SHARE WAVE 8 METHODOLOGY:

Collecting Cross-National Survey Data in Times of COVID-19

SHARE
SURVEY OF HEALTH, AGEING AND RETIREMENT IN EUROPE

SHARE
SURVEY OF HEALTH, AGEING AND RETIREMENT IN EUROPE

SHARE WAVE 8 METHODOLOGY:

Collecting Cross-National Survey Data in Times of COVID-19

SHARE Wave 8 Methodology:
Collecting Cross-National Survey Data in Times of COVID-19

Edited by:
Michael Bergmann
Axel Börsch-Supan

Authors:
Nora Angleys
Louis Arnault
Michael Bergmann
Arne Bethmann
Axel Börsch-Supan
Marika de Bruijne
Paolo Li Donni
Salima Douhou
Irene Ferrari
Fabio Franzese
Tessa-Virginia Hannemann
Imke Herold
Florence Jusot
Giuseppe De Luca
Howard Litwin
Diana López-Falcón
Maurice Martens
Sebastiaan Pennings
Yuri Pettinicchi
Jeny Tony Philip
Moslem Rashidi
Thomas Renaud
Gregor Sand
Annette Scherpenzeel
Daniel Schmidutz
Elena Sommer
Iggy van der Wielen
Melanie Wagner
Luzia Weiss
Yasemin Yilmaz

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: +49-89-38602-390
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:
denisismagilov – stock.adobe.com (coronavirus pandemia concept, planet hologram)

Suggested citation:
Bergmann, M. & Börsch-Supan, A. (2021). SHARE Wave 8 Methodology: Collecting cross-national survey data in times of COVID-19. Munich: MEA, Max Planck Institute for Social Law and Social Policy.

ISBN 978-3-00-069877-4

Sponsored by:

MAX-PLANCK-GESELLSCHAFT

Horizon 2020

This project has received funding from the European Union under grant agreements VS/2019/0332, VS/2020/0313 and the European Union's Horizon 2020 research and innovation programme under grant agreements No 870628, No 101015924.

Federal Ministry of Education and Research

NIH National Institute on Aging

TABLE OF CONTENTS

II SHARE CORONA SURVEY (CATI)

III ENHANCE SHARE DATA COLLECTION

CHAPTER 1

SHARE Wave 8 Methodology: Collecting Cross-National Survey Data in Times of COVID-19

01

1 SHARE WAVE 8 METHODOLOGY: COLLECTING CROSS-NATIONAL SURVEY DATA IN TIMES OF COVID-19

Michael Bergmann and 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 eighth wave of the Survey of Health, Ageing and Retirement in Europe (SHARE). SHARE is a research infrastructure aimed at better understanding and coping with the challenges and opportunities of population ageing (Börsch-Supan et al., 2013). The main objective of SHARE is to provide excellent data to study the effects of health, social, economic and environmental policies over the life-course of European citizens and beyond through a combination of: (a) transdisciplinarity, studying the interactions between biomedical and socio-economic factors; (b) longitudinality, combining a prospective panel structure and retrospective life histories; and (c) European coverage with strict cross-national comparability by the use of *ex ante* harmonised 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. Unless mentioned otherwise, the following chapters are based on preliminary SHARE wave 8 release 0 data (Börsch-Supan, 2020) that were available at the time of writing.

Compared to previous waves the eighth wave of SHARE was unique in many ways. While having a smooth start of fieldwork in October 2019, the COVID-19 pandemic started to hit Europe in early 2020 and even at the time of writing one and a half year later virtually affects all aspects of life – including survey research. Similar to other surveys SHARE had to suspend its regular face-to-face interviewing in all 28 participating countries in March 2020 (see Scherpenzeel et al., 2020 for an overview). The implementation of strict epidemiological control measures in nearly all participating countries made it infeasible to continue face-to-face fieldwork. Stopping the survey was particularly urgent considering SHARE's target population of people aged 50+, including very old respondents as well as retirement and nursing homes residents who face the highest risks from a possible infection. Against this background, SHARE switched the interview mode to telephone interviewing, using a special "SHARE Corona" questionnaire.

1.1 The impact of COVID-19 on data collection in SHARE

Around February 2020, COVID-19 was spreading quickly across Europe, leading to a suspension of SHARE fieldwork in all participating countries between March 10 and March 23. All stakeholders involved shared the opinion that SHARE data about the health and living situation of the 50+ population in Europe were now needed more than ever to shed light on the short- and long-term implications of this global pandemic. This led to the development of the SHARE Corona Survey that was developed and successfully conducted in all 28 countries between April and August 2020[1]. Together with the full wealth of background information on people aged 50 years or older from previous SHARE waves, this data collection offers huge potential for substantive analyses and cross-national comparisons regarding health, social and economic developments as well as outcomes.

After fieldwork was suspended, it soon became clear that a quick return to the normal face-to-face Computer Assisted Personal Interview (CAPI) was unlikely. Taking into account the variation in internet use across countries and especially among older age groups in SHARE, it was decided to resume interviewing with a Computer Assisted Telephone Interview (CATI), targeted to the COVID-19 living situation of people aged 50 or older. This decision additionally considers evidence from other studies, which have shown that mode effects on response behavior and measurement error tend to be larger between interviewer- and self-administered modes than between modes that are both interviewer-administered such as face-to-face and telephone (e.g. Couper, 2011).

This switch of the interview mode brought about various changes, a fact that is also reflected in the structure of this volume. It has three parts: First, we describe the methodology of the face-to-face part of Wave 8 that was suspended in March 2020. Second, we detail the methodology of the first SHARE Corona Telephone Survey in early summer 2020. Finally, we present three overarching methodological advances that took place during both surveys.

1 Austria could only start fieldwork later and finished fieldwork in September.

In this respect, the first part on the regular SHARE Wave 8 describes the usual preparations in the run-up to a new wave. Michael Bergmann, Arne Bethmann and Giuseppe De Luca outline the sampling procedure to draw inferences about the population of people aged 50 years or older across countries (Chapter 2). Melanie Wagner and Howie Litwin give an overview of changes in the regular Wave 8 questionnaire (Section 3.1). This includes new measures of healthcare utilization and access to care (Section 3.2 by Thomas Renaud, Louis Arnault and Florence Jusot) as well as new cognitive measures to identify interactions of biomedical and socio-economic conditions over the life course (Section 3.3 by Melanie Wagner and Salima Douhou). In addition, several new add-on modules have been implemented in Wave 8, including a module on saving regret to shed light on long-term economic decisions and its systematic failures (Section 3.4 by Axel Börsch-Supan and Irene Ferrari), a time expenditure module to measure how much time people aged 50+ spent on different activities such as cooking, shopping, watching TV, volunteering for charitable work, helping other people, caring for grandchildren, etc. (Section 3.5 by Annette Scherpenzeel and Jeny Tony Philip) and a module to ask eligible respondents for their consent to participate in the accelerometer study during the Wave 8 interview (Section 3.6 by Annette Scherpenzeel, Nora Angleys, Fabio Franzese and Luzia Weiss; see also Chapter 12). Additionally, Maurice Martens and Iggy van der Wielen discuss the technical design and programming of these instruments that already involved various updates to the existing systems and tools (Chapter 4) and in the following played a key role regarding the successful switch to another interview mode. In Chapter 5, first Gregor Sand gives a comprehensive overview of fieldwork monitoring as well as survey participation (Section 5.1) and then Michael Bergmann and Tessa-Virginia Hannemann describe the back-checking procedure to guarantee data quality in the regular SHARE Wave 8 (Section 5.2). Finally, Giuseppe De Luca, Paolo Li Donni and Moslem Rashidi provide a description of the weighting and imputation strategies used for dealing with problems of unit non-response, sample attrition and item non-response (Chapter 6).

The second part of this volume covers adaptions and innovations that have to be implemented for the SHARE Corona Survey. Michael Bergmann and Arne Bethmann describe the sampling strategy for the new CATI instrument. Hereby, a sample was selected in each country, which includes all eligible panel members, independent from having been interviewed before the suspension of fieldwork or not (Chapter 7). Yasemin Yilmaz, Melanie Wagner and Axel Börsch-Supan then present the motivation and the content of the SHARE Corona Survey (Chapter 8). It focusses on direct and indirect effects of the pandemic and its accompanying epidemiological control measures during the first phase of the pandemic, covering the most important life domains for SHARE's target population: health (physical and mental) and health behaviour; COVID-19-related infections for respondents and their social network; quality of healthcare; work and economic situation; and social relationships. Besides the content, the second key challenge of switching the interview mode from face-to-face to telephone was the development of the technical framework that allows conducting telephone interviews. Marika de Bruijne and Sebastiaan Pennings describe the underlying considerations as well as the necessary practical adaptions that resulted in a web-based questionnaire to support the telephone interviews of the SHARE Corona Survey (Chapter 9). Finally, Chapters 10 and 11 describe the adaptions regarding fieldwork monitoring (Section 10.1) and data quality back-checks (Section 10.2) as well as weights and imputations due to the changed circumstances during the pandemic.

The third part of this volume includes additional data that enhance the collected responses from our respondents in SHARE. In Chapter 12, Annette Scherpenzeel, Nora Angleys, Fabio Franzese and Luzia Weiss describe the accelerometer project that was conducted in the regular SHARE Wave 8 to measure the level of activity and sedentary behaviour of the elderly across countries. Imke Herold, Yuri Pettinicchi and Daniel Schmidutz describe the various legal, ethical, technical and organisational challenges SHARE faces when harmonising record linkage proceedings to enhance self-reported survey data with administrative data, such as data from public authorities, insurances or governmental institutions (Chapter 13). Finally, Diana López-Falcón introduces the Social Policy Archive in SHARE (SPLASH) that aims to overcome existing data limitations in order to foster comparative policy-oriented research (Chapter 14). In this respect, SPLASH supports longitudinal multilevel research by providing time-series data at the national (but also the regional) level in the fields of education, family, health (including government policies implemented throughout Europe during the COVID-19 pandemic) and migration, as well as work and retirement policies. By this, it substantially enriches the analytical potential of the SHARE microdata.

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,

Michael Radhuber and Theresa Weitzhofer-Yurtisik (Austria); Sergio Perelman, Xavier Flawinne, Stephanie Linchet, Jean-François Reynaerts, Laurent Nisen, Marine Maréchal, Benjamin Thiry (Belgium-FR) and Koen Decancq, Daniela Skugor, Martine Vandervelden, Jorik Vergauwen (Belgium-NL); Ekaterina Markova, Gabriela Yordanova, Vassil Kirov, Svetla Koleva, Alexey Pamporov and Diana Nenkova (Bulgaria); Šime Smolić and Ivan Čipin (Croatia); Nikolaos Theodoropoulos and Neophyta Empora (Cyprus); Radim Bohacek, Miroslava Federicova, Jan Kroupa, René Levinsky, Filip Pertold, Jiri Satava, Jaroslava Pospisilova and Tereza Rosikova (Czech Republic); Karen Andersen-Ranberg, Jørgen Lauridsen, Lasse Lybecker Scheel-Hincke and Astrid Roll Vitved (Denmark); Luule Sakkeus, Liili Abuladze, Jana Bruns, Kaia Laidra, Tiina Tambaum, (Estonia); Anna Rotkirch, Miika Mäki, Henri Mikkola, Amanda Pitkänen and Tiina Helamaa (Finland); Florence Jusot, Anne Laferrère, Louis Arnault, Emilie Bourgeat, Justine Bondoux, Adèle Lemoine, Benjamin Levy and Thomas Renaud (France); Annette Scherpenzeel, Arne Bethmann, Felizia Hanemann, Michael Bergmann, Imke Herold, Sabine Friedel and Alexander Schumacher (Germany); Antigone Lyberaki, Platon Tinios, Michail Chouzouris, Thomas Georgiadis, Tassos Philalithis (Greece); Anikó Bíró, Gábor Kézdi, Réka Branyiczki and Maria Volkov (Hungary); Sharon Shiovitz-Ezra, Howard Litwin, Ella Cohn-Schwartz, Noam Damri, Avi Cohen, Bracha Erlich, and Michal Levinsky (Israel); Agar Brugiavini, Guglielmo Weber, Elena Bassoli, Michele Belloni, Marco Bertoni, Andrea Bonfatti, Raluca Elena Buia, Danilo Cavapozzi, Martina Celidoni, Angela Crema, Chiara Dal Bianco, Giacomo Pasini, Caterina Pavese, Irene Simonetti and Nancy Zambon (Italy); Andrejs Ivanovs, Ieva Reine, Signe Tomsone and Diana Baltmane (Latvia); Antanas Kairys, Vida Jakutienė and Olga Zamalijeva (Lithuania); Maria Noel Pi Alperin, Gaetan de Lanchy, Sylviane Breulheid, Nada Magnoni, Marie-Josée München, Thierry Kreuten and Roland Maas (Luxembourg); Marvin Formosa, Caroline Pavia-Sciortino and Roberta Vella (Malta); Adriaan Kalwij and Marika de Bruijne (the Netherlands); Agnieszka Chłoń-Domińczak, Dorota Holzer-Żelażewska and Monika Oczkowska (Poland); Alice Delerue Matos, Pedro Pita Barros, Fátima Barbosa, Patrícia Silva, Gina Voss, Manuela André, Cláudia Cunha and Andreia Paiva (Portugal); Alin Marius Andrieș, Mircea Asandului and Daniela Viorică (Romania); Ľubica Knošková, Ivan Hlavatý, Ivan Simek and Dana Vokounová (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); Jürgen Maurer, Alberto Holly, Carmen Borrat-Besson, Robert Reinecke, Simon Seiler and Sarah Vilpert, (Switzerland); Boris Majcen, 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 Mary Beth Ofstedal, Jim Smith and David Weir from HRS; and to the members of the SHARE scientific monitoring board (Arie Kapteyn, chair, David Bell, Lisa Berkman, Kaare Christensen, Mick Couper, Michael Hurd, Daniel McFadden, David Meltzer, 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, Danilo Cavapozzi and Elena Raluca Buia. Guglielmo Weber (University of Padua) led the income and wealth area with Anne Laferrère, Andrea Bonfatti, Chiara Dal Bianco, Martina Celidoni and Nancy Zambon. The health area was led by Karen Andersen-Ranberg and assisted by Jens Elmelund Rise (University of Southern Denmark). Health care and health services utilization fell into the realm of Florence Jusot (University of Paris-Dauphine), who was assisted by Thomas Renaud, Emilie Bourgeat and Louis Arnault. The fifth area, family and social networks, was led by Howard Litwin from the Hebrew University of Jerusalem with assistance from Ella Schwartz and Michal Levinsky.

The coordination of SHARE entails a large amount of day-to-day work that is easily understated. We would like to thank Kathrin Axt, Corina Lica, Karl Riedemann and Andrea Oepen for managing SHARE ERIC and SHARE finances; Verena Coscia, Veronika Máté and Julia Knoblechner for their efforts in public relations; and Renate Eggenreich, Hannelore Henning and Stephanie Lasson at MEA in Munich for their administrative support throughout various phases of the project. Finally, Annette Scherpenzeel, as international coordinator, guided us through the first phase of the COVID-19 pandemic with all its challenges regarding data collection – thank you for all your effort in this turbulent times (and also before)! Preparing the data files for fieldwork, monitoring the survey agencies, testing the data for errors and consistency are all tasks that are essential to this project. Many thanks therefore go to Julia Amorim, Nora Angleys, Josefine Atzendorf, Michael Bergmann, Arne Bethmann, Tim Birkenbach, Carolina Brändle, Johanna Bristle, Salima Douhou, Nele Ehrenberg, Theresa Fabel, Fabio Franzese, Magdalena Gerum, Stefan Gruber, Tessa-Virginia Hannemann, Magdalena Hecher, Imke Herold, Maria José Mendoza Jiménez, Judith Kronschnabl, Diana López-Falcón, Yuri Pettinicchi, Senta-Melissa Pflüger, Jeny Tony Philip, Gregor Sand, Daniel Schmidutz, Karin Schuller, Alexander Schumacher, Elena Sommer, Stephanie Stuck, Barbara Thumann, Melanie Wagner, Luzia Weiss, Yasemin Yilmaz 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 Paolo Li Donni for weight calculations and imputations in Palermo.

