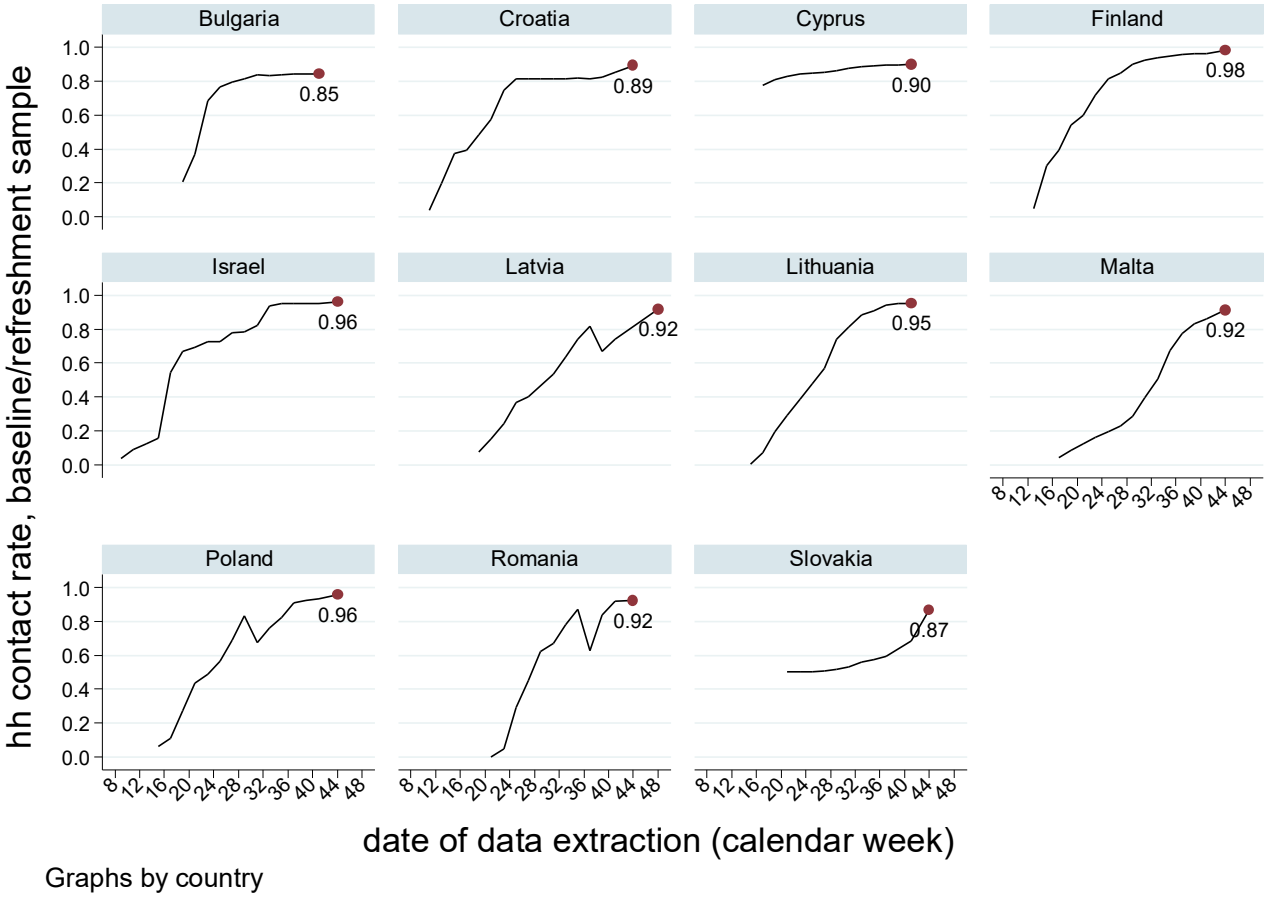


Figure 8.4: Contact rate of baseline/refreshment sample households by country over time

A picture emerges that is similar to the one above. With a contact rate of 85 percent, Bulgaria's interviewers had the most difficulties in contacting all their assigned households successfully. Slovakia had a constantly low contact rate during field-work but could achieve 87 percent by the end of the wave. All other countries are close to or beyond the 90 percent mark.

8.1.3.3.2 Household cooperation and response rate

Figure 8.5 shows the cooperation rate of refreshment/baseline samples by country (i.e., the rate of all contacted households with at least one completed interview).

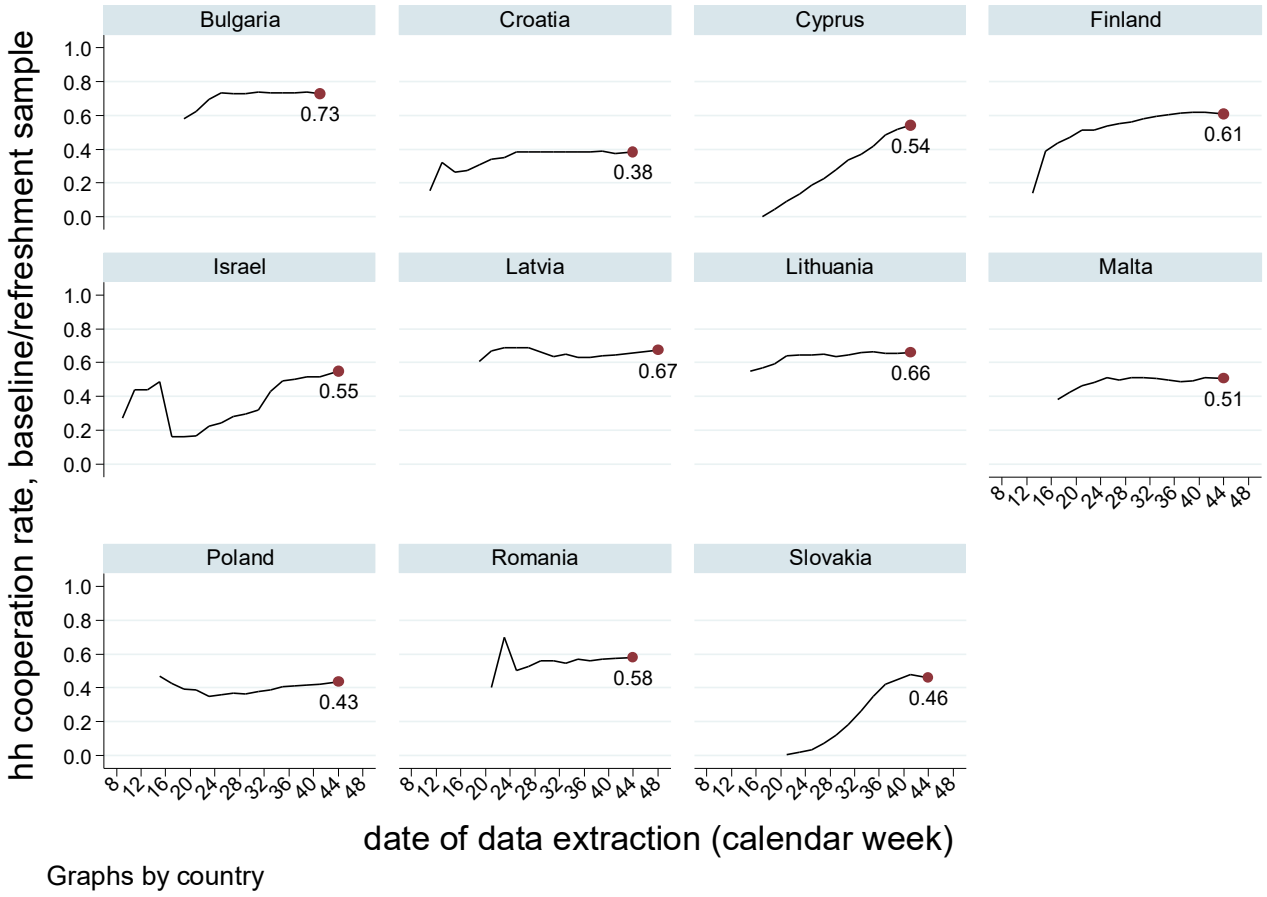


Figure 8.5: Cooperation rate of baseline/refreshment sample households by country over time

Cooperation rates are based only on sample units with a previous contact. The interpretation of cooperation rates becomes more meaningful as contact rates increase. This change is accompanied by a stabilisation of cooperation rates over the field-work period. Indeed, in most countries, the cooperation rates level out after the completion of the contact phase. Cyprus is an exception, with a linear increase in its cooperation rate. Despite having the lowest contact rate, Bulgaria attained the highest cooperation rate in the refreshment sample (73 percent). Among the new countries, cooperation was lowest in Slo-vakia, most likely because of a low contact rate throughout the major part of fieldwork.

Figure 8.6 shows the household response rate (i.e., the number of baseline/refreshment households with at least one complete interview divided by the total number of (estimated) eligible baseline/refreshment households).

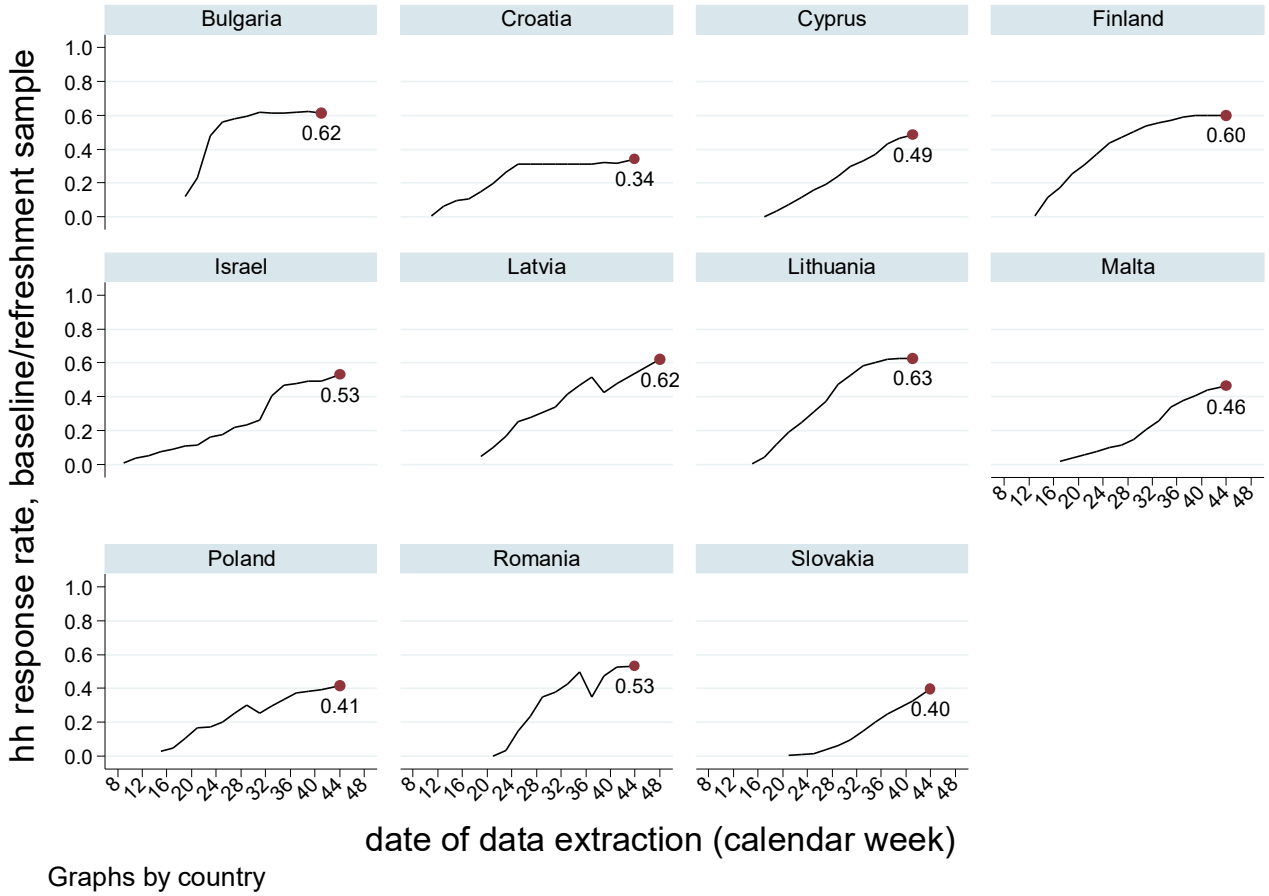


Figure 8.6: Response rate of baseline/refreshment sample households by country over time

Most countries have steadily increasing trajectories. Among the new countries, Bulgaria, Finland, Latvia, and Lithuania reached or surpassed the 60 percent mark. With a household response rate of 34 percent, Croatia is an example of how difficult it is to recover respondents for a second time in a panel survey. The low cooperation in Slovakia translated into the lowest household response among all new countries.

8.1.3.3.3 Individual participation of baseline/refreshment samples

Figure 8.7 shows the individual response rate of baseline/refreshment samples in Wave 7.

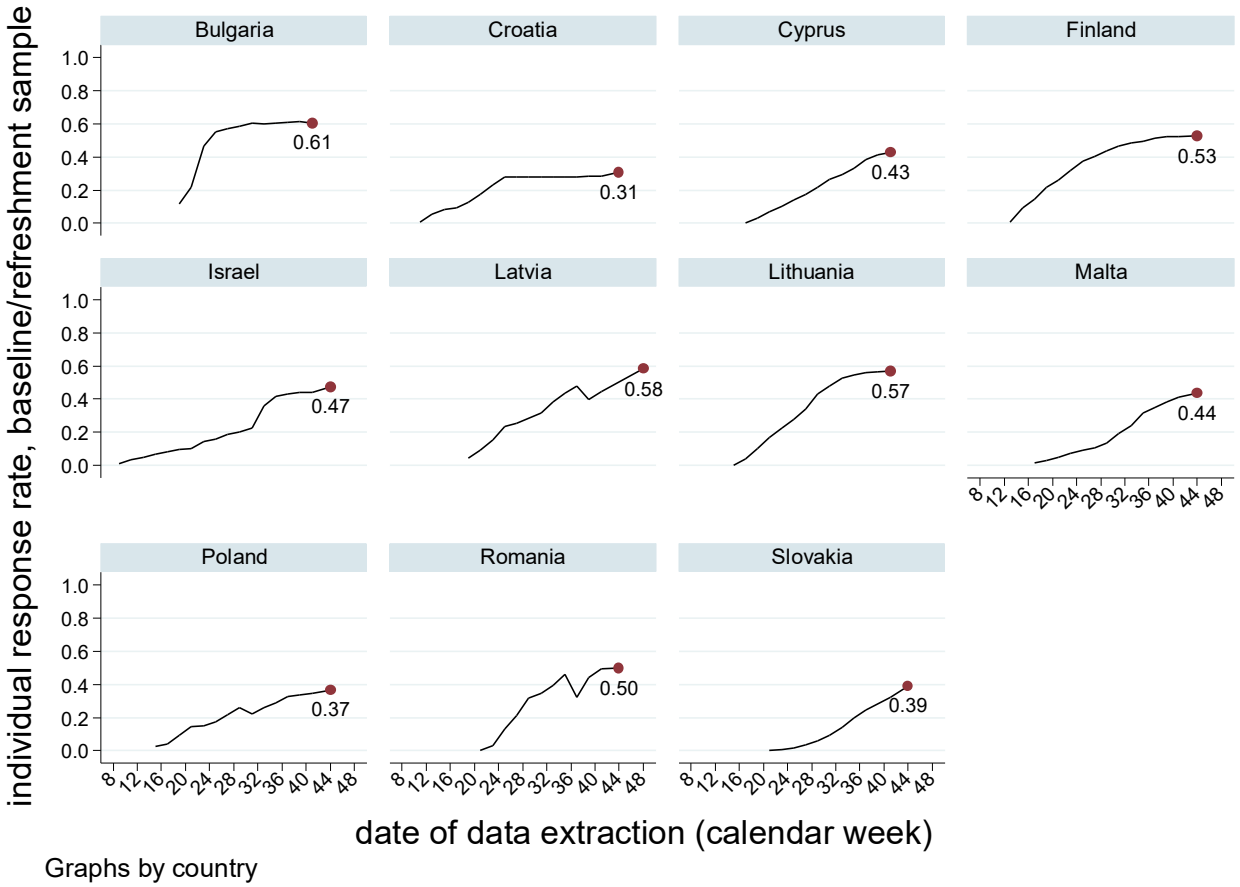


Figure 8.7: Individual response rate of baseline/refreshment respondents by country over time

The trajectories of the individual response rates in all countries are fairly similar to those of the household response rates, with Bulgaria having the highest individual response rate (61 percent). Individual participation is consistently lower than household responses because spouses or partners cannot always be convinced to cooperate. Among the new countries, all except Slovakia surpassed the desired minimum individual response rate of 40 percent.

8.1.3.3.4 Summary of baseline/refreshment samples

Figure 8.8 shows the final household contact, cooperation, and response rates at the end of fieldwork in Wave 7.

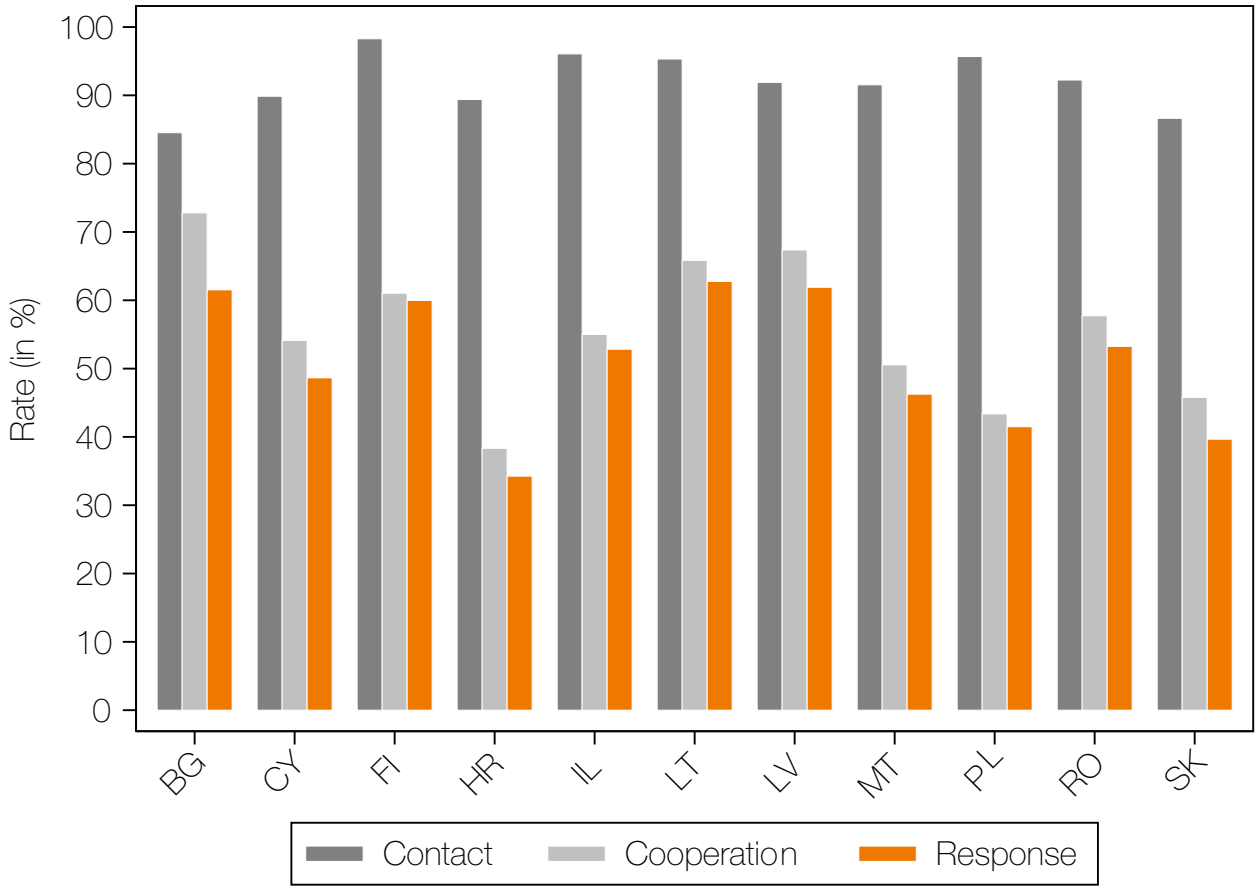


Figure 8.8: Contact, cooperation and response rates for baseline/refreshment samples

It can be seen that contact rates range from approximately 85 to almost 100 percent. From the countries with a refreshment sample (i.e., Croatia, Israel, Poland), Israel has the highest cooperation (55 percent) and response rates (53 percent) at the household level. In the baseline sample, Bulgaria, Lithuania, and Latvia share this position (all beyond 60 percent).

Figure 8.9 shows the final household and respondent-level response rates. The figure illustrates that household and individual response rates vary between approximately 30 and 50 percent among the three countries with refreshment samples. In all new countries with a baseline sample, they range between 40 and 60 percent. As mentioned above, the household response rates are always slightly higher than the individual response rates because of non-cooperation among some household members. The gap between both rates is smallest in Bulgaria and Slovakia, which means that in these countries, interviewers managed to convince the greatest share of all household members.

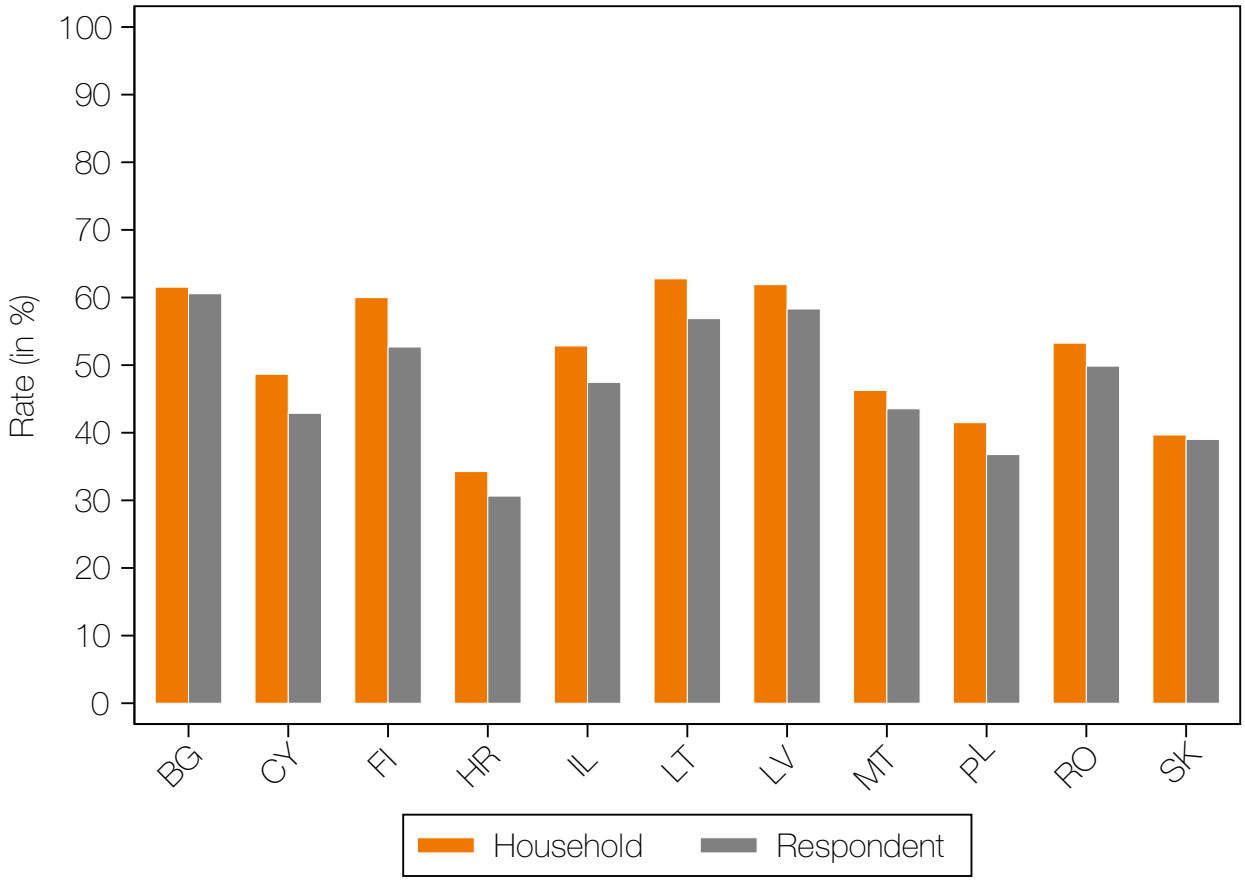


Figure 8.9: Household and respondent-level survey participation in baseline/refreshment samples

Figure 8.10 shows the absolute number of interviews per country in the baseline/refreshment samples at the end of fieldwork. While Croatia and Israel had a fairly small refreshment sample and therefore a considerably small number of interviews (less than 400), Poland had the largest refreshment sample and collected data from almost 3,200 interviews. There is some variation among the new countries, but most of them delivered approximately 2,000 interviews in their baseline samples.

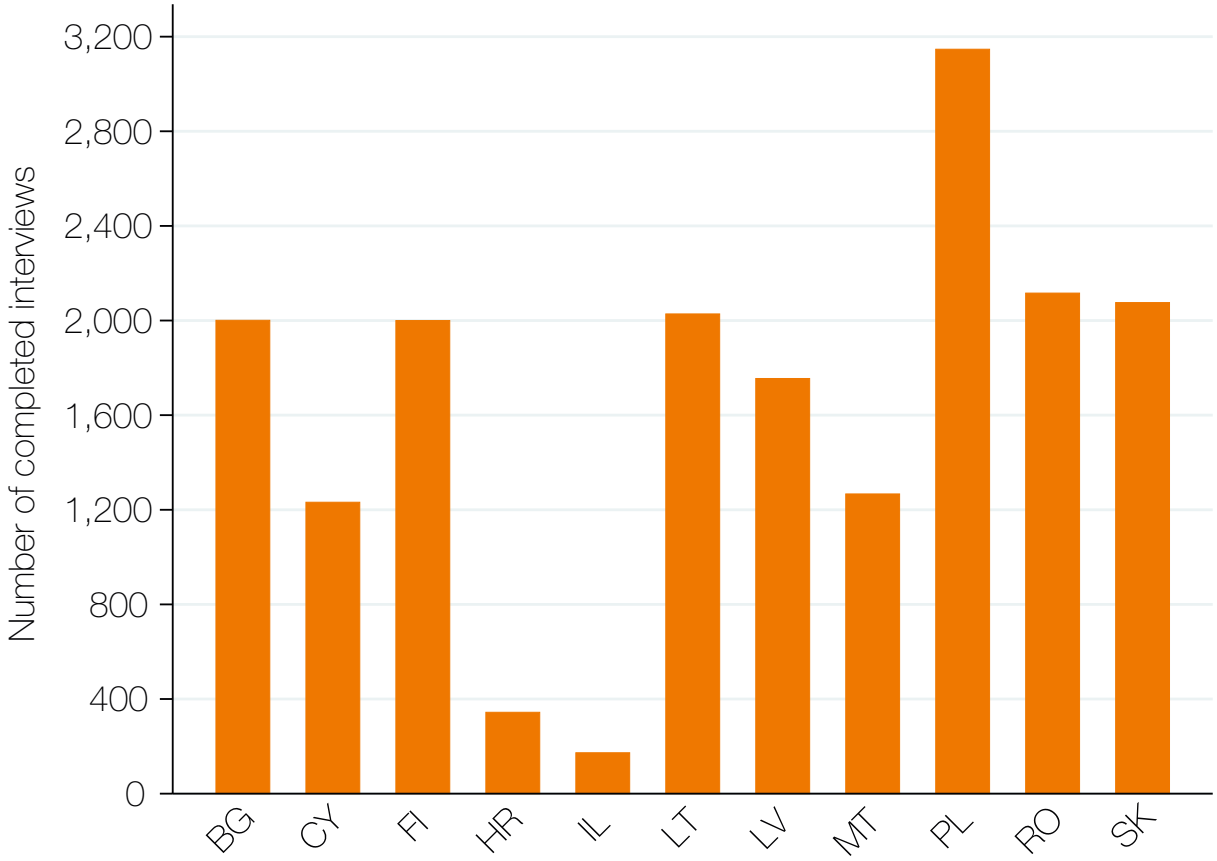


Figure 8.10: Absolute number of interviews in baseline/refreshment samples

8.1.3.4 Panel samples

In general, longitudinal samples can be divided into five subsamples at the individual level according to SHARE's eligibility rules. While Subsample A1 includes all respondents who participated in the previous wave and any other wave of the SHARE survey, Subsample A2 consists of respondents who live in households that participated for the first time in the previous wave (i.e., baseline or refreshment sample). They are usually the ones that take more time and effort to recuperate. Subsample B consists of respondents who have participated in SHARE, but not in the previous wave, and live in a household in which at least one household member participated in the previous wave. Respondents who have participated in any wave, but not in the previous wave, and do not live in a household in which at least one household member participated in the previous wave are subsumed under Subsample C. Finally, Subsample D comprises all missing and new partners who have not yet participated in SHARE.

Response rates are reported separately for these subsamples during fieldwork because sample composition is a crucial aspect particularly when comparing retention rates across countries. Individual-level retention is defined by the proportion of respondents in Subsamples A1 and A2. Additionally, response in Subsamples B and C depends on how well SHARE interviewers manage to bring respondents back who had already dropped out of the study for at least one wave. Finally, response in Subsample D relates to eligible persons in longitudinal households who have never been interviewed before (i.e., either new sample members or eligible sample members who finally participated after refusals in previous waves).

Figure 8.11 shows the size and composition of the longitudinal sample per country in Wave 7. At the household level, the size of the longitudinal gross sample is defined by the number of households with at least one age-eligible respondent interviewed in any previous SHARE wave. For the purpose of fieldwork monitoring, the longitudinal gross sample is determined by the number of households pre-loaded into the SMS. Households that must not be attempted again for legal reasons are dropped. Overall, the longitudinal gross samples of all countries contain almost exclusively eligible cases (98.6 percent).

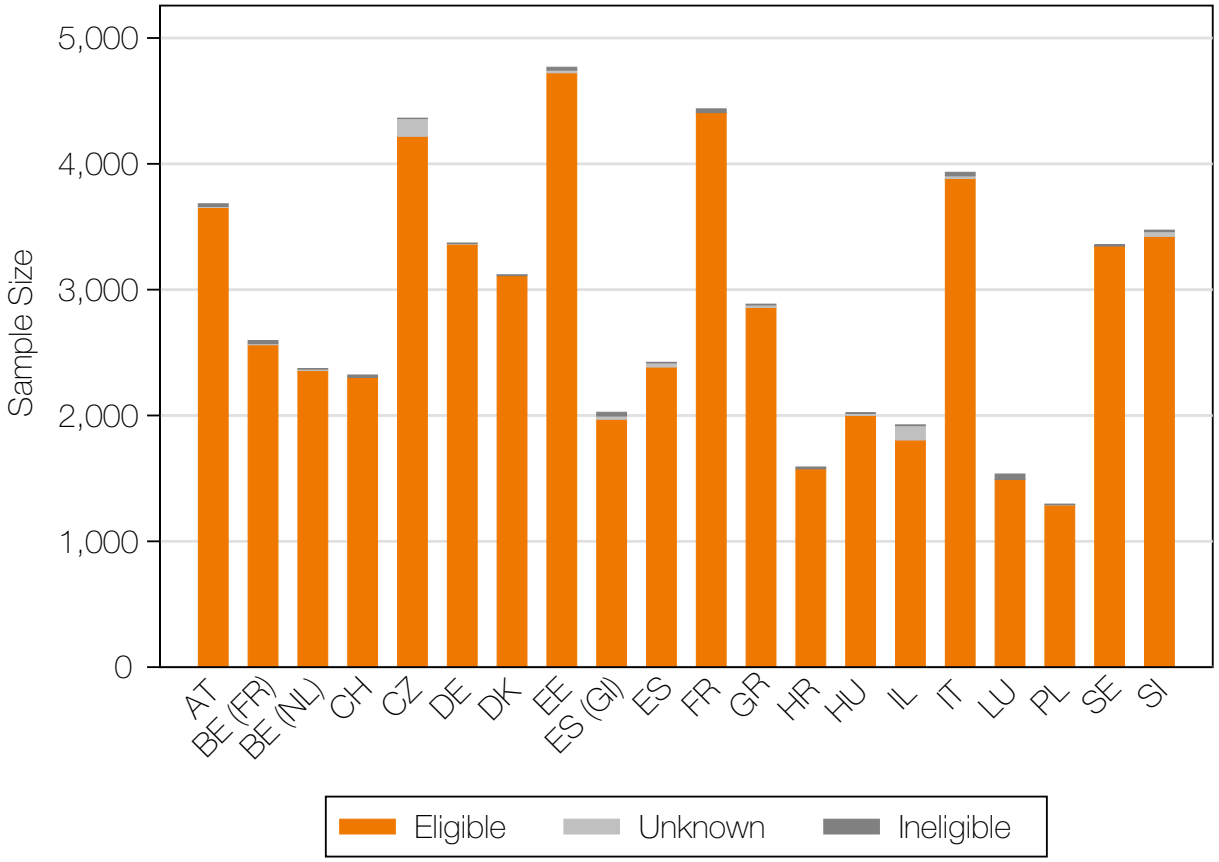


Figure 8.11: Panel samples by classification of sample units

Households in the longitudinal sample can only become ineligible for the following reasons: incarceration, moving abroad, and language barriers. Ineligibility applies to an average of 0.6 percent of all households in the longitudinal samples. Death does not lead to ineligibility. Instead, a proxy respondent is supposed to respond to an end-of-life interview about the deceased person. Households without any contact attempts are considered to be of unknown eligibility. On average and according to what was documented in the SMS, the eligibility of 0.1 percent of all longitudinal households was unknown in Wave 7.

8.1.3.4.1 Contacting households

Figure 8.12 shows the fraction of households in the longitudinal gross sample in which a contact was attempted (i.e., all households in which either an interviewer reports a contact attempt but was unable to actually contact anybody or a contact was successful). By definition, these criteria include households with one or more conducted interviews.

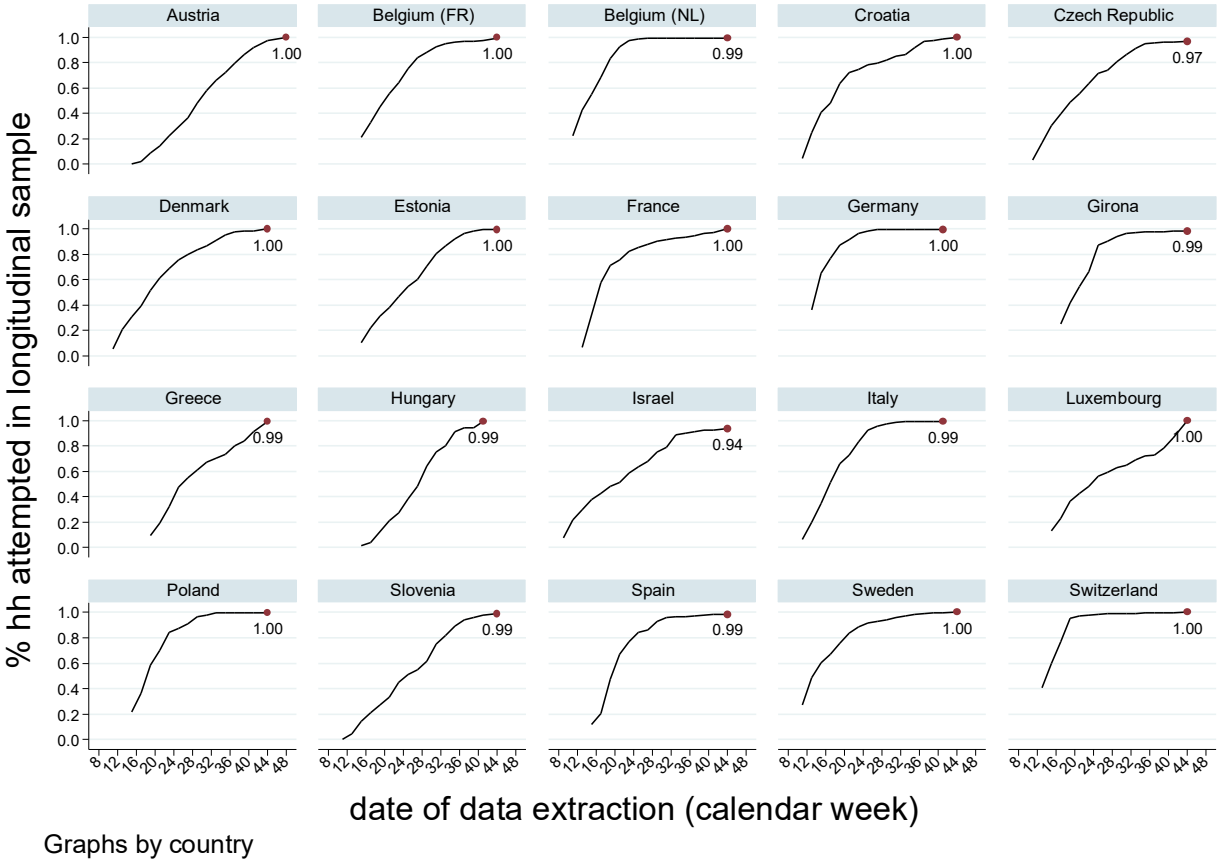


Figure 8.12: Fraction of panel households with contact attempts by country over time

In Wave 7, all countries except the Czech Republic and Israel managed at least one contact attempt in almost every panel household. It can be seen that most countries have a steep increase that levels out over time (i.e., interviewers were quick in attempting to contact the majority of all households). All other countries show a rather linear trend, possibly due to a different contact strategy.

Figure 8.13 shows country breakdowns of household contact rates over time. This figure contains contact attempts that resulted in an actual contact (i.e., at least one household member was reached). By definition, this criterion may also include households with at least one complete interview.

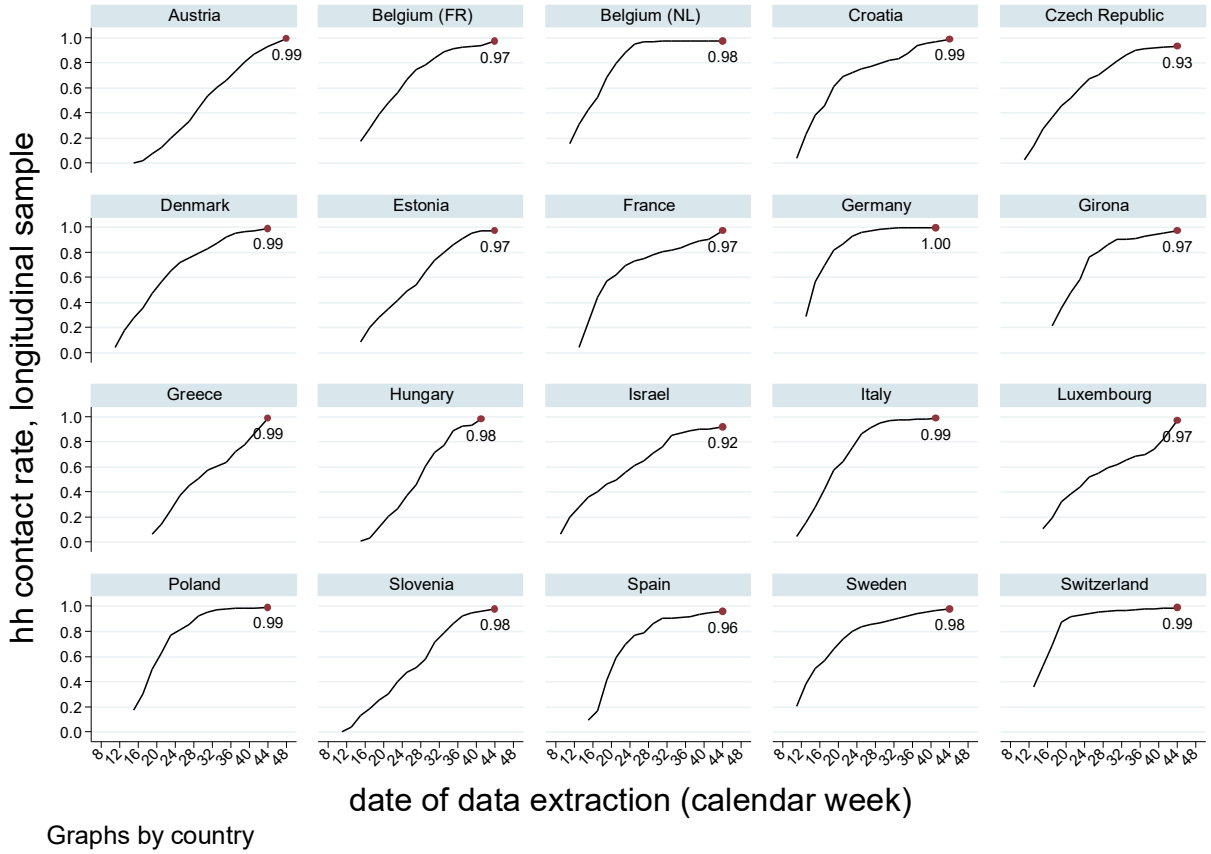


Figure 8.13: Contact rate of panel households by country over time

The trajectories of contact rates are similar to the rates on attempted households reported above. With contact attempt rates being the logical ceiling to contact rates, the Czech Republic and Israel have the "lowest" contact rates. Nevertheless, these countries have maintained rates of well above 90 percent.

8.1.3.4.2 Household cooperation and response rate

Figure 8.14 shows the cooperation rate of panel samples by country (i.e., the rate of all contacted households that have at least one completed interview). Similar to Waves 5 and 6, France was among the countries with the largest gross samples and the lowest cooperation rates (58 percent), followed by Hungary (61 percent) and Luxembourg (63 percent). Very high contact rates do not necessarily correspond with household cooperation. France had a considerably large C sample, the sample type that contained the respondents who had to be regained for the study. In the case of Hungary, this finding is also not surprising, since interviewers had to recover respondents from SHARE Wave 4. In most countries, cooperation rates continued to increase at different slopes until they reached a plateau of approximately 80 percent. The steepest slope can be observed in Greece. The head start followed by a sharp decline in Slovenia results from having achieved cooperation in all of the small number of contacted households at the beginning of fieldwork.

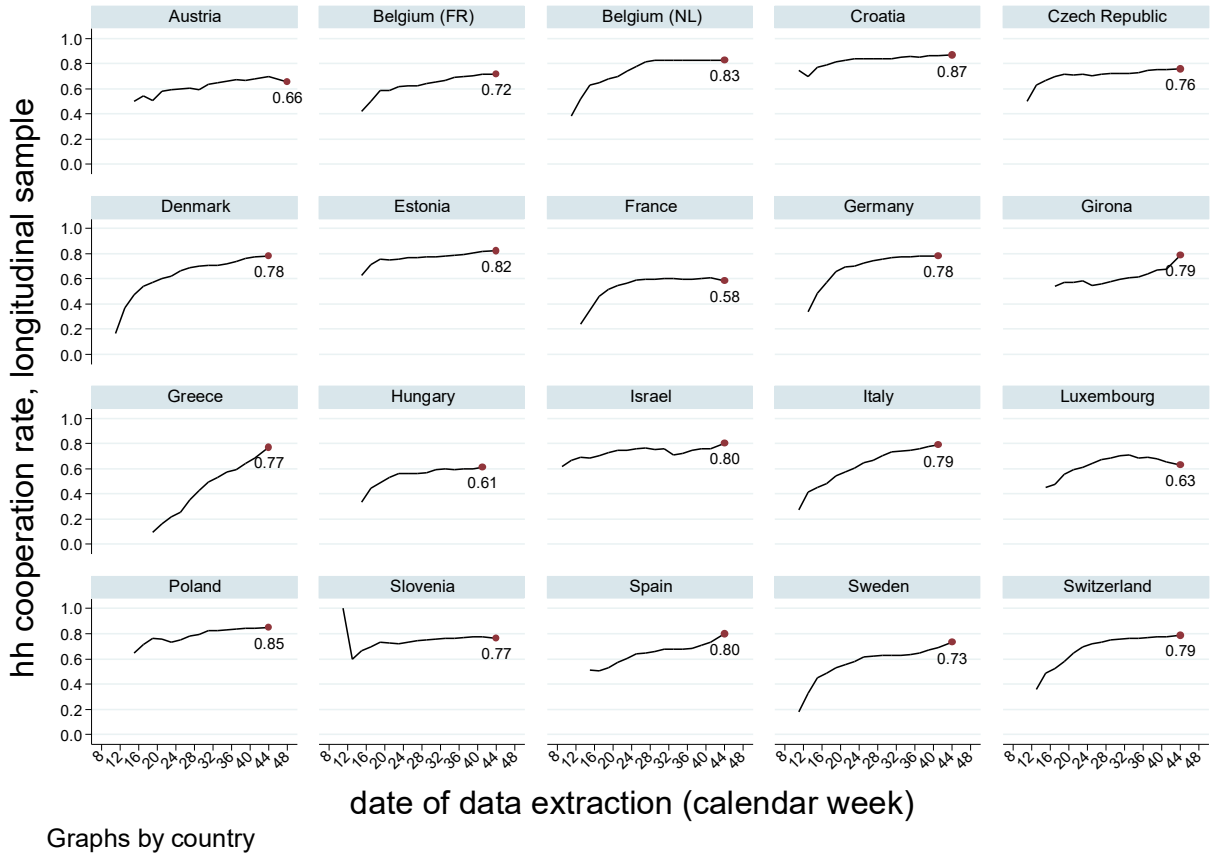


Figure 8.14: Cooperation rate of panel households by country over time

Figure 8.15 shows panel household retention rates (i.e., the number of panel households with at least one complete interview divided by the total number of (estimated) eligible panel households). Most countries have a steadily increasing trajectory that levels out over time in some countries. Due to their low cooperation rates, France, Hungary, and Luxembourg have the lowest household retention rates (keeping in mind that cooperation and contact rates represent the logical ceiling to the final retention rate).

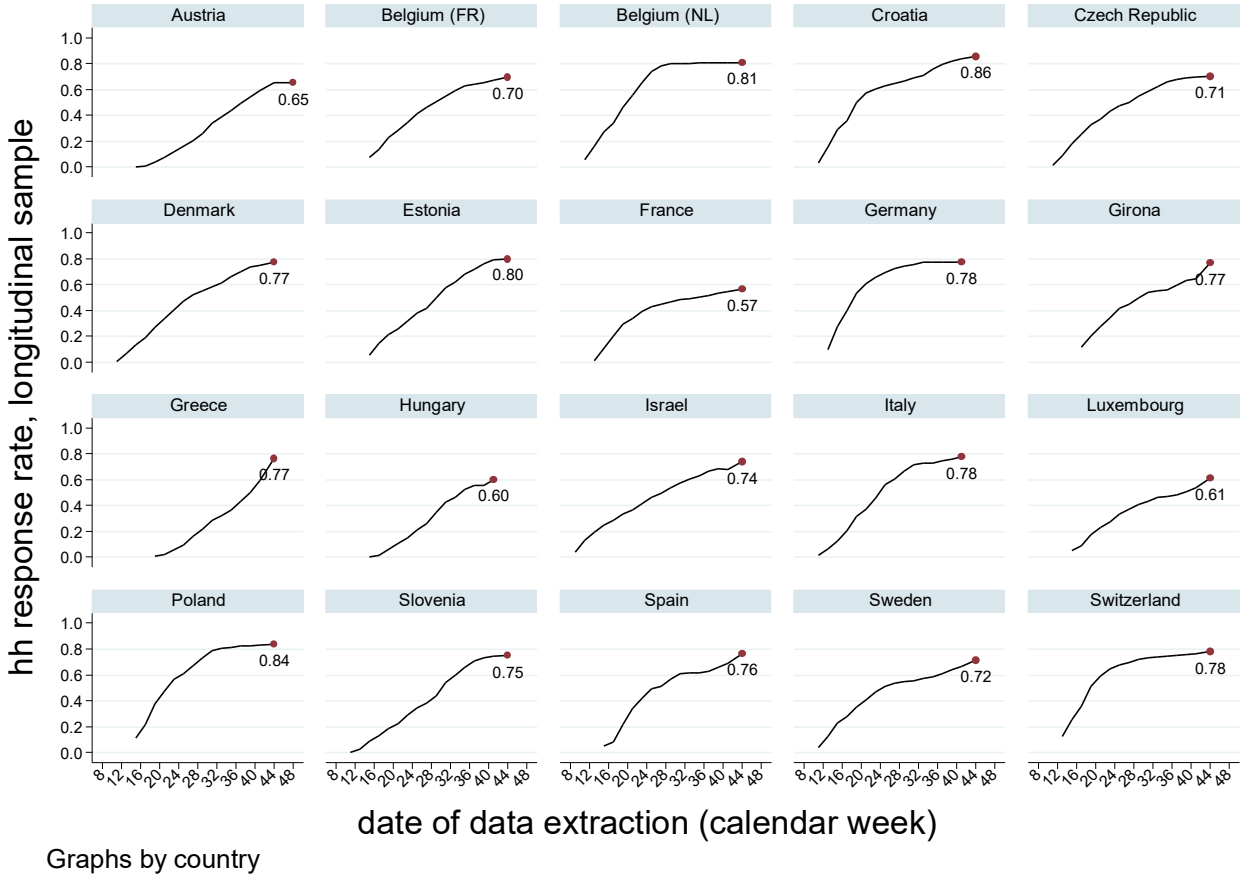


Figure 8.15: Retention rate of panel households by country over time

8.1.3.4.3 Individual participation of panel samples

Figure 8.16 shows the individual retention rate of subsamples A1 and A2. As noted before, Subsample A1 includes all respondents who participated in Wave 6 and at least one other previous wave; Subsample A2 includes all respondents who were part of a baseline or refreshment sample and participated for the first time in SHARE Wave 6. SHARE stipulates the continued inclusion of at least 85 percent of respondents in Subsample A1 and 75 percent of respondents in Subsample A2 in the current wave. Survey agencies were incentivised for rates exceeding these thresholds.

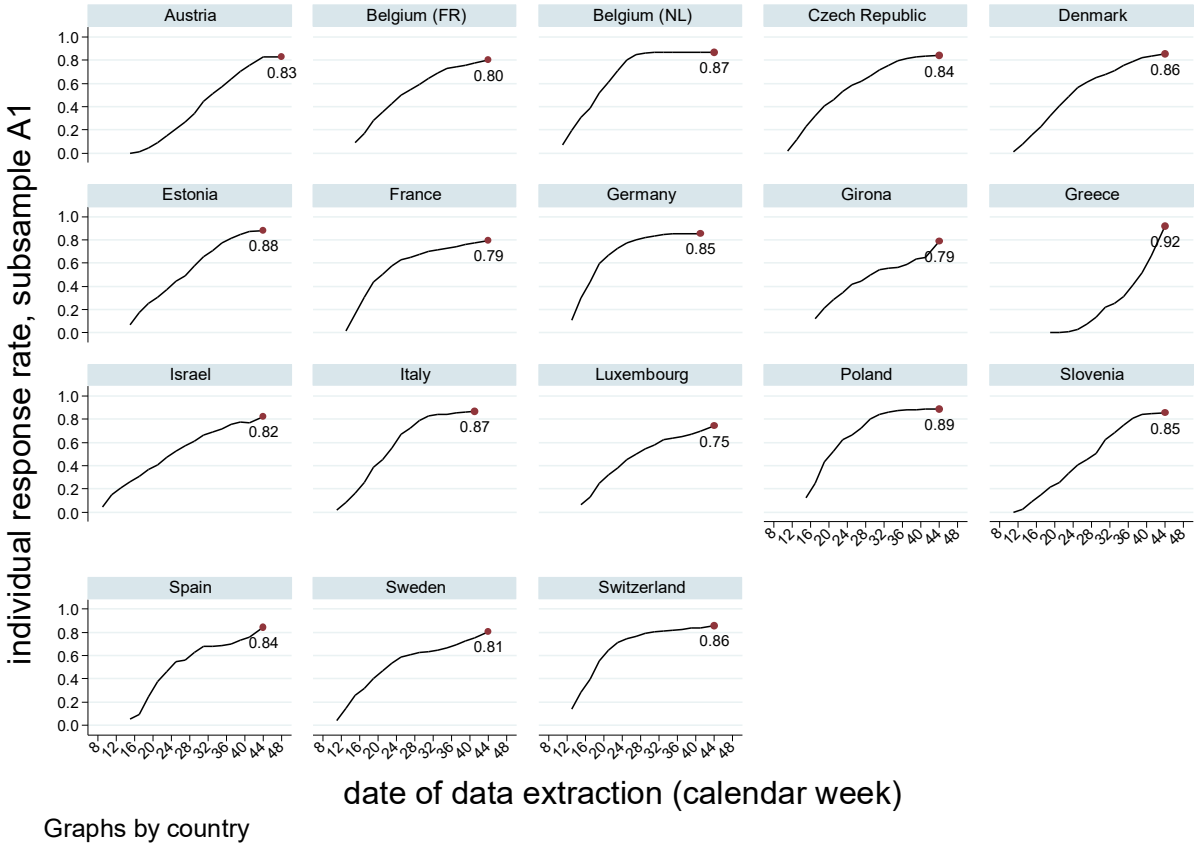


Figure 8.16: Individual retention rates in Subsamples A1 by country over time

While 9 out of 19 countries achieved a retention rate of 85 percent or more in Subsample A1, Luxembourg finished SHARE Wave 7 with the lowest retention rate among all countries (75 percent).

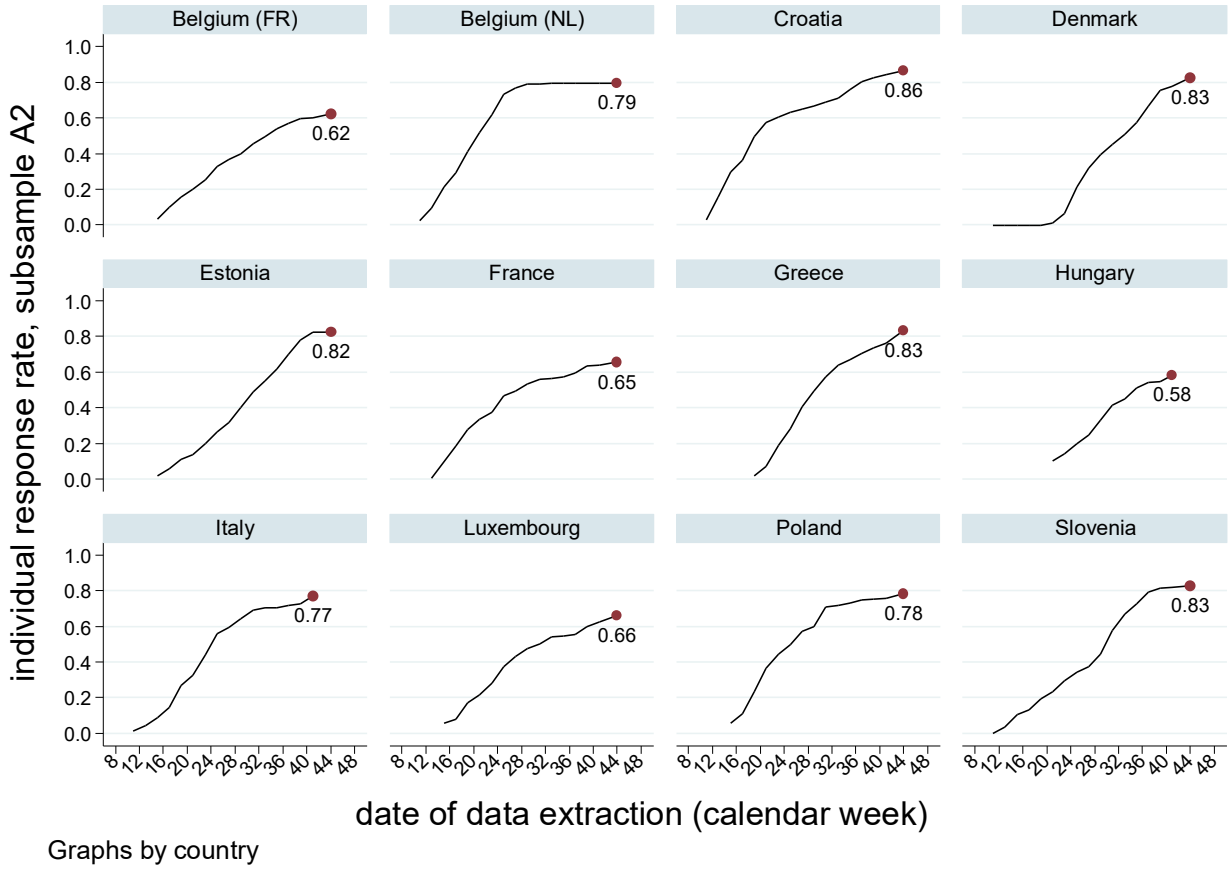


Figure 8.17: Individual retention rates in Subsamples A2 by country over time

With a final rate of 86 percent, Croatia is the frontrunner of Wave 7 in terms of individual retention among respondents who participated for the first time in the previous wave (see Figure 8.17), followed by Denmark, Greece, and Slovenia (83 percent each). The French part of Belgium, France, Hungary, and Luxembourg did not reach the expected minimum retention rate of 75 percent in Subsample A2. In France and Luxembourg, the individual-level retention rates in both subsamples are higher than the household retention rates, suggesting that a significant number of known eligible spouses or partners could be convinced to cooperate.