Programming and software development for the SHARE survey was done by CentERdata in Tilburg. We want to thank Eric Balster, Marika de Bruijne, Marcel Das, Maurice Martens, Marije Oudejans, Sebastiaan Pennings, Maarten Streefkerk, Iggy van der Wielen, Nick Wondergem, 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 and Antwerp University (Belgium), GfK (Bulgaria), IPSOS (Croatia), RAI (Cyprus), SC&C (Czech Republic), DST Survey (Denmark), Statistics Estonia (Estonia), Taloustutkimus (Finland), KANTAR (formerly known as TNS Sofres) (France), KANTAR (Germany), KAPA research (Greece), TARKI (Hungary), Cohen Institute (Israel), IPSOS (Italy), Institute of Sociological Research (Latvia), KANTAR (Lithuania), LISER (Luxembourg), EMCS (Malta), I&O Research (Netherlands), TNS (Poland), DOMP (Portugal), GfK (Romania), Go4insight & ACRC (Slovakia), IPSOS (Slovenia), KANTAR (Spain) and IPSOS (in the 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 2,500 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 SHARE data collection has been funded by the European Commission through the 5th framework programme (project QLK6-CT-2001-00360 in the thematic programme Quality of Life). Further support by the European Commission through the 6th framework programme (projects SHARE-I3, RII-CT-2006-062193, as an Integrated Infrastructure Initiative, COMPARE, CIT5-CT-2005-028857, as a project in Priority 7, Citizens and Governance in a Knowledge Based Society, and SHARE-LIFE (CIT4-CT-2006-028812)), through the 7th framework programme (SHARE-PREP (No 211909), SHARE-LEAP (No 227822) and M4 (No 261982) and through Horizon 2020 (SHAREDEV3 (No 676536), SERISS (No 654221), SSHOC (No 823782), RItrain (No 654156), and ERIC Forum (No 823798)) as well as by DG Employment, Social Affairs & Inclusion through VS 2015/0195, VS 2016/0135, VS 2018/0285, VS 2019/0332, and VS 2020/0313 is gratefully acknowledged. The SHARE-DEV 3 project financed the international coordination outside of Germany; with the help of the SHARE-COHESION project innovation and long-term sustainability of SHARE are supported. We thank the director generals Robert-Jan Smits and Jean-ERIC Paquet and, in alphabetical order Peter Droell, Ales Fiala, Philippe Froissard, Johannes Klumpers, Jean-David Malo, Anna Panagopoulou, Dominik Sobczak, Maria Theofilatou, Harry Tuinder and Adam Tyson in DG Research for their continuing support of SHARE. DG Employment, Social Affairs, and Equal Opportunities was instrumental to cover all EU Member States in SHARE. We are grateful for the support by Dana-Carmen Bachmann, Dario Fiorucci, Georg Fischer, Katarina Ivankovic-Knezevic, Ralf Jacob, Bettina-Maria Kromen, Ettore Marchetti and Vladis Zagorskis.

Substantial co-funding for add-ons such as the intensive training and retention program and the collection of HRS-harmonised biomarkers was granted by the US National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, R21 AG025169, Y1-AG-4553-01, IAG BSR06-11, OGHA 04-064, BSR12-04 and R01AG052527-02), further funding was granted for the development of a Harmonized Cognitive Assessment Protocol (HCAP) (R01 AG056329-02). We thank John Haaga, Georgeanne Patmios and John Phillips for their enduring support and intellectual 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 Minister Anja Karliczek, Head of Divison Ulrich Schüller, Angelika Willms-Herget, Klaus Schindel and Maike Koops who served as chairs of the SHARE-ERIC Council, and, in alphabetical order, Gisela Helbig, Hans Nerlich, Monika van Ooyen 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 President Martin Stratmann, Vice Presidents Angela Friederici and Ulman Lindenberger, Secretary-General Rüdiger Willems, as well as Maximilian Prugger, Sabine Mellinghoff and Corinna Heel.

The bulk of funding of Wave 8 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., Schaan, B., Stuck, S., & Zuber, S. (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

Börsch-Supan, A. (2020). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8. Release version: 0. SHARE-ERIC. Preliminary data set.

Couper, M. P. (2011). The future of modes of data collection. Public Opinion Quarterly, 75(5), 889-908. https://doi.org/10.1093/poq/nfr046

Scherpenzeel, A., Axt, K., Bergmann, M., Douhou, S., Oepen, A., Sand, G., Schuller, K., Stuck, S., Wagner, M., & Börsch-Supan, A. (2020). Collecting survey data among the 50+ population during the COVID-19 outbreak: The Survey of Health, Ageing and Retirement in Europe (SHARE). Survey Research Methods, 14(2), 217–221. https://doi.org/10.18148/srm/2020.v14i2.7738

I
SHARE WAVE 8 (CAPI)

CHAPTER 2

Sampling Design in SHARE Wave 8 and Recruitment of Refreshment Samples until the Suspension of Fieldwork

02

2 SAMPLING DESIGN IN SHARE WAVE 8 AND RECRUITMENT OF REFRESHMENT SAMPLES UNTIL THE SUSPENSION OF FIELDWORK

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), and Giuseppe De Luca – University of Palermo

2.1 Introduction

The aim of the SHARE survey design is to be able to draw inferences about the population of people aged 50 years or 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 have to be internationally comparable. As in previous waves, this chapter documents the sampling design adopted in the eighth wave of SHARE that had to be suspended due to COVID-19 in March 2020. While the pandemic had severe consequences for data collection (especially regarding the drawn refreshment samples), the regular sampling procedure in the run-up to Wave 8 was not affected. Starting with a definition of the SHARE target population, we describe the protocol for harmonizing and documenting the sampling procedure and present the sampling frames used by the countries that recruited a baseline or refreshment sample in Wave 8. We then discuss some important aspects of the SHARE sampling designs, such as stratification, clustering, variation in selection probabilities and sample composition. Finally, we provide additional information about the sampling variables provided in the SHARE release.

2.2 The SHARE Target Population

The target population of SHARE consists of all persons born in 1969 or before at the time of sampling (i.e. 2019 in Wave 8) who have their regular domicile in the respective SHARE country. Persons are excluded if they were incarcerated, hospitalised or out of the country throughout the entire survey period, unable to speak the country's language(s)[2], could not 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 or 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 2.4 and Schanze, 2017). Additionally, respondents are followed when entering a nursing or residential home. Further information on eligibility for the study can be found in the SHARE Release Guide that is publicly available on the SHARE website (www.share-project.org).

2.3 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 completed Sampling Design Form (SDF), containing a full description of both the chosen sampling frame and the associated sampling design. In the second stage, the sampling proposal is carefully evaluated by the SHARE Central coordination team in Munich. Open questions and uncertainties are clarified on a bilateral basis with the country team and/or the survey agency before the sampling design is finally 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 the gross sample via the completed Gross Sample Template (GST). This template contains all selected persons or households, the associated sampling frame information needed for the computation of selection probabilities (e.g. household-level and population-level information about stratification and clustering), household-level information about NUTS and LAU codes and (if any) additional auxiliary variables that could be used for ex

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

post compensation of non-sampling errors. After another round of checks by SHARE Central to detect inconsistencies with the proposed sampling design, the GST forms the baseline/refreshment sample part (in addition to a possible panel sample) of the SampleCTRL software (see Chapter 4) that is used to assign new respondents to the interviewers' laptops.

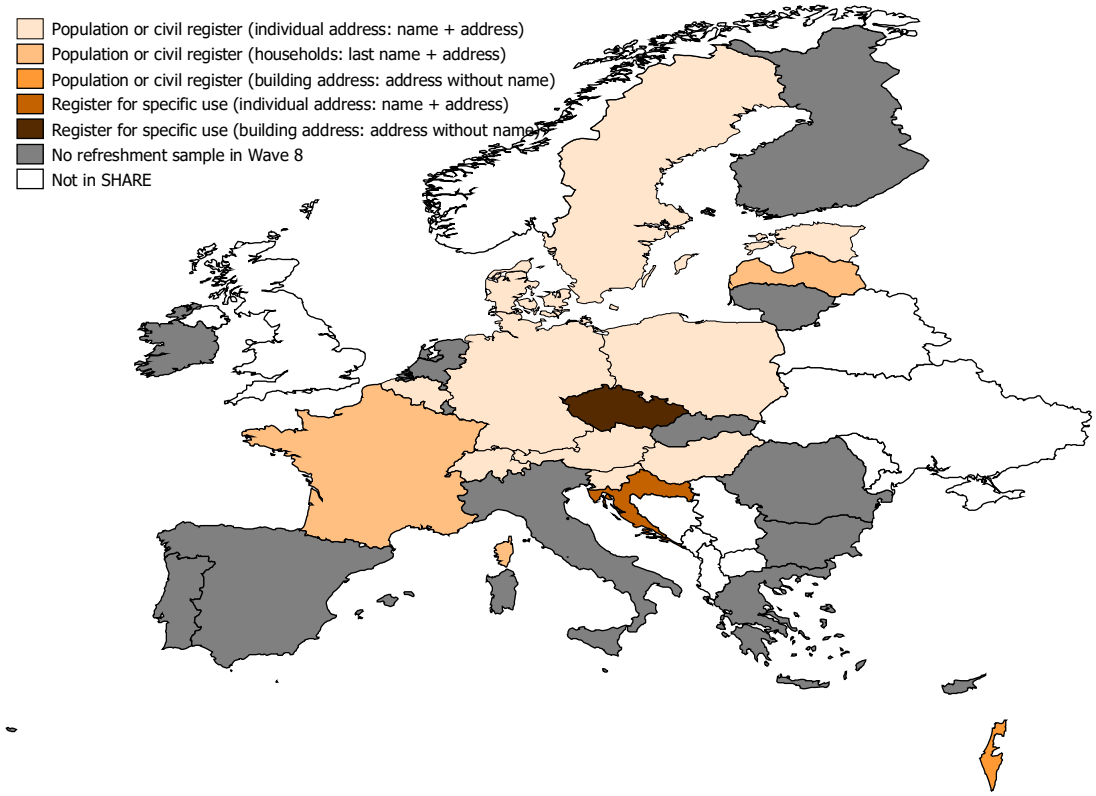
2.4 Sampling Frames and Population Coverage

In the ideal case scenario, all countries included in SHARE would have a probability-based sample from an official person register covering the population of interest. The availability of population registers that can be used as sampling frames varies a lot, however, 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 that may jeopardize the standard properties of sample-based inference. Therefore, SHARE Central provides a template letter that can be customised to the country-specific situation in order to facilitate access to an official register for sampling (Scherpenzeel, 2018a, 2018b).

All countries in SHARE that draw a baseline/refreshment sample are requested to use the best sampling frame available in each wave, implying that sampling frames can in principle be different between waves and/or countries. For the target population of SHARE, a key feature any sampling frame has to 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 CaseCTRL software (see Chapter 4) has to be applied before starting fieldwork in order to identify sample members aged 50 years or older.

Figure 2.1 shows which countries recruited a refreshment sample and on what type of sampling frame these samples were based. Refreshment samples were drawn in Austria, Belgium, Switzerland, the Czech Republic, Germany, Denmark, Estonia, Spain, Finland, France, Croatia, Hungary, Israel, Latvia, Poland, Portugal, Sweden and Slovenia. While in most countries (central or local) population registers with individual information on age were available, the Czech Republic, France, Israel and Latvia used a register that only allowed households with at least one person aged 50 years or older to be selected. Thus, although a screening procedure had to be applied, the size of the drawn gross sample could be minimised here. All other countries participating in Wave 8 interviewed only their longitudinal samples and did not draw a refreshment sample.

Figure 2.1: Types of Sampling Frames Used in SHARE Countries for Refreshment Samples in Wave 8



In Europe, 2.7 million people are aged 65 or older and live in a retirement home, nursing home or a healthcare institution (Laferrère et al., 2013). Within the group of Europeans aged 85 or older, 12.6 per cent live in an institution. As described above, persons aged 50 years or 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 errors might be either due to 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 targeted measures to help interviewers gain access to nursing home respondents.

2.5 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 8, as in all foregoing waves, is the same as that applied by all advanced population-based survey programmes 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 being 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, a number of pieces of 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 (Bethmann et al., 2019) as well as bilateral discussions with the SHARE Central coordination team. We summarize these important aspects of the sampling design in the following sections.

2.6 Stratification

The most frequently used sampling design in SHARE is a multistage stratified sampling design. Regional stratification schemes are recommended in order 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 – countries are advised to also use these for stratification.

2.7 Clustering

SHARE aims to use sampling schemes with a minimum variation of selection probabilities and a minimum amount of clustering. However, designing 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 then has to 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:

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

where n is the sample size and w_i are design weights defined as the inverse of the selection probabilities.

Other studies (e.g. the European Social Survey, see Lynn et al., 2018) have shown that $Deff_p$ 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, $Deff_p$ is smaller than this 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 or older or a couple containing 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 8, all countries had access to some form of official register and sample 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 concerns statistical efficiency. For any estimator $\hat{\theta}$ of a parameter θ, the design effect due to clustering can be measured by

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

where $\text{Var}(\hat{\theta})_c$ and $\text{Var}(\hat{\theta})_s$ 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. Stratification of the population of clusters can help to contrast this efficiency loss and was hence strongly advised. Further, the countries were instructed to choose a mean cluster size as small as possible and to select as many PSUs as possible (see Section 2.10 for an overview of the sampling design variables included in the released SHARE data set).

2.8 Selection Probabilities

The calculation of selection probabilities in SHARE is subject to three difficulties: first, these probabilities must take into account the aforementioned 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 the marital status, partnership or age of the spouse/partner that is required to compute selection probabilities of couples with two age-eligible persons; third, as the panel goes ahead, many countries attempt to maintain the representation of the younger age 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 non-zero probability of being selected in the refreshment subsample and the elements of the refreshment subsample may have, in turn, a non-zero probability of being selected in the longitudinal sample. Further, the sampling frame information needed to compute these non-zero "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 in 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. More precisely, we deal with the lack of sampling frame information about the spouse/partner of each sample member by using the household composition data collected through the Coverscreen (CV) module, which is asked about 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 non-zero 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 estimation of the unknown cross-selection probabilities using 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.

2.9 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. The choice of conducting a refreshment sample is mainly made by the countries, because they have to apply for their own funding to their national funding agencies. Since funding and sampling resources vary across participating countries, SHARE does not define a minimum net sample size. Instead, SHARE advises countries to maximize their net sample size with the available budget. In Wave 8, all countries that drew a refreshment sample except Switzerland, Denmark, Estonia, Portugal and Sweden selected the full age range of people born in 1969 or earlier to compensate for the effect of panel attrition on all age cohorts. Where possible, these full-range refreshment samples included an oversampling of the youngest cohorts that were not age-eligible in the previous refreshment samples to maintain the representation of younger age cohorts. In contrast to other studies (e.g. Labour Force Survey), no panel rotation method was used in order to maximize the sample size available for longitudinal analyses. In other words, all units in the panel sample were considered eligible for an interview in the eighth wave, including non-responding partners of panel members who

were interviewed in a previous wave. Table 2.1 gives an overview of all countries that ever participated in SHARE up to Wave 8 and the composition of their samples in the respective wave(s).

Table 2.1: Sample Type by Wave and Country

Country	Wave 1	Wave 2		Wave 3	Wave 4		Wave 5		Wave 6		Wave 7		Wave 8	
	Baseline	Panel	Refreshment /Baseline	Panel	Panel	Refreshment /Baseline	Panel	Refreshment /Baseline	Panel	Refreshment /Baseline	Panel	Refreshment /Baseline	Panel	Refreshment /Baseline
AT	≤1954	✓		✓	✓	≤1960	✓		✓		✓		✓	≤1969
BE_FR	≤1954	✓	≤1956	✓	✓	≤1960	✓	≤1962	✓	≤1964	✓		✓	≤1969
BE_NL	≤1954	✓		✓	✓	≤1960	✓	≤1962	✓	≤1964	✓		✓	≤1969
BG												≤1966	✓	
CH	≤1954	✓	≤1956	✓	✓	≤1960	✓		✓		✓		✓	[1962-1965]
CY												≤1966	✓	
CZ			≤1956	✓	✓	≤1960	✓	≤1962	✓		✓		✓	≤1969
DE	≤1954	✓	≤1956	✓	✓		✓	≤1962	✓		✓		✓	≤1969
DK	≤1954	✓	≤1956	✓	✓	[1957-1960]	✓	≤1962	✓	[1963-1964]	✓		✓	[1967-1969]
EE						≤1960	✓		✓	[1963-1964]	✓		✓	[1965-1969]
EG								≤1962	✓		✓		✓	
ES	≤1954	✓	≤1956	✓	✓	≤1960	✓		✓		✓		✓	≤1969
FI												≤1966	✓	≤1969
FR	≤1954	✓	≤1956	✓	✓	≤1960	✓		✓	≤1964	✓		✓	≤1969
GR	≤1954	✓	≤1956	✓					✓	≤1964	✓		✓	
HR										≤1964	✓	≤1966	✓	≤1969
HU						≤1960					✓		✓	≤1969
IE			≤1956	✓										
IL	≤1954	✓	≤1956				✓	[1953-1962]	✓		✓	≤1966	✓	≤1969
IT	≤1954	✓	≤1956	✓	✓	≤1960	✓	≤1962	✓	≤1964	✓		✓	
LT												≤1966	✓	
LU								≤1962	✓	≤1964	✓		✓	
LV												≤1966	✓	≤1969
MT												≤1966	✓	
NL	≤1954	✓	≤1956	✓	✓	≤1960	✓	≤1962	✓		✓		✓	
PL			≤1956	✓	✓				✓	[1963-1964]	✓	≤1966	✓	≤1969
PT						≤1960			✓		✓		✓	[1961-1969]
RO												≤1966	✓	
SE	≤1954	✓	≤1956	✓	✓		✓	≤1962	✓		✓		✓	[1955-1969]
SI						≤1960	✓	≤1962	✓	≤1964	✓		✓	≤1969
SK												≤1966	✓	

Note:

≤1966	Baseline sample
≤1969	Full-range refreshment sample
[1967-1969]	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 i) to maintain the representation of the younger age cohorts of the target population in Wave 8 that were not age-eligible in the previous waves (i.e. people born between 1967 and 1969), and ii) to compensate the reduction in the size of the panel sample due to attrition. In Wave 8, refreshment samples were recruited in Austria, Belgium, Switzerland, the Czech Republic, Germany, Denmark, Estonia, France, Croatia, Hungary, Israel, Latvia, Poland, Sweden and Slovenia until fieldwork had to be stopped in March 2020 due to the outbreak of the COVID-19 pandemic. In Spain, Finland and Portugal, the drawn refreshment samples could not be fielded anymore. Table 2.2 presents an overview of the collected longitudinal and refreshment interviews until the stopping of fieldwork. It is planned to continue interviewing the started/drawn refreshment samples in Wave 9, provided that face-to-face interviews (CAPI) are possible again and to release them together with the Wave 9 data.