Figure 8.18 shows the individual retention rate (or, more precisely, the "recovery rate") of Subsample C (i.e., the percentage of panel respondents who participated in any previous SHARE wave, but not in the last (i.e., Wave 6), and who do not live in a household where at least one household member participated in Wave 6). In Wave 7, new cleaning rules were applied that primarily affected Subsample C by deleting households without any interview in the last three waves and persons for which no end-of-life interview could be realized in the last two waves, respectively.

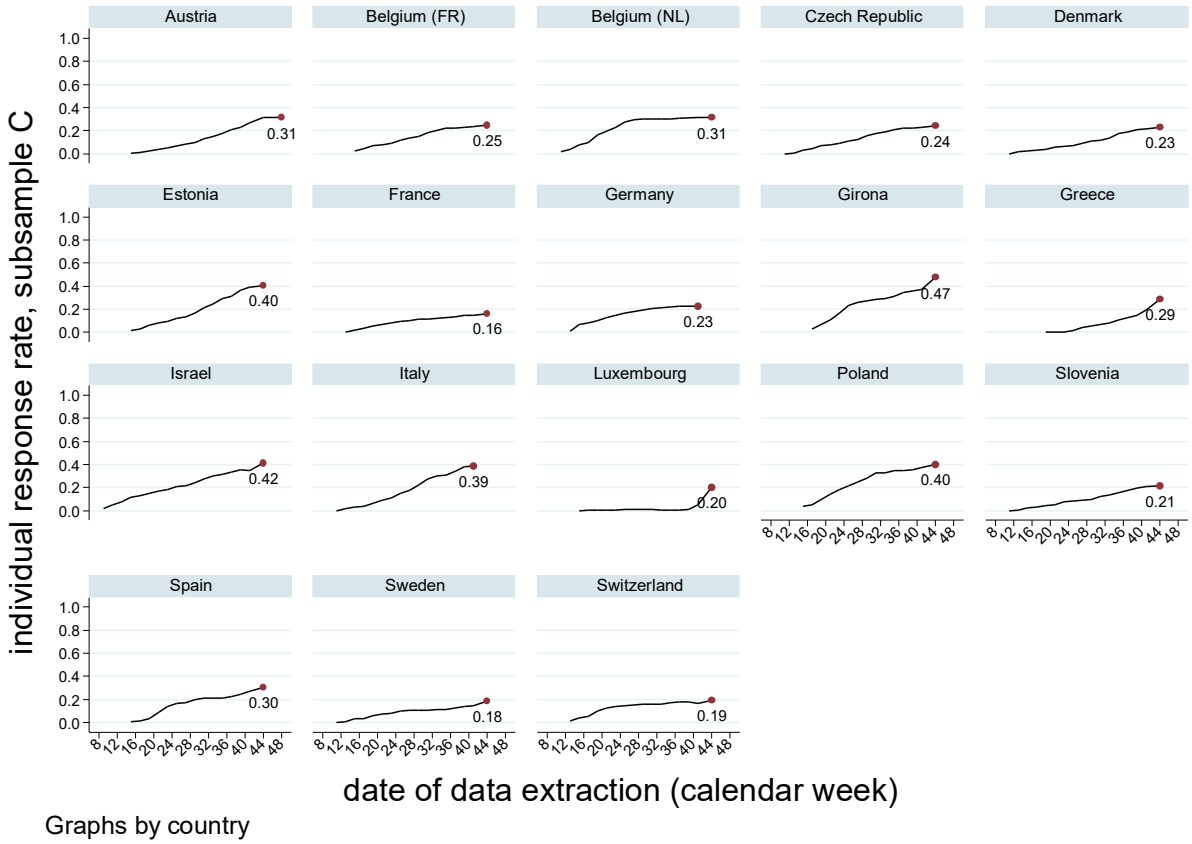


Figure 8.18: Individual retention (recovery) rates in Subsample C by country over time

It can be seen that at the end of fieldwork, the region of Girona focused on recovering respondents and therefore showed the best performance in recuperating as many "lost" respondents as possible.

8.1.3.4.4 Summary of panel samples

Figure 8.19 shows the final household-level contact, cooperation, and retention rates of the panel samples at the end of fieldwork in Wave 7.

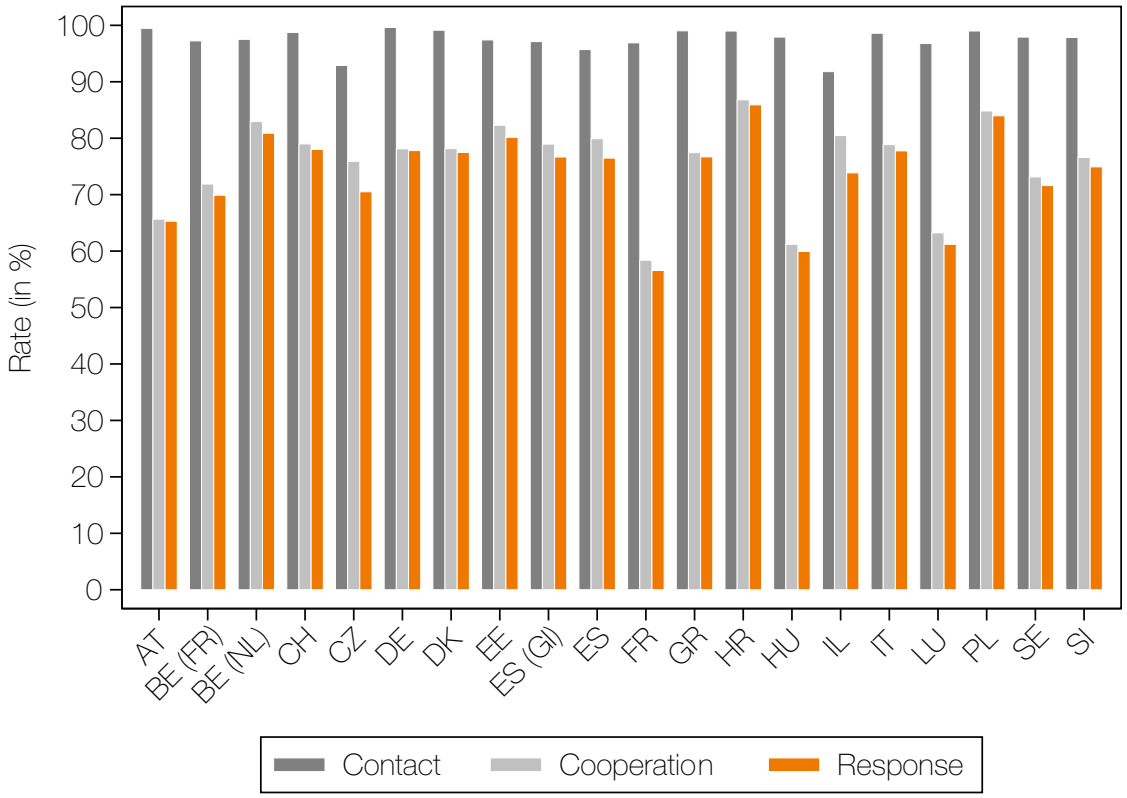


Figure 8.19: Contact, cooperation and retention rates for panel households

It can be seen that contact rates are close to 100 percent in most countries. The varying cooperation rates represent the ceiling of the final response/retention rates. Except for Austria, France, Hungary, and Luxembourg, the final household retention rates range between 70 and 86 percent in all other countries. In this respect, it should be noted that variation in rates can, at least partly, be related also to differences in the panel sample composition. Thus, when making comparisons across countries it should be considered when and how often refreshment samples were drawn, but also whether an oversampling of younger cohorts was applied that additionally affects the age composition of the sample and thus also might affect retention rates to a certain degree.

Figure 8.20 shows the final individual retention rates by subsample. Apart from the above-defined Subsamples A, B, and C, Subsample D includes all nonresponding spouses or partners and new spouses or partners that have not yet participated in any previous SHARE wave. For Croatia and Hungary, both of which had only participated in one previous wave (Wave 6 and Wave 4, respectively), the classification into Subsamples B and C is not yet applicable.

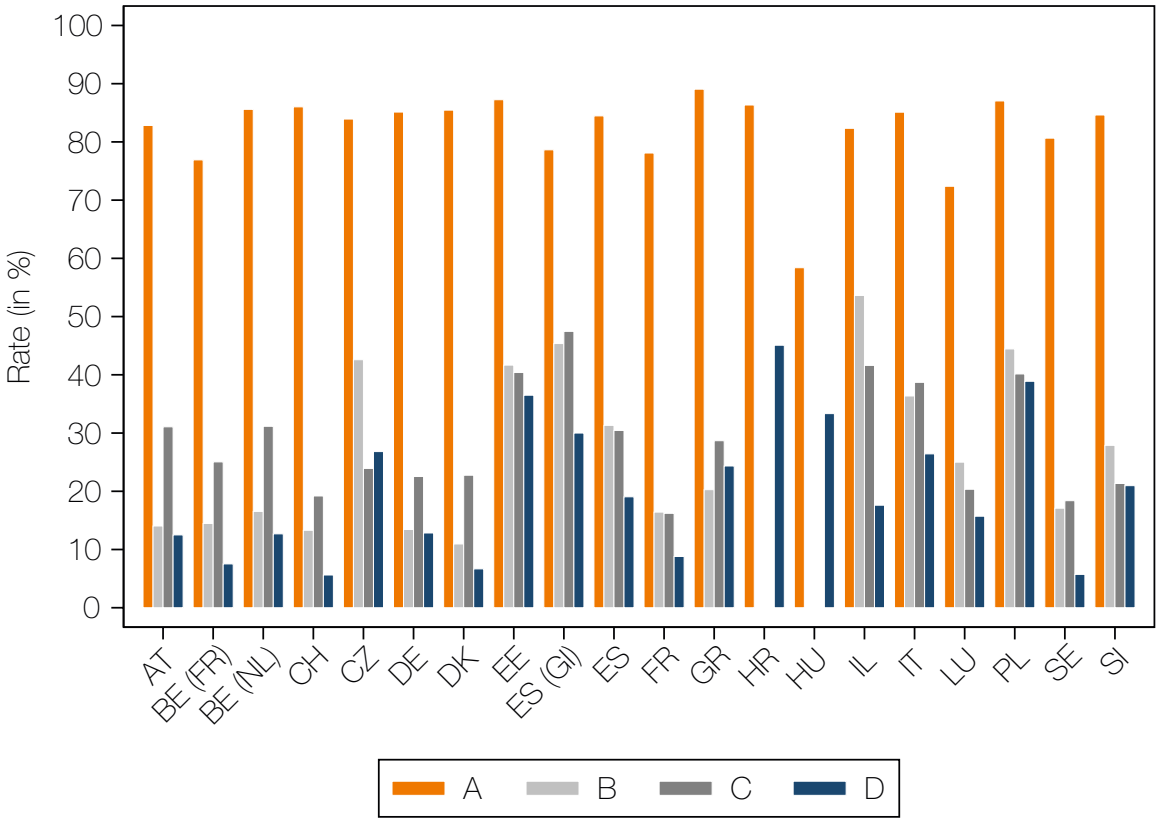


Figure 8.20: Respondent-level retention and recovery for panel households

Overall, Subsample A (i.e., A1 and A2 combined) retention is approximately 80 percent, except in Hungary. In some countries, it can be seen that many respondents who did not participate in the previous wave cooperated in Wave 7 (e.g., Subsample B and C respondents in the Czech Republic, Estonia, the region of Girona, Israel, Poland). With over 40 percent, Croatia managed to conduct the highest share of interviews among spouses and partners who had never taken part in SHARE.

Figure 8.21 displays individual retention by Subsamples A1 and A2. These rates were compiled after the end of fieldwork of Wave 7. While Subsample A1 includes all respondents who participated in the previous wave and any other wave of the SHARE survey, Subsample A2 consists of respondents who live in households that participated for the first time in the previous wave (i.e., baseline or refreshment sample). All countries missing the bar for Subsample A2 did not have a refreshment sample in Wave 6.

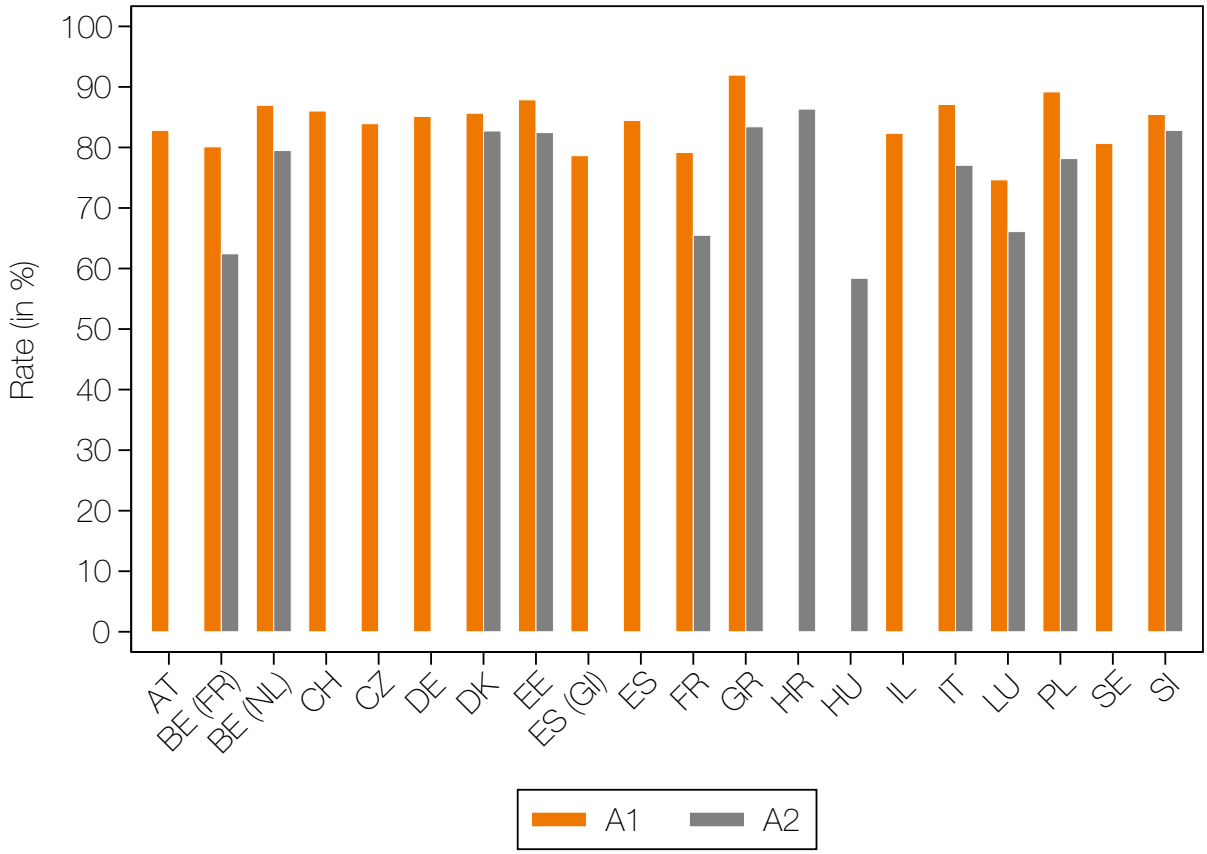


Figure 8.21: Respondent-level retention and recovery for Subsamples A1 and A2

In general, A1 retention is always higher than A2, mostly because it is usually more difficult to win back households that have participated only once. While A1 rates vary between approximately 75 and 92 percent, A2 rates range from approximately 58 to 86 percent.

Figure 8.22 shows the absolute number of panel interviews per country at the end of fieldwork in Wave 7. Detailed breakdowns can be found in the appendix of this chapter. The number of completed interviews varies largely across countries. While the Czech Republic, Estonia, Germany, Italy, and Slovenia conducted 4,000 or more interviews, most other countries finished fieldwork around the mark of 2,000 interviews (e.g., Belgium, Switzerland, the region of Girona, Croatia, Hungary, Israel).

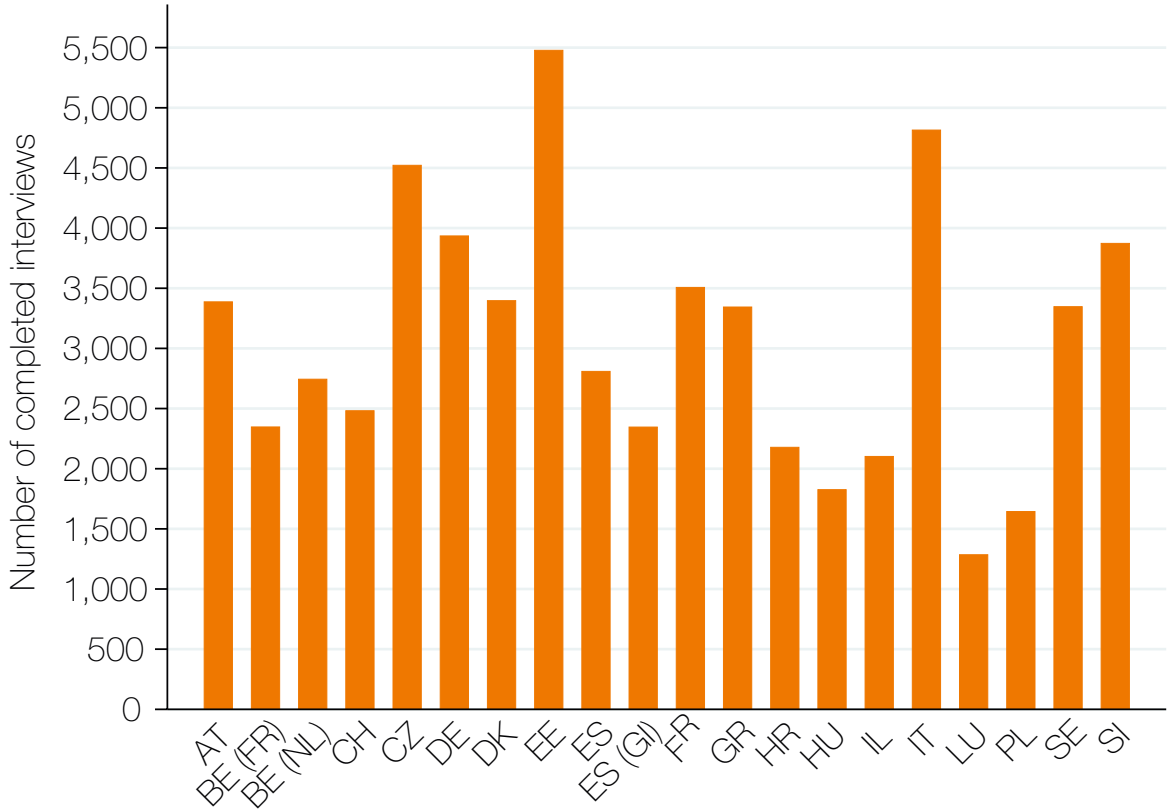


Figure 8.22: Absolute numbers of interviews in panel samples

8.1.4 Concluding remarks

Overall, all survey agencies managed to collect more than 80,000 interviews with the help of approximately 2,000 interviewers across all 28 countries, pushing the overall numbers to approximately 140,000 respondents and 330,000 interviews.

SHARE Central still benefits from large gains in efficiency by building on the conceptual framework established before Wave 5. To maximise data quality and cross-study comparability, we maintained the fieldwork monitoring procedures from previous waves and made some improvements. All numbers and rates are calculated biweekly based on formulas set by AAPOR. This standardised way of computing fieldwork outcomes allows for transparency for survey agencies and comparability with other studies. In addition, the provision of interviewer-specific rates can help improve interviewers' motivation to establish cooperation among new respondents ("response rate") and repeated participation among panel respondents ("retention rate"). In Wave 7, we implemented additional procedures to improve data quality by detecting fake interviews during fieldwork (see Chapter 8.2 on interviewer monitoring). Furthermore, we continued compiling and publishing the so-called SHARE "Compliance Profiles" after the end of fieldwork. This document is a short evaluation report that describes all operative tasks per wave in all participating countries. It contains all relevant quality indicators regarding the development of fieldwork, interviewer trainings, data transfers, and the final response and retention rates. Regarding the actual outcomes of the fieldwork of Wave 7, we summarised all findings in the "traffic-light table" below (see Table 8.4).

A positive fieldwork outcome is characterised by high contact rates, which are necessary to maximise respondent cooperation and therefore response and retention rates. In Table 8.4, it can be seen that interviewers in almost all countries worked

off their entire household samples and managed to establish successful contacts in most of them (i.e., at least 99 percent of all households were attempted, of which at least 95 percent could be successfully contacted). Similar to previous waves, the majority of countries did not reach our goal of approaching non-reached panel households six or more times. One explanation may be that many interviewers did not enter their contact attempts properly into the SMS.

The SHARE model contract stipulates that a minimum of 85 percent of respondents and a minimum of 75 percent of respondents must be re-interviewed in panel Subsample A1 and in panel Subsample A2, respectively. For baseline samples or refreshment samples, the document stipulates a minimum of 40 percent of eligible households to be interviewed. Table 8.4 shows whether countries passed or failed these contractual standards. It can be seen that half of all countries with a panel Subsample A1 surpassed the limit, whereas 8 out of 12 countries reached their goal in their A2 Subsample. Most countries – among them all new ones – reached or surpassed the target rate in their baseline/refreshment samples.

Table 8.4: "Traffic-light" summary of fieldwork outcomes in SHARE Wave 7 countries

	Panel sample					**Baseline/refreshment sample**		
	HH attempt rate	HH contact rate	Median # of contact attempts in HH without interview	Retention rate in subsample A1	Retention rate in subsample A2	HH attempt rate	HH contact rate	Individual response rate
Cut-off	99%	95%	6	85%	75%	95%	85%	40%
AT	🟢	🟢	🔴	🔴	---	---	---	---
BE (FR)	🟢	🟢	🟢	🔴	🔴	---	---	---
BE (NL)	🟢	🟢	🔴	🟢	🟢	---	---	---
BG	---	---	---	---	---	🔴	🟢	🟢
CH	🟢	🟢	🟢	🟢	---	---	---	---
CY	---	---	---	---	---	🟢	🟢	🟢
CZ	🔴	🔴	🔴	🔴	---	---	---	---
DE	🟢	🟢	🟢	🟢	---	---	---	---
DK	🟢	🟢	🔴	🟢	🟢	---	---	---
FI	---	---	---	---	---	🟢	🟢	🟢
EE	🟢	🟢	🔴	🟢	🟢	---	---	---
ES	🟢	🟢	🔴	🔴	---	---	---	---
ES (GI)	🟢	🟢	🔴	🔴	---	---	---	---
FR	🟢	🟢	🟢	🔴	🔴	---	---	---
GR	🟢	🟢	🔴	🟢	🟢	---	---	---
HR	🟢	🟢	🔴	---	🟢	🟢	🟢	🔴
HU	🟢	🟢	🔴	---	🔴	---	---	---
IL	🔴	🔴	🔴	🔴	---	🟢	🟢	🟢

	Panel sample					Baseline/refreshment sample		
IT	●	●	●	●	●	---	---	---
LT	---	---	---	---	---	●	●	●
LU	●	●	●	●	●	---	---	---
LV	---	---	---	---	---	●	●	●
MT	---	---	---	---	---	●	●	●
RO	---	---	---	---	---	●	●	●
PL	●	●	●	●	●	●	●	●
PT	●	●	●	●	---	---	---	---
SE	●	●	●	●	---	---	---	---
SI	●	●	●	●	●	---	---	---
SK	---	---	---	---	---	●	●	●

In sum, the overall fieldwork performance was remarkable in most countries, especially new ones. However, national interviewer trainings should stress the importance of (proper) documentation of contact attempts in the software. In addition, to increase retention in Subsample A1 – the subsample that is most important in sustaining SHARE's longitudinal dimension – agencies are advised to focus more on panel care.

References

American Association for Public Opinion Research (2016). Standard definitions: Final dispositions of case codes and outcome rates for surveys (9th edition). Retrieved from https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf.

Bergmann, M., Kneip, T., De Luca, G., & Scherpenzeel, A. (2017). Survey Participation in the Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 1-6. SHARE Working Paper Series (31-2017). Munich: Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC).

Das, M., de Bruijne, M., Janssen, J., & Kalwij, A. (2017). Experiment: Internet interviewing in the sixth wave of SHARE in the Netherlands. In F. Malter & A. Börsch-Supan (Eds.), SHARE Wave 6: Panel innovations and collecting dried blood spots (pp. 151-162). Munich: Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC).

Kneip, T. (2013). Survey participation in the fourth wave of SHARE. In F. Malter & A. Börsch-Supan (Eds.), SHARE Wave 4: Innovations & methodology (pp. 140-155). Munich: Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC).

Kneip, T., Malter, F., & Sand, G. (2015). Fieldwork monitoring and survey participation in fifth wave of SHARE. In F. Malter & A. Börsch-Supan (Eds.), SHARE Wave 5: Innovations & methodology (pp. 102-159). Munich: Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC).

Malter, F., & Sand, G. (2017). Fieldwork monitoring and survey participation in sixth wave of SHARE. In F. Malter & A. Börsch-Supan (Eds.), SHARE Wave 6: Panel innovations and collecting dried blood spots (pp. 95-125). Munich: Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC).

Appendix: Final outcomes of SHARE Wave 7 by country

Table A.1 Austria

Longitudinal sample	
Gross sample:	3688
Households attempted:	3679
Households contacted:	3668
Households estimated to be eligible:	3661.95
Households with completed Coverscreen Interview:	2358
Households with at least one complete interview:	2391
Percentage of Households attempted:	99.76 %
Contact rate:	99.46 %
Cooperation rate:	65.65 %
Household response rate:	65.29 %
Refusal rate:	22.09 %
Other non-interview rate:	12.07 %
Individual interviews:	3387
Sample A:	2786
Sample A1:	2786
Sample A2:	0
Sample B:	8
Sample C:	562
Sample D:	31
Estimated average number of eligibles in hh:	1.49
Individual response rate:	61.91 %
Sample A:	82.82 %
Sample A1:	82.82 %
Sample A2:	-
Sample B:	14.04 %
Sample C:	31.07 %
Sample D:	12.50 %
Median number of attempts for non-contacted hh:	3

Table A.2 Belgium (FR)

Longitudinal sample	
Gross sample:	2604
Households attempted:	2595
Households contacted:	2533
Households estimated to be eligible:	2570.91
Households with completed Coverscreen Interview:	1774
Households with at least one complete interview:	1797
Percentage of Households attempted:	99.65 %
Contact rate:	97.24 %
Cooperation rate:	71.88 %
Household response rate:	69.90 %
Refusal rate:	22.64 %
Other non-interview rate:	4.71 %
Individual interviews:	2347
Sample A:	2145
Sample A1:	1831
Sample A2:	314
Sample B:	22
Sample C:	149
Sample D:	31
Estimated average number of eligibles in hh:	1.53
Individual response rate:	59.59 %
Sample A:	76.91 %
Sample A1:	80.10 %
Sample A2:	62.43 %
Sample B:	14.47 %
Sample C:	25.04 %
Sample D:	7.52 %
Median number of attempts for non-contacted hh:	8

Table A.3 Belgium (NL)

Longitudinal sample	
Gross sample:	2376
Households attempted:	2364
Households contacted:	2317
Households estimated to be eligible:	2368.98
Households with completed Coverscreen Interview:	1895
Households with at least one complete interview:	1916
Percentage of Households attempted:	99.49 %
Contact rate:	97.51 %
Cooperation rate:	82.94 %
Household response rate:	80.88 %
Refusal rate:	13.00 %
Other non-interview rate:	3.63 %
Individual interviews:	2744
Sample A:	2560
Sample A1:	2126
Sample A2:	434
Sample B:	21
Sample C:	129
Sample D:	34
Estimated average number of eligibles in hh:	1.60
Individual response rate:	72.30 %
Sample A:	85.59 %
Sample A1:	86.95 %
Sample A2:	79.49 %
Sample B:	16.54 %
Sample C:	31.16 %
Sample D:	12.69 %
Median number of attempts for non-contacted hh:	1

Table A.4 Bulgaria

Baseline / refreshment sample	
Gross sample:	3000
Households attempted:	2782
Households contacted:	2537
Households estimated to be eligible:	2189.22
Households with completed Coverscreen Interview:	1396
Households with at least one complete interview:	1347
Percentage of Households attempted:	92.73 %
Contact rate:	84.53 %
Cooperation rate:	72.79 %
Household response rate:	61.53 %
Refusal rate:	16.58 %
Other non-interview rate:	6.42 %
Individual interviews:	2001
Estimated average number of eligibles in hh:	1.51
Individual response rate:	60.56 %
Median number of attempts for non-contacted hh:	2

Table A.5 Switzerland

Longitudinal sample	
Gross sample:	2323
Households attempted:	2323
Households contacted:	2294
Households estimated to be eligible:	2304.00
Households with completed Coverscreen Interview:	1780
Households with at least one complete interview:	1797
Percentage of Households attempted:	100.00 %
Contact rate:	98.74 %
Cooperation rate:	78.99 %
Household response rate:	77.99 %
Refusal rate:	19.49 %
Other non-interview rate:	1.26 %
Individual interviews:	2482
Sample A:	2344
Sample A1:	2344
Sample A2:	0
Sample B:	23
Sample C:	95
Sample D:	20
Estimated average number of eligibles in hh:	1.63
Individual response rate:	66.23 %
Sample A:	86.02 %
Sample A1:	86.02 %
Sample A2:	-
Sample B:	13.29 %
Sample C:	19.19 %
Sample D:	5.62 %
Median number of attempts for non-contacted hh:	11

Table A.6 Cyprus

Baseline / refreshment sample	
Baseline / refreshment sample	
Gross sample:	2498
Households attempted:	2498
Households contacted:	2266
Households estimated to be eligible:	1736.88
Households with completed Coverscreen Interview:	861
Households with at least one complete interview:	845
Percentage of Households attempted:	100.00 %
Contact rate:	89.87 %
Cooperation rate:	54.13 %
Household response rate:	48.65 %
Refusal rate:	23.71 %
Other non-interview rate:	17.51 %
Individual interviews:	1232
Estimated average number of eligibles in hh:	1.65
Individual response rate:	42.87 %
Median number of attempts for non-contacted hh:	6

Table A.7 Czech Republic

Longitudinal sample	
Gross sample:	4367
Households attempted:	4226
Households contacted:	4057
Households estimated to be eligible:	4359.90
Households with completed Coverscreen Interview:	2986
Households with at least one complete interview:	3074
Percentage of Households attempted:	96.77 %
Contact rate:	92.89 %
Cooperation rate:	75.90 %
Household response rate:	70.51 %
Refusal rate:	15.78 %
Other non-interview rate:	6.61 %
Individual interviews:	4522
Sample A:	4046
Sample A1:	4046
Sample A2:	0
Sample B:	46
Sample C:	371
Sample D:	59
Estimated average number of eligibles in hh:	1.54
Individual response rate:	67.56 %
Sample A:	83.92 %
Sample A1:	83.92 %
Sample A2:	-
Sample B:	42.59 %
Sample C:	23.92 %
Sample D:	26.82 %
Median number of attempts for non-contacted hh:	2

Table A.8 Germany

Longitudinal sample	
Gross sample:	3377
Households attempted:	3373
Households contacted:	3364
Households estimated to be eligible:	3365.99
Households with completed Coverscreen Interview:	2611
Households with at least one complete interview:	2619
Percentage of Households attempted:	99.88 %
Contact rate:	99.61 %
Cooperation rate:	78.11 %
Household response rate:	77.81 %
Refusal rate:	18.72 %
Other non-interview rate:	3.09 %
Individual interviews:	3935
Sample A:	3718
Sample A1:	3718
Sample A2:	0
Sample B:	18
Sample C:	146
Sample D:	53
Estimated average number of eligibles in hh:	1.65
Individual response rate:	70.86 %
Sample A:	85.12 %
Sample A1:	85.12 %
Sample A2:	-
Sample B:	13.43 %
Sample C:	22.53 %
Sample D:	12.83 %
Median number of attempts for non-contacted hh:	10

Table A.9 Denmark

Longitudinal sample	
Gross sample:	3126
Households attempted:	3126
Households contacted:	3099
Households estimated to be eligible:	3111.00
Households with completed Coverscreen Interview:	2404
Households with at least one complete interview:	2410
Percentage of Households attempted:	100.00 %
Contact rate:	99.13 %
Cooperation rate:	78.15 %
Household response rate:	77.47 %
Refusal rate:	20.09 %
Other non-interview rate:	1.58 %
Individual interviews:	3397
Sample A:	3179
Sample A1:	2978
Sample A2:	201
Sample B:	23
Sample C:	165
Sample D:	30
Estimated average number of eligibles in hh:	1.64
Individual response rate:	66.69 %
Sample A:	85.46 %
Sample A1:	85.65 %
Sample A2:	82.72 %
Sample B:	10.95 %
Sample C:	22.76 %
Sample D:	6.67 %
Median number of attempts for non-contacted hh:	5

Table A.10 Estonia

Longitudinal sample	
Gross sample:	4779
Households attempted:	4757
Households contacted:	4656
Households estimated to be eligible:	4744.89
Households with completed Coverscreen Interview:	3740
Households with at least one complete interview:	3803
Percentage of Households attempted:	99.54 %
Contact rate:	97.41 %
Cooperation rate:	82.28 %
Household response rate:	80.15 %
Refusal rate:	14.88 %
Other non-interview rate:	2.38 %
Individual interviews:	5478
Sample A:	4896
Sample A1:	4365
Sample A2:	531
Sample B:	40
Sample C:	492
Sample D:	50
Estimated average number of eligibles in hh:	1.49
Individual response rate:	77.72 %
Sample A:	87.24 %
Sample A1:	87.86 %
Sample A2:	82.45 %
Sample B:	41.67 %
Sample C:	40.39 %
Sample D:	36.50 %
Median number of attempts for non-contacted hh:	5

Table A.11 Spain – Region of Girona

Longitudinal sample	
Gross sample:	2034
Households attempted:	2008
Households contacted:	1976
Households estimated to be eligible:	1995.61
Households with completed Coverscreen Interview:	1526
Households with at least one complete interview:	1530
Percentage of Households attempted:	98.72 %
Contact rate:	97.11 %
Cooperation rate:	78.95 %
Household response rate:	76.67 %
Refusal rate:	16.09 %
Other non-interview rate:	4.36 %
Individual interviews:	2346
Sample A:	1999
Sample A1:	1999
Sample A2:	0
Sample B:	49
Sample C:	271
Sample D:	27
Estimated average number of eligibles in hh:	1.65
Individual response rate:	71.05 %
Sample A:	78.64 %
Sample A1:	78.64 %
Sample A2:	-
Sample B:	45.37 %
Sample C:	47.46 %
Sample D:	30.00 %
Median number of attempts for non-contacted hh:	4

Table A.12 Spain

Longitudinal sample	
Gross sample:	2427
Households attempted:	2394
Households contacted:	2323
Households estimated to be eligible:	2417.93
Households with completed Coverscreen Interview:	1833
Households with at least one complete interview:	1849
Percentage of Households attempted:	98.64 %
Contact rate:	95.70 %
Cooperation rate:	79.90 %
Household response rate:	76.47 %
Refusal rate:	13.69 %
Other non-interview rate:	5.54 %
Individual interviews:	2809
Sample A:	2566
Sample A1:	2566
Sample A2:	0
Sample B:	31
Sample C:	200
Sample D:	12
Estimated average number of eligibles in hh:	1.59
Individual response rate:	72.95 %
Sample A:	84.46 %
Sample A1:	84.46 %
Sample A2:	-
Sample B:	31.31 %
Sample C:	30.44 %
Sample D:	19.05 %
Median number of attempts for non-contacted hh:	3

Table A.13 Finland

Baseline / refreshment sample	
Gross sample:	2400
Households attempted:	2400
Households contacted:	2360
Households estimated to be eligible:	2319.00
Households with completed Coverscreen Interview:	1397
Households with at least one complete interview:	1391
Percentage of Households attempted:	100.00 %
Contact rate:	98.28 %
Cooperation rate:	61.04 %
Household response rate:	59.98 %
Refusal rate:	34.15 %
Other non-interview rate:	4.14 %
Individual interviews:	2000
Estimated average number of eligibles in hh:	1.64
Individual response rate:	52.68 %
Median number of attempts for non-contacted hh:	7

Table A.14 France

Longitudinal sample	
Gross sample:	4452
Households attempted:	4452
Households contacted:	4315
Households estimated to be eligible:	4407.00
Households with completed Coverscreen Interview:	2478
Households with at least one complete interview:	2493
Percentage of Households attempted:	100.00 %
Contact rate:	96.89 %
Cooperation rate:	58.38 %
Household response rate:	56.57 %
Refusal rate:	31.15 %
Other non-interview rate:	9.17 %
Individual interviews:	3507
Sample A:	3061
Sample A1:	2856
Sample A2:	205
Sample B:	33
Sample C:	383
Sample D:	30
Estimated average number of eligibles in hh:	1.54
Individual response rate:	51.56 %
Sample A:	78.09 %
Sample A1:	79.18 %
Sample A2:	65.50 %
Sample B:	16.42 %
Sample C:	16.21 %
Sample D:	8.80 %
Median number of attempts for non-contacted hh:	6

Table A.15 Greece

Longitudinal sample	
Gross sample:	2893
Households attempted:	2875
Households contacted:	2865
Households estimated to be eligible:	2877.96
Households with completed Coverscreen Interview:	2109
Households with at least one complete interview:	2207
Percentage of Households attempted:	99.38 %
Contact rate:	99.03 %
Cooperation rate:	77.44 %
Household response rate:	76.69 %
Refusal rate:	20.36 %
Other non-interview rate:	1.98 %
Individual interviews:	3344
Sample A:	3026
Sample A1:	2066
Sample A2:	960
Sample B:	15
Sample C:	257
Sample D:	46
Estimated average number of eligibles in hh:	1.58
Individual response rate:	73.59 %
Sample A:	89.05 %
Sample A1:	91.94 %
Sample A2:	83.41 %
Sample B:	20.27 %
Sample C:	28.68 %
Sample D:	24.34 %
Median number of attempts for non-contacted hh:	5

Table A.16 Croatia

Baseline / refreshment sample	
Gross sample:	862
Households attempted:	839
Households contacted:	785
Households estimated to be eligible:	680.19
Households with completed Coverscreen Interview:	237
Households with at least one complete interview:	233
Percentage of Households attempted:	97.33 %
Contact rate:	89.39 %
Cooperation rate:	38.32 %
Household response rate:	34.26 %
Refusal rate:	48.81 %
Other non-interview rate:	6.32 %
Individual interviews:	344
Estimated average number of eligibles in hh:	1.65
Individual response rate:	30.63 %
Median number of attempts for non-contacted hh:	2

Table A.16 (continued)

Longitudinal sample	
Gross sample:	1597
Households attempted:	1597
Households contacted:	1581
Households estimated to be eligible:	1576.00
Households with completed Coverscreen Interview:	1353
Households with at least one complete interview:	1354
Percentage of Households attempted:	100.00 %
Contact rate:	98.98 %
Cooperation rate:	86.79 %
Household response rate:	85.91 %
Refusal rate:	10.41 %
Other non-interview rate:	2.66 %
Individual interviews:	2178
Sample A:	2123
Sample A1:	0
Sample A2:	2123
Sample B:	0
Sample C:	0
Sample D:	55
Estimated average number of eligibles in hh:	1.63
Individual response rate:	84.71 %
Sample A:	86.34 %
Sample A1:	-
Sample A2:	86.34 %
Sample B:	-
Sample C:	-
Sample D:	45.08 %
Median number of attempts for non-contacted hh:	3

Table A.17 Hungary

Longitudinal sample	
Gross sample:	2033
Households attempted:	2016
Households contacted:	1991
Households estimated to be eligible:	2018.99
Households with completed Coverscreen Interview:	1216
Households with at least one complete interview:	1210
Percentage of Households attempted:	99.16 %
Contact rate:	97.92 %
Cooperation rate:	61.20 %
Household response rate:	59.93 %
Refusal rate:	26.10 %
Other non-interview rate:	11.89 %
Individual interviews:	1826
Sample A:	1782
Sample A1:	0
Sample A2:	1782
Sample B:	0
Sample C:	0
Sample D:	44
Estimated average number of eligibles in hh:	1.57
Individual response rate:	57.50 %
Sample A:	58.39 %
Sample A1:	. %
Sample A2:	58.39 %
Sample B:	.
Sample C:	.
Sample D:	33.33 %
Median number of attempts for non-contacted hh:	3

Table A.18 Israel

Baseline / refreshment sample	
Gross sample:	285
Households attempted:	280
Households contacted:	275
Households estimated to be eligible:	229.02
Households with completed Coverscreen Interview:	128
Households with at least one complete interview:	121
Percentage of Households attempted:	98.25 %
Contact rate:	96.06 %
Cooperation rate:	55.00 %
Household response rate:	52.83 %
Refusal rate:	27.07 %
Other non-interview rate:	16.16 %
Individual interviews:	173
Estimated average number of eligibles in hh:	1.59
Individual response rate:	47.45 %
Median number of attempts for non-contacted hh:	1

Table A. 18 (continued)

Longitudinal sample	
Gross sample:	1932
Households attempted:	1818
Households contacted:	1774
Households estimated to be eligible:	1918.56
Households with completed Coverscreen Interview:	1425
Households with at least one complete interview:	1417
Percentage of Households attempted:	94.10 %
Contact rate:	91.79 %
Cooperation rate:	80.47 %
Household response rate:	73.86 %
Refusal rate:	13.29 %
Other non-interview rate:	4.64 %
Individual interviews:	2102
Sample A:	1668
Sample A1:	1668
Sample A2:	0
Sample B:	74
Sample C:	335
Sample D:	25
Estimated average number of eligibles in hh:	1.62
Individual response rate:	67.79 %
Sample A:	82.33 %
Sample A1:	82.33 %
Sample A2:	-
Sample B:	53.62 %
Sample C:	41.61 %
Sample D:	17.61 %
Median number of attempts for non-contacted hh:	2

Table A.19 Italy

Longitudinal sample	
Gross sample:	3941
Households attempted:	3920
Households contacted:	3886
Households estimated to be eligible:	3904.85
Households with completed Coverscreen Interview:	3016
Households with at least one complete interview:	3036
Percentage of Households attempted:	99.47 %
Contact rate:	98.60 %
Cooperation rate:	78.86 %
Household response rate:	77.75 %
Refusal rate:	16.90 %
Other non-interview rate:	3.94 %
Individual interviews:	4815
Sample A:	4291
Sample A1:	3522
Sample A2:	769
Sample B:	32
Sample C:	413
Sample D:	79
Estimated average number of eligibles in hh:	1.66
Individual response rate:	74.28 %
Sample A:	85.11 %
Sample A1:	87.09 %
Sample A2:	77.05 %
Sample B:	36.36 %
Sample C:	38.71 %
Sample D:	26.42 %
Median number of attempts for non-contacted hh:	8

Table A.20 Lithuania

Baseline / refreshment sample	
Gross sample:	8000
Households attempted:	8000
Households contacted:	7633
Households estimated to be eligible:	2444.14
Households with completed Coverscreen Interview:	1579
Households with at least one complete interview:	1534
Percentage of Households attempted:	100.00 %
Contact rate:	95.32 %
Cooperation rate:	65.84 %
Household response rate:	62.76 %
Refusal rate:	29.65 %
Other non-interview rate:	2.90 %
Individual interviews:	2028
Estimated average number of eligibles in hh:	1.46
Individual response rate:	56.87 %
Median number of attempts for non-contacted hh:	6

Table A.21 Luxembourg

Longitudinal sample	
Gross sample:	1537
Households attempted:	1537
Households contacted:	1489
Households estimated to be eligible:	1492.00
Households with completed Coverscreen Interview:	911
Households with at least one complete interview:	913
Percentage of Households attempted:	100.00 %
Contact rate:	96.78 %
Cooperation rate:	63.23 %
Household response rate:	61.19 %
Refusal rate:	27.82 %
Other non-interview rate:	7.77 %
Individual interviews:	1285
Sample A:	1111
Sample A1:	842
Sample A2:	269
Sample B:	11
Sample C:	83
Sample D:	80
Estimated average number of eligibles in hh:	1.67
Individual response rate:	51.48 %
Sample A:	72.38 %
Sample A1:	74.65 %
Sample A2:	66.09 %
Sample B:	25.00 %
Samples A+B combined:	
Sample C:	20.34 %
Sample D:	15.72 %
Median number of attempts for non-contacted hh:	2

Table A.22 Latvia

Baseline / refreshment sample	
Gross sample:	2934
Households attempted:	2931
Households contacted:	2697
Households estimated to be eligible:	2106.61
Households with completed Coverscreen Interview:	1326
Households with at least one complete interview:	1304
Percentage of Households attempted:	99.90 %
Contact rate:	91.90 %
Cooperation rate:	67.36 %
Household response rate:	61.90 %
Refusal rate:	28.01 %
Other non-interview rate:	1.99 %
Individual interviews:	1755
Estimated average number of eligibles in hh:	1.43
Individual response rate:	58.30 %
Median number of attempts for non-contacted hh:	6

Table A.23 Malta

Baseline / refreshment sample	
Gross sample:	2500
Households attempted:	2485
Households contacted:	2349
Households estimated to be eligible:	1729.38
Households with completed Coverscreen Interview:	869
Households with at least one complete interview:	800
Percentage of Households attempted:	99.40 %
Contact rate:	91.54 %
Cooperation rate:	50.54 %
Household response rate:	46.26 %
Refusal rate:	41.11 %
Other non-interview rate:	4.16 %
Individual interviews:	1267
Estimated average number of eligibles in hh:	1.68
Individual response rate:	43.53 %
Median number of attempts for non-contacted hh:	3

Table A.24 Poland

Baseline / refreshment sample	
Gross sample:	5754
Households attempted:	5735
Households contacted:	5529
Households estimated to be eligible:	5163.12
Households with completed Coverscreen Interview:	2176
Households with at least one complete interview:	2142
Percentage of Households attempted:	99.67 %
Contact rate:	95.68 %
Cooperation rate:	43.36 %
Household response rate:	41.49 %
Refusal rate:	47.80 %
Other non-interview rate:	6.39 %
Individual interviews:	3147
Estimated average number of eligibles in hh:	1.66
Individual response rate:	36.77 %
Median number of attempts for non-contacted hh:	4

Table A. 24 (continued)

Longitudinal sample	
Gross sample:	1299
Households attempted:	1298
Households contacted:	1286
Households estimated to be eligible:	1291.00
Households with completed Coverscreen Interview:	1072
Households with at least one complete interview:	1084
Percentage of Households attempted:	99.92 %
Contact rate:	98.99 %
Cooperation rate:	84.82 %
Household response rate:	83.97 %
Refusal rate:	10.53 %
Other non-interview rate:	4.49 %
Individual interviews:	1644
Sample A:	1557
Sample A1:	1285
Sample A2:	272
Sample B:	4
Sample C:	55
Sample D:	28
Estimated average number of eligibles in hh:	1.55
Individual response rate:	82.10 %
Sample A:	87.03 %
Sample A1:	89.17 %
Sample A2:	78.16 %
Sample B:	44.44 %
Sample C:	40.15 %
Sample D:	38.89 %
Median number of attempts for non-contacted hh:	5

Table A.25 Romania

Baseline / refreshment sample	
Gross sample:	3800
Households attempted:	3584
Households contacted:	3506
Households estimated to be eligible:	2649.85
Households with completed Coverscreen Interview:	1443
Households with at least one complete interview:	1411
Percentage of Households attempted:	94.32 %
Contact rate:	92.24 %
Cooperation rate:	57.73 %
Household response rate:	53.25 %
Refusal rate:	34.72 %
Other non-interview rate:	4.28 %
Individual interviews:	2116
Estimated average number of eligibles in hh:	1.60
Individual response rate:	49.85 %
Median number of attempts for non-contacted hh:	3

Table A.26 Sweden

Longitudinal sample	
Gross sample:	3359
Households attempted:	3359
Households contacted:	3289
Households estimated to be eligible:	3346.00
Households with completed Coverscreen Interview:	2331
Households with at least one complete interview:	2396
Percentage of Households attempted:	100.00 %
Contact rate:	97.91 %
Cooperation rate:	73.14 %
Household response rate:	71.61 %
Refusal rate:	22.89 %
Other non-interview rate:	3.41 %
Individual interviews:	3347
Sample A:	3129
Sample A1:	3129
Sample A2:	0
Sample B:	40
Sample C:	150
Sample D:	28
Estimated average number of eligibles in hh:	1.62
Individual response rate:	61.78 %
Sample A:	80.64 %
Sample A1:	80.64 %
Sample A2:	-
Sample B:	17.09 %
Sample C:	18.38 %
Sample D:	5.73 %
Median number of attempts for non-contacted hh:	10

Table A.27 Slovenia

Longitudinal sample	
Gross sample:	3482
Households attempted:	3444
Households contacted:	3407
Households estimated to be eligible:	3460.87
Households with completed Coverscreen Interview:	2582
Households with at least one complete interview:	2593
Percentage of Households attempted:	98.91 %
Contact rate:	97.84 %
Cooperation rate:	76.58 %
Household response rate:	74.92 %
Refusal rate:	19.97 %
Other non-interview rate:	2.95 %
Individual interviews:	3873
Sample A:	3559
Sample A1:	2465
Sample A2:	1094
Sample B:	29
Sample C:	150
Sample D:	135
Estimated average number of eligibles in hh:	1.63
Individual response rate:	68.64 %
Sample A:	84.62 %
Sample A1:	85.44 %
Sample A2:	82.82 %
Sample B:	27.88 %
Sample C:	21.34 %
Sample D:	20.96 %
Median number of attempts for non-contacted hh:	5

Table A.28 Slovakia

Baseline / refreshment sample	
Gross sample:	5661
Households attempted:	5661
Households contacted:	4906
Households estimated to be eligible:	3242.31
Households with completed Coverscreen Interview:	1404
Households with at least one complete interview:	1286
Percentage of Households attempted:	100.00 %
Contact rate:	86.64 %
Cooperation rate:	45.78 %
Household response rate:	39.66 %
Refusal rate:	9.62 %
Other non-interview rate:	37.35 %
Individual interviews:	2076
Estimated average number of eligibles in hh:	1.64
Individual response rate:	39.01 %
Median number of attempts for non-contacted hh:	1

8.2 Improving the efficiency of data quality back checks: A new procedure to prevent curbstoning

Michael Bergmann, Karin Schuller – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC) and the Technical University of Munich (TUM)

The purpose of this chapter[17] is to describe our efforts in introducing a procedure that improves the efficiency of detecting fake interviews in the Survey of Health, Ageing and Retirement in Europe (SHARE; Börsch-Supan, et al., 2013) during fieldwork. There had been two key innovations in Wave 7: first, the development of a technical script to identify interview fabrication; second, the solicitation of feedback on back checks in a more formalised and systematic way.

Interviewer falsification ("fake interviews") is a problem in all interviewer-conducted surveys. While there are many variations and different reasons for interviewers deviating from properly administering the survey (for an overview see Murphy et al., 2016), we only address the most extreme form of deviation: interviewers' fabrication of entire interviews. For brevity sake, we refer to this issue with the term "curbstoning", coined by the U.S. Census Bureau (Ericksen & Kadane, 1985; Werker, 1981). In the original sense, curbstoning refers to "sitting on a curbstone and completing questionnaires, rather than interviewing respondents" (Koczela et al., 2015, p. 414). In contrast, we do not address any ideas on minimum interview quality, for example, minimum levels or constellations of data that would constitute a cut-off for releasing an interview or not. The latter is very difficult to conceptualise and particularly much less enforceable in a complex panel survey such as SHARE.

Until Wave 7, specific (statistical) procedures of identifying cheating interviewers as well as the consequences of detection were under-specified. This is not to say that there were no back checks at all in SHARE. Survey agencies were asked to verify a minimum of 20 percent of each interviewer's complete interviews by supervisory personnel of the survey agencies in each country. These general controls were based on a random sample of interviews that was checked by calling the selected respondents and asking them if they had done the interview at all and if they had answered some questions in the recorded way. However, there was quite some variation in the verification mechanisms applied among the participating countries in SHARE, and thus, the need to harmonise procedures emerged to achieve the goal of comparability. Moreover, one aim was to improve the current procedures of random back checks, which can be seen as the gold standard of data quality checks regarding time and cost efficiency.

The remainder of this chapter is structured as follows: Section 8.2.1 describes the technical procedure, which was implemented during the fieldwork of Wave 7 to identify interview fabrication. Section 8.2.2 addresses the implementation of this technical procedure and describes the efforts to solicit feedback on back checks in a more standardised way. Finally, the outcome of this new procedure is described in Section 8.2.3, followed by some concluding remarks in Section 8.2.4.

8.2.1 Development of a technical procedure to identify interview fabrication

The most common way to identify curbstoning is by re-contacting interviewed households and asking them if they have done the interview (Murphy et al., 2016). Since it is too expensive to re-interview the whole sample, a random selection of the sample is usually chosen. However, it would be much more efficient to find a way to filter a selection of the sample according to a suspicion to generate a focused sample that should be checked (Bredl et al., 2013). In the scholarly debate of curbstoning, it is assumed that fabricated data from falsifying interviewers deviate from the data of real respondents in certain aspects. We use this idea to develop a procedure that generates a subsample that should be checked according to a suspicion. Possible indicators can be gained from CAPI (computer-assisted personal interviewing) data (e.g., follow-up questions, item nonresponse) or from paradata, which is frequently available as a by-product of the data collection process (e.g., interview length or performance indicators).

Theoretical assumptions – if applied at all – are frequently based on the satisficing model developed by Krosnick and Alwin (1987), which has been further evolved in the context of curbstoning. The basic idea of this refinement is that falsifying interviewers want to save time and effort while simultaneously trying to minimise the risk of being detected (Menold et al., 2013). Based on these considerations theory and on previous findings, we expect that in-

17 This subchapter is a shortened version of the article "Preventing curbstoning in the Survey of Health, Ageing and Retirement in Europe (SHARE)" that is currently under review in Longitudinal and Life Course Studies.

terviewers fabricating their interviews would have fewer contact attempts, fewer interviewer notes, a shorter interview duration and fewer asked items, as well as fewer "other" or "code all that apply" answers and fewer follow-up questions. Similarly, we expect more duplicates and more straight-lining because these approaches also save time and effort. On the other hand, we expect that falsifiers have a lower item-nonresponse rate, less extreme answer patterns and a lower level of primacy effects than do honest interviewers, since they want to avoid detection. For the same reason, we assume that data falsifiers more often pretend to use proxy respondents and to implement grip strength measurements. In this respect, we also expect less rounding of numeric values because inputting an invented but realistic number for the grip strength test does not require any additional effort. As a by-product of this behaviour but also because of the dominant payment structure of European survey agencies disbursing their interviewers per completed interview, we assume that falsifying interviewers perform a larger average number of interviews per day in the field and show higher cooperation rates (i.e., indicators of a better performance). Finally, in contrast to previous research in this field, the data used here offer the possibility of including panel information from previous waves. We expect a higher probability for interview fabrication if there are (unrealistic) large deviations for a respondent's answer between two successive waves. To test this assumption, we included an indicator that measures the absolute deviation in the bodyweight of respondents compared to that measured in the last interview. Table 8.5 presents an overview of the indicators used in our analyses as well as our assumptions for each of them.[18]

Table 8.5: Indicators and hypotheses for curbstoning

Paradata		CAPI data	
Number of contact attempts	-	Duplicates	+
Interviewer notes	-	Straight-lining	+
Interview duration	-	Other answers	-
Number of asked items	-	Code all that apply answers	-
Number of interviews per days in field	+	Follow-up questions	-
Cooperation rate	+	Item nonresponse	-
Cooperation rate of partner	+	Extreme answers	-
		Primacy effect	-
		Proxy respondents	+
		Grip strength: test done	+
		Grip strength: rounding	-
		Deviation from last wave	+

Note: A minus (plus) sign besides the variable indicates that we assume less (more) of this respective indicator for fabricated interviews; e.g., we expect a lower number of contact attempts for falsifying interviewers.