Table 2.2: Number of Realised Interviews from Longitudinal and Refreshment Samples of Wave 8 until the Suspension of Fieldwork by Country

Country	Individuals			Households		
	Longitudinal	Refreshment	Total	Longitudinal	Refreshment	Total
AT	1738	383	2121	1254	301	1555
BE_fr	798	131	929	625	115	740
BE_nl	1401	136	1537	1008	101	1109
BG	993		993	679		679
CH	1995	186	2181	1479	145	1624
CY	570		570	386		386
CZ	2968	328	3296	2048	232	2280
DE	2963	979	3942	2006	757	2763
DK	2305	193	2498	1676	153	1829
EE	3475	380	3855	2519	279	2798
EG	809		809	606		606
ES	1504	0	1504	1018	0	1018
FI	1182	0	1182	817	0	817
FR	2601	532	3133	1871	391	2262
GR	3267		3267	2160		2160
HR	1337	835	2172	869	573	1442
HU	941	427	1368	661	300	961
IL	1077	476	1553	756	375	1131
IT	2376		2376	1534		1534
LT	1536		1536	1098		1098
LU	963		963	698		698
LV	836	421	1257	592	303	895
MT	834		834	494		494
NL	2067		2067	1373		1373
PL	2307	773	3080	1531	498	2029
PT	0	0	0	0	0	0

Country	Individuals			Households		
	Longitudinal	Refreshment	Total	Longitudinal	Refreshment	Total
RO	1407		1407	900		900
SE	2535	179	2714	1831	151	1982
SI	2724	542	3266	1742	378	2120
SK	1036		1036	636		636
Total	**50545**	**6901**	**57446**	**34867**	**5052**	**39919**

Data: SHARE Wave 8, end of fieldwork.

2.10 Sampling Variables in the Released SHARE Data

The upcoming release will include a generated module (gv_weights) with variables providing information about the sampling design in each country. Hence, the variable "subsample" identifies the various subsamples drawn in a specific country and wave of the SHARE panel sample, while the indicators "psu", "ssu", "stratum1" and "stratum2" provide information about stratification and clustering in each subsample. In addition, the gv_housing module will contain regional information (so called NUTS areas; due to data protection, only the NUTS1 level can be released) about the interviewed respondents that are also part of the GST (see "Bethmann et al., 2019" for further information). Table 2.3 provides an overview of these variables that are necessary to construct appropriate weights addressing problems of unit non-response and attrition (see Chapter 6 on weighting and imputation strategies).

Table 2.3: 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	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 Center for the Economics of Aging (MEA).

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. First results from SHARE after the economic crisis (pp. 253–264). De Gruyter.

Lynn, P. (2009). Methods for longitudinal surveys. In P. Lynn (Ed.), Methodology of longitudinal surveys (pp. 1–19). Wiley.

Lynn, P., Scherpenzeel, A., Liberts, M., Kaminska, O., & Goedeme, T. (2018). European Social Survey round 9 sampling guidelines: Principles and implementation. The ESS Sampling and Weighting Expert Panel. ESS ERIC Headquarters.

Schanze, J.-L. (2017). Report on sampling practices for the institutionalized 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. Available at: www.seriss.eu/resources/deliverables

Scherpenzeel, A. (2018a). Background to the letter to request access for sampling for European survey infrastructures. Deliverable 2.3a of the SERISS project funded under the European Union's Horizon 2020 research and innovation programme GA No: 654221. Available at: www.seriss.eu/resources/deliverables

Scherpenzeel, A. (2018b). Letter to request access for sampling for European survey infrastructures. Deliverable 2.3b of the SERISS project funded under the European Union's Horizon 2020 research and innovation programme GA No: 654221. Available at: www.seriss.eu/resources/deliverables

Smith, P., Lynn, P., & Elliot, D. (2009). Sample design for longitudinal surveys. In P. Lynn (Ed.), Methodology of longitudinal surveys (pp. 21-33). Wiley.

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. https://doi.org/10.18148/srm/2014.v8i3.5818

CHAPTER 3

Questionnaire Innovations

03

3 QUESTIONNAIRE INNOVATIONS

3.1 Introduction to Questionnaire Innovations

Melanie Wagner – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC) and Howard Litwin – Hebrew University of Jerusalem

The eighth wave of SHARE holds several novelties with regard to the questionnaire structure and its content. The first novelty is that Wave 8 is the first regular SHARE wave for the eight new countries of Finland, Latvia, Lithuania, Slovakia, Bulgaria, Romania, Cyprus and Malta. They started in Wave 7 with a questionnaire on people's life histories, the so-called SHARELIFE survey[3], and only received a few regular SHARE questions in addition. Wave 8 therefore contains for the first time regular SHARE data for all new countries and all respondents who did the SHARELIFE interview in Wave 7. The second major innovation in Wave 8 is a clear distinction between the core questionnaire and the so-called "add-on modules". The core questionnaire consists of the usual panel modules, which are intended to be stable over the waves. Add-on modules are either designed as experimental modules, which may be integrated into the core questionnaire in a later wave, or as one-time-only modules.

3.1.1 Changes in the Core Questionnaire

Changes in the core questionnaire were kept to a minimum for the sake of panel stability and only administered if indispensable for the improvement of the data quality (e.g. wording changes to reduce ambiguity). An exception is the Health Care (HC) module, which underwent larger changes. Louis Arnault, Florence Jusot and Thomas Renaud will lay down the rationale for these changes and explain the content of the new HC module in detail in Section 3.2 of this volume.

Also, the Social Network (SN) module was administered for the third time in the course of the SHARE project. As in the two previous administrations in Waves 4 and 6, respondents were asked to name up to six persons with whom they discussed important matters, with the option of naming a seventh person who was important to them for any reason. Then, additional information was requested on each named person (gender, relationship type, geographic proximity, frequency of contact and emotional closeness).

The second administration of the SN module in Wave 6 allowed researchers to take into account changes that may have occurred in their networks (e.g. increase or decrease in size, composition, closeness, etc. since the previous administration of the module). The third administration of the module, in Wave 8, widens the possibility of examining changes in the interpersonal environments of the SHARE participants. Accordingly, the network change variables in the database have been updated to reflect the network dynamics that were observed across the three waves in which the module was administered (Waves 4, 6 and 8). Thus, three new generated variables were added, reflecting changes in the following dimensions: the size of the social network, the number of children in the network and the number of friends.

3.1.2 New Add-on Modules

The add-on part of the Wave 8 questionnaire consists of four new modules on cognition, saving regrets, time use and a collection of accelerometry data. The background and content of these modules are described in more detail in Sections 3.3 to 3.6.

In Section 3.3, Melanie Wagner and Salima Douhou explain the rationale for the inclusion of new tests on cognitive functioning (CF add-on). Together with the existing tests and questions on cognitive functioning, it will help identify which interactions of biomedical and socio-economic conditions over the life course affect cognition in later life.

Axel Börsch-Supan and Irene Ferrari will present the Saving Regret (SR) module in Section 3.4. The aim of this module is to measure the extent of saving regret in a way that exploits the SHARE data from earlier waves and the life histories in order to link insufficient or excessive saving to its potential causes in varying European institutional environments created by the welfare state and private markets. The module includes questions asking about life course events that had positive or negative effects on a household's financial

3 For further information see Philip & Wagner (2019).

situation, whether the respondents wish that they had saved more or less in the past and about areas on which they wish they had spent less.

In Section 3.5, Annette Scherpenzeel and Jeny Tony Philip explain the Time Expenditure (TE) module. SHARE data on time use offer valuable insights into how active the elderly still are, how much they contribute to paid and unpaid work, and at what ages important changes in the activity patterns occur. Moreover, the data make it possible to analyse how time use patterns before and after retirement vary among different countries in Europe.

Fabio Franzese, Annette Scherpenzeel, Luzia Weiss and Nora Angleys will present the accelerometer project in more detail in Section 3.6. The accelerometer project collects measurements of physical activity based on sensor data. A short CAPI module was designed to ask eligible respondents for their consent to participate in the accelerometer study during the regular SHARE Wave 8 interview.

References

Philip, J. T., Wagner, M. (2019). SHARELIFE. In M. Bergmann, A. Scherpenzeel & A. Börsch-Supan (Eds.), SHARE Wave 7 Methodology: Panel innovations and life histories (pp. 109–153). MEA, Max Planck Institute for Social Law and Social Policy.

3.2 Modernising Measures of Healthcare Utilisation and Access to Care

Thomas Renaud and Louis Arnault – Laboratoire d'Economie de Dauphine (LEDa), Université Paris Dauphine, PSL and Florence Jusot – Laboratoire d'Economie de Dauphine (LEDa), Université Paris Dauphine and Institut de Recherche et Documentation en Economie de la Santé (IRDES)

3.2.1 Motivation and Implementation

The in-depth revision of the Health Care (HC) module in the SHARE core questionnaire responds to the need to improve the measurement of access to care, by better identifying unmet needs and barriers to access to care, and by refining our comprehension of healthcare utilisation in its various dimensions: ambulatory vs hospital care, primary vs secondary care, preventive vs curative care, short-term vs long-term care.

A precise understanding of the demand for care and of healthcare utilisation is required to accurately assess the impact of use of care on other related aspects, like subjective, mental and functional health, economic and relational vulnerability, labour market and social participation, etc. Hence, researchers do need comparable measures of healthcare utilisation beyond the diversity of national healthcare and health insurance systems, although such heterogeneity makes it especially challenging to design questions that are equally applicable and relevant for all countries.

Revision of the Health Care (HC) module mainly consists of enrichments in the following five topics: 1) routine preventive care and screening examinations carried out; 2) distinction between primary and secondary care in reporting the number of contacts with doctors; 3) distinction between emergency and programmed hospital care; 4) precise assessment of unmet needs through the refinement of questions on forgone care; 5) measure of health literacy. In the end, the HC module is comprised of a total of 34 questions in Wave 8 (against 41 in Wave 7), of which 24 are left unchanged. Eight questions deviate at least partly from previous waves, two questions are entirely new and six questions from Wave 7 were dropped. The median duration to answer this module was 3.6 minutes in the main fieldwork of SHARE Wave 8 (Release version: 0).

3.2.2 Preventive Care

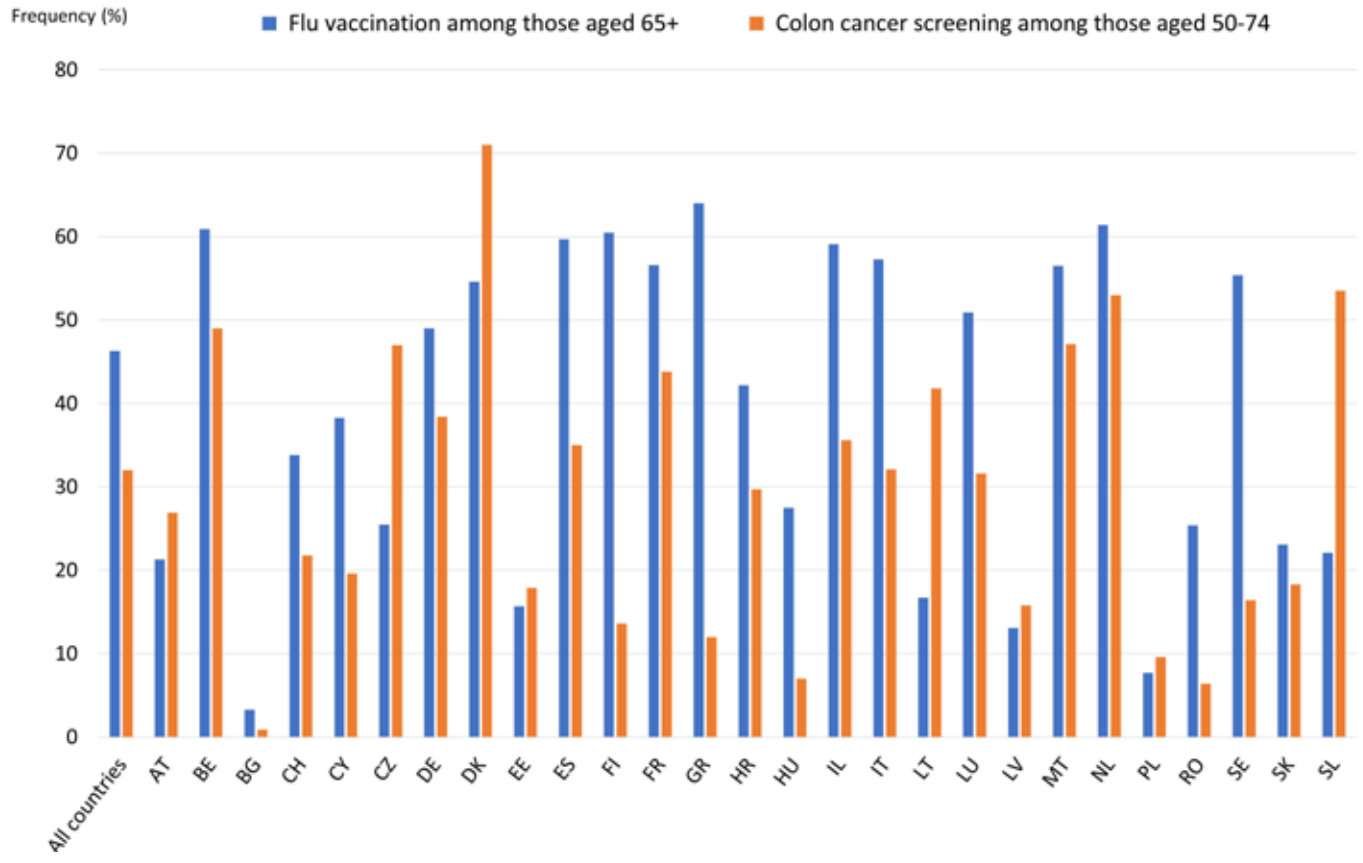
In the EU-27, around 645,000 of the 1.5 million deaths of persons aged less than 75 years in 2016 could have been prevented through better public health interventions (Eurostat, 2020). In this respect, developing routine vaccination (primary prevention) or cancer screening (secondary prevention) programmes is one of the 10 essential public health operations (EHPOs) led and supported by the World Health Organisation to help countries to improve the global health of populations.

The contents of questions introduced in Wave 8 about preventive care originates from both pre-existing questions in SHARE and from international consensus on high-priority prevention interventions. The first one refers to the annual flu vaccination, which has been proved to be successful in preventing flu-related complications and premature mortality among the elderly (Demicheli et al., 2018). This information had already been collected in the past in SHARE as part of drop-off questionnaires in Waves 1 and 2. Two questions relate to screening for two of the cancers with the highest incidence: breast cancer and colorectal cancer. It seems appropriate to describe what the examination is about in the wording of these two questions, as the exact name of the screening test may not be known or remembered by the respondent. Several questions related to breast cancer screening (mammograms) had already been asked in the SHARELIFE life course questionnaire in Wave 3. One additional question asks if the respondent has had an eye examination performed recently, which is useful for detecting age-related eye diseases at an early stage (cataract, glaucoma, macular degeneration). Along with the question on flu vaccine and the existing question about routine visits to the dentist, kept unchanged in SHARE since Wave 1, it provides an insight into the respondents' primary prevention practices.

HC884	In the last year, that is since {date1}, did you have a flu vaccination? *Yes/No*
HC885	In the last two years, that is since {date2}, have you had an eye exam performed by an eye care professional such as an ophthalmologist or optometrist? *Yes/No*
HC886	In the last two years, that is since {date2}, have you had a mammogram (x-ray of the breast)? *Yes/No*
HC887	Some health care providers do tests such as test for detecting hidden blood in your stool, sigmoidoscopy or colonoscopy to check for colon cancer. In the past two years, that is since {date2}, have you had any of these tests? *Yes/No*

In SHARE Wave 8, 46 per cent of respondents aged 65 and over have been vaccinated against flu the year before the interview and 32 per cent of people aged between 50 and 74 have been screened for colon cancer in the previous two years (see Figure 3.1). These proportions are remarkably low in Bulgaria (3 per cent for flu vaccination and 1 per cent for colon cancer screening) and at their highest in Denmark (71 per cent and 64 per cent, respectively).

Figure 3.1: Colon Cancer Screening (Among 50- to 74-year-olds) and Flu Vaccination (Among those Aged 65+) by Country



Data: SHARE Wave 8, Release version: 0 (weighted).
Note: Respondents aged 50 to 74 who answered the question about colon cancer screening (N = 29,550); 65+ respondents who answered the question about flu vaccination (N = 30,814).