Based on these assumptions, we build a model that predicts curbstoning. More precisely, we used the indicators presented in Table 8.5 in a multivariate cluster analysis[19] to distinguish between two groups of interviewers: those who honestly interviewed their assigned respondents and hence produced valid interviews on one side and those who fabricated their interview data on the other side. During the fieldwork, new interview data were synchronised every two weeks. We started to run the cluster analysis as soon as at least ten interviewers were in the field and a minimum of 500 interviews were conducted in one country to ensure that our analyses were based on a sufficiently large database and thus were robust to outliers. Based on these data, we applied cluster analyses to distinguish completed interviews in two cluster groups for each country that fulfilled the above-mentioned criteria: fabricated versus valid interviews. In a next step, we used these two clustered groups (0: valid interviews, 1: fabricated interviews) as dependent variables in a logistic regression based on all indicators of Table 8.5 as independent variables. To reduce the number of possible false alarms, we only flagged the most suspicious interviews, which had a predicted probability of being fabricated above 95 percent. Furthermore, an interviewer was only flagged as "at risk" when more than 50 percent of his/her completed interviews were flagged as suspicious by our procedure.

8.2.2 Implementation of the new procedure in SHARE Wave 7

We started fieldwork preparations for the seventh wave of SHARE beginning in February 2017; this fieldwork lasted for approximately eight months. The implementation of the procedures described above during fieldwork was communicated in advance as an attempt to determine how well a focused sample works compared to random back checks in identifying fabricated interviews. The statistical procedure and the indicators used were shared beforehand with the survey agencies to increase their willingness to cooperate.

18 Further details on the indicators used can be found in the Appendix of this subchapter.
19 All our analyses were performed with Stata 14. For the cluster analysis we used the *kmeans* command that is implemented in many other software packages and can easily applied by other researches, too; regression analyses are based on the *logistic* command. All scripts are available on request from the authors.

We sent a list of anonymised interviewer IDs that had been flagged as at risk by the cluster analysis to the respective survey agency, which was then requested to check at least three interviews of every listed interviewer by re-contacting the respective interviewed households. If the survey agency detected any irregularities, all interviews of a certain interviewer had to be checked. These checks were performed mostly via telephone in addition to the classical random back checks. The results of all back checks had to be documented in a template that was provided to the survey agencies (see Table 8.6). This template required detailed information and helped to document the outcome of back checks in a more standardised way. The whole procedure was repeated roughly every four weeks during the survey's fieldwork period with the cumulative dataset. As a consequence, it was possible that the same interviewers were identified as at risk several times.

Table 8.6: Standardised template for documentation of back check results

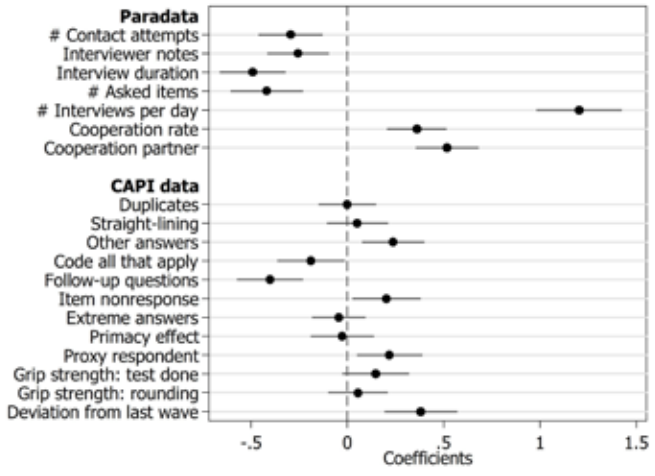
Laptop ID	
Day of the interview	
Month of the interview	
Year of the interview	
Interviewer ID	
ID of the household	
ID of the individual	
Date of birth	
Gender	
Contact results	1 - successful contact
	2 - no phone number
	3 - wrong phone number
	4 - no one answered
	5 - respondent did not want to talk
	6 - deceased
Has an interviewer interviewed you for the study "50+ in Europe"?	1 - yes
	2 - no
	3 - DK
How long was the interview?	minutes
How/where was the interview conducted?	1 - at respondent's home
	2 - at nursing home
	3 - on telephone
	4 - at another place
Did the interviewer use a laptop during the interview?	1 - yes
	2 - no
	3 - DK
Did the interviewer use showcards during the interview?	1 - yes
	2 - no
	3 - DK
Did the interviewer us a device to measure the strength of your hands?	1 - yes
	2 - no
	3 - DK
Did the interviewer behave in a proper way?	1 - yes
	2 - no
	3 - DK
If no: What was wrong with the interviewer's behaviour?	1 - interview ok
	2 - interview not ok/fake
	3 - not sure/DK

8.2.3 Outcome of the new back check procedure

The first time we ran our identification procedure to detect fabricated interviews was the end of March 2017, approximately four weeks after the start of fieldwork in most countries. At that time, three countries had at least ten interviewers in the field that together conducted 500 interviews or more. Based on this sample of 1,621 interviews from 137 active interviewers, we derived two clusters sized 48.4 and 51.6 percent. Based on our hypotheses regarding the presence of certain indicators predicting interview falsification, we can assume that the smaller cluster was predicted to be the fabricated cluster, while the larger cluster consisted of valid cases. After regressing the derived cluster solution on the above-mentioned indicators, we identified 88 suspicious interviews from nine interviewers in the field as at risk for curbstoning. We hence informed the concerned survey agencies and provided them with a list of anonymised IDs of these interviewers. Based on this list, the survey agencies had to check at least three interviews of these at-risk interviewers to determine whether the initial suspicion could be corroborated.

After the first round of focused back checks, our suspicions were not confirmed. While this finding could be interpreted as a good sign that curbstoning did not take place, the survey agencies' back checks suffer from two problems. First, despite their effort to contact all concerned respondents – some survey agencies even conducted personal visits when a certain number of contact attempts by telephone were not successful – several back checks could still not be realised. Second, in some cases, (particularly much older) respondents could not remember being interviewed, while the respective interviewer insisted on having conducted the interview. In both cases, our suspicion could not be verified, and we had to count all interviews from the listed interviewers as valid.

Approximately every four weeks, we repeated the identification procedure with the steadily increasing sample of conducted interviews and provided the survey agencies with anonymised IDs of at-risk interviewers. In this respect, the logistic regression offers the possibility to more closely investigate the patterns as well as the predictive power of each indicator to identify fabricated interviews. Figure 8.23 shows the results of the logistic regression based on the whole sample of 70,133 interviews at the end of fieldwork in Wave 7 in October 2017. As seen, indicators derived from paradata more effectively distinguish between the two clusters than do indicators from the CAPI interview. Apart from being highly significant in differentiating between fabricated and valid interviews, the former also confirm our hypotheses: interviewers stating that they contact their assigned respondents less often, make fewer notes and have shorter interviews with fewer asked items are more likely to deliver fabricated interviews. Moreover, those interviewers show a much better performance, which can be seen with respect to the average number of interviews conducted per day in the field, and achieve higher cooperation rates, both for the assigned respondents and their partners. This is an interesting finding because it shows that a very good performance (sometimes too good to be true) in terms of cooperation/response rates should be carefully evaluated. Additionally, this finding clearly holds some potential for conflict because high response rates are in the interest of both the survey agencies and the public or scientific institutions that run the surveys.



Note: Logistic regression coefficients with 95 percent confidence intervals; SHARE Wave 7 data (end of fieldwork).

Figure 8.23: Predictive power of used indicators to distinguish between fabricated and valid interviews

In contrast, only some indicators that are directly derived from the CAPI data can help to distinguish between fabricated and valid interviews, particularly for follow-up questions (fewer follow-up questions are correlated with a lower probability for curbstoning) and deviations in recorded answers compared to the last wave (more deviations are related to a higher probability for curbstoning). Both findings support our hypotheses. In particular, the last indicator using panel information from previous waves shows the potential of applying our identification procedure based on a longitudinal survey. To a lesser degree, our hypotheses were also confirmed for "code all that apply" answers and the use of proxy respondents. Two other indicators – the selection of "other" in questions with several answers possible and the number of missing data points for sensitive questions – do not support our hypotheses, although they are slightly significant. The other indicators, while largely in line with our hypotheses, do not reach the 95 percent level of significance.

In addition, Table 8.7 shows the result of the cluster analysis at the end of fieldwork in Wave 7. Based on our identification procedure, the survey agencies checked 1226 suspicious interviews from at-risk interviewers out of 70,133 interviews overall by re-contacting the concerned households. From these, two flagged interviewers could be convicted for curbstoning. Overall, 52 interviews from these interviewers could be verified as fabricated. The random back checks identified four interviewers with 67 fabricated interviews out of a total of 28,719 checked interviews. Therefore, it must be stated that our model has a rather low sensitivity (43.7 percent, i.e., we only identified approximately half of all fabricated interviews that could be verified by the survey agencies). However, our targeted back check procedure seems to be more efficient compared to the random back checks when taking into account the number of interviews that actually have been checked based on the respective procedure. The rather low sensitivity, of course, is partly due to our conservative approach of only flagging the most suspicious interviews to reduce the effort on the part of the survey agencies of checking valid interviews with a lower probability of being fabricated. Hence, the rate of false positives, i.e., interviews that turned out to be valid while being flagged as suspicious based on our identification procedure, is very low (only about two percent). However, it must be noted that the number of fabricated interviews that have been verified by the survey agencies is also very low: only 119 interviews from six interviewers could be verified as curbstoning.

Table 8.7: Sensitivity and specificity of the identification procedure in Wave 7

		True state of interview	
		Fabricated	*Valid*
State of interview according to identification procedure	*Fabricated*	43.7 %	1.7 %
	Valid	56.3 %	98.3 %

Note: SHARE Wave 7 data (end of fieldwork).

8.2.4 Concluding remarks

Curbstoning, i.e., the fabrication of an entire interview, is a very rare event in SHARE but can nevertheless lead to unpleasant consequences regarding the panel sample, such as a loss in sample size or the need for time-consuming data corrections of collected information from previous waves. As a consequence, we developed a technical procedure to identify interview fabrication and address this issue during ongoing fieldwork in the seventh wave of SHARE. Overall, we can summarise that our identification procedure, based on a multivariate cluster analysis, is able to identify fabricated interviews, but additional random back checks are useful in increasing the number of detected curbstoning cases. We found that in our case, paradata works better than CAPI data in predicting interview fabrication. The variables that perform best are mostly performance indicators. Specifically, at-risk interviewers show a significantly better (perhaps too good) performance in terms of realising cooperation with both an assigned respondent and a possible partner, as well as conducting more interviews in a shorter time period. As interviewers are paid per completed interview in SHARE, the incentive structure seems rather straightforward here. In line with this finding, shorter interviews with fewer items asked (probably due to the avoidance of time-consuming follow-up questions) are further powerful predictors for suspicious interviews. In addition, the use of panel information from previous waves also significantly helps to distinguish between fabricated and valid interviews.

There are at least two constraints that limit the generalisability of our findings. First, the small number of (detected) fabrications makes it more difficult for any statistical identification procedure to identify curbstoning. The number of verified interview fabrications in our case was not large enough. This, of course, is good news, as one could argue that data quality in SHARE is not severely affected by curbstoning. Although not clearly verifiable, we believe that the mere announcement of detailed interview back checks, both towards survey agencies and interviewers during the training sessions before the beginning of fieldwork, has contributed to this result. This is not to say that we can be perfectly sure that curbstoning is not an issue at all in SHARE, but at least there are no obvious signs of large-scale interview fabrications. Second, we followed a very conservative approach and only determined the most suspicious cases for back checks by the survey agencies. Further investigations showed that fabricated interviews that were not detected by our cluster analysis but rather by the additional random back checks of the survey agencies exhibit a probability slightly below our cut-off criteria of 95 percent. Thus, it might be helpful to reduce the threshold, although this approach leads to a higher chance for false negatives and thus higher costs.

What seems clear is that survey agencies play a key role in this respect, as they – at least in SHARE – are the only entities that can legally contact their interviewers. Hence, it is of utmost importance that a close collaboration is developed among all involved partners (i.e., survey agencies, including interviewers, on one side and scientific institutions responsible for the survey on the other side) and a sincere commitment to the provision of the highest data quality possible.

References

American Association for Public Opinion Research (2016). Standard definitions: Final dispositions of case codes and outcome rates for surveys (9th edition). Retrieved from https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf.

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., et al. (2013). Data resource profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). International Journal of Epidemiology, 42(4), 992-1001. https://doi.org/10.1093/ije/dyt088.

Bredl, S., Storfinger, N., & Menold, N. (2013). A literature review of methods to detect fabricated survey data. In P. Winker, N. Menold & R. Porst (Eds.), Interviewers' deviations in surveys. Impact, reasons, detection and prevention (pp. 3-24). Frankfurt a. M.: Peter Lang.

Ericksen, E. P., & Kadane, J. B. (1985). Estimating the population in a census year: 1980 and beyond: Rejoinder. Journal of the American Statistical Association, 80(389), 98-109. https://doi.org/10.2307/2288062.

Koczela, S., Furlong, C., McCarthy, J., & Mushtaq, A. (2015). Curbstoning and beyond: Confronting data fabrication in survey research. Statistical Journal of the IAOS, 31(3), 413-422. https://doi.org/10.3233/SJI-150917.

Krosnick, J. A., & Alwin, D. F. (1987). An evaluation of a cognitive theory of response-order effects in survey measurement. Public Opinion Quarterly, 51(2), 201-219. https://doi.org/10.1086/269029.

Menold, N., Winker, P., Storfinger, N., & Kemper, C. J. (2013). A method for ex-post identification of falsifications in survey data. In P. Winker, N. Menold & R. Porst (Eds.), Interviewers' deviations in surveys. Impact, reasons, detection and prevention (pp. 25-47). Frankfurt a. M.: Peter Lang.

Murphy, J., Biemer, P., Stringer, C., Thissen, R., Day, O., & Hsieh, Y. P. (2016). Interviewer falsification: Current and best practices for prevention, detection, and mitigation.

Statistical Journal of the IAOS, 32(3), 313-326. https://doi.org/10.3233/SJI-161014.

Werker, H. F. (1981). Results of the 1980 US Census challenged. Population and Development Review, 7(1), 155-167. https://doi.org/10.2307/1972793.

Appendix

Table A1: Description of used indicators

Variable	Mean	SD	Operationalisation
Number of contact attempts	2.12	2.36	Frequency of all contact attempts for one specific interview that have been recorded by the interviewer via telephone, in person or by other means
Interviewer notes	.18	.38	Dichotomous variable indicating whether an interviewer made at least one note (either regarding the person(s) living in the household or a specific question) or not
Interview duration	57.22	22.28	Duration of the complete interview (in minutes) based on all CAPI modules that have been asked
Number of asked items	324.45	94.72	Number of items that have been asked by the interviewer within the whole interview
Number of interviews per days in field	.58	.42	Number of completed interviews conducted by an interviewer divided by the number of days passed since his/her first interview
Cooperation rate	.76	.16	Number of completed interviews divided by the number of interviews (complete plus partial) plus the number of non-interviews that include the contact with an eligible respondent (refusal and break-off plus other) (see AAPOR, 2016; COOP1)
Cooperation rate of partner	.62	.15	Dichotomous variable indicating whether an interview with the partner was done or not
Duplicates	2.58	1.09	Number of identical answers for all CAPI modules; an interview is marked as a duplicate if the questions in at least one module show the same answer pattern
Straight-lining	.48	.14	Frequency of selecting the same answer category across all items in three multi-item sets; this value is standardized by the number of items, taking into account that identical answer patterns are more likely when based on fewer items
Other answers	.52	.74	Frequency of items across all questions for which an "other" category is available in the questionnaire
Code all that apply answers	.92	.57	Frequency of selecting more than one answer option based on five items for which this is possible
Follow-up questions	1.35	.63	Frequency of choosing "no" in four filter questions with follow-up questions
Item nonresponse	1.39	5.57	Number of missing values across all substantial items in the presented questionnaire
Extreme answers	.16	.36	Dichotomous variable indicating whether the (absolute) extreme values on two 11-point scales were chosen
Primacy effect	.50	.78	Frequency of choosing the first answer category in a list of possible answer options based on four variables offering such lists
Proxy respondent	.05	.22	Dichotomous variable indicating whether a respondent is assisted by a so-called proxy respondent in case physical and/or cognitive limitations make it too difficult for him/her to complete the interview by him-/herself
Grip strength: test done	.90	.30	Dichotomous variable indicating whether the grip strength measurement was conducted or not

Variable	Mean	SD	Operationalisation
Grip strength: rounding	.25	.25	Dichotomous variable indicating whether multiples of 5 and 10 have been recorded by the interviewer when the test was done
Panel information	1.43	3.21	Absolute deviation in the measured bodyweight (in kg) of the respondent between Wave 6 and Wave 7

Note: SHARE Wave 7 data (end of fieldwork); SD=standard deviation.

Table A2: Interview length in Wave 7 by subgroups[20]

Interview version	Sample	Mean	Median	SD	Min	Max	N
Single	Baseline/refreshment	77.19	74.27	23.88	48.66	134.38	205
Single	Panel	72.63	68.62	25.04	23.07	182.61	804
Couple, first respondent	Baseline/refreshment	69.55	67.13	21.65	35.65	136.28	272
Couple, first respondent	Panel	69.32	65.94	22.29	25.53	176.94	1133
Couple, second respondent	Baseline/refreshment	59.10	55.18	22.01	24.19	141.03	235
Couple, second respondent	Panel	57.70	54.48	21.24	20.89	166.56	840

Note: SHARE Wave 7 data (Release 0); SD=standard deviation, N=Number of observations.
The sample is restricted to completed interviews. End-of-life interviews and nursing home interviews are excluded; cases with missing keystroke information could not be considered. All analyses are restricted to individual interviews of at least 20 minutes and 200 minutes most. Analyses on couple level only include couple interviews where both interviews were released and both were either longitudinal or baseline/refreshment (i.e., households with new spouses were dropped).

Table A3: Interview length in Wave 7 by countries and subgroups

Country	Interview version	Sample	Mean	Median	SD	Min	Max	N
AT	Single	Baseline/refreshment	122.13	127.72	26.54	93.23	145.43	3
	Single	Panel	75.55	71.93	24.75	24.28	184.88	1026
	Couple, first respondent	Baseline/refreshment	69.58	69.00	10.28	53.85	86.32	12
	Couple, first respondent	Panel	69.93	66.15	22.29	29.57	197.97	1021
	Couple, second respondent	Baseline/refreshment	61.29	58.32	27.33	21.48	123.12	33
	Couple, second respondent	Panel	56.27	52.08	21.48	20.28	189.72	799
BE (fr)	Single	Baseline/refreshment	94.92	94.92		94.92	94.92	1
	Single	Panel	89.70	84.19	27.65	35.70	195.90	698
	Couple, first respondent	Baseline/refreshment	66.72	67.97	13.93	40.58	87.53	8
	Couple, first respondent	Panel	82.38	77.49	24.73	35.30	194.28	850
	Couple, second respondent	Baseline/refreshment	74.98	67.28	23.89	34.38	121.67	28
	Couple, second respondent	Panel	68.95	63.54	25.76	27.43	195.33	462

20 We greatly appreciate the help received from Arne Bethmann in calculating the interview durations for Tables A2 and A3.

Country	Interview version	Sample	Mean	Median	SD	Min	Max	N
BE (nl)	Single	Baseline/refreshment	83.35	83.35		83.35	83.35	1
	Single	Panel	81.60	77.41	26.01	24.02	181.18	624
	Couple, first respondent	Baseline/refreshment	92.40	92.13	15.78	72.15	122.98	10
	Couple, first respondent	Panel	76.06	72.52	22.85	33.83	189.62	1032
	Couple, second respondent	Baseline/refreshment	77.41	70.68	25.19	38.45	151.10	36
	Couple, second respondent	Panel	59.98	56.06	21.27	20.55	144.45	664
BG	Single	Baseline/refreshment	74.51	70.36	31.32	20.15	193.72	602
	Couple, first respondent	Baseline/refreshment	73.35	69.16	29.91	22.78	190.18	658
	Couple, second respondent	Baseline/refreshment	66.53	62.96	29.01	21.53	198.15	616
CH	Single	Baseline/refreshment	100.01	82.01	45.83	69.18	166.85	4
	Single	Panel	88.61	80.53	30.21	35.23	195.12	610
	Couple, first respondent	Baseline/refreshment	83.83	80.55	20.44	64.92	130.73	9
	Couple, first respondent	Panel	81.99	77.82	24.84	30.57	188.32	988
	Couple, second respondent	Baseline/refreshment	74.17	66.83	23.26	34.42	133.97	31
	Couple, second respondent	Panel	70.54	67.25	26.03	22.13	189.45	581
CY	Single	Baseline/refreshment	61.03	59.65	20.39	22.50	148.75	279
	Couple, first respondent	Baseline/refreshment	62.31	59.10	21.85	20.27	174.75	527
	Couple, second respondent	Baseline/refreshment	58.32	53.88	23.01	21.23	199.98	367
CZ	Single	Baseline/refreshment	64.22	63.57	5.44	59.13	69.95	3
	Single	Panel	75.23	72.18	23.78	20.48	181.68	1294
	Couple, first respondent	Baseline/refreshment	68.08	64.67	22.31	42.20	131.23	21
	Couple, first respondent	Panel	70.14	68.05	19.59	24.32	177.35	1447
	Couple, second respondent	Baseline/refreshment	57.08	52.10	23.58	21.45	155.58	63
	Couple, second respondent	Panel	59.86	57.70	20.56	20.22	175.37	1233
DE	Single	Baseline/refreshment	110.64	110.50	22.09	90.57	131.00	4
	Single	Panel	80.97	78.17	25.74	32.10	190.25	835
	Couple, first respondent	Baseline/refreshment	80.55	76.03	32.08	38.02	170.93	19
	Couple, first respondent	Panel	76.81	73.02	22.21	31.73	196.75	1561
	Couple, second respondent	Baseline/refreshment	68.34	65.72	20.70	29.73	128.50	53
	Couple, second respondent	Panel	63.63	60.87	21.57	21.80	198.48	1135
DK	Single	Baseline/refreshment	60.28	60.28	39.77	32.15	88.40	2
	Single	Panel	84.84	79.58	26.63	22.50	199.12	760
	Couple, first respondent	Baseline/refreshment	86.22	91.12	16.45	59.08	113.05	9
	Couple, first respondent	Panel	77.81	74.47	21.33	33.93	166.48	1395
	Couple, second respondent	Baseline/refreshment	75.09	68.45	27.45	27.67	160.25	39
	Couple, second respondent	Panel	63.36	58.70	22.73	21.45	176.12	871
EE	Single	Baseline/refreshment	48.72	48.72		48.72	48.72	1
	Single	Panel	73.63	68.27	29.10	21.08	197.05	1723
	Couple, first respondent	Baseline/refreshment	67.92	68.44	17.82	43.77	93.20	10
	Couple, first respondent	Panel	69.83	65.48	25.15	20.08	191.97	1561
	Couple, second respondent	Baseline/refreshment	50.52	47.35	21.72	20.20	114.07	37
	Couple, second respondent	Panel	58.14	54.60	24.22	20.05	198.67	1331

Country	Interview version	Sample	Mean	Median	SD	Min	Max	N
ES	Single	Baseline/refreshment	73.66	71.32	26.67	34.40	137.48	24
	Single	Panel	63.84	60.61	21.29	21.90	170.82	632
	Couple, first respondent	Baseline/refreshment	61.90	59.90	15.02	29.10	97.33	31
	Couple, first respondent	Panel	62.55	60.41	19.44	20.17	181.83	832
	Couple, second respondent	Baseline/refreshment	52.46	52.43	14.46	27.30	102.02	47
	Couple, second respondent	Panel	53.67	52.91	18.63	21.37	158.40	698
ES (Girona)	Single	Panel	60.75	57.00	24.23	20.12	181.70	455
	Couple, first respondent	Baseline/refreshment	50.25	49.78	20.78	27.70	73.72	4
	Couple, first respondent	Panel	61.61	57.98	21.41	20.00	197.28	803
	Couple, second respondent	Baseline/refreshment	56.26	56.43	13.02	23.22	81.57	22
	Couple, second respondent	Panel	54.92	53.10	20.47	20.25	159.00	648
FI	Single	Baseline/refreshment	99.88	94.67	32.43	29.32	193.73	461
	Couple, first respondent	Baseline/refreshment	88.63	85.50	26.92	37.75	199.10	896
	Couple, second respondent	Baseline/refreshment	80.34	78.03	24.36	29.48	183.23	594
FR	Single	Baseline/refreshment	98.03	98.03	14.05	88.10	107.97	2
	Single	Panel	81.24	77.32	26.84	20.63	192.70	1034
	Couple, first respondent	Baseline/refreshment	70.93	64.40	18.91	44.33	111.03	21
	Couple, first respondent	Panel	74.98	71.30	23.24	27.20	197.92	1133
	Couple, second respondent	Baseline/refreshment	64.08	60.90	23.48	30.47	143.02	43
	Couple, second respondent	Panel	62.15	58.85	22.08	22.68	196.13	826
GR	Single	Baseline/refreshment	88.39	80.63	21.23	72.13	112.42	3
	Single	Panel	79.41	75.61	29.85	20.28	190.67	810
	Couple, first respondent	Baseline/refreshment	77.88	79.57	28.15	29.77	134.93	11
	Couple, first respondent	Panel	77.65	73.95	29.50	21.27	177.20	1105
	Couple, second respondent	Baseline/refreshment	64.48	61.60	27.54	20.45	116.70	37
	Couple, second respondent	Panel	61.89	58.18	25.97	20.18	199.13	936
HR	Single	Baseline/refreshment	72.07	67.49	25.56	33.47	152.02	84
	Single	Panel	65.79	62.62	21.81	22.40	192.93	439
	Couple, first respondent	Baseline/refreshment	66.19	64.82	23.39	21.00	150.47	157
	Couple, first respondent	Panel	65.51	63.67	17.79	22.00	172.92	781
	Couple, second respondent	Baseline/refreshment	55.20	53.52	17.19	20.13	111.10	146
	Couple, second respondent	Panel	56.26	54.72	15.97	20.68	149.62	707
HU	Single	Panel	69.64	65.25	23.17	22.83	163.38	471
	Couple, first respondent	Baseline/refreshment	84.62	87.29	15.69	65.27	98.62	4
	Couple, first respondent	Panel	68.05	64.34	22.33	23.10	169.22	510
	Couple, second respondent	Baseline/refreshment	54.07	49.48	22.46	26.58	126.32	38
	Couple, second respondent	Panel	57.91	56.72	18.72	20.75	140.48	414
IL	Single	Baseline/refreshment	59.74	54.54	20.48	25.32	128.07	72
	Single	Panel	55.31	50.15	22.23	20.50	150.18	449
	Couple, first respondent	Baseline/refreshment	55.95	52.97	20.19	25.25	142.28	100
	Couple, first respondent	Panel	58.85	54.28	24.55	20.10	189.37	724
	Couple, second respondent	Baseline/refreshment	48.52	42.22	18.86	20.40	102.73	74
	Couple, second respondent	Panel	49.05	42.75	22.91	20.38	163.67	535

Country	Interview version	Sample	Mean	Median	SD	Min	Max	N
IT	Single	Baseline/refreshment	113.67	113.67		113.67	113.67	1
	Single	Panel	61.39	58.32	23.30	20.07	174.28	923
	Couple, first respondent	Baseline/refreshment	63.62	63.25	14.74	33.30	96.03	20
	Couple, first respondent	Panel	59.32	56.58	20.84	20.85	198.97	1756
	Couple, second respondent	Baseline/refreshment	59.81	56.89	19.40	25.22	110.80	88
	Couple, second respondent	Panel	49.56	46.32	19.49	20.05	144.98	1430
LT	Single	Baseline/refreshment	72.03	67.02	25.51	26.82	199.13	791
	Couple, first respondent	Baseline/refreshment	65.44	61.33	22.55	21.22	195.87	697
	Couple, second respondent	Baseline/refreshment	56.30	53.42	20.29	20.22	191.30	477
LU	Single	Panel	77.77	74.97	27.34	27.18	189.33	247
	Couple, first respondent	Baseline/refreshment	62.88	59.73	25.24	22.50	105.67	13
	Couple, first respondent	Panel	71.40	70.72	21.09	25.45	134.67	589
	Couple, second respondent	Baseline/refreshment	52.18	43.60	32.38	20.57	195.75	61
	Couple, second respondent	Panel	55.95	54.29	17.47	23.47	108.45	266
LV	Single	Baseline/refreshment	67.41	61.32	28.41	21.60	180.52	709
	Couple, first respondent	Baseline/refreshment	66.62	60.34	28.61	22.50	191.55	544
	Couple, second respondent	Baseline/refreshment	52.53	49.21	20.79	20.23	180.35	428
MT	Single	Baseline/refreshment	58.74	55.37	21.87	24.08	142.58	227
	Couple, first respondent	Baseline/refreshment	61.97	56.97	21.61	24.02	167.63	520
	Couple, second respondent	Baseline/refreshment	48.04	44.80	17.57	20.35	134.17	427
PL	Single	Baseline/refreshment	63.89	62.02	21.21	20.67	184.57	708
	Single	Panel	58.84	56.83	16.68	22.27	139.17	422
	Couple, first respondent	Baseline/refreshment	64.21	62.12	19.41	23.43	160.92	1411
	Couple, first respondent	Panel	58.80	56.88	16.70	20.82	120.92	510
	Couple, second respondent	Baseline/refreshment	56.16	53.72	18.27	20.92	150.42	1046
	Couple, second respondent	Panel	46.73	43.89	16.82	21.03	134.33	406
PT	Single	Panel	49.50	46.10	22.63	21.78	159.75	97
	Couple, first respondent	Baseline/refreshment	50.77	50.77		50.77	50.77	1
	Couple, first respondent	Panel	49.53	44.69	22.69	21.58	148.88	184
	Couple, second respondent	Baseline/refreshment	34.56	34.56	3.22	32.28	36.83	2
	Couple, second respondent	Panel	39.99	36.52	18.47	20.03	125.95	166
RO	Single	Baseline/refreshment	52.33	48.80	22.12	20.05	150.80	511
	Couple, first respondent	Baseline/refreshment	53.69	50.98	21.61	20.17	166.23	797
	Couple, second respondent	Baseline/refreshment	46.99	44.29	18.98	20.20	136.05	632
SE	Single	Baseline/refreshment	66.66	64.87	9.81	57.87	77.23	3
	Single	Panel	90.99	86.87	30.54	20.13	198.45	819
	Couple, first respondent	Baseline/refreshment	79.05	77.88	17.42	46.45	102.50	17
	Couple, first respondent	Panel	83.90	79.67	25.70	36.02	194.12	1325
	Couple, second respondent	Baseline/refreshment	71.17	68.78	24.44	26.27	132.03	43
	Couple, second respondent	Panel	74.70	69.72	28.71	20.70	196.13	820

Country	Interview version	Sample	Mean	Median	SD	Min	Max	N
SI	Single	Baseline/refreshment	59.88	59.88		59.88	59.88	1
	Single	Panel	62.84	59.29	21.33	20.50	192.95	904
	Couple, first respondent	Baseline/refreshment	63.10	58.67	16.01	41.55	98.50	27
	Couple, first respondent	Panel	61.61	58.49	19.24	22.57	147.37	1426
	Couple, second respondent	Baseline/refreshment	49.00	45.03	17.99	23.63	107.20	102
	Couple, second respondent	Panel	52.43	49.33	17.15	20.25	123.67	1036
SK	Single	Baseline/refreshment	75.53	70.91	32.58	21.50	197.23	432
	Couple, first respondent	Baseline/refreshment	69.06	59.43	35.99	20.02	196.63	803
	Couple, second respondent	Baseline/refreshment	60.96	50.33	32.45	20.30	183.20	741

Note: SHARE Wave 7 data (Release 0); SD=standard deviation, N=Number of observations.
The sample is restricted to completed interviews. End-of-life interviews and nursing home interviews are excluded; cases with missing keystroke information could not be considered. All analyses are restricted to individual interviews of at least 20 minutes and 200 minutes most. Analyses on couple level only include couple interviews where both interviews were released and both were either longitudinal or baseline/refreshment (i.e., households with new spouses were dropped).

CHAPTER 9

Weights and imputations

09

9 WEIGHTS AND IMPUTATIONS

Giuseppe De Luca – University of Palermo, Claudio Rossetti – University of Naples Federico II and CSEF

9.1 Introduction

Nonresponse is a serious problem that always affects empirical studies based on survey data. A distinction can be made between at least two types of nonresponse. The first – unit nonresponse – occurs when eligible sample units fail to participate in a survey because of noncontact or explicit refusal to cooperate. The second – item nonresponse – emerges when responding units do not provide useful answers to particular items of the questionnaire, as is often the case with income, wealth and consumption expenditure items. Longitudinal studies are also subject to a third source of nonresponse – sample attrition – when sample units interviewed up to a given wave of the panel drop out in some subsequent wave. These different sources of nonresponse errors may have similar implications, namely, selectivity bias and loss of precision. The key difference is that for unit nonresponse and attrition, all items of the questionnaire are missing, while for item nonresponse, missing values are confined to specific items of the questionnaire. Such a distinction has therefore relevant implications for the auxiliary information that can be used in ex-post adjustment procedures. For unit nonresponse in the baseline sample, the auxiliary information is necessarily confined to that obtained from the sampling frame or the data collection process (in SHARE, this information includes age, gender and regional NUTS-1 indicators). For sample attrition, one may also exploit the information collected in previous waves. Finally, for item nonresponse, one may exploit the additional information collected in the interview process of the current wave.

This chapter provides a description of the weighting and imputation strategies used to address problems of unit nonresponse, sample attrition and item nonresponse in the seventh wave of SHARE. We first describe the composition of the Wave 7 sample resulting from an innovative data collection process that combines a SHARELIFE interview for all respondents who did not participate in Wave 3 and a standard interview for all respondents who had already completed a SHARELIFE interview in Wave 3. We then discuss a variety of calibrated weights available for handling problems of unit nonresponse in the complete cross-sectional sample of Wave 7 (both regular and SHARELIFE respondents and problems of sample attrition in the wave combinations of the SHARE panel. Finally, we describe the multiple imputation procedures used for dealing with missing values due to nonresponse on specific items of the standard and SHARELIFE questionnaires.

9.2 Composition of the Wave 7 sample: SHARELIFE and regular subsamples

The data collection process of Wave 7 consists of a SHARELIFE interview for all respondents who did not participate in Wave 3 and a standard interview for all respondents who had already answered a SHARELIFE interview in Wave 3. Table A. 1 in the Appendix shows a breakdown of the Wave 7 sample by country, interview type and unit of analysis. In total, there are 74,548 individual interviews for 51,233 households over 27 countries.[21] Approximately 82 percent of the respondents answered the SHARELIFE interview, and the remaining 18 percent answered the standard interview. In what follows, we refer to these two subsamples as the SHARELIFE subsample and the regular subsample of Wave 7. Weighting and imputation strategies were tailored to these two subsamples to take into account the different nature of the data collected in the two types of interviews and the strategies that were used to combine the information available in the first seven waves of the SHARE panel.

The SHARELIFE subsample includes 60,832 respondents: 30,166 respondents from 15 countries that did not participate in Wave 3 (i.e., Israel plus 14 other countries that joined SHARE between Wave 4 and Wave 7), 30,166 respondents from the refreshment samples drawn between Wave 4 and Wave 6 in the other 12 countries, and 409 nonresponding partners (NRPs) from Wave 3. Notice that the latter group of SHARELIFE respondents includes eligible partners from the baseline and refreshment samples of Waves 1 and 2 who, unlike their responding partners, did not answer the SHARELIFE interview of Wave 3. Thus, in Wave 7, we find 408 households in which one partner received the standard interview and the other partner received the SHARELIFE interview. Similar considerations apply to "third respondents" and "new spouses/partners" who did not answer the SHARELIFE interview of Wave 3.[22] The regular subsample includes 13,716 respondents who already answered the

21 All numbers are based on SHARE Wave 7 date (Release 0).

22 Third respondents are singles living with a couple, e.g. parents or relatives. Usually, these respondents are entered in the sample at the time of Wave 1, when all household members over 50 years were considered eligible for the interview. The term new spouses/partners refers instead to new spouses and new partners of respondents who were interviewed in same previous wave of the panel. By the SHARE design, the new spouses/partners are always eligible for the interview irrespective of their age.

SHARELIFE interview of Wave 3 and hence received a standard interview in Wave 7. Note that the size of this subsample amounts to 48 percent of the sample observed in Wave 3. Another 14,772 respondents who answered the SHARELIFE interview of Wave 3 were lost due to the cumulative effects of sample attrition between Wave 3 and Wave 7.

9.3 Weighting strategies

Under the ideal situation of complete response, the availability of design weights that compensate for unequal selection probabilities of the sample units allow the users to account for the randomness of probability sampling when estimating the population parameters of interest. Unfortunately, many theoretical properties of inferential procedures based on the sampling design weights depend crucially on the assumption of complete survey response, which is almost never satisfied in the practical implementation of surveys. SHARE is not an exception to this common situation. The baseline and refreshment samples drawn in each wave of the study suffer from unit nonresponse. Moreover, the longitudinal samples from previous waves are subject to attrition at each follow-up. Due to the presence of these sources of non-sampling errors, we usually discourage users from relying on sampling design weights for standard analyses of the SHARE data. These weights are included in the public release of the SHARE data only to favour the implementation and the comparison of alternative statistical procedures that account for nonresponse errors.

The strategy adopted by SHARE for handling unit nonresponse and attrition errors relies on the calibration approach proposed by Deville and Särndal (1992). This approach is preferred to other weighting methods because of its simplicity in achieving the alignment of the sample and the population distributions of some benchmark variables without the need to specify an explicit model for the nonresponse mechanism. Under the assumption that the missing data mechanism is missing-at-random, calibrated weights may help reduce the potential selection bias associated with nonresponse errors. Thus, unless these non-sampling errors are controlled for in other ways, this is the type of weight that we generally recommend for use in standard analyses of the SHARE data. In the remainder of this section, we first discuss the key methodological advantages and limitations of the calibration procedure. Next, we describe the various types of calibrated cross-sectional and longitudinal weights available in Release 7.0.0 of the SHARE data.

9.3.1 The calibration procedure

Let $U=\{1,\ldots,i,\ldots,N\}$ be a finite population of N elements, from which a probability sample $s=\{1,\ldots,i,\ldots,n\} \subseteq U$ of size $n \leq N$ is drawn according to the sampling design $p(\cdot)$. Unless otherwise specified, we shall assume that the inclusion probability $\pi_i=Pr(i \in s)$ and the associated design weights $w_i=\pi_i^{-1}$ are known and strictly positive for all population units. The sampling design weights w_i account for the randomness due to probability sampling. For example, if we wish to estimate the population total $t_y=\sum_{i \in U} y_i$ of a study variable y, then the following Horvitz-Thompson estimator is known to be unbiased:

$$\hat{t}_y=\sum_{i \in s} w_i y_i \qquad (1)$$

That is, it is found that $E_p(\hat{t}_y)=\hat{t}_y$, where $E_p(\cdot)$ denotes the expectation with respect to the sampling design.

Next, we assume that additional information is available to construct a class of more efficient estimators. Let $x_i=(x_{i1},\ldots,x_{iq})^T$ be a q-vector of categorical variables, for which we know the corresponding vector of population totals $t_x=\sum_{i \in U} x_i$ from either the sampling frame or other external sources, such as census data and administrative archives. We refer to the auxiliary variables x_i as calibration variables and to the population totals t_x as calibration margins. The basic idea of calibration is to determine a new set of calibrated weights w_i^* that are as close as possible to the design weights w_i (on average with respect to a distance function), while also satisfying the following constraints:

$$\sum_{i \in s} w_i^* x_i = t_x. \qquad (2)$$

Thus, given a distance function $G(w_i^*,w_i)$, calibration consists of minimising the aggregate distance $\sum_{i \in s} G(w_i^*,w_i)$ with respect to w_i^* subject to the q equality constraints in (2). Note that this constrained optimisation problem does not necessarily admit a solution. However, if a solution exists, then, under certain regularity conditions on the distance function $G(w_i^*,w_i)$, this solution is unique and can always be written as follows:

$$w_i^* = w_i F(\eta_i), i=1,\ldots,n, \qquad (3)$$

where $\eta_i=x_i^T\lambda$ is a linear combination of the calibration variables x_i, $\lambda=(\lambda_1,\ldots,\lambda_q)^T$ is a q-vector of Lagrangian multipliers associated with the constraints (2), and $F(\cdot)$ is a calibration function uniquely determined by the distance function $G(w_i^*,w_i)$.

A key feature of this approach is that many traditional re-weighting procedures, such as post-stratification, raking, and generalised linear regression (GREG), correspond to special cases of the calibration estimator for particular choices of the calibration function $F(\cdot)$, or, equivalently, of the distance function $G(\cdot,\cdot)$, and the vector of calibration variables x_i:

$$\hat{t}_y^* = \sum_{i \in s} w_i^* y_i \qquad (4)$$

Alternative specifications of the distance function $G(w_i^*,w_i)$

and the associated calibration function $F(\eta_i)$ are listed in Table 1 of Deville and Särndal (1992). For example, the chi-square distance function $G(w_i^*,w_i) = (w_i^*-w_i)^2/2w_i$, which leads to the widely used GREG estimator, has the advantage of ensuring a closed-form solution for the calibrated weights w_i^*. The main drawback of this specific function, which is unbounded, is that, depending on the chosen set of calibration variables, the resulting weights can be negative or extremely large. Other specifications of the calibration function allow us to avoid these issues, but a solution to the calibration problem may not exist, and the associated Lagrange multipliers must be obtained by some iterative procedure. In SHARE, we rely on the following logit specification:

$$G(w_i^*, w_i) \propto \left(\frac{w_i^*}{w_i} - l\right) \ln\left(\frac{w_i^*/w_i - l}{1-l}\right) + \left(u - \frac{w_i^*}{w_i}\right) \ln\left(\frac{u - w_i^*/w_i}{u-1}\right),$$

This specification leads to a calibrated function of the form as follows:

$$F(\eta_i; u, l) = \frac{l(u-1) + u(1-l)\exp(a\eta_i)}{u - 1 + (1-l)\exp(a\eta_i)},$$

with $\alpha = [(1-l)(u-1)]^{-1}(u-l)$. Unlike other specifications, this distance function has the advantage of restricting ex-ante the range of feasible values for the calibrated weights by suitable choices of the lower bound l and the upper bound u. Specifically, if a solution exists, then $w_i l \leq w_i^* \leq w_i u$.

As argued by Deville and Särndal (1992), the effectiveness of the calibrated weights depends crucially on the correlation between the study variable y and the calibration variables x. In the extreme case in which y is a linear combination of x, the calibrated estimator $\hat{t}_y^*$ gives an exact estimate of t_y for every realised sample s. They also show that under suitable regularity conditions, the calibration estimator $\hat{t}_y^*$ has desirable asymptotic properties. Moreover, the calibration estimators resulting from alternative specifications of the distance function are asymptotically equivalent to the GREG estimator resulting from the chi-squared specification. Thus, in large samples, calibrated weights are robust to arbitrary choices of the calibration function F(·). Unfortunately, this robustness property does not necessarily extend to the more realistic setting in which survey data are affected by nonresponse errors. In this case, the statistical properties of the calibration estimator can be different from those achieved in the complete response setting due to the additional randomness and possible selection effects generated by the nonresponse process. Previous studies by Lundström and Särndal (1999) and Haziza and Lesage (2016) also suggest that, unlike the complete response setting, alternative specifications of the calibration function F(·) correspond to the imposition of different parametric models on the relationship between the response propensity and the calibration variables.

As an alternative to the calibration approach discussed thus far, a number of more robust weighting methods have been proposed.[23] These methods involve a propensity-score approach for the nonresponse mechanism. Unfortunately, the implementation of these propensity-score methods in SHARE would lead to two main difficulties. First, they would require special effort in modelling the nonresponse mechanisms for the national subsamples of different waves and countries. A second and more important issue is that these methods would also require knowledge of selection probabilities and auxiliary variables for both respondents and non-respondents, which are not available for all countries participating in SHARE. The calibration approach is much simpler than the propensity-score approach because parametric assumptions on the nonresponse process are somehow implicit in the choice of the calibration function F(·). Finally, in addition to external information about the population margins, the calibration approach requires the knowledge of selection probabilities and auxiliary variables only for respondents.

9.3.2 Calibrated cross-sectional weights

As in the previous releases of the SHARE weights, Release 7.0.0 contains a set of calibrated cross-sectional weights for each wave of the study. In particular, the calibrated cross-sectional weights of Wave 7 are defined for the complete sample of 74,548 regular and SHARELIFE respondents. In addition, we provide a set of calibrated cross-sectional weights for both the SHARELIFE and the regular sample.

These three sets of calibrated cross-sectional weights (SHARELIFE+regular, SHARELIFE only, regular interview only) are computed separately by country to match the size of national target populations of individuals in 2017 (i.e., the national populations of individuals who are 50 years or older in 2017). In each country, we use a logit specification of the calibration function F(·) and a set of calibration margins for the size of the target population across eight gender-age groups (i.e., males and females in the age classes ([50-59], [60-69], [70-79], [80+]) and across NUTS-1 regional areas. External information on the population calibration margins is taken from the Central Bureau of Statistics for Israel and from the EUROSTAT regional database for all other countries. Additional information on these external data sources can be found in De Luca and Rossetti (2018). Table A. 2 shows the eight gender-age calibration margins used for the three sets of calibrated cross-sectional weights of Wave 7.

23 See, for example, Brick (2013), Molenberghs et al. (2015), Vermeulen and Vansteelandt (2015), and Haziza and Lesage (2016).

As usual, calibrated cross-sectional weights are computed at the individual level for inference to the target population of individuals and at the household level for inference to the target population of households. At the individual level, we assign an individual-specific weight to each 50+-year-old respondent that depends on the household design weight and the respondent's set of calibration variables. At the household level, we assign instead a common calibrated weight to all interviewed household members that depends on the household design weight and the calibration variables of all 50+ respondents in that household. By construction, calibrated cross-sectional weights are missing for respondents younger than 50 years (i.e., age-ineligible partners of an age-eligible respondent), for those with missing information on the calibration variables (i.e., year of birth, gender and NUTS-1 code), and for those with missing sampling design weights (i.e., respondents with missing sampling frame information).

9.3.3 Calibrated longitudinal weights

In addition to cross-sectional calibrated weights, Release 7.0.0 also includes calibrated longitudinal weights for longitudinal data analyses. The theoretical framework of these weights is similar to the framework described in the previous section for cross-sectional weights. Nonetheless, there are two important differences. First, calibrated longitudinal weights are defined for the balanced subsample of respondents who participated in at least two waves of the study. Second, since mortality is a source of attrition that affects both the sample and the population, calibrated longitudinal weights account for the mortality of the original target population across waves. The target population for panel data analysis is therefore defined as the target population at the beginning of a reference time period that survives up to the end of the period considered (see, e.g., Lynn, 2009).

Note that one can compute many different types of calibrated longitudinal weights depending on the selected combination of waves and the basic unit of analysis (either individuals or households). To simplify the structure of the public release of the data, we provide calibrated longitudinal weights only for selected wave combinations of its panel. The wave combinations available in Release 7.0.0 are the 6 possible couples of any two adjacent waves (i.e., wave combinations 1-2, 2-3, 3-4, 4-5, 5-6, and 6-7) and the fully balanced panel (i.e., wave combination 1-2-3-4-5-6-7).

The calibrated longitudinal weights of the generic wave combination t-…-s are computed separately by country to represent the national populations of Wave t that survive up to the interview year of Wave s. For example, in most countries, the wave combination 1-2 represents the national population of people aged 50+ in 2004 who survive until 2006. The fully balanced panel instead allows for the representation of the national population of people aged 50+ in 2004 who survive until 2017. Specifically, in each country, we use a logit specification of the calibration function $F(\cdot)$ and a set of calibration margins for the size of the target population across eight gender-age groups (i.e., males and females with age at the time of the starting wave in the classes ([50-59], [60-69], [70-79], [80+]). Compared to calibrated cross-sectional weights, we do not control for NUTS-1 regional areas due to the relatively lower sample size in this case. Moreover, we account for mortality in the population by subtracting from each calibration margin the corresponding number of deaths that occurred between the interview years of Wave t and Wave s. The calibration margins of all longitudinal weights are reported in Tables A. 3-9.

Similar to calibrated cross-sectional weights, calibrated longitudinal weights are available both at the individual and at the household level. For the individual weights, the balanced sample consists of respondents interviewed in each wave of the chosen wave combination. For the household weights, the balanced sample consists instead of households with at least one eligible member interviewed in each wave of the selected wave combination. Note that these definitions imply that the balanced sample of households is larger than the corresponding balanced sample of individuals. For example, couples with one partner participating in Wave 6 and the other partner participating in Wave 7 belong to the balanced sample of households for the wave combination 6-7, even if neither partner belongs to the corresponding balanced panel of individuals.

9.3.4 Structure of the SHARE weights in Release 7.0.0

Release 7.0.0 contains 14 datasets of calibrated weights. The first seven datasets contain the cross-sectional weights of any single wave from 1 to 7 (named sharewt_rel7-0-0_gv_ weights for the generic Wave t=1,…,7) and are provided together with the other modules of each corresponding wave. The remaining datasets are delivered as separate download files: seven datasets with the calibrated longitudinal weights of the selected wave combinations 1-2, 2-3, 3-4, 4-5, 5-6, 6-7, 1-2-3-4-5-6-7 (named sharewX_rel7-0-0_gv_longitudinal_weights_wt_ws for the generic wave combination t-…-s). Tables A. 10-11 provide a description of the variables included in these 14 datasets. In addition to the individual, household and country identifiers, each dataset includes the following:

- one variable for the relevant sampling design weights
- two variables for the underlying type of calibrated weights at the household and the individual levels

- one indicator that identifies the various subsamples drawn in any specific country and wave of the SHARE panel
- four indicators for the information about stratification and clustering in each subsample

Longitudinal datasets also contain a binary indicator that allows for the identification of a balanced sample of individuals within the balanced sample of households in the chosen wave combination. In this case, sampling design information (e.g., design weights, strata and clusters) refers to the initial wave of the chosen wave combination.

9.3.5 Supplementary material and user guide on calibrated weights

Since the SHARE panel now consists of seven waves, one can in principle compute (2^7-1)*2=254 different types of calibrated longitudinal weights, depending on the balanced sample of the selected combination of waves and the basic unit of analysis (either individuals or households). In addition, one can compute many different types of calibrated cross-sectional weights for specific subsamples of the data available in each wave (e.g., the respondents to the vignette questionnaires of Waves 1 and 2 or the drop-off questionnaires of Waves 1 to 6). These considerations make it clear why the strategy of providing all possible calibrated cross-sectional and longitudinal weights is not feasible, especially in the future, when additional waves will be available. For cross-sectional studies based on specific subsamples and longitudinal studies based on other wave combinations, users are required to control for the potential selection effects of unit nonresponse and attrition by computing their own calibrated weights or by implementing some alternative correction method.

To support users in this nontrivial methodological task, we provide a set of Stata do-files and ado-files that illustrate step-by-step how to compute calibrated cross-sectional and longitudinal weights and one data file with information on population size and number of deaths by year, gender, age and NUTS-1 regional areas. Registered users can download this supplementary material on calibrated weights from the SHARE data dissemination website. A discussion of the step-by-step operations can also be found in the accompanying user guide "Computing Calibrated Weights".

9.4 Imputations of missing values due to item nonresponse

Due to the different nature of the data collected in the two types of interviews in Wave 7, missing values due to item nonresponse in the regular and SHARELIFE subsamples were imputed separately. In the regular subsample, we imputed the missing values on the key survey variables using a model that is similar to those used in the most recent waves of SHARE (see, e.g., De Luca et al. 2015). Compared to Waves 4-6, there are two major differences. First, the imputation model for the monetary variables relies on a slightly more parsimonious specification to account for the smaller sample size of the regular subsample of Wave 7. Second, to compute the total household income of households with a mixed interview type among the household members, we treated the respondents to the SHARELIFE interview as NRPs in the regular interview. The problem is that to reduce the length of the interview process, detailed questions on current income sources were asked only of respondents of the regular interview. Thus, like the standard NRP problem, incomes of all household members who answered the SHARELIFE interview are unobservable, and focusing attention on the observable incomes of the regular respondents may introduce a severe downward bias in the aggregated measure of total household income. The basic idea of our imputation procedure is to address this problem of unit nonresponse by exploiting the discrepancies in the distribution of total household income for couples with and without NRPs and the additional information obtained from a one-shot question on monthly household income (HH017).

Unfortunately, the imputation model currently used in SHARE is unsuitable for producing reliable imputations of the missing values on life history data collected in the SHARELIFE interview. This issue explains why we have not thus far provided imputations of the SHAREFILE data collected in Wave 3. The development of a suitable imputation model for this type of data is a challenging task, which we have left for future research. As a preliminary step in the construction of complete imputation datasets for the SHARELIFE data of Waves 3 and 7, we now provide imputations of those variables that have a well-defined interpretation in terms of retrospective cross-sectional data.

9.4.1 Imputations of variables collected in the regular interview

The imputation model for the variables collected in the standard interview of Wave 7 is similar to the imputation models used in the previous regular waves of SHARE. As usual, the model was adapted to the specific features of the Wave 7 interview in terms of branching, skip patterns, proxy interviews, country-specific deviations from the generic version of the questionnaire, and availability of partial information from the sequence of unfolding bracket questions. However, we preserved as much as possible the comparability of the imputations across different waves of the SHARE panel. Our procedure draws upon a hot-deck imputation method for various types of variables affected by negligible fractions

of missing values (usually, much less than 5 percent) and a fully conditional specification (FCS) method for monetary variables affected by more relevant fractions of missing data.

The hot-deck method involves replacing the missing values of one or more variables for a non-respondent (called the recipient) with the observed values from a respondent (called the donor) who is "similar" to the recipient according to some metric (see, e.g., Andridge & Little, 2010). The FCS method (van Buuren et al., 1999) is instead a Gibb sampling algorithm that imputes multiple variables jointly and iteratively via a sequence of (univariate) regression models. Specifically, at each step of this iterative process, one imputes the missing values on the j-th variable (j=1,…,J) by estimating a model that includes the most updated imputed values of the other J-1 variables as predictors. Fully observed variables, if available, can also be used as predictors. The imputed values for the j-th variable are then obtained as draws from the predictive distribution implied by the estimated model. The same procedure is applied sequentially to the set of J variables affected by missing data and is repeated in a cyclical manner by overwriting, at each iteration, the imputed values from the previous iteration. Despite a lack of rigorous theoretical justification (see, e.g., Arnold et al., 1999, 2001; van Buuren, 2007), the FCS method has become one of the most popular multivariate imputation procedures due to its flexibility in handling complicated data structures and its ability to preserve the correlations of the imputed variables (Raghunathan et al. 2001; van Buuren et al. 2006). Comparisons of the FCS method with other multivariate imputation techniques can be found in Lee and Carlin (2010).

To allow data users to take into account the additional variability generated by the imputation process, we always provide five imputations of the missing values. These multiple imputations were constructed through five independent replicates of either the hot-deck or the FCS imputation procedures. Notice that neglecting this additional source of uncertainty by selecting only one of the five available replicates may result in misleadingly precise estimates.

As in previous waves, we performed hot-deck imputations at an early stage, separately by country, and according to a convenient order of the variables that accounts for branching and skip patterns in the Wave 7 questionnaire. In our implementation of this technique, the donors were selected randomly from imputation classes constructed on auxiliary variables that are also observed for the recipients. We first imputed few missing values on basic socio-demographic characteristics, such as age and education, so that these variables could then be used as auxiliary variables to impute the other variables. More precisely, the set of auxiliary variables typically included country, gender, five age classes ([-49], [50-59], [60-69], [70-79], [80+]), five groups for years of education, and two groups for self-reported health. For some variables, we used a larger set of auxiliary variables. For example, we also used the number of children for imputation of the number of grandchildren and an indicator for being hospitalised overnight during the last year for imputation of health-related variables. Variables that were known to be logically related, such as respondent's weight, height and body mass index, were imputed jointly. In total, the Wave 7 imputation dataset contains multiple hot-deck imputations for more than 70 variables collected in the standard interview.