3.2.3 Contacts with General Practitioners and Specialist Doctors

The question measuring the overall number of contacts with doctors (*HC602*), which already existed in the previous Waves 6 and 7 of SHARE, was preserved in Wave 8 for the sake of longitudinal comparability. Its scope is broad since it includes contacts with several categories of health professionals, from doctors to qualified nurses, for different types of consulta-

tions, from ambulatory care to emergency department or outpatient clinic visits. Nevertheless, it excludes dental visits and hospital stays, which are the main focus of other questions in the HC module. The question is phrased (*"how many times in total have you seen or talked to…"*) to embrace the many ways in which a healthcare professional can be met: visits at the medical practice or at home, in-person or remotely, visits undergone by the respondent himself or by a relative on his/her behalf over the phone, etc.

Then, for those who reported at least one contact, two distinct questions about the number of contacts with general practitioners (GPs) on the one hand (*HC876*), and specialist doctors on the other (*HC877*), have been reintroduced in Wave 8, based on questions already asked in Waves 1, 2 and 4. The role assigned to GPs and the fees charged by both generalists and specialists vary widely across Europe, depending on the healthcare systems in place and the healthcare pathways recommended in the different countries to improve care coordination. It is therefore crucial to be able to distinguish between the care provided by each of these two types of professionals to understand and analyse the utilisation of healthcare by respondents in more depth.

HC602	During the last 12 months, that is since {date1}, about how many times in total have you seen or talked to a medical doctor or qualified/registered nurse about your health? Please exclude dentist visits and hospital stays, but include emergency room or outpatient clinic visits.
(IF HC602 > 0): HC876	How many of these contacts were with a general practitioner or with a doctor at your health care center?
(IF HC602 > 0): HC877	How many of these contacts were with a specialist, excluding dentist and emergency visits?

3.2.4 Planned vs Emergency Hospital Care

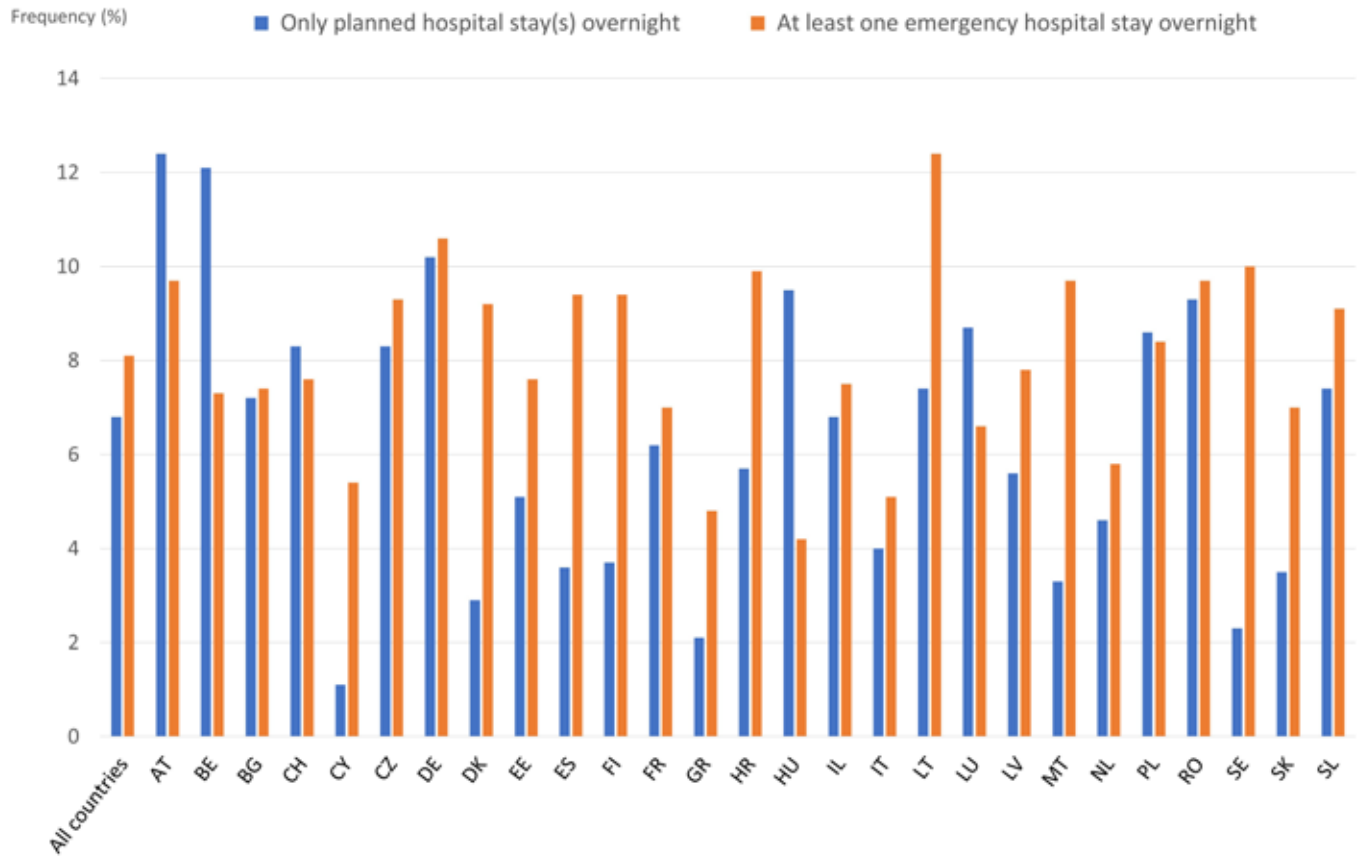
It is usual to distinguish in the activity of hospitals between planned care (also known as "elective", "scheduled" or "programmed" hospitalisations) and emergency ("unplanned") hospitalisations that respond to urgent and acute health issues that cannot be foreseen. A comparatively higher proportion of unplanned hospitalisations within a population or territory is seen as a reflection of some ineffectiveness of the healthcare system, in particular of poor quality and coordination in upstream ambulatory care and of a lack of emphasis on preventive care. To a certain extent, the rate of unplanned hospital admissions, especially among the elderly, can be seen as a crude performance indicator (Reed et al., 2015) – all the more so as the frequency of this phenomenon seems very heterogeneous across territories otherwise similar in age structures and prevalence rates of chronic diseases (Busby et al., 2015).

From the very start of SHARE, the respondent was continuously asked how many times he/she had been a patient in a hospital overnight (*HC013*). In Wave 8, two new questions have been added to distinguish between planned and emergency hospitalisations: a binary question (*HC888*) as to whether the respondent reports having been hospitalised only once in the last 12 months (*HC013* = 1) and a three-item question (*Planned, Emergency, Both*) regarding whether the respondent has been hospitalised more than once (*HC013* > 1).

HC013	How many times have you been a patient in a hospital overnight during the last twelve months?
(IF HC013 = 1): HC888	Was this stay in hospital planned or was it an emergency?
(IF HC013 > 1): HC890	Were these stays in hospital all planned, or were they all emergencies, or both? *Planned / Emergency / Both*

Among the 50+ in Europe, 15 per cent had been a patient in a hospital overnight during the 12 months before the interview: 7 per cent for planned hospital stays only and 8 per cent for at least one unplanned or emergency hospitalisation (see Figure 3.2).

Figure 3.2: Overnight Hospital Stays During the Last 12 Months by Country



Data: SHARE Wave 8, Release version: 0 (weighted).
Note: Respondents aged 50+ who answered the questions about having been a patient in a hospital overnight during the last 12 months and the type of hospital stays if any (N=43,718).

3.2.5 Unmet Healthcare Needs

SHARE has collected information on unmet care needs on a few occasions, for Wave 1, and also for Wave 7 in both regular and SHARELIFE questionnaires to determine if such events had occurred over the life course.

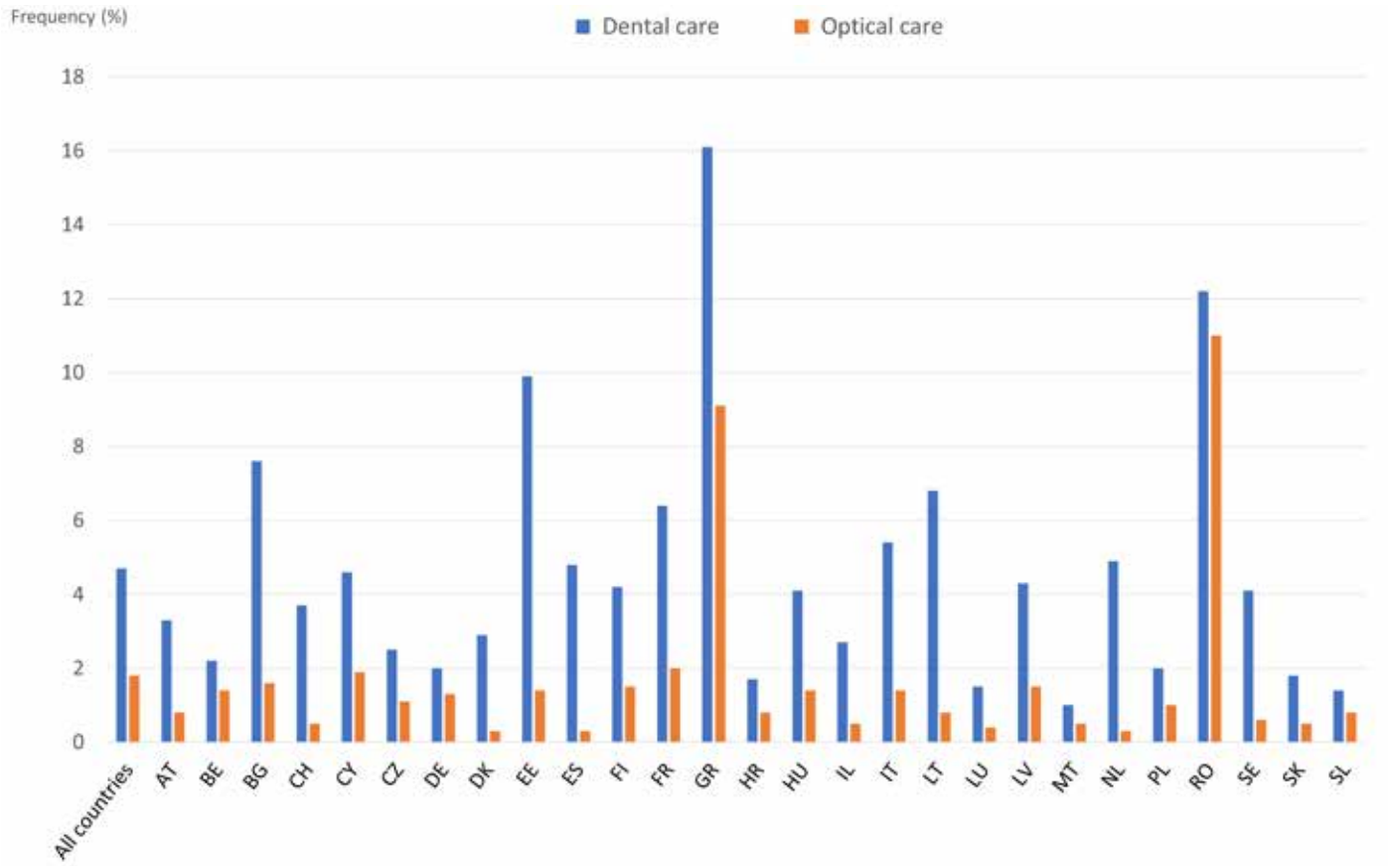
The measurement of unmet need is of great importance since it is widely recognised at the same time as being a valid indicator of the difficulties in accessing healthcare services, a threat to the equity of healthcare systems (Gibson et al., 2019) and a risk factor for deteriorating health, in terms of quality of life (Ko, 2016) and avoidable mortality (Alonso et al., 1997). Usually, the measurement of unmet healthcare needs differentiates between at least two main reasons why a person might not be able to fulfil his/her need for healthcare: financial reasons and reasons of availability/accessibility. The former refers to situations where people forgo healthcare they cannot afford, while the latter applies to people who relinquish care because of excessive distances or waiting times.

Two standard questions on forgone care due to cost and availability have already been reintroduced from Wave 5 onwards, after having been asked only once in SHARE Wave 1. These questions needed to be refined with the precise types of care, which respondents are led to forgo. The list of different types of care subject to unmet needs administered in Wave 8 is a short version of the list used in the Wave 1 questionnaire.

HC841	Please look at card [X]. During the last twelve months, which of the following types of care did you forgo because of the costs you would have to pay, if any? Code all that apply *Care from a general practitioner / Care from a specialist physician / Drugs / Dental care / Optical care / Home care / Paid home help / None of these*
HC843	Please look at card [X]. During the last twelve months, which of the following types of care did you forgo because they were not available or not easily accessible, if any? Code all that apply. *Care from a general practitioner / Care from a specialist physician / Drugs / Dental care / Optical care / Home care / Paid home help / None of these*

A share of 4.7 per cent of those aged 50+ have been forgoing dental care because they could not bear the costs, and 1.8 per cent optical care for the same reason (see Figure 3.3). These figures are remarkably high in Greece (16.1 per cent for dental care and 9.1 per cent for optical care) and in Romania (12.2 per cent and 11.0 per cent, respectively), whereas they are under 2 per cent and 0.5 per cent, respectively, in countries with the lowest levels of unmet needs (Malta, Slovakia, Luxembourg, Slovenia, Croatia). We carefully checked any translation issues, without observing problems.

Figure 3.3: Forgone Dental and Optical Care Due to Costs during the Last 12 Months by Country



Data: SHARE Wave 8, Release version: 0 (weighted).
Note: Respondents aged 50+ who answered the questions about forgone care because of costs (N = 43,693).

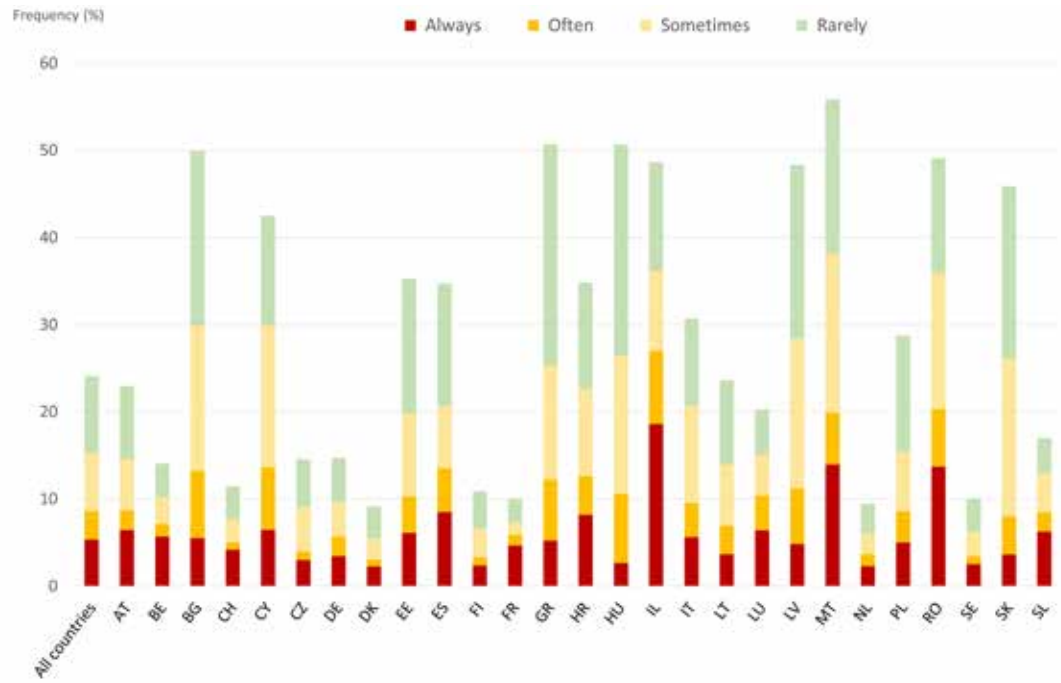
3.2.6 Health Literacy

Health literacy has emerged as a key concept over the last 20 years since the pioneering works of Williams (1995) and Nutbeam (2000). Originally, health literacy referred to the ability to perform health-related tasks requiring reading and computational skills. But the scope of this concept has evolved over the years and can now be defined more generally as "the degree to which an individual has the capacity to obtain, communicate, process, and understand basic health information and services to make appropriate health decisions"[4]. In this respect, it has a direct relevance to contemporary challenges of health democracy and patient empowerment. Health literacy has been shown to be consistently associated with education, ethnicity and age (Paasche-Orlow et al., 2005) as well as cognitive abilities (Wolf et al., 2012). Patients with limited health literacy are at greater risk of poor health outcomes, including poorer health-related knowledge, lower use of preventive care, and higher risks of hospitalisation and decease (Berkman et al., 2011; Bostock & Steptoe, 2012). Given the size constraint of the SHARE questionnaire, one single question is added, which asks the respondent to self-assess his/her health literacy on a five-item verbal rating scale.

HC889	How often do you need to have someone help you when you read instructions, pamphlets, or other written material from your doctor or pharmacy? *Always / Often / Sometimes / Rarely / Never*

In light of the risks of misinterpretation that arose during the development, translation and testing phases of the questionnaire, additional instructions are provided to the interviewers to ensure accuracy and consistency in this measure. In particular, it has been stressed that the "help" being referred to in this question is the help provided by "someone" and not "something" that could ease reading, like glasses. Some 15.3 per cent of individuals aged 50 or over report that, to some extent, they need someone to help them read instructions, pamphlets or other written material from their doctor or pharmacy: 6.7 per cent declare that this help is needed "sometimes", 3.2 per cent "often" and 5.4 per cent that this help is "always" needed (see Figure 3.4).

Figure 3.4: Need to Have Someone Helping When Reading Instructions from Doctor or Pharmacy (Health Literacy Score) by Country



Data: SHARE Wave 8, Release version: 0 (weighted).
Note: Respondents aged 50+ who answered the question about health literacy (N = 43,636).

4 Centers for Disease Control and Prevention (CDC) website. Available at: https://www.cdc.gov/healthliteracy.

3.2.7 Conclusion

These enrichments of the HC module allow a more detailed understating of healthcare utilisation, through new questions on the types of doctors visited (GP vs specialist) or the distinction between emergency and planned hospital care. They are of great value in grasping the individual patterns and mechanisms in using or forgoing healthcare, especially in relation to health status, health insurance coverage, individual preferences or the characteristics of the national healthcare systems. This information is all the more crucial as data collection for SHARE Wave 8 was conducted just prior to the outbreak of COVID-19, which resulted in periods of significant rationing and prioritisation of care. Furthermore, two additional questions provide indirect insight into general public health concerns, revealing large disparities between countries in the level of comprehension of health-related terms and concepts (health literacy) on the one hand, and preventive care (individual behaviours or "collective" prevention campaigns) on the other.

References

Alonso, J., Orfila, F., Ruigómez, A., Ferrer, M., & Antó, J. M. (1997). Unmet health care needs and mortality among Spanish elderly. American Journal of Public Health, 87(3), 365–370. https://doi.org/10.2105/ajph.87.3.365

Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J., & Crotty, K. (2011). Low health literacy and health outcomes: An updated systematic review. Annals of Internal Medicine, 155(2), 97-107. https://doi.org/10.7326/0003-4819-155-2-201107190-00005

Bostock, S., & Steptoe, A. (2012). Association between low functional health literacy and mortality in older adults: Longitudinal cohort study. BMJ, 344(e1602). https://doi.org/10.1136/bmj.e1602

Busby, J., Purdy, S., & Hollingworth, W. (2015). A systematic review of the magnitude and cause of geographic variation in unplanned hospital admission rates and length of stay for ambulatory care sensitive conditions. BMC Health Services Research, 15(324). https://doi.org/10.1186/s12913-015-0964-3

Demicheli, V., Jefferson, T., Di Pietrantonj, C. D., Ferroni, E., Thorning, S., Thomas, R. E., & Rivetti, A. (2018). Vaccines for preventing influenza in the elderly. Cochrane Database of Systematic Reviews, 2(2). https://doi.org/10.1002/14651858.CD004876.pub4

Eurostat (2020). Preventable and treatable mortality statistics. Eurostat Statistics Explained. Available at: https://ec.europa.eu/eurostat/statistics-explained/index.php/Preventable_and_treatable_mortality_statistics

Gibson, G., Grignon, M., Hurley, J., & Wang, L. (2019). Here comes the SUN: Self-assessed unmet need, worsening health outcomes, and health care inequity. Health Economics, 28(6), 727–735. https://doi.org/10.1002/hec.3877

Ko, H. (2016). Unmet healthcare needs and health status: Panel evidence from Korea. Health Policy, 120(6), 646–653. https://doi.org/10.1016/j.healthpol.2016.04.005

Nutbeam, D. (2000). Health literacy as a public health goal: A challenge for contemporary health education and communication strategies into the 21st century. Health Promotion International, 15(3), 259–267. https://doi.org/10.1093/heapro/15.3.259

Paasche-Orlow, M. K., Parker, R. M., Gazmararian, J. A., Nielsen-Bohlman, L. T., & Rudd, R. R. (2005). The prevalence of limited health literacy. Journal of General Internal Medicine, 20(2), 175–184. https://doi.org/10.1111/j.1525-1497.2005.40245.x

Reed, R. L., Isherwood, L., & Ben-Tovim, D. (2015). Why do older people with multi-morbidity experience unplanned hospital admissions from the community: A root cause analysis. BMC Health Services Research, 15(525). https://doi.org/10.1186/s12913-015-1170-z

Williams, M. V., Parker, R. M., Baker, D. W., Parikh, N. S., Pitkin, K., Coates, W. C., & Nurss, J. R (1995). Inadequate functional health literacy among patients at two public hospitals. Journal of the American Medical Association, 274(21), 1677-1682. https://doi.org/10.1001/jama.1995.03530210031026

Wolf, M. S., Curtis, L. M., Wilson, E. A. H., Revelle, W., Waite, K. R., Smith, S. G., Weintraub, S., Borosh, B., Rapp, D. N., Park, D. C., Deary, I. C., & Baker, D. W. (2012). Literacy, cognitive function, and health: Results of the LitCog Study. Journal of General Internal Medicine, 27(10), 1300–1307. https://doi.org/10.1007/s11606-012-2079-4

3.3 Cognitive Measures

Melanie Wagner and Salima Douhou – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC)

3.3.1 Integration of New Cognitive Measures into SHARE

A great innovation was the integration of new cognitive measures into the SHARE questionnaire. Several competing goals had to be traded off against each other for the selection of tests. The first goal was to harmonise with the HRS sister study in the United States, which is why most tests were selected from that sister study. Second, we wanted to cover different cognitive domains and especially those that were not covered yet with existing tests. Third, the tests needed to be easy to roll out in the SHARE countries. This means that they were neither copyrighted tests nor language dependent. Fourth, the tests should not be too long. Finally, but importantly, some of the tests will be used for a new study in SHARE – the Harmonized Cognitive Assessment Protocol (SHARE-HCAP, which will be conducted on the back of Wave 9). In order to select the most relevant tests, we consulted the SHARE-HCAP project advisory board, which is a group of specialists from neuroscience, gerontology, survey methodology and cognitive science, who ranked the tests according to their relevance for SHARE and SHARE-HCAP.

After pretesting, comprehensive quality checks were conducted. Qualitative interviewer feedback was collected via the survey agencies during their debriefing sessions. Interviewers were asked if they had problems with administrating the questions, how the respondents perceived the new tests and if any difficulties in understanding occurred. Next, duration and variance in duration were compared among countries to detect possible difficulties in understanding the items. For countries with more than one language, the analyses were done at the language level. Also, data analysis was conducted to check for inconsistencies in response behaviour among countries (or languages) like distribution of answers and item non-response. Whenever possible, the results were compared to the results in the HRS study. For the drawing exercises, described below, we contacted the test developers directly and asked for guidance on scoring. Ultimately, the following tests were selected.

3.3.2 Self-rated Memory Change

A self-report of memory change is relevant for detecting impending problems concerning memory before clinical tests can detect them. It is known that people expressing subjective memory complaints have a higher risk of developing Alzheimer's disease (Reid & MacLullich, 2006). This is especially true for people with high levels of cognitive abilities since these people would still score high on cognitive tests despite an onset of cognitive decline. However, memory complaints are also associated with mental health difficulties, such as anxiety disorders or depression (see, for example, Kindermann & Brown, 1997), which are assessed in the SHARE questionnaire as well such that these two concepts can be analysed. The wording of this question was adopted from the HRS survey. This single-item question of about 11 seconds, on average over all languages, neatly fits into the existing questions on cognitive abilities.

3.3.3 Backward Count

Backward count is a speed test under (assumed) time pressure. It measures processing speed and attention. In normal aging, there will be subtle or no decline, whereas for people with Alzheimer's disease, there will be strong decline (see, for example, Driscoll et al., 2006). The test is part of the Telephone Interview for Cognitive Status (TICS; Brandt et al., 1988). We use the HRS version, which is slightly adjusted by counting fewer numbers and by integrating the element of time pressure. The respondent is asked to count backwards as quickly as possible from the number 20 (as far as 10).

During the pretest, it became apparent that we needed to program a more intuitive start signal for starting the backward count. Apart from that, the question was well received. The test took slightly more than half a minute on average.

3.3.4 Object Naming Test

The object naming test measures semantic memory by asking the respondents to recall three simple words. There is typically no decline in semantic memory for the normal aging brain but a strong decline for patients with Alzheimer's disease (see, for example, Petersen et al., 2010; Spaan et al., 2003). The test is only given to respondents aged 65 and over, whereas the other tests are given to respondents aged 60 or older. These age cut-offs were chosen to increase the likelihood of detecting early stages of cognitive decline. The questions were taken from TICS. Lasting less than half a minute on average, this was a quick test. Pretest results

showed high refusal rates, up to 6.5 per cent, and there was high variation across countries in two of the three questions. One such question was "What do people usually use to cut paper?" During interviewer debriefings, it was discovered that in some countries paper knives are commonly used, but this was originally not considered a correct answer. The interviewer instructions were therefore changed to also count culturally appropriate answers as correct. The question "Who is the President of the United States right now?" was dropped after the first pretest as it does not measure semantic memory but rather knowledge about USA politics, which varies across countries. This question was replaced by the question from the CSI-D questionnaire (Prince et al., 2011), which has been developed for application in different cultural settings ("Where do people usually go to buy medicine?"). An adaptation of the president question to the president of each participating country was discarded due to country differences in the knowledge of the presidents or other heads of state in European countries.

3.3.5 Drawing Exercises

Drawing exercises, also known as "constructional praxis tests", measure visuospatial skills. There is only a subtle decline in the normal aging brain but a strong decline for people with Alzheimer's disease (see, for example, Iachini et al., 2009). The respondent is asked to copy two intersecting infinity loops, a three-dimensional cube, and to draw a clock face with numbers and place the hands correctly at ten past five. The drawing of the clock is a very common screening test for cognitive impairment in memory clinics and different versions of this test exist (Shulman, 2000). The test was taken from the non-licensed Addenbrooke's Cognitive Examination-III (ACE-III, Hsieh et al., 2013) battery, which is a paper-and-pencil test and therefore had to be partly adapted to the CAPI mode. As a result, the interviewers have to do the scoring right away, while in the ACE-III test, a clinician scores the drawings after the completion of the test.

It was the longest test, lasting more than two minutes on average. Interviewer debriefings revealed that the test was liked by most respondents as it was a change to the typical question-answer format. However, it was the most difficult test for the interviewers because they had to rate the quality of the drawings. After the second test run, four countries sent the drawings that had been done during the interviews to SHARE Central, which allows for an in-depth analysis of the drawings. The interviewer ratings were compared to ratings by the questionnaire development team at SHARE. The drawings for which the ratings differed most were analysed to find possible reasons for inconsistencies. At the end of this exercise, test developers were contacted to clarify the remaining open questions. They provided valuable feedback on how to score these cases correctly. These insights were then used to develop comprehensive fieldwork materials (see below).

3.3.6 Proxy Cognition Questions

Proxy (or informant) interviews are a key source of information for a better diagnosis of mild cognitive impairment and dementia in clinical settings (see, for example, Mackinnon & Mulligan, 1998). They provide information on the respondent's prior and current cognitive and physical function. The questions measure decline in memory, decline in other mental abilities (like temporal orientation, learning, decision-making, handling financial matters and reasoning) and problematic behaviour (like wandering off). The questions were taken from the HRS questionnaire. HRS used the short form of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE; Jorm, 1994, 2004). The original scale was adapted slightly by changing the time frame for changes in mental abilities from 10 years to two years. A particular strength is that the IQCODE is relatively unaffected by education and pre-morbid ability or by proficiency in the culture's dominant language (Jorm, 1994, 2004). The original IQCODE questionnaire takes 4.46 minutes (in the HRS version) and was shortened to half that before it was included in SHARE. The need to shorten the scale was mainly driven by the restriction that the respondent should not be present while the proxy answers questions about the respondent's mental abilities. Therefore, questions that capture similar constructs or questions that ask about infrequent behaviour were dropped completely. For example, a question about learning new things in general was kept, but a question on learning to use a new gadget or machine around the house was dropped. For seven of the remaining 14 questions, the follow-up questions "Is it much worse or a bit worse?" and "Is it much improved or a bit improved?" were dropped as well for time constraints. This resulted in a reduction from a five-point to a three-point response scale: Improved – Not much changed – Got worse instead of Much improved – A bit improved – Not much changed – A bit worse – Much worse. The questions took a bit more than two minutes, on average, in the end.

Fifty-six proxy respondents answered questions about the respondent's mental abilities, which was 1.6 per cent of all pretest interviews. At the beginning of the section, an interviewer instruction states that the questions should only be answered not in the presence of the respondents. The respondent had left the room in only 45 out of 56 cases during the pretest. It was also checked whether respondents answered some of the cognitive functioning questions by themselves, which was in fact the case. For example, 20 out of 56 respondents answered questions on memory (1 excellent, 1 very good, 1 good, 5 fair, 12 poor) and 10 respondents completed the animal naming test. However, the

item response rate decreased over the course of the SHARE Cognitive Functioning (CF) module, which might indicate that proxy questions were primarily asked when respondents displayed strong cognitive deficits.

3.3.7 Training

Training is essential for conducting these tests. The majority of tests were developed for use in clinical settings where they are administered by professionals (psychologists, nurses or psychiatrists). During the Train-The-Trainer (TTT) sessions, each test was explained in detail, giving background information and practical information for interviewers for correct administration and scoring. For the drawing exercises, a quiz was provided (with solutions) on correct ratings for the National Training Sessions (NTS).

3.3.8 Overall Assessment

The qualitative analysis revealed that both the interviewers and most respondents appreciated the new tests in general. Interviewers reported that respondents were stressed (mentioned twice) or frustrated (once) by the tests and that respondents with hearing difficulties, as well as older and less educated people, found it hard to do the tests. The test selection was carefully considered and only one question remains where country differences might bias results ("What do you call the kind of prickly plant that grows in the desert?").

By introducing these new cognitive measures, SHARE can help identify which interactions of biomedical and socio-economic conditions over the life course affect cognition in later life. The understanding of such life course pathways to first mild cognitive impairment and then, possibly, dementia will help in developing preventions and early interventions.

References

Brandt, J., Spencer, M., & Folstein, M. (1988). The telephone interview for cognitive status. Neuropsychiatry, Neuropsychology and Behavioral Neurology, 1(2), 111–117.

Driscoll, I., Resnick, S. M., Troncoso, J. C., An, Y., O'Brien, R., & Zonderman A. B. (2006). Impact of Alzheimer's pathology on cognitive trajectories in nondemented elderly. Annals of Neurology, 60(6), 688–695. https://doi.org/10.1002/ana.21031

Hsieh S., Schubert S., Hoon C., Mioshi E., & Hodges J. R. (2013). Validation of the Addenbrooke's Cognitive Examination III in frontotemporal dementia and Alzheimer's disease. Dementia and Geriatric Cognitive Disorders, 36, 242–250. https://doi.org/10.1159/000351671

Iachini, T., Iavarone, A., Senese, V. P., Ruotolo, F. & Ruggiero, G. (2009). Visuospatial memory in healthy elderly, AD and MCI: A review. Current Aging Science, 2(1), 43–59. https://doi.org/10.2174/1874609810902010043

Jorm, A. F. (1994). A short form of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): Development and cross-validation. Psychological Medicine, 24(1), 145–153. https://doi.org/10.1017/S003329170002691X

Jorm, A. F. (2004). The Informant Questionnaire on cognitive decline in the elderly (IQCODE): A review. International Psychogeriatrics, 16(3), 275–293. https://doi.org/10.1017/S1041610204000390

Kindermann, S. S., & Brown, G. G. (1997). Depression and memory in the elderly: A meta-analysis. Journal of Clinical and Experimental Neuropsychology, 19(5), 625–642. https://doi.org/10.1080/01688639708403749

Mackinnon, A., & Mulligan, R. (1998). Combining cognitive testing and informant report to increase accuracy in screening for dementia. The American Journal of Psychiatry, 155(11), 1529–1535. https://doi.org/10.1176/ajp.155.11.1529

Petersen R. C., Aisen P. S., Beckett L. A., Donohue M. C., Gamst A. C., Harvey D. J., Jack C. R. Jr, Jagust W. J., Shaw L. M., Toga A. W., Trojanowski J. Q., & Weiner M. W. (2010). Alzheimer's Disease Neuroimaging Initiative (ADNI): Clinical characterization. Neurology, 74(3), 201–209. https://doi.org/10.1212/WNL.0b013e3181cb3e25

Prince M., Acosta D., Ferri C. P., Guerra M., Huang Y., Jacob K. S., Llibre Rodriguez, J. J., Salas A., Sosa, A. L., Williams, J. D., & Hall, K. S. (2011). A brief dementia screener suitable for use by non-specialists in resource poor settings – the cross-cultural derivation and validation of the brief Community Screening Instrument for Dementia. International Journal of Geriatric Psychiatry, 26(9), 899–907. https://doi.org/10.1002/gps.2622

Reid, L. M., & MacLullich, A. M. J. (2006). Subjective memory complaints and cognitive impairment of older people. Dementia and Geriatric Cognitive Disorders, 22, 471–485. https://doi.org/10.1159/000096295

Shulman, K. I. (2000). Clock-drawing: Is it the ideal cognitive screening test? International Journal of Geriatric Psychiatry, 15(6), 548–561. https://doi.org/10.1002/1099-1166(200006)15:6<548::AID-GPS242>3.0.CO;2-U

Spaan, P. E. J., Raaijmakers, J. G. W. & Jonker, C. (2003). Alzheimer's disease versus normal ageing: A review of the efficiency of clinical and experimental memory measures. Journal of Clinical and Experimental Neuropsychology, 25(2), 216–233. https://doi.org/10.1076/jcen.25.2.216.13638

3.4 Saving Regret

Axel Börsch-Supan and Irene Ferrari – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC)

3.4.1 Motivation

In many countries, past savings represent a large proportion of retirement income. Past saving decisions are therefore a significant determinant of economic well-being in retirement. How to plan for retirement and how much to save for it is difficult for most people. Managing intertemporal decisions – how much to consume today vs how much to save for the future – is complex and aggravated by the inherent uncertainties, for example, about health events that require additional and potentially expensive care or about income changes around retirement (Beshears et al., 2013; McCarthy, 2011; Strömbäck et al., 2017). People may end up feeling that they may have saved too little or too much, and thus voice regret over their past saving and spending decisions. This is the topic of this module: saving regret is the wish in hindsight that one had saved differently earlier in life.