The second stage of the imputation process addresses the more worrisome issue of item nonresponse on monetary variables, such as income from various sources, assets, and consumption expenditures, which were collected through retrospective and open-ended questions that are sensitive and difficult to answer precisely. To illustrate, we show in Figure 9.1 the item nonresponse rates of two monetary variables: value of the house (HO024) and amount in bank account (AS003). For the first variable, the percentage of missing values among the eligible respondents ranges from a minimum of 4 percent in Denmark to a maximum of 54 percent in Spain. The percentage of missing values becomes even more dramatic for questions that are likely to be very sensitive for the respondents, for example, "About how much do you and your partner currently have in bank accounts, transaction accounts, saving accounts or postal accounts?" In 10 out of 12 countries, more than 20 percent of the respondents either refused or did not know how to answer this question. The resulting cross-country average of the item nonresponse rate is equal to 30 percent.

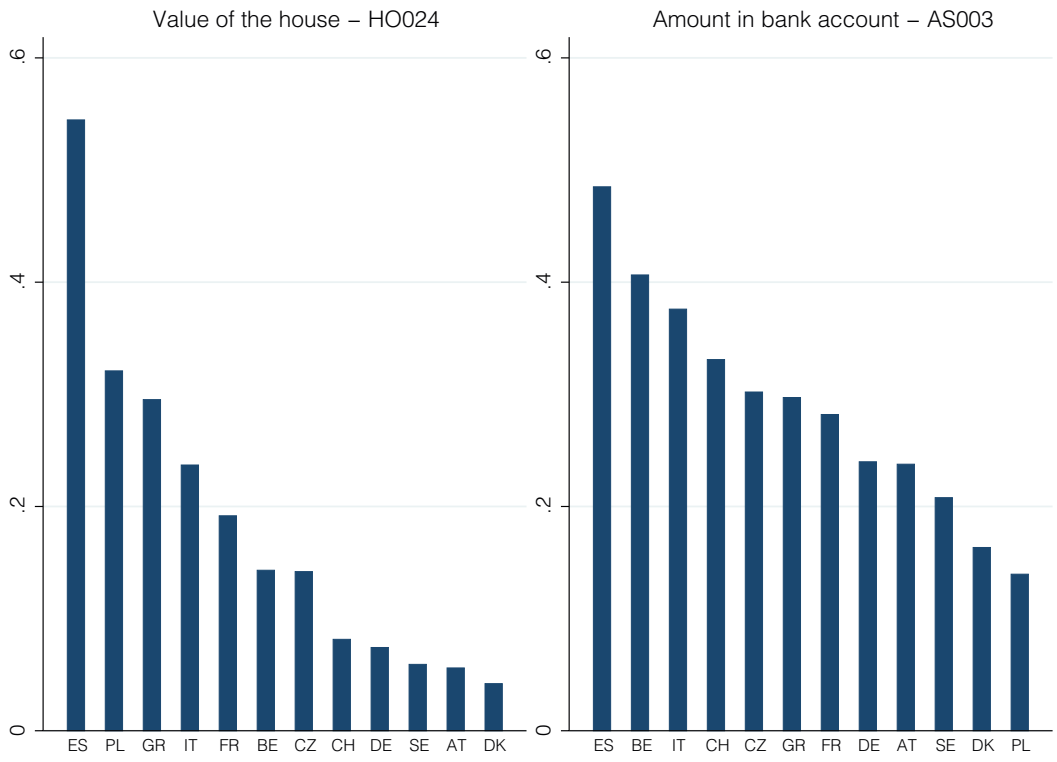


Figure 9.1: Percentage of missing values for some monetary variables by country

Starting from Wave 1, we handled these large fractions of missing values by FCS imputations. In the previous regular waves, FCS imputations were performed separately by country and household type to account for heterogeneity across these different groups. Specifically, we considered three household types: singles and third respondents (sample 1), couples with both partners interviewed (sample 2), and all couples with and without NRPs (sample 3). In addition to heterogeneity across household types, the distinction between the first two samples was motived by the possibility of using characteristics of the partner of the designated respondent as predictors to impute the missing monetary amounts within couples. Furthermore, the overlapping partition of the last two samples was motivated by the need to properly impute total household income in couples with NRPs. An extended discussion of these issues can be found in De Luca et al. (2015). In Wave 7, we introduced two simplifications due to the lower sample size of the regular subsample. First, we pooled the first two samples to increase the number of observations available in the estimation step of the FCS method. In what follows, we shall refer to the pooling of these two samples as sample 1-2. Second, for all couples within sample 1-2 and sample 3, we limited the subset of predictors referring to characteristics of the partner of the designed respondent to age and year of education only. Except for these two simplifications, the FCS imputation model of Wave 7 is identical to those employed in the previous waves.

In principle, after aggregating the original income, wealth and consumption expenditure items as in the previous waves (see De Luca et al., 2015), we imputed by the FCS method approximately 40 monetary variables. In practice, however, the exact list of variables used in the Gibbs sampling algorithm was country- and sample-specific because of the additional requirement of having at least 100 observations for estimating the regression model of each variable in each subsample. Monetary variables that did not satisfy this requirement were imputed first and then used as observed predictors in the imputation of the other variables.

For most variables, we used a two-part model that involved a probit model for ownership and a regression model for the amount conditional on ownership.[24] Conditional on eligibility and ownership, non-zero values of monetary variables were converted (if needed) in annual euro amounts to avoid differences in the time reference period of each question and the national currencies of non-euro countries. Furthermore, to account for skewness in the right tails of the conditional distribution of each monetary variable, we transformed all

24 For few variables with no ownership question, such as food at home consumption and total household income, we used a linear regression model.

strictly positive variables in logarithms and all variables that may also take negative values (such as income from self-employment, bank account, and value of own business) according to the inverse hyperbolic sine transformation. The set of observed predictors (i.e., the variables imputed previously by the hot-deck method) included gender, age, years of education, self-perceived health, number of children, number of chronic diseases, score of the numeracy test, employment status and willingness to answer. For respondents belonging to sample 1-2, we also used a binary indicator for being part of a couple and its interactions with age and years of education of the partner. For respondents belonging to sample 3, which consists of couples only, we instead used only age and years of education of the partner. In few cases where the number of observations available for estimation was lower than 30, we exploited a smaller subset of predictors based on gender, age, years of education and self-reported health only. Imputed monetary amounts were constrained to fall within individual-level bounds that summarised the partial information available on the missing values (e.g., country-specific thresholds used to trim outliers in the tails of the observed distribution of each monetary variable, individual-level bounds obtained from the sequence of unfolding bracket questions, and lower bounds based on the observed components of aggregated monetary variables).

In the Gibbs sampling algorithm for the FCS imputations, the sequence of imputations for the j-th monetary variable was performed in a similar fashion by including the most updated imputations of the other J-1 monetary variables as additional predictors. This additional set of predictors was excluded from the regression model only in the first iteration of the Gibbs sampling algorithm to initialise the starting values. The convergence of the algorithm was assessed by the Gelman-Rubin criterion (Gelman & Rubin, 1992; Gelman et al., 2004) applied to the mean, the median and the 90th percentile of the five imputed distributions of each monetary variable.

Like the interviews of Waves 4-6, the information collected in the standard interview of Wave 7 allows us to define two alternative measures of total household income. The first measure (thinc) is obtained by a suitable aggregation at the household level of all individual income components, while the second (thinc2) is obtained by the one-shot question on monthly household income (HH017). As argued in De Luca et al. (2015), it is difficult to find strong arguments to prefer one measure to the other. Moreover, the availability of these two alternative measures may greatly improve the imputation process because each measure could contribute relevant information on the missing values of the other measure. The procedure adopted in Wave 7 to impute the two available measures of total household income consisted of the following two stages.

- Stage 1 – Imputations for sample 1-2:
  We first imputed all monetary variables by the FCS method discussed before. At the end of each iteration, we then computed the first measure of total household income (thinc), household net worth (hnetw) and total household expenditure (thexp) by suitable aggregations of the imputed income, wealth and expenditure items. We finally imputed the second measure of total household income (thinc2), using as predictors the first measure of total household income (thinc), household net worth (hnetw), total household expenditure (thexp), and characteristics of the household respondent. The imputed values of thinc2 were constrained to fall in the bounds derived from the sequence of UB questions for HH017.

- Stage 2 – Imputations for couples with NRPs in sample 3
  We first selected from sample 1-2 all couples in which both partners were interviewed, including the imputations of their missing monetary amounts computed in the previous stage. In stage 2, these couples were included in sample 3 only as donors to impute the missing monetary amounts in couples with NRPs. Similar to stage 1, we first imputed all monetary variables by the FCS method applied to the responding partners of couples with NRPs. At the end of each iteration, we also imputed the second measure of total household income (thinc2) using household net worth (hnetw), total household expenditure (thexp), and characteristics of the responding partner as predictors. Notice that, unlike stage 1, we no longer used thinc as a predictor of thinc2. In fact, the first measure of total household income was considered to be missing due to the unobserved incomes of the NRPs. What could be computed by summing the incomes of the responding partner was only a lower bound. Thus, in stage 2, thinc was imputed as the latest variable using the second measure of total household income (thinc2), household net worth (hnetw), total household expenditure (thexp) and characteristics of the responding partner as predictors, while couples with two partners interviewed served as donors and the total income of the responding partner served as the lower bound.

The full list of variables included in the imputation dataset of Wave 7 is presented in Table A. 12. For each imputed variable, we also provide a flag variable (named variableame_f), which summarises the status of the imputation process, as illustrated in Table A. 13.

9.4.2 Imputations of variables collected in the SHARELIFE interview

The imputation procedure for the variables collected in the

SHARELIFE interview draws mainly on the hot-deck method, as the fraction of missing values on these variables imputed is generally low (less than 5 percent). As explained above, we imputed by the hot-deck method only those variables that have a well-defined interpretation in terms of retrospective cross-sectional data. These variables include, for instance, basic socio-demographic characteristics (e.g., age and education), objective and subjective measures of current health status (e.g., maximum grip strength, body mass index and self-reported health status), measures of current cognitive abilities (e.g., score of words list learning test), socio-economic and health conditions in childhood or adolescence (e.g., number of books at age 10 and health status at age 15), Big Five personality trait measures, and conditions of the interview process.

Similar to the variables collected in the standard interview, we performed hot-deck imputations separately by country and according to a convenient order of the variables that accounts for branching and skip patterns in the SHARELIFE interview. Imputation classes for the implementation of this method were also based on the same set of auxiliary variables, namely, country, gender, five age classes ([-49], [50-59], [60-69], [70-79], [80+]), five groups for years of education, and two groups for self-reported health. Moreover, variables that were known to be logically related, such as respondent's weight, height and body mass index, were imputed jointly. In total, the Wave 7 imputation dataset contains multiple hot-deck imputations for approximately 40 variables collected in the SHARELIFE interview.

In addition to hot-deck imputations, we also provide FCS imputations of four monetary items that were asked of household respondents independently of the interview type. Specifically, from the CO module of the SHAREFILE interview, we have information about the consumption of food at home (CO002), the consumption of food outside the home (CO003), the consumption of food produced in the home (CO011) and total household income in a typical month (HH017). Figure 9.2 illustrates the item nonresponse rates on food-at-home consumption and total household income by country. The cross-country averages of the item nonresponse rates on these two variables are equal to 12 and 17 percent, respectively. In some countries, such as Spain, Portugal and Israel, the item nonresponse rates are considerably larger than the corresponding cross-country averages.

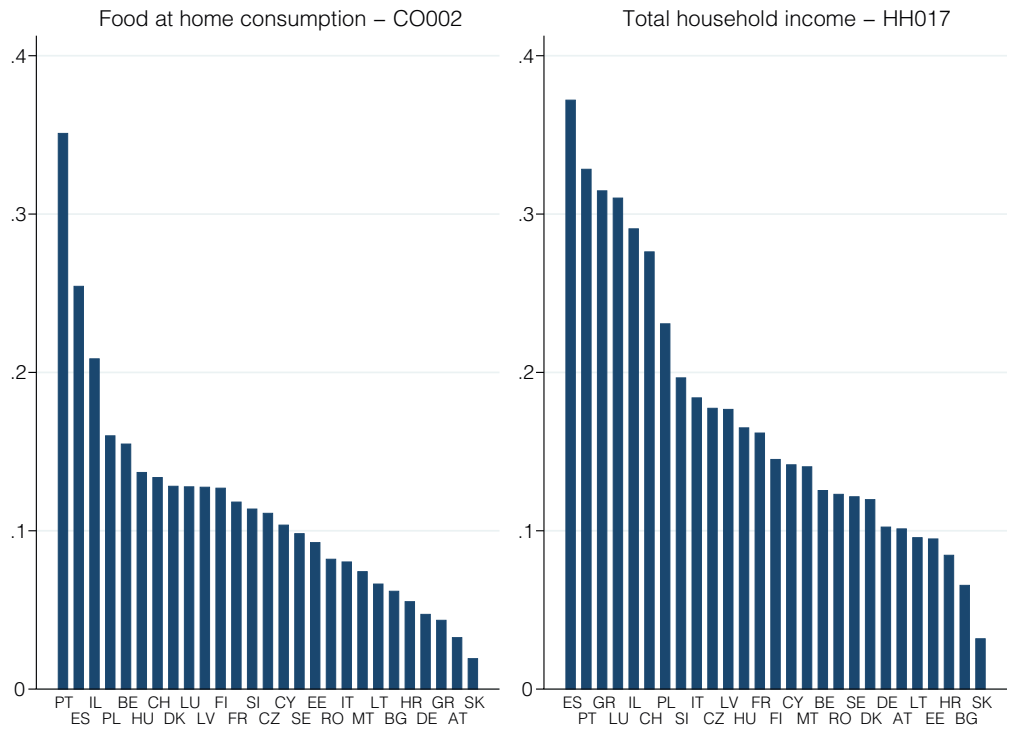


Figure 9.2: Percentage of missing values for some monetary variables of the SHARELIFE subsample by country

FCS imputations of the four monetary variables were again performed separately by country and household type. Since the SHARELIFE interview does not provide information on the first measure of total household income (thinc), there was no need to distinguish between couples with or without NRPs. In this case, given the relatively larger sample size of the SHARELIFE subsample, we considered only two household types: singles and third respondents (sample 1) and all couples (sample 3). In the Gibbs sampling algorithm, we used a linear regression model for the consumption of food at home and total household income (thinc2) and a two-part model for the consumption of food outside the home and home-produced consumption. The four variables were transformed into logarithms of their annual euro amounts. The set of observed predictors included socio-demographic characteristics of the household respondent (e.g., gender, age, years of education, self-reported health, number of children, number of chronic diseases, and willingness to answer) for the sample of singles and third respondents, plus the age of the partner of the household respondent for the sample of couples. Similar to the variables collected in the standard interview, we provide five imputations of the missing values using five independent replicates of either the hot-deck or the FCS imputation procedures.

References

Andridge, R., & Little, R. (2010). A review of hot deck imputation for survey non-response. International Statistical Review, 78(1), 40-64. https://doi.org/10.1111/j.1751-5823.2010.00103.x.

Arnold, B. C., Castillo, E., & Sarabia, J. M. (1999). Conditional specification of statistical models. New York: Springer.

Arnold, B. C., Castillo, E., & Sarabia, J. M. (2001). Conditionally specified distributions: An introduction. Statistical Science, 16(3), 249-274. https://doi.org/10.1214/ss/1009213728.

Brick, J. M. (2013). Unit nonresponse and weighting adjustments: A critical review. Journal of Official Statistics, 29(3), 329–353. https://doi.org/10.2478/jos-2013-0026.

De Luca, G., Celidoni, M., & Trevisan, E. (2015). Item nonresponse and imputation strategies in SHARE Wave 5. In F. Malter & A. Börsch-Supan (Eds.), SHARE Wave 5: Innovations & methodology. Munich: Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC).

De Luca, G., & Rossetti, C. (2018). Stata program to produce calibrated survey weights. Deliverable 2.10 of the SERISS project funded under the European Union's Horizon 2020 research and innovation programme GA No: 654221. Retrieved from www.seriss.eu/resources/deliverables.

Deville, J.-C., & Särndal, C.-E. (1992). Calibration estimators in survey sampling. Journal of the American Statistical Association, 87(418), 376-382. https://doi.org/10.1080/01621459.1992.10475217.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. Statistical Science, 7(4), 457-511. https://doi.org/10.1214/ss/1177011136.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). Bayesian data analysis (2nd ed.). Boca Raton: Chapman and Hall.

Haziza, D., & Lesage, E. (2016). A discussion of weighting procedures for unit nonresponse. Journal of Official Statistics, 32(1), 129–145. https://doi.org/10.1515/jos-2016-0006.

Lee, K. J., & Carlin, J. B. (2010). Multiple imputation for missing data: Fully conditional specification versus multivariate normal imputation. American Journal of Epidemiology, 171(5), 624-632. https://doi.org/10.1093/aje/kwp425.

Lundström, S., & Särndal, C. E. (1999). Calibration as a standard method for treatment of nonresponse. Journal of Official Statistics, 15(2), 305-327. https://doi.org/10.3233/SJI-160996.

Lynn, P. (2009). Methods for longitudinal surveys. In P. Lynn (Ed.), Methodology of longitudinal surveys (pp. 1-19). Chichester: Wiley. https://doi.org/10.1002/9780470743874.ch1.

Molenberghs, G., Fitzmaurice, G., Kenward, M. G., Tsiatis, A., & Verbeke, G. (2015). Handbook of missing data methodology. New York: CRC Press.

Raghunathan, T. E., Lepkowski, J. M., Van Hoewyk, J., & Solenberger P. (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. Survey Methodology, 27(1), 85–95.

Van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. Statistical Methods in Medical Research, 16(3), 219-242. https://doi.org/10.1177/0962280206074463.

Van Buuren, S., Boshuizen, H. C., & Knook, D. L. (1999). Multiple imputation of missing blood pressure covariates in survival analysis. Statistics in Medicine, 18(6), 681–694. https://doi.org/10.1002/(SICI)1097-0258(19990330)18:6<681::AID-SIM71>3.0.CO;2-R.

Van Buuren, S., Brands, J. P. L., Groothuis-Oudshoorn, C. G. M., & Rubin, D. B. (2006). Fully conditional specification in multivariate imputation. Journal of Statistical Computation and Simulation, 76(12), 1049-1064. https://doi.org/10.1080/10629360600810434.

Vermeulen, K., & Vansteelandt, S. (2015). Biased-reduced doubly robust estimation. Journal of the American Statistical Association, 110(511), 1024-1036. https://doi.org/10.1080/01621459.2014.958155.

Appendix

Table A. 1: Composition of the Wave 7 sample by interview type and country

Country	Respondents			Households			
	SHARELIFE	Regular	Total	SHARELIFE	Regular	Mixed	Total
AT	2665	482	3147	1864	348	17	2229
DE	2967	829	3796	1967	556	36	2559
SE	2121	1055	3176	1469	753	60	2282
ES	3242	1248	4490	2111	832	28	2971
IT	2889	1538	4427	1783	944	76	2803
FR	2148	1122	3270	1534	772	39	2345
DK	1940	1280	3220	1357	903	25	2285
GR	1142	1871	3013	732	1253	17	2002
CH	1621	748	2369	1155	542	28	1725
BE	3272	1541	4813	2402	1083	33	3518
IL	2035	0	2035	1382	0	0	1382
CZ	3196	901	4097	2157	598	31	2786
PL	3463	1101	4564	2341	733	18	3092
LU	1226	0	1226	881	0	0	881
HU	1467	0	1467	998	0	0	998
PT	441	0	441	275	0	0	275
SI	3376	0	3376	2294	0	0	2294
EE	5015	0	5015	3536	0	0	3536
HR	2333	0	2333	1478	0	0	1478
LT	2034	0	2034	1538	0	0	1538
BG	2003	0	2003	1348	0	0	1348
CY	1232	0	1232	845	0	0	845
FI	2003	0	2003	1393	0	0	1393
LV	1754	0	1754	1307	0	0	1307
MT	1225	0	1225	774	0	0	774
RO	2080	0	2080	1390	0	0	1390
SK	1942	0	1942	1197	0	0	1197
Total	60832	13716	74548	41508	9317	408	51233

Table A. 2: Gender-age national calibration margins for the cross-sectional weights of Wave 7

Country	Men				Women				Total
	[80+]	[70-79]	[60-69]	[50-59]	[80+]	[70-79]	[60-69]	[50-59]	
AT	149.995	326.330	445.147	652.650	280.216	401.622	486.192	655.688	3.397.840
BE	222.326	377.930	627.529	793.963	396.256	455.070	657.480	789.397	4.319.951
BG	118.515	255.874	445.240	483.067	214.522	380.255	540.512	495.346	2.933.331
CH	153.177	296.776	441.553	616.375	263.690	347.816	461.019	601.676	3.182.082
CY	11.536	26.540	43.632	52.480	16.639	30.130	45.731	54.272	280.960
CZ	138.779	345.095	665.306	671.034	281.757	473.789	749.044	667.467	3.992.271
DE	1.687.679	3.737.040	4.609.580	6.523.318	3.041.524	4.502.051	4.924.360	6.470.087	35.495.639
DK	92.574	229.599	335.339	384.233	151.184	256.717	345.868	381.015	2.176.529
EE	16.818	37.998	66.436	82.919	50.276	71.424	90.883	93.043	509.797
ES	1.027.475	1.606.242	2.365.630	3.239.265	1.758.700	1.953.613	2.550.809	3.295.406	17.797.140
FI	97.309	206.094	364.669	370.591	186.172	252.265	388.254	372.020	2.237.374
FR	1.375.228	2.124.016	3.808.024	4.280.140	2.531.521	2.588.040	4.190.633	4.493.330	25.390.932
GR	287.191	441.770	594.737	690.217	413.820	539.416	660.778	762.100	4.390.029
HR	64.558	144.837	252.127	296.592	137.655	212.073	286.234	310.871	1.704.947
HU	124.395	302.404	575.647	594.265	296.536	495.600	727.102	647.375	3.763.324
IL	98.205	176.424	334.558	386.246	149.143	214.867	374.857	409.392	2.143.692
IT	1.455.925	2.550.154	3.512.422	4.327.588	2.593.132	3.085.625	3.818.736	4.531.162	25.874.744
LT	40.606	86.088	131.111	201.987	111.267	164.084	186.727	235.949	1.157.819
LU	8.259	15.820	27.441	41.774	14.578	18.540	27.121	38.821	192.354
LV	23.977	60.305	94.143	131.075	74.437	120.449	135.759	153.176	793.321
MT	6.482	16.021	28.827	29.550	11.378	18.904	29.669	29.318	170.149
PL	479.800	922.771	2.219.127	2.530.401	1.083.676	1.419.100	2.644.887	2.657.049	13.956.811
PT	218.465	400.416	576.516	693.270	395.988	532.209	669.085	768.010	4.253.959
RO	295.399	599.761	1.111.504	1.180.879	539.248	898.070	1.337.533	1.217.547	7.179.941
SE	193.233	411.791	569.465	619.461	308.417	444.855	578.113	605.198	3.730.533
SI	31.976	69.904	127.656	153.748	70.567	91.736	132.361	150.226	828.174
SK	51.981	128.250	292.991	363.768	119.109	205.763	349.134	377.478	1.888.474

Table A. 3: Gender-age national calibration margins for the longitudinal weights of Waves 1-2

Country	Men				Women				Total
	[80+]	[70-79]	[60-69]	[50-59]	[80+]	[70-79]	[60-69]	[50-59]	
AT	69.969	224.051	386.023	468.330	189.183	337.541	434.087	484.501	2.593.685
BE	102.098	333.874	454.387	658.100	231.058	467.188	503.189	660.348	3.410.242
CH	82.107	203.892	332.010	476.661	172.350	285.867	364.536	478.180	2.395.603
DE	720.789	2.444.070	4.879.081	4.891.028	1.978.719	3.575.138	5.298.005	4.938.152	28.724.982
DK	53.502	140.433	252.283	374.767	113.971	182.327	267.299	372.256	1.756.838
ES	450.016	1.382.243	1.800.800	2.386.459	915.872	1.865.342	2.033.395	2.481.907	13.316.034
FR	679.410	1.848.651	2.468.372	3.907.267	1.480.420	2.644.646	2.788.596	4.059.177	19.876.539
GR	123.975	394.427	524.243	630.792	201.278	497.624	589.631	651.904	3.613.874
IL	54.965	119.368	181.236	329.188	85.969	165.138	212.248	358.337	1.506.449
IT	677.199	2.030.644	3.002.310	3.550.634	1.473.120	2.836.074	3.392.788	3.707.950	20.670.719
NL	129.584	405.236	696.262	1.099.655	302.602	552.990	726.556	1.077.018	4.989.903
SE	128.661	275.961	433.327	620.698	243.897	353.064	448.606	612.179	3.116.393

Table A. 4: Gender-age national calibration margins for the longitudinal weights of Waves 2-3

Country	Men				Women				Total
	[80+]	[70-79]	[60-69]	[50-59]	[80+]	[70-79]	[60-69]	[50-59]	
AT	78.107	226.890	397.094	489.228	206.108	322.003	445.591	505.984	2.671.005
BE	116.406	338.856	463.693	685.334	252.273	463.819	508.359	688.698	3.517.438
CH	88.010	213.013	353.754	483.304	182.343	289.159	382.106	485.164	2.476.853
CZ	70.769	249.032	451.127	748.978	171.670	395.698	544.350	788.428	3.420.052
DE	803.302	2.702.187	4.703.656	5.153.123	2.105.265	3.682.619	5.099.632	5.206.970	29.456.754
DK	55.576	143.721	279.421	365.280	114.835	181.704	292.424	364.479	1.797.440
ES	505.063	1.440.713	1.863.722	2.477.856	1.007.157	1.918.691	2.085.462	2.566.455	13.865.119
FR	761.382	1.889.624	2.517.711	4.103.902	1.639.326	2.665.648	2.815.456	4.312.972	20.706.021
GR	135.452	411.546	521.232	645.130	217.456	519.939	582.821	670.763	3.704.339
IE	28.952	87.074	155.122	233.919	57.557	107.540	157.438	230.261	1.057.863
IT	750.765	2.107.676	2.981.889	3.676.779	1.607.866	2.872.616	3.342.246	3.842.749	21.182.586
NL	140.437	423.069	737.358	1.131.781	318.473	556.700	761.275	1.114.293	5.183.386
PL	221.585	848.235	1.254.423	2.547.910	572.230	1.446.968	1.654.989	2.820.724	11.367.064
SE	133.206	275.151	481.137	602.413	247.821	344.217	491.642	597.014	3.172.601

Table A. 5: Gender-age national calibration margins for the longitudinal weights of Waves 3-4

Country	Men				Women				Total
	[80+]	[70-79]	[60-69]	[50-59]	[80+]	[70-79]	[60-69]	[50-59]	
AT	87.054	228.892	417.500	507.011	215.022	312.507	467.219	524.909	2.760.114
BE	127.833	344.792	492.344	698.124	269.174	460.812	533.375	702.388	3.628.842
CH	94.049	221.836	378.434	494.415	191.021	292.315	403.955	493.520	2.569.545
CZ	78.209	248.765	512.596	735.395	185.550	386.330	609.790	767.350	3.523.985
DE	910.825	2.962.247	4.492.556	5.449.544	2.189.281	3.867.307	4.853.094	5.497.431	30.222.285
DK	57.989	150.100	305.961	354.716	115.582	185.420	317.378	354.839	1.841.985
ES	564.720	1.478.066	1.936.317	2.627.444	1.099.310	1.949.933	2.145.377	2.707.324	14.508.491
FR	830.918	1.913.784	2.769.395	4.101.339	1.761.759	2.635.810	3.056.858	4.340.937	21.410.800
IT	824.693	2.166.838	3.113.481	3.671.956	1.724.199	2.891.832	3.454.976	3.842.342	21.690.317
NL	152.853	443.216	832.373	1.111.907	331.426	565.298	850.818	1.099.956	5.387.847
PL	245.965	868.894	1.341.288	2.666.830	633.823	1.464.285	1.740.379	2.944.979	11.906.443
SE	135.823	281.829	523.648	588.619	248.755	344.291	531.978	583.160	3.238.103