The extent of saving regret is interesting for economists who want to assess how good people are at making intertemporal decisions. It is a fundamental challenge for social policy since the welfare state wants to strike an optimal balance between self-support by its citizens and social programmes such as public pensions, healthcare and long-term care insurance (Laibson, 1998; Madrian & Shea, 2001). Data on saving regret are also important for psychologists since saving decisions are not only driven by sober economic thinking but also by optimism, cheerfulness and self-confidence (Barberis & Thaler, 2003; Shepperd et al., 2017; Weinstein, 1980). Many people overestimate their ability to control things and underestimate the probability of negative life events such as unemployment, bad health or divorce. Moreover, even the smartest of decisions does not help if it is not properly carried out. Saving over a life course requires perseverance, but often people procrastinate – and may thus regret their failures later in life. The aim of the saving regret module is to measure the extent of saving regret in a way that exploits the SHARE data from earlier waves and the life histories in order to link insufficient or excessive saving to its potential causes.

3.4.2 Content

The SHARE Saving Regret (SR) module is short and has three sections. It is modelled after the saving regret module in the American Life Panel (Börsch-Supan et al., 2018) and administered to respondents aged between 60 and 79 addressing the hindsight view of the questions. The first section asks about life course shocks in a systematic way, complementing the life history data from Waves 3 and 7. It distinguishes positive and negative shocks, each relating to the respondent's household finances. Positive shocks include, for example, unexpected pay rises, an inheritance, luck in business, or investment decisions. Negative shocks include unemployment, divorce, bad health, business or investment failures.

Only then, in a second section, do we ask whether the respondent wished, thinking back to when they were about 45 years old, that they had saved more or less. This sequence is important in order not to bias the answers about positive and/or negative shocks.

Finally, in the third section, we want to test whether respondents were serious when they said that they wished they had saved more. We probe the respondents and ask them which categories they wish they had spent less on. The rationale here is that one can only save more if one spends less. At this point, we allow respondents to retract their earlier answer and correct themselves, or name one or several spending items (e.g. food, clothing, holidays, cars or education).

3.4.3 Conclusions

The SR module is targeted at the intersection of economics and psychology. It is designed to shed light on one of the most important and at the same time difficult long-term economic decisions and its systematic failures. Saving decisions are made in an institutional environment created by the welfare state and private markets – public pensions, healthcare and long-term care but also private pensions, potentially very large out-of-pocket costs for health and care, and private insurance. Here, the international variation among the 28 SHARE countries comes into play. The data from this module will be very informative regarding how the SHARE countries compare in providing an institutional environment that maximises the economic well-being of its citizens and minimises regret over past private saving decisions.

The data will be linked to SHARE life histories, to SHARE data on financial knowledge and cognitive skills, to SHARE data on personality traits, and to data on the income and wealth situation in which the respondents found themselves after their retirement.

References

Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. Handbook of the Economics of Finance, 1, 1053–1128. https://doi.org/10.1016/S1574-0102(03)01027-6

Beshears, J., Choi, J., Laibson, D., & Madrian, B. C. (2013). Simplification and saving. Journal of Economic Behavior and Organization, 95, 130–145. https://doi.org/10.1016/j.jebo.2012.03.007

Börsch-Supan, A., Bucher-Koenen, T., Hurd, M. D., & Rohwedder, S. (2018). Saving regret. NBER Working Paper No. w25238.

Laibson, D. (1998). Life-cycle consumption and hyperbolic discount functions. European Economic Review, 42(3–5), 861–871. https://doi.org/10.1016/S0014-2921(97)00132-3

McCarthy, Y. (2011). Behavioural characteristics and financial distress. ECB Working Paper No. 1303.

Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: Inertia in 401(k) Participation and savings behavior. The Quarterly Journal of Economics, 116(4), 1149–1187. https://doi.org/10.1162/003355301753265543

Shepperd, J. A., Pogge, G., & Howell, J. L. (2017). Assessing the consequences of unrealistic optimism: Challenges and Recommendations. Consciousness and Cognition, 50, 69–78. https://doi.org/10.1016/j.concog.2016.07.004

Strömbäck, C., Lind, T., Skagerlund, K., Västfjäll, D., & Tinghög, G. (2017). Does self-control predict financial behavior and financial well-being? Journal of Behavioral and Experimental Finance, 14, 30-38. https://doi.org/10.1016/j.jbef.2017.04.002

Weinstein, N. D. (1980). Unrealistic optimism about future life events. Journal of Personality and Social Psychology, 39(5), 806-820. https://doi.org/10.1037/0022-3514.39.5.806

3.5 Time Expenditure

Annette Scherpenzeel – Netherlands Institute for Health Services Research (Nivel) and Jeny Tony Philip – Munich Center for the Economics of Aging (MEA) at the Max-Planck-Institute for Social Law and Social Policy (MPISOC)

3.5.1 Introduction

SHARE Wave 8 included a Time Expenditure (TE) module as an add-on module. Time expenditure refers to activities as diverse as cooking, shopping, watching TV, volunteering for charitable work, helping other people, caring for grandchildren, etc. SHARE data on time expenditure offer valuable insights into how active the elderly still are, how much they contribute to paid and unpaid work and at what ages important changes in the activity patterns occur. Moreover, the data make it possible to analyse how time expenditure patterns before and after retirement vary among different countries in Europe.

The purpose of the SHARE Wave 8 TE module was, therefore, to measure how much time people aged over 50 in different countries spent on different activities and to keep track of changes in their time expenditure, after retirement or after changes in their living situation or health. In the following, we first describe what method for measuring time expenditure was chosen for the SHARE module. Afterwards, the considerations underpinning this choice and the design aspects are taken into account.

3.5.2 Method

A number of methods have been developed in survey research to measure people's time expenditure, each of which result in very different data and level of detail, and serve very different aims. The most frequently used measures we distinguish are:

1. **Time use surveys,** such as the Harmonised European Time Use Survey (HETUS; Eurostat, 2020) or the American Time Use Survey (ATUS; United States Department of Labor, 2011). These surveys use diaries that respondents fill in for one or two days, or for one week. The diaries consist of episodes of one and the same activity, varying in time, which respondents describe in free text and which are coded afterwards.

2. **Kahneman's Day Reconstruction Method** (Kahneman et al., 2004), which asks respondents to look back at the past or foregoing day and report all activities they did that day in a questionnaire. The respondents are asked to define their own episodes of one and the same activity. The activity questions are sometimes accompanied by experience sampling questions about emotions in the same episodes (i.e. questions to record feelings and activities at randomly selected moments or at predetermined times; see Scollon et al., 2009).

3. **Time expenditure modules,** implemented in panel surveys such as the Socio-Economic Panel (SOEP; Goebel et al., 2019), *European Quality of Life Survey* (EQLS; Eurofound, 2017), Longitudinal Internet Studies for the Social Sciences (LISS; Scherpenzeel & Das, 2010) and Consumption and Activities Mail Survey (CAMS; Health and Retirement Study, 2019). In these survey modules, respondents are requested to fill in the number of hours they spent on a range of general activity categories, during the last week or during a normal day.

A true time use diary (method one above) is inappropriate for a CAPI interview. It could be implemented in a paper-and-pencil drop-off, but that format poses a very large burden for respondents.

The DRM method (method two above) was in fact tested in SHARE Wave 5 as a drop-off for a small subselection of respondents. However, this test clearly showed that this method, having many repetitive questions for each generated episode, is very time-consuming and also burdensome and annoying for respondents (Binswanger, 2014). Furthermore, the pretests indicated that many SHARE respondents had difficulty understanding the concept of episodes and generating such episodes themselves (Binswanger, 2014).

A Time Expenditure module (method three above) had been implemented in many panel studies before and had thus proven to be feasible for use in a large survey. Hence, it was the best choice for SHARE.

As such, in SHARE, we included the Time Expenditure questionnaire module as it is used in many other panel surveys, implementing it as a CAPI add-on module in Wave 8. Such a module is relatively simple to implement in CAPI. It generally consists of a list of predefined activity categories that can be shown in a tabular format on one screen, together with a question text asking how much time was spent on different activities and a single standardised answering scale for all activities (see Figure 3.5 below for an illustration of the question format).

Figure 3.5: Implementation of a Time Expenditure Question in the SHARE Wave 8 Instrument

3.5.3 Design Aspects

Some additional design choices then had to be made, as not all panel studies had implemented the time expenditure method in exactly the same way. These choices concerned: 1) the reference period to ask about (e.g. time spent on one day, two days or in a week); 2) the level of detail for reporting the time spent (hours, minutes); and 3) the list of activities people spend time on. In summary, the following choices were made:

1. The reference period selected to ask about was one day, defined as "yesterday" in the interview.

2. The time spent was considered in hours and minutes.

3. The activity lists as used in the LISS panel (Scherpenzeel & Das, 2010) and in CAMS (Health and Retirement Study, 2019) were chosen as the basis on which the SHARE list was developed.

We will now briefly describe each of these three design aspects and the arguments behind our choices. As described above, in each of the panel studies that include a Time Expenditure module, respondents are requested to fill in the time they spent on a range of general activity categories. However, the reference period varies across the studies: while CAMS (Health and Retirement Study, 2019) and LISS (Scherpenzeel & Das, 2010) ask about the past seven days, SOEP (Goebel et al., 2019) asks about a "normal" or "typical" day. The method of asking about the past seven days requires respondents to retrieve all seven days of the week for each of the activities in the list, and perform a mental calculation of the time spent on each day in an overall sum. This is likely to be easier to do in an online interview among the general population. For the SHARE CAPI interview, it will take too much time and may lead to imprecision or memory problems for part of the SHARE target population of respondents aged over 50 (see, e.g., Hurd & Rohwedder, 2008). Therefore, we have chosen to ask about the time spent on one day only, and decided to take yesterday as the reference day. This is in concordance with Kahneman's DRM and should hence lead to a more reliable day reconstruction from memory (Kahneman et al., 2004).

Since time expenditure may substantially differ between weekdays and weekend days for working people, we considered repeating the whole list of activities again for one more day: a randomly selected weekday or weekend day, depending on what day yesterday was. However, this made the Time Expenditure module not only much too long for an add-on module but also very repetitive and annoying for respondents. Therefore, the design was restricted to only yesterday as reference.

A point of criticism of CAMS and SOEP has been that they ask only about entire hours spent on different activities. Therefore, van Soest and Vermeulen (2008) proposed asking about both hours and minutes in the LISS time expenditure module. Since Cherchye et al. (2012) have shown that this worked well in practice and resulted in useful data, this choice was also made for the SHARE module.

The activity lists used in CAMS and LISS are similar. The list used in SOEP is shorter, less detailed and constitutes very broad categories with little explanation or definition. Since the data obtained with the former list have resulted in some outstanding publications (e.g. Cherchye et al., 2012; see above), we considered this list accurate and sensitive enough for the purpose of observing changes in time expenditure patterns over time.

During the questionnaire development stage and after the pretest in the field, the original list as well as the definition of certain activities were adapted a few times, to make the list of activities more comprehensible and complete. These changes are described in the following section.

3.5.4 Question Format and List of Activities

The Time Expenditure module was placed after the Activities (AC) module in SHARE and asked only to panel respondents not resident in nursing homes. On the basis of the pretest results, the following changes in form and content of the Time Expenditure module were made before the fielding of the main study:

1. Adjustment of the routing to exclude vulnerable respondents like those in nursing homes and new respondents. The intention here was to make the interview less burdensome.

2. Three questions, one on time spent on further schooling, the others on yardwork/gardening and praying/meditation, were dropped to reduce the interview length. In addition, further schooling was considered a non-in-

formative item since almost none of the pretest respondents had spent any time on it. The items yardwork/gardening and praying/meditation were not part of the original list of Cherchye et al. (2012) and were therefore left out when the module length had to be shortened.

A transcript of the final questionnaire as fielded in SHARE Wave 8 is appended below:

TE001_intro
The following questions are about how you spent your time yesterday.
1. Continue

TE002_Weekday
Do Not Read out.
IWER: Please note what day YESTERDAY was.

1. Monday
2. Tuesday
3. Wednesday
4. Thursday
5. Friday
6. Saturday
7. Sunday

TE003_YesterdaySpecial
Please think about **YESTERDAY**, which was [<Fill in day of the week from te002>], from the morning until the end of the day. Think about where you were, what you were doing, who you were with and how you felt. Was yesterday a normal day for you or did something unusual, bad or good happen?
IWER: Read out
1. Yes – just a normal day
2. No – my day included unusual bad or stressful things
3. No – my day included unusual good things

TE004_Chores_INTRO
Continue to think about yesterday, from the morning until the end of the day, and the amount of time you spent on diverse activities over the course of the day.
How much time did you spend yesterday on household chores like cleaning, laundry, shopping, cooking, gardening, etc.? Please do NOT include personal care or care for children, parents or other family members.
IWER: If respondent is not sure, then ask him/her to estimate the amount of time as best as he/she can.
If respondent did not spend any time on a certain activity, enter 0 in both fields.
If respondent spent for example an hour and a half on a certain activity, then enter 1 hour and 30 minutes.
If respondent spent 40 minutes on a certain activity, then enter 0 hours and 40 minutes.

TE005_Chores_Hrs
Hours ☐

TE006_Chores_Mts
Minutes ☐

TE010_PersonalCare_Intro
How much time did you spend yesterday on personal care, such as washing, dressing, visiting the hairdresser, seeing the doctor, etc.?

TE011_PersonalCare_Hrs
Hours ☐

TE012_PersonalCare_Mts
Minutes ☐

TE013_Children_Intro
How much time did you spend yesterday on activities with your children, grandchildren, children you baby-sit or any other children you look after?
This can include washing, dressing, playing, taking to school/other activities, helping with homework etc.
IWER: Please exclude adult children.

TE014_Children_Hrs
Hours ☐

TE015_Children_Mts
Minutes ☐

TE016_HelpParents_Intro
How much time did you spend yesterday on helping your **parents** or **parents-in-law**?
This can include assistance with administrative chores, washing, dressing, taking them to see the doctor etc.
IWER: Please include time spent with step parents and adoptive parents too.

TE017_HelpParents_Hrs.
Hours ☐

TE018_HelpParents_Mts
Minutes ☐

TE019_HelpPartner_Intro[5]
How much time did you spend yesterday on helping **[your husband/ your wife/ your partner/ your partner]?**
This can include assistance with administrative chores, washing, dressing, taking [him/her] to see the doctor etc.

TE020_HelpPartner_Hrs
Hours ☐

TE021_HelpPartner_Mts
Minutes ☐

TE022_HelpOther_IntroTE022
How much time did you spend yesterday on **helping other family or other people you know?**
DO NOT include helping [your husband/ your wife/ your partner/ your partner] or parents and kids that you have already mentioned here.
IWER: If necessary repeat: for instance assistance with administrative chores, washing, dressing, taking someone to see the doctor, etc.

TE023_HelpOther_Hrs
Hours ☐

TE023_HelpOther_Hrs
Minutes ☐

TE025_Leisure_Intro
How much time did you spend yesterday on leisure time activities?
This can include watching TV, social media, sports, hobbies, talking with friends or family, going out etc.

TE026_Leisure_Hrs
Hours ☐

TE027_Leisure_Mts
Minutes ☐

5 Filter exists so that only those with partner (either within or outside the household) get this question.

TE031_Admin_Intro
How much time did you spend yesterday on administrative chores and own family finances?

TE032_Admin_Hrs
Hours ☐

TE033_Admin_Mts
Minutes ☐

TE034_PaidWork_Intro
How much time did you spend yesterday on paid work? Paid work can be in employment or as self-employed. Please, do NOT include the time spent traveling to and from work, but do count overtime hours.

TE035_PaidWork_Hrs
Hours ☐

TE036_PaidWork_Mts
Minutes ☐

TE037_VoluntaryWork_Intro
How much time did you spend yesterday on voluntary work?
Please, do NOT include household chores, helping family members, care for children, and other activities you have already just mentioned.
IWER: Examples are voluntary work for religious, educational, political, health-related or other charitable organisations

TE038_VoluntaryWork_Hrs
Hours ☐

TE039_VoluntaryWork_Mts
Minutes ☐

TE040_Travel_Intro[6]
Continue to think about yesterday, from the morning until the end of the day.
How much time did you spend yesterday on traveling to and from work or voluntary work?
IWER: Enter zero if the respondent did not work on the previous day.

TE041_Travel_Hrs
Hours ☐

TE042_Travel_Mts
Minutes ☐

TE046_Napping_Intro
How much time did you spend yesterday on napping and resting during daytime? Do not include sleeping at night time.

TE047_Napping_Hrs
Hours ☐

TE048_Napping_Mts
Minutes ☐

TE049_Sleeping_Intro
How much time did you spend yesterday on sleeping at night time?

TE050_Sleeping_Hrs
Hours ☐

TE051_Sleeping_Mts
Minutes ☐

6 Only respondents who do work (paid or voluntary) get this question.

TE052_OtherActivities
Did you spend time yesterday on other activities which we have not asked about yet?
1. Yes
2. No
[IF TE020_OtherActivities=1]

TE053_WhatActivities
What other activity was that or what other activities were those?

TE054_TimeOtherActivities_Intro
[IF TE020_OtherActivities=1]
How much time did you spend yesterday on this activity or these activities?
IWER: If more than one other activity was mentioned, sum up the time spent on each of these other activities.

TE055_TimeOtherActivities_Hrs
Hours ☐

TE056_TimeOtherActivities_Mts
Minutes ☐

[if partner and if te025_Leisure >0 hours+minutes]

TE057_PartnerActivities_Intro
You mentioned that you spent *[fill in hours given at question te013]* hours and *[fill in minutes given at question te013]* minutes on leisure time activities, yesterday.
How much of that time did you spend together with your partner? If none at all, please enter 0.
IWER: If respondent did not spend any time on leisure activities together with partner, enter 0.

TE058_PartnerActivities_Hrs
Hours ☐

TE059_PartnerActivities_Mts
Minutes ☐

TE060_IntCheck
IWER:
CHECK:
Who answered the questions in this section?
1. Respondent only
2. Respondent and proxy
3. Proxy only

References

Binswanger, J. (2014). DRM. SHARE Meeting Krakow. Munich Center for the Economics of Ageing (MEA).

Cherchye, L., De Rock, B., & Vermeulen, F. (2012). Married with children: A collective labor supply model with detailed time use and intrahousehold expenditure information. The American Economic Review, 102(7), 3377–3405. https://www.jstor.org/stable/41724638

Eurofound (2017). European Quality of Life Surveys (EQLS). Available at: https://www.eurofound.europa.eu/surveys/european-quality-of-life-surveys

Eurostat (2020). Harmonised European Time Use Surveys (HETUS). Available at: https://ec.europa.eu/eurostat/web/time-use-surveys/overview

Goebel, J., Grabka, M. M., Liebig, S., Kroh, M., Richter, D., Schröder, C., & Schupp, J. (2019). The German Socio-Economic Panel (SOEP). Jahrbücher für Nationalökonomie und Statistik, 239(2), 345–360. https://doi.org/10.1515/jbnst-2018-0022

Health and Retirement Study. (2019). RAND HRS CAMS Data 2017 (V1). Available at: https://hrsdata.isr.umich.edu/data-products/rand-hrs-cams-data-2017

Hurd, M. D., & Rohwedder, S. (2008). Time use in the older population: Variation by socioeconomic status and health. In R. F. Belli, F. P. Stafford & D. F. Alwin (Eds.), Calendar and time diary methods in life course research (pp. 207-224). Sage Publications, Inc.

Kahneman, D., Krueger, A. B., Schkade, D., Schwarz, N., & Stone, A. (2004). The Day Reconstruction Method (DRM): Instrument documentation. Available at: https://dornsife.usc.edu/assets/sites/780/docs/drm_documentation_july_2004.pdf

Scollon, C. N., Kim-Prieto,C., & Diener, E. (2009). Experience sampling: Promises and pitfalls, strengths and weaknesses. In E. Diener (Ed.), Assessing well-being. (pp. 157–180). Springer.

Scherpenzeel, A., & Das, M. (2010). 'True' longitudinal and probability-based internet panels: Evidence from the Netherlands. In M. Das, P. Ester & L. Kaczmirek (Eds.), Social and behavioral research and the Internet: Advances in applied methods and research strategies (pp. 77–103). Taylor & Francis.

United States Department of Labor. Bureau of Labor Statistics (2011). American Time Use Survey (ATUS). Inter-university Consortium for Political and Social Research. https://doi.org/10.3886/ICPSR34453.v1

Van Soest, A. & Vermeulen, F. (2008). Proposal for a time use and consumption module in LISS.

3.6 Asking for Participation in the SHARE Accelerometer Study

Annette Scherpenzeel – Netherlands Institute for Health Services Research (Nivel), Nora Angleys, Fabio Franzese and Luzia Weiss – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC)

SHARE Wave 8 comprised the accelerometer study, which was aimed at collecting measurements of physical activity based on sensor data (for a detailed description see Chapter 12). In ten countries, a subsample of the panel respondents was asked to wear an accelerometer on their upper thigh for eight consecutive days. Eligible respondents for the accelerometer study were randomly drawn before fieldwork of Wave 8 and the related CAPI module was routed accordingly. This Accelerometer (AX) module was designed to ask eligible respondents for their consent to participate in the accelerometer study during the regular SHARE Wave 8 interview.

The AX module started with an introduction to the SHARE accelerometer project, describing the aim and mode of the study. It was explained by the interviewer that participation is voluntary and only movement, and not location, is recorded (i.e. the device has no GPS tracking). The respondent was shown a picture and the interviewer illustrated the small size of the device. Occasionally interviewers used little wooden toy blocks that were similar in shape, size and weight to the accelerometer device to provide the respondents with an actual model of the device. Additionally, an instruction booklet was shown that provided visual and written instructions on how to attach the device to the upper thigh. Furthermore, the booklet specified in detail what the respondents should expect, if they were to agree to participate in the study. The respondent was then asked for their consent to participate in the accelerometer study *(AX002)*. If the respondent was not willing to participate, the reason for refusal was noted in the CAPI *(AX003)*. The refusal reasons that interviewers could select were based on the most frequently given answers in the test runs (pretest and field rehearsal), which included an open-ended question. The categories in the questionnaire were:

- Respondent believes he/she is not active enough.
- Respondent considers him/herself too old.
- Respondent cannot participate because of work or hobby.
- Respondent will be absent for a longer time.
- Respondent considers it too complicated or too burdensome.
- Respondent considers it a violation of privacy.
- Respondent prefers not to participate because of allergy/sensitive skin.
- Other (specify).

In addition, supplementary information screens were developed, which were presented following the coding of the refusal. The additional information shown on the follow-up screen was tailored to the refusal reason the respondent had selected. For example, when a respondent selected that they believed they were not active enough to participate in such a study, the additional information on the interviewer's screen explained why it is important that both active and not active persons are represented in the study (for all respondents' concerns and respective interviewers' clarification, see Table 3.1). The screens were introduced in the field rehearsal (second test run) after many interviewers reported the respondents' uncertainty when confronted with this new approach to data collection in the pretest. The information given was aimed at preventing refusals that were based on misunderstanding or false beliefs, as well as enabling respondents to make an informed decision. Interviewers would not pressure or persuade respondents to participate in the measurement. This approach had been successfully used before in the LISS panel accelerometer study (Scherpenzeel, 2017). After the information was read out, the respondent was asked again if they wished to participate.

If a respondent was willing to participate, the interviewer explained that due to a limited number of devices it might take several weeks for the respondent to receive a device. The limited number also meant that not every respondent could be sent a device during the fieldwork period, which means, effectively, that consent did not guarantee selection for participation.

Table 3.1: Concerns and Corresponding Clarifications on the Interviewers' Information Screen

Concern	Clarification
Respondent believes he/she is not active enough	I understand that you consider yourself not active enough. For this study, it is important that not only active people participate but also people who are not so active or even disabled. Only then can the researchers get a complete picture of health and activity in the [country's] population. If we included only active persons, the health and activity of the ageing population would appear better than it actually is. Your participation in the study is hence very valuable.
Respondent considers him/herself too old	I understand that you consider yourself too old. For this study, it is important that not only younger persons participate but also older persons. Only then can the researchers get a complete picture of health and activity in the [country's] population. If we included only healthy and younger persons, the health and activity of the ageing population would appear better than it actually is. Your participation in the study is hence very valuable.
Respondent cannot participate because of work or hobby	I understand that you consider the device to be limiting in pursuing your work or hobbies. The device is very small, lightweight and easy to wear. You do not have to worry about it during your daily activities. You don't have to turn it on or off, it works by itself. Most people who wear it almost forget they are wearing it.
Respondent will be absent for a longer time	You indicated that you will be absent for a longer time. You only have to wear the accelerometer for eight consecutive days. Only if you will be absent all the time until June 2020 can you not participate. If you will be absent for a few weeks, for example for holidays or work, we can take that into account and send you the accelerometer another week.
Respondent considers it too complicated or too burdensome	I understand that you consider participation too complicated or burdensome. The only thing you have to do is to wear the accelerometer for eight consecutive days. It is very small and lightweight, easy to attach and you do not have to worry about it. You don't have to turn it on or off, or do anything else with it while wearing it, it works by itself. Most people who wear it almost forget they are wearing it. After the eight days, you can simply return it to us in the envelope that comes with it. You will get [country incentive] for your participation.
Respondent considers it a violation of privacy	If I understand correctly, you think that wearing the device would give access to your private and sensitive data. The device is not GPS-enabled, so your exact location cannot be recorded. The device simply measures the number and intensity of your movements and nothing else. It is not possible to track your whereabouts or to measure any other information.
Respondent prefers not to participate because of allergy/ sensitive skin	If I understand correctly, you think the tape with which you attach the device will damage your skin or provoke an allergic reaction. The tape is specially designed for longer wear time on the skin. It is used for medical purposes where people have to wear devices directly on the skin for a longer period than eight days. The risk of an allergic reaction is therefore very small.

References

Scherpenzeel, A. (2017). Mixing online panel data collection with innovative methods. In S. Eifler & F. Faulbaum (Eds.), Methodische Probleme von Mixed-Mode-Ansätzen in der Umfrageforschung (pp. 27–49). Springer Fachmedien. https://doi.org/10.1007/978-3-658-15834-7_2

CHAPTER 4

Software Innovations: Architecture Upgrade

04

4 SOFTWARE INNOVATIONS: ARCHITECTURE UPGRADE

Maurice Martens and Iggy van der Wielen – CentERdata

4.1 Introduction

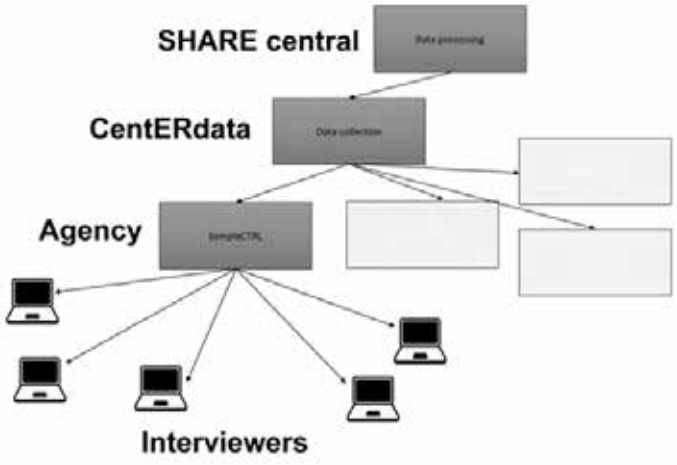
To bring the SHARE software environment up to a level needed to support the challenges Wave 8 presented, major changes were implemented. The questionnaire software migrated to a completely new system. This change triggered various updates to all existing systems and forced us to replace several tools. This chapter will discuss the software and migration.

4.2 Questionnaire Software

Since Wave 1, CentERdata has scripted the SHARE questionnaire in Blaise 4. Over the years we developed several tools that connect to these questionnaires. For the eighth wave of SHARE, the questionnaire was migrated from Blaise version 4.8 to version 5.4. This migration forced us to adapt many of the existing tools. To understand the changes and impact, we should look into the overall software architecture that supported the SHARE study over the years.

The SHARE study is *ex ante* harmonised. We develop a single source questionnaire and define a data model that is used in all participating countries. Only the translations may differ per installation. An overview of the SHARE dataflow is depicted in the figure below.

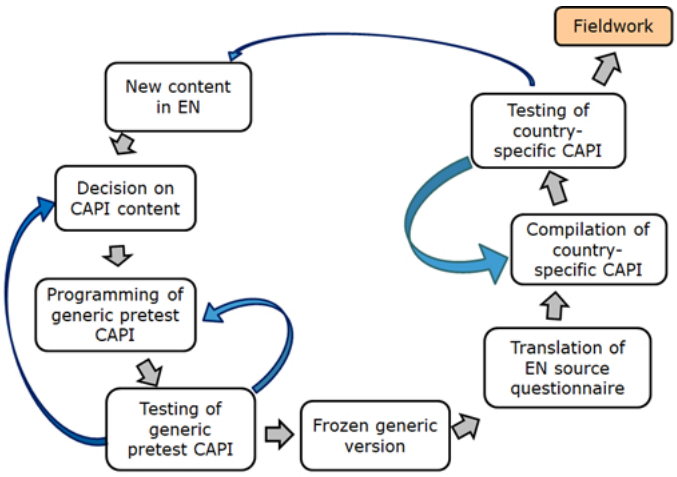
Figure 4.1: Overall Architecture



The SHARE development workflow starts with a document that describes the questionnaire definition for a new wave. CentERdata reviews this and builds a first version of the questionnaire in Blaise; this is tested. Any issues that arise are adapted in the questionnaire definition or in the programmed source tool. Improvement cycles are implemented until everybody is happy with the programmed source questionnaire. At that stage the development of the source questionnaire is frozen. Next, the questionnaire is transferred to the Translation Management Tool (TMT). The TMT is an online tool that coordinates the translation process in SHARE. It can be configured to support various translation processes, including the TRAPD (Translation, Review, Adjudication, Pretest, Documentation) process.

Up to Wave 7 of SHARE, the TMT used the Blaise API to load in a compiled Blaise source questionnaire. The import script extracts texts from the compiled questionnaire and compares them with a previous version that has already been imported before. Any changes that are detected are flagged in the translation environment. This allows translators to focus on the changed items. The translatable elements are shown in context, together with texts used in other tools, like the Sample Management System (SMS), and are translated by professional translators. During the country-specific testing phase, it can happen that we find country- or language-specific requirements that force an adaptation in the source questionnaire. If the issue is problematic, it might mean that an adaptation to the original questionnaire definition is needed. This would trigger the development loop from the start for the involved items. Ultimately, tools are generated that can be used to do fieldwork with (see Figure 4.2). SHARE has three such development cycles per wave: one for a pretest, where new items are tested, next for the field rehearsal, which is aimed at generating a final instrument to be used in the main fieldwork phase, and lastly the version for the main fieldwork phase; in general in this phase there are only bug fixes and possibly some questions removed.

Figure 4.2: Questionnaire Development Process

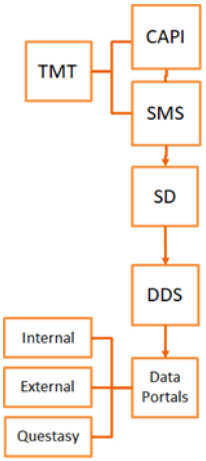


The workflow described above worked fine for SHARE over the years, however at the start of Wave 8, CentERdata and the SHARE Central coordination team decided that to safeguard the technical future of SHARE some changes were needed. While the concept worked, several of the tools that supported this were becoming out of date. For example, the version of the Blaise software platform was no longer supported, so we needed to migrate from Blaise 4.8 to a Blaise 5 version. In addition, the SMS and Sample Distributor (SD) were built on legacy software. While we were updating these systems to support this process, we decided to try to reduce the number of cycles needed in our development process. With faster and better tools, our turnaround could be quicker. With stricter version management and better scheduling the overall process could be lifted to a higher level. Before, especially on non-Latin instruments, issues were often reported where people encountered non-readable text, particularly when fills were used. This triggered many unnecessary improvement loops. So one major goal was to arrange, once and for all, for Unicode support in all tools.

We concluded that, although it was a big task, we needed to push the migration through. Luckily, the Blaise software migration was low-risk. Version 4.8 and version 5 are highly compatible, and the questionnaire script itself did not change too much. We could easily import previous wave questionnaire definitions and routing information. Many of the special features that are integrated in SHARE, however, needed an alternative solution. The new Blaise 5 XML export made it more convenient to import the translation toward the translation environment. The overall architecture remained the same, but most communication protocols and software tools were replaced. The newly developed infrastructure provides a solid basis for the coming waves of SHARE.

The communication flow of the SHARE tools up to Wave 7 is depicted in Figure 4.3. We use the TMT to load in texts into both the CAPI and the SMS. The CAPI is, in this context, the Blaise questionnaire. The CAPI communicates with the SMS and the SD, which in turn transfers to the Data Delivery System (DDS). Finally, the data are ported to the Data Portal, which has three views on it: Internal, External and Questasy.