Table A. 6: Gender-age national calibration margins for the longitudinal weights of Waves 4-5

Country	Men				Women				Total
	[80+]	[70-79]	[60-69]	[50-59]	[80+]	[70-79]	[60-69]	[50-59]	
AT	96.914	249.191	421.121	527.233	221.286	329.084	470.809	544.287	2.859.925
BE	139.552	351.642	517.966	718.044	285.877	460.424	554.609	723.659	3.751.773
CH	100.662	231.542	401.088	514.460	200.294	297.027	424.708	510.127	2.679.908
CZ	85.886	252.156	561.574	709.751	197.939	383.909	660.649	735.498	3.587.362
DE	1.051.550	3.256.342	4.325.819	5.657.169	2.280.578	4.137.951	4.648.004	5.701.196	31.058.609
DK	60.758	158.989	321.242	353.012	115.960	192.016	333.271	353.270	1.888.518
EE	9.505	34.197	54.420	80.370	33.343	70.729	80.449	96.850	459.863
ES	631.602	1.429.864	2.112.277	2.754.728	1.205.485	1.873.737	2.327.029	2.825.705	15.160.427
FR	905.231	1.935.863	3.032.734	4.078.273	1.877.588	2.605.520	3.326.360	4.336.918	22.098.487
IT	898.645	2.281.518	3.177.529	3.703.808	1.834.771	2.971.727	3.506.774	3.894.492	22.269.264
NL	167.169	470.129	896.450	1.124.236	346.092	579.900	910.727	1.116.374	5.611.077
SE	138.814	295.186	554.039	579.502	249.532	350.797	563.114	572.445	3.303.429
SI	17.023	59.346	97.114	153.366	46.827	89.758	108.170	148.036	719.640

Table A. 7: Gender-age national calibration margins for the longitudinal weights of Waves 5-6

Country	Men				Women				Total
	[80+]	[70-79]	[60-69]	[50-59]	[80+]	[70-79]	[60-69]	[50-59]	
AT	105.944	281.011	407.189	562.878	224.868	361.047	454.248	577.325	2.974.510
BE	152.395	342.369	553.322	744.107	302.441	441.052	590.664	749.387	3.875.737
CH	105.285	241.883	415.318	540.334	201.957	304.569	438.064	534.146	2.781.556
CZ	93.631	265.363	597.183	679.950	209.570	392.492	699.225	697.712	3.635.126
DE	1.076.448	3.497.831	4.139.802	5.796.901	2.286.922	4.392.882	4.471.983	5.873.783	31.536.552
DK	63.388	170.713	330.591	356.917	116.410	202.288	343.250	357.008	1.940.565
EE	11.015	35.307	55.236	81.990	36.558	72.042	80.410	96.932	469.490
ES	701.269	1.445.496	2.181.034	2.920.932	1.307.876	1.862.242	2.394.723	2.990.197	15.803.769
FR	978.949	1.910.312	3.308.097	4.092.009	1.977.593	2.510.517	3.643.706	4.355.354	22.776.537
IL	72.814	142.272	283.534	364.941	114.788	182.074	318.469	394.706	1.873.598
IT	985.286	2.327.434	3.188.920	3.803.444	1.954.255	2.981.835	3.520.032	4.026.454	22.787.660
LU	5.620	13.407	22.893	35.007	10.962	17.092	23.047	33.703	161.731
SE	141.561	311.691	571.702	579.173	248.986	359.378	581.837	571.564	3.365.892
SI	19.880	62.093	104.440	153.323	51.088	89.604	114.242	149.372	744.042

Table A. 8: Gender-age national calibration margins for the longitudinal weights of Waves 6-7

Country	Men				Women				Total
	[80+]	[70-79]	[60-69]	[50-59]	[80+]	[70-79]	[60-69]	[50-59]	
AT	112.458	297.300	412.100	605.077	225.939	377.364	458.589	615.818	3.104.645
BE	164.453	342.412	584.415	764.663	314.844	431.797	621.223	768.289	3.992.096
CH	112.728	260.869	423.985	575.538	209.035	319.400	447.298	565.599	2.914.452
CZ	99.716	287.185	621.907	662.766	216.951	416.181	722.580	673.538	3.700.824
DE	1.147.882	3.603.195	4.228.830	6.124.073	2.293.027	4.489.624	4.583.858	6.152.732	32.623.221
DK	66.614	189.604	332.406	365.329	118.042	220.551	345.391	365.122	2.003.059
EE	12.060	34.950	58.167	82.679	38.872	70.235	83.747	95.550	476.260
ES	767.408	1.455.872	2.259.273	3.044.029	1.396.361	1.848.692	2.475.109	3.117.852	16.364.596
FR	1.044.706	1.945.160	3.521.652	4.147.767	2.063.824	2.488.635	3.906.581	4.401.486	23.519.811
GR	210.246	411.661	567.816	680.282	311.555	522.618	637.565	747.073	4.088.816
HR	43.028	137.662	221.196	300.715	99.734	211.944	262.169	315.580	1.592.028

Table A. 8 (continued)

Country	Men				Women				Total
	[80+]	[70-79]	[60-69]	[50-59]	[80+]	[70-79]	[60-69]	[50-59]	
IL	75.821	148.304	311.842	372.150	118.912	186.706	352.143	399.516	1.965.394
IT	1.069.222	2.401.338	3.316.424	4.077.619	2.055.350	3.019.949	3.638.350	4.300.069	23.878.321
LU	6.133	14.036	24.691	37.925	11.544	17.424	24.654	36.153	172.560
PL	342.081	843.684	1.900.851	2.606.181	827.699	1.381.564	2.342.874	2.791.164	13.036.098
PT	156.796	365.224	547.426	675.613	302.902	503.547	648.524	747.807	3.947.839
SE	144.956	345.101	571.159	589.022	248.034	387.659	582.464	579.542	3.447.937
SI	22.330	65.221	112.695	152.339	54.674	90.770	119.774	149.022	766.825

Table A. 9: Gender-age national calibration margins for the longitudinal weights of Waves 1-7

Country	Men				Women				Total
	[80+]	[70-79]	[60-69]	[50-59]	[80+]	[70-79]	[60-69]	[50-59]	
AT	10.200	99.831	300.472	417.572	43.419	181.211	379.686	457.343	1.889.734
BE	14.033	144.806	347.184	586.606	54.967	256.152	433.659	620.133	2.457.540
CH	15.647	97.705	271.048	439.673	47.610	165.884	323.979	456.572	1.818.118
DE	147.760	1.048.393	3.739.515	4.366.750	451.882	1.862.721	4.575.565	4.654.324	20.846.910
DK	6.215	56.552	190.372	333.570	25.356	89.800	221.502	345.703	1.269.070
ES	69.717	633.630	1.403.850	2.149.229	223.611	1.096.152	1.818.291	2.381.800	9.776.280
FR	145.880	901.487	1.960.320	3.487.244	467.076	1.630.406	2.492.902	3.863.405	14.948.720
IT	89.793	907.328	2.376.667	3.249.391	352.400	1.606.079	2.994.729	3.538.087	15.114.474
SE	18.719	121.091	345.518	570.697	57.127	188.024	387.408	579.355	2.267.939

Table A. 10: Calibrated cross-sectional weights of Wave t=1,…,7 (sharewt_rel7-0-0_gv_weights)

Variable	Description	Unit of analysis
dw_wt	Design weight - Wave t	Household & individual
cchw_wt	Calibrated cross-sectional household weight - Wave t	Household
cciw_wt	Calibrated cross-sectional individual weight - Wave t	Individual
cchw_wt_SHL	Calibrated cross-sectional household weight, SHARELIFE only - Wave t	Household
cciw_wt_SHL	Calibrated cross-sectional individual weight, SHARELIFE only - Wave t	Individual
cchw_wt_REG	Calibrated cross-sectional household weight, regular interview only - Wave t	Household
cciw_wt_REG	Calibrated cross-sectional individual weight, regular interview only - Wave t	Individual
subsample	Subsamples within country	Household & individual

Table A. 10 (continued)

Variable	Description	Unit of analysis
stratum1	First stratum	Household & individual
stratum2	Second stratum	Household & individual
psu	Primary sampling unit	Household & individual
ssu	Secondary sampling unit	Household & individual

Table A. 11: Calibrated longitudinal weights of wave combination t-…-s (sharewX_rel7-0-0_gv_longitudinal_weights_wt_ws)

t - … - s	Variable	Description	Unit of analysis
1- 2	dw_w1	Design weight - Wave 1	Household & individual
	clhw_wc_12	Calibrated longitudinal household weight	Household
	cliw_wc_12	Calibrated longitudinal individual weight	Individual
2 - 3	dw_w2	Design weight - Wave 2	Household & individual
	clhw_wc_23	Calibrated longitudinal household weight	Household
	cliw_wc_23	Calibrated longitudinal individual weight	Individual
3 - 4	dw_w3	Design weight - Wave 3	Household & individual
	clhw_wc_34	Calibrated longitudinal household weight	Household
	cliw_wc_34	Calibrated longitudinal individual weight	Individual
4 - 5	dw_w4	Design weight - Wave 4	Household & individual
	clhw_wc_45	Calibrated longitudinal household weight	Household
	cliw_wc_45	Calibrated longitudinal individual weight	Individual
5 - 6	dw_w5	Design weight - Wave 5	Household & individual
	clhw_wc_56	Calibrated longitudinal household weight	Household
	cliw_wc_56	Calibrated longitudinal individual weight	Individual
6 - 7	dw_w6	Design weight - Wave 6	Household & individual
	clhw_wc_67	Calibrated longitudinal household weight	Household
	cliw_wc_67	Calibrated longitudinal individual weight	Individual
1 - … - 7	dw_w1	Design weight - Wave 1	Household & individual
	clhw_wc_1_7	Calibrated longitudinal household weight	Household
	cliw_wc_1_7	Calibrated longitudinal individual weight	Individual

Table A. 11 (continued)

t - … - s	Variable	Description	Unit of analysis
ALL	subsample	Subsamples within country	Household & individual
	stratum1	First stratum	Household & individual
	stratum2	Second stratum	Household & individual
	psu	Primary sampling unit	Household & individual
	ssu	Secondary sampling unit	Household & individual
	panel_resp	Respondent participation in the selected panel	Individual

Table A. 12: List of variables included in the imputation dataset of Wave 7

Variable name	Description	Questionnaire	Interview type: R-Regular S-SHARELIFE
mergeid	Person ID		R-S
implicat	Implicat number		R-S
hhidcom7	Household ID Wave 5		R-S
cvid	Wave specific person identifier		R-S
cvidp	Wave specific person identifier of spouse/partner		R-S
country	Country identifier		R-S
language	Language of questionnaire		R-S
htype	Household type		R-S
fam_resp	Family respondent		R-S
fin_resp	Financial respondent		R-S
hou_resp	Household respondent		R-S
excrate	Exchange rate		R-S
nursinghome	Living in nursing home	MN024	R-S
hhsize	Household size		R-S
single	Single		R-S
couple	Couple		R-S
partner	Partner in the couple		R-S
p_nrp	Partner of nonresponding partner		R-S
sample1	Imputation sample for singles		S
sample_1_2	Imputation sample for singles and couples with two partners interviewed		R
sample3	Imputation sample for all couples		R-S
ydip	Earnings from employment	EP205	R
yind	Earnings from self-employment	EP207	R

Table A. 12 (continued)

Variable name	Description	Questionnaire	Interview type: R-Regular S-SHARELIFE
ypen1	Annual old age, early retirement pensions, survivor and war pension	EP078_1-2-3-7-8-9	R
ypen2	Annual private occupational pensions	EP078_11-16	R
ypen3	Annual disability pension and benefits	EP078_4-5	R
ypen4	Annual unemployment benefits and insurance	EP078_6	R
ypen5	Annual payment from social assistance	EP078_10	R
yreg1	Other regular payments from private pensions	EP094_1-2-5	R
yreg2	Other regular payment from private transfer	EP094_3-4	R
ylsum1	Lump sum payments for old age, early retirement, survivor and war pension	EP082_1-2-3-7-8-9	R
ylsum2	Lump sum payments for private occupational pension	EP082_11-16	R
ylsum3	Lump sum payments for disability pension and benefits	EP082_4-5	R
ylsum4	Lump sum payments for unemployment benefits and insurance	EP082_6	R
yslum5	Lump sum payments for social assistance	EP082_10	R
yslum6	Lump sum payments for other private pension	EP209_1-2-5	R
yslum7	Lump sum payments for other private transfer	EP209_3-4	R
rhre	Annual rent and home-related expenditures	HO005, HO008	R
home	Value of main residence	HO024	R
mort	Mortgage on main residence	HO015	R
ores	Value of other real estate – Amount	HO027	R
ysrent	Annual income from rent or sublet	HO074, HO030	R
yaohm	Annual income from other household members	HO002, HO011	R
fahc	Annual food at home consumption	CO002	R-S
fohc	Annual food outside home consumption	CO003	R-S
hprc	Annual home produced consumption	CO011	R-S
bacc	Bank accounts	AS003	R
bsmf	Bond, stock and mutual funds	AS007, AS011, AS017	R
slti	Savings for long-term investments	AS021, AS023, AS27, AS030	R
vbus	Value of own business	AS042	R
sbus	Share of own business	AS044	R
car	Value of cars	AS051	R
liab	Financial liabilities	AS055	R
yibacc	Interest income from bank accounts		R

Table A. 12 (continued)

Variable name	Description	Questionnaire	Interview type: R-Regular S-SHARELIFE
yibsmf	Interest income from bond, stock and mutual funds		R
thinc	Total household net income - version A		R
thinc2	Total household net income - version B	HH017	R-S
thexp	Total household expenditure (rhre+fahc+fohc+hprc)		R
hrass	Household real assets (home*per-ho/100+vbus*sbus/100+car+ores-mor)		R
hgfass	Household gross financial assets (back+bsmf+slti)		R
hnfass	Household net financial assets (hgfass-liab)		R
hnetw	Household net worth		R
gender	Gender	DN042	R-S
age	Age in 2010	DN003	R-S
age_p	Age of partner in 2010	DN003	R-S
yeduc	Year of education	DN041	R-S
yeduc_p	Year of education of partner	EX102	R
sphus	Self-perceived health - US scale	PH003	R-S
mstat	Marital status	DN014	R
nchild	Number of children	CH001	R-S
ngcchild	Number of grandchildren	CH201	R-S
gali	Limitation with activities	PH005	R-S
chronic	Number of chronic diseases	PH006	R-S
bmi	Body mass index	PH012, PH013	R-S
weight	Weight	PH012	R-S
height	Height	PH013	R-S
mobility	Mobility limitations	PH048	R-S
adl	Limitations with activities of daily living	PH049_1	R-S
iadl	Limitations with instrumental activities of daily living	PH049_2	R-S
esmoked	Ever smoked daily	BR001	R
phactiv	Physical inactivity	BR015	R
orienti	Score of orientation in time test	CF003 - CF006	R
memory	Score of memory test	CF103	R
wllft	Score of words list learning test - trial 1	CF104_* - CF107_*	R-S
wllst	Score of words list learning test - trial 2	CF113_* - CF116_*	R-S
fluency	Score of verbal fluency test	CF010	R

Table A. 12 (continued)

Variable name	Description	Questionnaire	Interview type: R-Regular S-SHARELIFE
numeracy1	Score of first numeracy test	CF012 - CF015	R
numeracy2	Score of second numeracy test	CF108 - CF112	R
eurod	EURO depression scale	MH002 - MH017	R
doctor	Seen/talked to medical doctor	HC002	R-S
hospital	In hospital last 12 months	HC012	R-S
thospital	Times being patient in hospital	HC013	R-S
nhospital	Total nights stayed in hospital	HC014	R-S
cjs	Current job situation	EP005	R
pwork	Did any paid work	EP002	R
empstat	Employee or self-employed	EP009	R
lookjob	Looking for job	EP337	R
rhfo	Received help from others (how many)	SP002, SP005, SP007	R
ghto	Given help to others (how many)	SP008, SP011, SP013	R
ghih	Given help in the household (how many)	SP018	R
rhih	Received help in the household (how many)	SP020	R
gfg	Number of given financial gifts 250 or more	FT002, FT007_*	R
rfg	Number of received financial gifts 250 or more	FT009, FT014_*	R
otrf	Owner, tenant or rent free	HO002	R
perho	Percentage of house owned	HO070	R
fdistress	Household able to make ends meet	CO007	R-S
lifesat	Life satisfaction	AC012	R-S
lifehap	Life happiness	AC022	R-S
naly	Number of activities last year	AC035_*	R-S
saly	Satisfied with no activities	AC038	R-S
big5_1	Big Five – Reserved	AC701	R-S
big5_2	Big Five – Trust	AC702	R-S
big5_3	Big Five – Lazy	AC703	R-S
big5_4	Big Five – Relaxed	AC704	R-S
big5_5	Big Five – Few Interests	AC705	R-S
big5_6	Big Five – Outgoing	AC706	R-S
big5_7	Big Five – Find Fault	AC707	R-S
big5_8	Big Five – Thorough Job	AC708	R-S
big5_9	Big Five – Nervous	AC709	R-S

Table A. 12 (continued)

Variable name	Description	Questionnaire	Interview type: R-Regular S-SHARELIFE
big5_10	Big Five – Imagination	AC710	R-S
big5_11	Big Five – Kind	AC711	R-S
willans	Willingness to answer	IV004	R-S
clarify	Respondent asked for clarifications	IV007	R-S
undersq	Respondent understood questions	IV008	R-S
hnrsc	Help needed to reed showcards	IV018	R-S
nomxyear	Nominal exchange rate		R-S
pppxyear	PPP adjusted exchange rates		R-S
Currency	Currency in which amounts are denominated		R-S

Table A. 13: Description of flag variables associated with imputations

Value	Label	Description
-99	Missing by design	Missing values depends from skip patterns in the questionnaire
1	Not designed resp	Missing values depends on the type of respondents designed to respond
2	No ownership	No declared ownership
3	Regular obs.	Regular observation
4	Imp: ub point	Imputation based on specific declared amounts in the unfolding brackets routing
5	Imp: ub range	Imputation is based on unfolding brackets range information
6	Imp: ub incomplete	Imputation is based on unfolding brackets partial information
7	Imp: ub uniformative	Unfolding brackets uninformative
8	Imp: ownership	Ownership has been imputed
9	Imp: amount	Imputed amount
10	Imp: outlier LB	Imputed value if lower than LB
11	Imp: outlier UB	Imputed value if lower than UB
12	Imp: aggregate	Imputation of the corresponding aggregate variable, see table 2
13	Imp: NRP	(only for thinc)
14	Imp: missing value	(only for explanatory variables imputed ex-ante by hot-deck)

CHAPTER 10

The SHARE Data & Documentation Tool

10

10 THE SHARE DATA & DOCUMENTATION TOOL

Edwin de Vet – CentERdata, Maurice Martens – CentERdata; Tilburg University, Stefan Gruber – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC)

10.1 Background

SHARE is a very rich but also complex data set. There are several reasons why its data structure is more complex than that of other surveys. First, the large number of countries participating in SHARE, and the resulting large number of questionnaire languages, make not only the harmonisation and operational tasks but also the handling and documentation of the data and metadata more complicated than that in national surveys administered in a single country. Second, SHARE covers a wide range of topics, from basic demographics on health and health behaviour, health care, job history, and socio-economic status to social and family networks. To structure the various topics, the SHARE interview consists of thematic blocks or modules. Overall, the data sets released thus far contain more than 40 different interview modules. To address contemporary matters and due to time constraints, not every module is part of every data collection wave. Additionally, the questionnaire is designed in such a way that not every respondent answers every question to avoid collecting repeated information and to keep the time burden of the interview as low as possible. Some modules (or parts of a module) are restricted to certain subgroups of respondents. Selected household members serve as family, financial or household respondents. They answer specific questions, e.g., on children, financial aspects and household features, on behalf of the couple or the whole household. A third aspect that adds to its complexity is the panel structure of SHARE. Basically, two types of questionnaires can be distinguished: the baseline questionnaire for respondents who participate in a SHARE interview for the first time and the longitudinal questionnaire for respondents who participated in at least one previous wave of data collection. In Wave 7, however, this situation was even more complicated (see Chapters 2.1 and 2.2). Some information is only collected once in the baseline interview (e.g., educational degree), whereas other information is collected differently depending on baseline or longitudinal questionnaire (e.g., smoking habits). This approach leads to a relatively complex routing scheme.

For such elaborate survey data to be successfully used in secondary analysis, a comprehensive and accessible documentation of the data and metadata is necessary (Vardigan et al., 2016). To facilitate the use of SHARE data, a set of different documentation files is provided to the researchers. The SHARE data resource profile, published in the International Journal of Epidemiology, provides a compact overview of the structure and content of SHARE (Börsch-Supan et al., 2013). Additionally, wave- and country-specific questionnaires are provided to the research community. The SHARE release guide is specifically directed to researchers working with the data. It contains general information on the database, such as the naming of variables, the missing data code scheme, the merging of modules and/or waves, and wave-specific information, such as important questionnaire innovations and new procedures introduced between waves. The release guide also describes the content and structure of the generated variable modules provided to the users to facilitate their ability to work with the data. Because of its divergence from the regular panel waves, the life history data of Wave 3 (SHARELIFE) is documented in a separate release guide. Methodological aspects and advancements are described in wave-specific methodology volumes. The cross-wave comparison document is another important source of information. It contains an overview of the generic version of the questionnaire to easily track deviations between waves. The purpose of the scales and multi-item indicators manual is to provide an informative overview of all scales and item groups that are included in the questionnaire. The manual covers literature based on the definition and content of the respective scale and its operationalisation in SHARE. Finally, the FAQ section on the SHARE website contains useful information for researchers.

The new SHARE Data & Documentation Tool (available at https://www.share-datadocutool.org) adds to the existing comprehensive documentation material. It is a web application developed by CentERdata, Institute for Data Collection and Research, in cooperation with SHARE Central at the Munich Center for the Economics of Aging (MEA). For researchers, this management system is intended to be a fast, customisable, easy-to-use web interface for browsing and searching the SHARE (meta)data. Additionally, researchers who plan to use SHARE data can obtain a first impression regarding content and case numbers. The following section provides a description of the technical implementation before the key features of the new tool are introduced.

10.2 Technical implementation

To give data users adequate insight into the SHARE data, it was deemed necessary to develop an online system in

which one could find the proper background (metadata and paradata) that shows which questions were fielded in what waves and how a specific question was asked in the different languages. We needed a tool that could help the data users find the questions and variables they needed: the SHARE Data & Documentation Tool.

Instead of developing this tool from scratch, we decided to review whether the existing Questasy code base, also developed by CentERdata, could be further developed. Questasy is a web application for managing the dissemination of data and metadata for survey projects. It was originally developed for the LISS Data Archive but was designed to be repurposed for other surveys as well. Questasy manages both metadata and survey data and provides an easy-to-use data entry module for administrators to create metadata. The external web interface allows researchers to browse and search both the survey data and the metadata. The Questasy system also manages files, tracks downloads and creates web pages for viewing documentation. Due to the longitudinal nature of SHARE, the ability to track questions and variables across waves was a key requirement of the system. To support this function, DDI 3 was chosen as the basis for the structure of the application. The Data Documentation Initiative (DDI) is an international standard for describing the data produced by surveys and other observational methods in the social, behavioural, economic, and health sciences. DDI is a free standard that can document and manage different stages in the research data lifecycle, such as conceptualisation, collection, processing, distribution, discovery, and archiving. Documenting data with DDI facilitates understanding, interpretation, and use – by people, software systems, and computer networks. "The choice for DDI 3 was initially not an obvious one. The main reason that DDI 3 piqued our interest was its support for longitudinal studies" (De Bruijne & Amin, 2010: 11).

The Questasy tool fits the requirements we drew up to a certain degree. Questasy divides and structures questions and question elements in a manner similar to that of the translation management tool (TMT; see https://seriss.centerdata.nl/#tmtvideo), which would allow for an easy import. Questasy did support multilingualism, but only up to two languages. Therefore, the integration of further languages was a central task of adapting Questasy to SHARE needs. Additionally, it was necessary to focus on processes in bulk due to the size of the existing datasets as well as the need to constantly integrate new data releases. A new data release should be uploaded, and changes should be adapted automatically.

Given this list of adaptations, we developed a new tool: the SHARE Data & Documentation Tool. In this tool, the longitudinal concept was further developed, and questions were linked over multiple waves. Linking questions to their variables was not always straightforward but was nonetheless performed.

10.3 Main features

The SHARE Data & Documentation Tool combines the questionnaire level with the data level. This creates the possibility of generating wave- and module-specific codebooks, one of the main features of the new tool. It also contains all the relevant documentation material as well as a list of SHARE-based publications, thereby enabling users to easily search in all relevant fields. As shown in Figure 10.1, on the main level of the website, one of the following options can be chosen: Getting Started, Browse SHARE Waves, Search, and Browse Publications.

Figure 10.1: Basic features of the SHARE Data & Documentation Tool

The Getting Started section serves as an introduction to the tool. It explains the purpose of the tool and contains a manual style description of its basic features. By means of screenshots, the user is introduced step-by-step into the potentials and features of the new tool. In the Browse SHARE Waves section, users can choose one of the currently released waves via three options: (1) Questionnaire Map, (2) Variables and Datasets, and (3) Documentation. The tab Questionnaire Map represents the questionnaire level. After providing some general information on the wave of interest (e.g., participating countries and fieldwork period), users can choose one of the questionnaire modules to receive a list of all questions that are included in the module. Figure 10.2 shows the mental health (MH) module of the Wave 6 generic questionnaire as an example.

Figure 10.2: Questionnaire map for the MH module of the Wave 6 generic questionnaire

Choosing a specific item/question gives an overview of ques-

tion text, answer options, active filters and related questions, the latter indicating in which other waves the question was asked. The column "Cross-Country Deviation" on the right side has not yet been completed for Wave 5 onwards (work in progress). For the waves prior to Wave 5, this column contains a "Y" for known deviations between countries in either the question text or the answer options. Those deviations may be intended (e.g., in the questions about educational degrees, where the difference in national educational systems must be accounted for). Nevertheless, in some cases, those deviations are unintended, e.g., due to translation errors.

At the top of each questionnaire module page, users can generate a module-specific codebook in PDF format. The generated codebook contains country-specific frequencies when choosing a specific language, whereas the generic codebook contains frequencies including all participating countries of the respective wave. The module routing is stored at the very end of each codebook. Apart from the codebook, the routing can also be displayed in HTML format using the Show Routing tab. One main advantage of this feature is that filter questions are linked, which makes it easier to retrace the routing steps. The Hover-over function of the linked questions immediately shows the question text and answer options of the filter questions, making the routing more comprehensible for users.

The Variables and Datasets tab represents the data level. By choosing one module, a list of all module-specific variables, as well as their respective variable labels, appears. Response options as well as descriptive statistics both for the whole SHARE sample and for each country separately can be displayed via the selection of a specific variable. Therefore, researchers interested in using SHARE data can obtain a first impression regarding case numbers without needing to download the data set first. Finally, the Documentation tab contains links to all existing documentation files.

The Search function is intended to help users quickly find the information they are seeking. The Advanced Search function provides the option either to search in all fields or to focus the search on data, questionnaires or publications. This tool can also be helpful for scientists who are not yet SHARE users but who want to determine whether SHARE contains the information they need for their research project.

Users are requested to provide references to all papers based on SHARE data to the SHARE co-ordination team. The Browse Publications tab on the main level contains all publications that have been reported to SHARE Central, including journal articles, books, book chapters, and other types of publications, such as discussion papers or theses.

References

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., et al. (2013). Data resource profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). International Journal of Epidemiology, 42(4), 992-1001. https://doi.org/10.1093/ije/dyt088.

De Bruijne, M., & Amin, A. (2010): Questasy: Online survey data dissemination using DDI 3. IASSIST Quarterly, 33(1), 10-15. https://doi.org/10.29173/iq645.

Vardigan, M., Granda, P., & Hoelter L. (2016). Documenting survey data across the life cycle. In C. Wolf, D. Joye, T. W. Smith & Y. Fu (Eds.), The SAGE Handbook of Survey Methodology (pp. 443–459). Thousand Oaks: SAGE Publications.

ISBN 978-3-00-062956-3