Figure 4.3: Previous Wave Tools



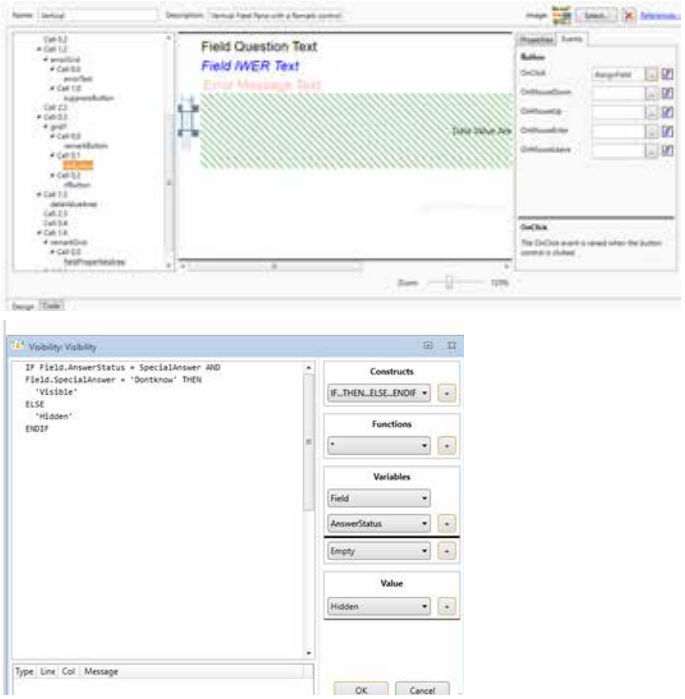
4.3 Software Migration

Many of these tools used the Blaise API or interfaced with the questionnaire's database, and might have to be adapted. We first drew up a migration plan to set up SHARE Blaise 5 CAPI, so we could estimate how feasible a migration would be:

1. Import Wave 6/Wave 7 questionnaire; several items from the Wave 6 and Wave 7 questionnaires reoccurred in the Wave 8 questionnaire. We imported them in Blaise 5.
2. Update to Wave 8 questionnaire; in the Blaise 5 environment, we adapted the imported questionnaire to the Wave 8 definition.
3. Design default interface; because the questionnaire layout as used under previous waves, with the split screen, was no longer available, we had to build an interface that still supported several "old" features, like keyboard navigation.
4. Implement non-default features; in SHARE there are some questions that need counters, show wordlists or need more complex lookup tables. In phase 4 these were addressed.
5. Export to TMT; the export to the TMT needed to be redefined, and we decided to review this once we knew which of the features would be possible or when it would be clear if there were new concepts in the questionnaire.
6. Import from TMT; in earlier waves we copy-pasted the translated texts in the source code. Ideally there would be better processes we could use this time to generate translated questionnaires.
7. Connect our SMS to the Blaise API.

The importing of the previous wave's questionnaire worked without any problems. Scripting the questionnaire worked perfectly. The new environment was very useful. It sped up development time and the background parsing, and helped in finding problems immediately. Since the interviewers were used to a split-screen interface, which had the navigation paths displayed on the bottom of the screen, we tried our best to mimic this in the Blaise 5 environment ourselves. This failed; we tried several ideas but concluded it was not feasible to implement. Another feature the interviewers liked was the icons that showed when a "Don't Know" (DK), "Refusal (RF)" or "Remark" was attached to a question (see Figure 4.4). We added these icons to a field pane and determined their visibility in the status of the field.

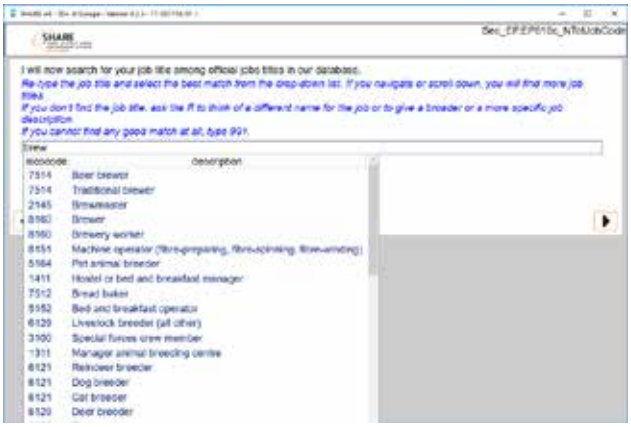
Figure 4.4: DK and RF



Several non-default features needed to be explored:

- Jobcoder: SHARE invested a lot of effort in setting up large databases of already classified job titles; in previous waves we wrote external apps to show these, so we could have full control over the algorithms and behaviour. We hoped in Blaise 5 the way lookups behaved would be improved, to allow for alternative algorithms, and we hoped to show a list that also accepts whatever is typed in, so we could detect that there was no match and still collect the response. This was unfortunately not possible; we ended up using the lookup as they are available in Blaise 5 with trigram search (see Figure 4.5), but we still hope this will be further improved.

Figure 4.5: Lookup Table



- Wordlists/counter: In SHARE, there are some cognitive measures, like a timed word recall, with questions where a timer and a stopwatch are shown. In previous waves, videos were used to display these. Unfortunately, at the time we developed the Wave 8 questionnaire, videos were not yet supported in Blaise 5. We invested quite some time in using the timer, somehow feeding an array of texts to the timer that would then change the text of a label at each tick. This, however, did not work; changing labels at runtime via an action is not possible, and this also makes it impossible to develop a feature like a stopwatch. We ended up using animated GIF files and disabled navigation (see Figure 4.6), but this solution was not ideal since the GIF restarts when enter is pressed. If only the text property of labels could be changed by clicking a button or by a tick from the timer, we would have developed this differently.

Figure 4.6: Disabled Navigation during the Word Recall Question

- Keyboard navigation: A key functionality, the complete questionnaire is in CAPI mode; the interviewers work quicker if they do not need to point and click. In version 5.4, this was not yet implemented. To get this working we introduced enumerationTextBoxes and setTextBoxes, which gets the focus when a field pane is activated; one can type in the responses and the attached checkboxes or radio buttons will be checked (see Figure 4.7). In addition, when the interviewer checks the radio buttons or checkboxes, the textboxes will show the value, as in earlier waves (see Figure 4.8).

Figure 4.7: Check for Values in Radio Buttons

Figure 4.8: Keyboard Navigation to Select Multiple Response Options

- Text roles: We like the newly introduced text role features in Blaise 5 very much. In earlier waves of SHARE this was already done implicitly, interviewer instructions and questions text were already defined as separate translatable items and during import we gave them different fonts and colours. Now we can easily set this up because this is provided natively.
- Unicode: Under Blaise 4, CentERdata developed several hacks to support the SHARE questionnaire in non-Western scripts like Arabic, Hebrew or Russian. These hacks caused us frustration and needed thorough checking, especially for fills that were cut off at 256 characters without any error message when fed through procedures. We really appreciate the full support of Unicode in Blaise 5.
- Images: Wave 8 of SHARE had some questions that used images on show cards; since the use of images was so straightforward, we also show them in the responses the interviewer sees.
- Child grid: Since the split screen as common in Blaise CAPI questionnaires in Blaise 4 is not implemented in Blaise 5, it was not convenient to present a full overview of all children in a household, or at least not one compatible with the previous questionnaire routing and field. We chose to rephrase the questions and develop a child overview on screen as a separate area in the field pane. We match children that were preloaded from a previous wave or from the responses of the partner and children we may have detected in the Social Network module, and possibly add children that are not already mentioned. In this interface, the list of children is shown on the right, and will change to green and will add a check mark when a child is confirmed; if one of the children is mentioned twice or should not be in the list, they are coloured red and a cross is added (see Figure 4.9).

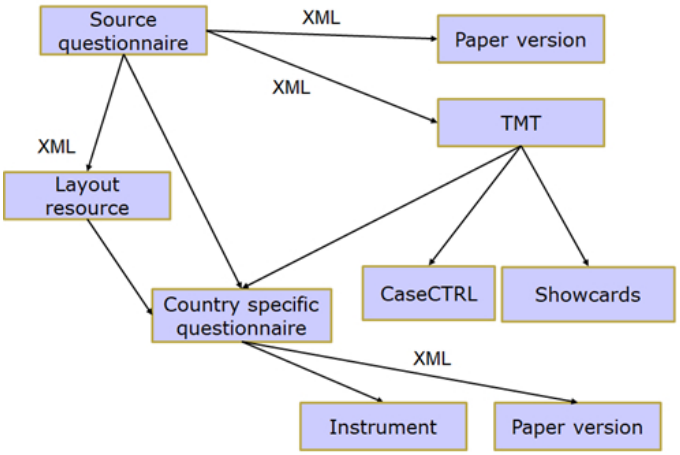
Figure 4.9: Children Overview

The audit trail functionality works differently from before. We gained better insight into loading times on the question level. This created quite a discussion; since the definition of interview duration depends on this, do we only measure the time a question is shown on screen? Or should we also include the time between when the response is submitted and the next question is shown. And to which question belongs that time. This is maybe not relevant for most surveys, but for a survey that on average takes over an hour, this definition can be quite relevant.

In earlier waves, we would use the API to walk through a compiled questionnaire and determined the questionnaire structure from that, for example to import into the translation environment. Since Blaise 5 exports XML definition, we

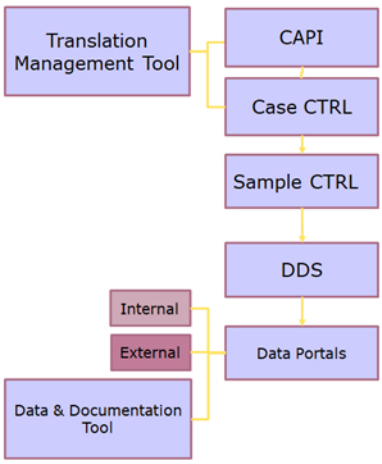
appreciate that there is no need to call the API (see Figure 4.10); simply parsing XML made our systems more stable, as did paper version interface, Stata scripts, various Excel files with metadata overviews.

Figure 4.10: Use of Metadata XML



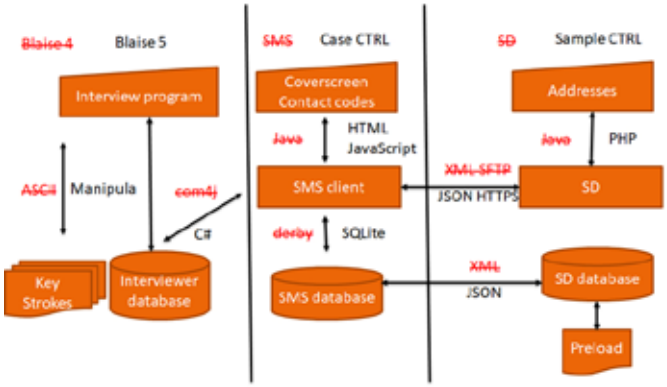
An extra feature we hoped for was to support the use of tablets for the SHARE system. Since Blaise 5 promised to support this, we knew that our SMS system would become a bottleneck in any future support. In addition, the software stack used by the SMS system was at the end of its live and not ready to be ported to native mobile operating systems. To overcome this, a new SMS was developed, under the name CaseCTRL. This lightweight program can run on any device, it has a web interface and can link to various different questionnaire engines. Linked to this we also replaced the SD in our systems and replaced it with SampleCTRL. Where SD was a single-user desktop application running on a server, SampleCTRL gives us a fully web-enabled server application. This resulted in the tools described in Figure 4.11.

Figure 4.11: SHARE Wave 8 Tools



This resulted in the following changes for the SHARE Wave 8 architecture (see figure 4.12 below):

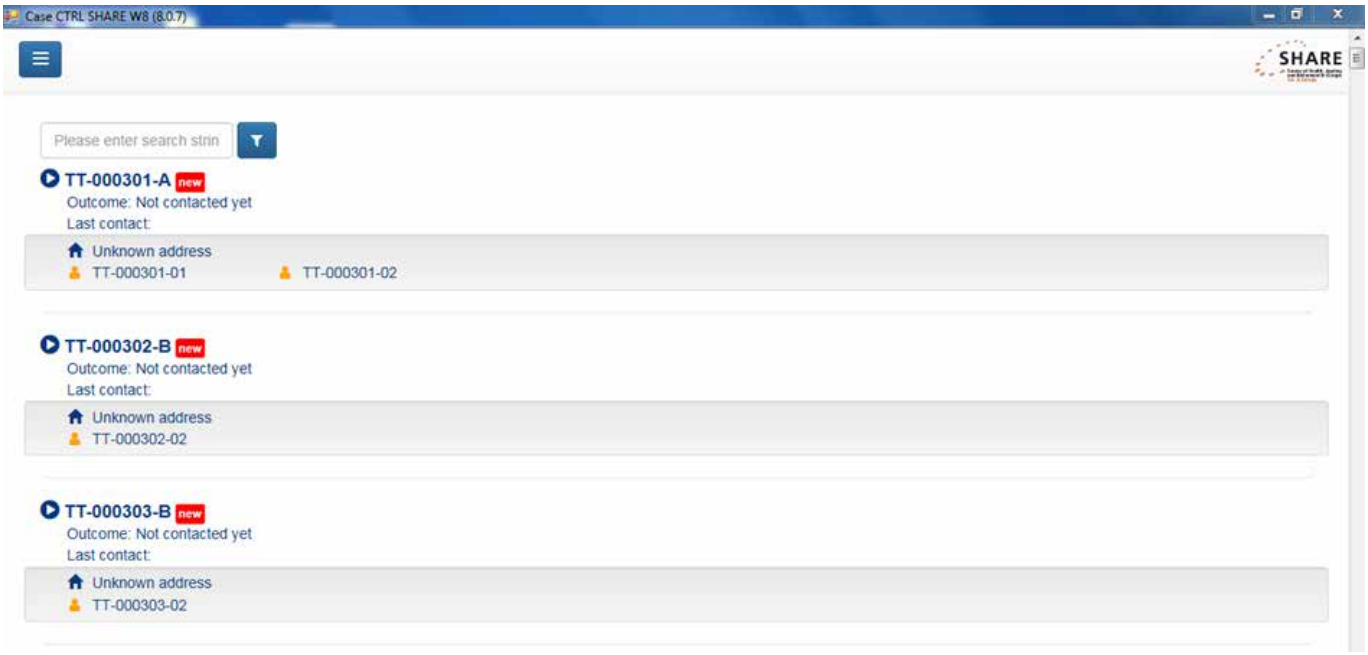
Figure 4.12: Changes for Wave 8



4.4 CaseCTRL and SampleCTRL

Before we can conduct a questionnaire, some administration is needed. We need to assign a household to an interviewer. We need to know the household composition, including possible split households, and we should arrange for an appointment or register contacts. To manage this, the case management tool CaseCTRL has been developed. An interviewer uses this to manage the local subsample as shown in Figure 4.13.

Figure 4.13: Subsample in the CaseCTRL



This tool registers contacts and contact attempts, it has a built-in agenda and runs a household grid. When the interviewer has an Internet connection, they can synchronize their collected data from CaseCTRL with an installation of SampleCTRL. SampleCTRL is an application that runs on a server at the agency that is hired to coordinate the fieldwork in a country or region. Using SampleCTRL, an agency can monitor the individual interviewers, and can assign and (re-)distribute subsamples.

They are requested to synchronize their collected data with a central server every two weeks. In this second synchronisation step, the data are anonymised and sent over to a Data Collection server. CentERdata hosts this server. In this data collection environment, the data are merged together and some basic validation is done on the data. After the data pass the validation step they are pushed to the SHARE Central server in Munich, Germany, where during fieldwork the data are processed into fieldwork progress reports and ultimately disseminated, cleaned and published.

CaseCTRL is a lightweight multimode case management system, developed by CentERdata. It is compatible with Blaise questionnaires but can also connect to various other questionnaire engines. Built as a single-based application, the CaseCTRL can be deployed as a Windows, Android or IOS application. When we needed to introduce the SHARE Corona Survey in the midst of Wave 8 fieldwork, this was quite convenient. With little effort, a CATI questionnaire could be linked to the CaseCTRL system. This will be further discussed in Chapter 9.

4.5 Concluding Remarks

The Blaise 5 environment has greatly improved compared to the Blaise 4 system. Its programming interface works much faster, which is convenient when developing a complex, long questionnaire like the SHARE questionnaire. The separation of layout and questionnaire definition is very useful. The introduction of text roles and the compatibility with Unicode solved many of the problems we had before. Thus, the metadata Blaise XML export really gave us a solid reusable definition we used in various exports in pdf, html, Excel and Stata script format. It would be perfect if from this definition a questionnaire could be built. This would allow us to generate instruments as an automatic integral part of the translation cycle.

Overall, the new CaseCTRL and SampleCTRL tools provide an up-to-date environment. This solution is flexible and scalable. It allows for any future challenges in questionnaire modes, be it offline vs online, household vs individual or self-completion vs interviewer. It also could support mixed and multimode strategies. The introduction of CaseCTRL and SampleCTRL enables the use of different questionnaire engines. With the need to quickly develop a CATI instrument, this functionality has already proven its worth. Based on the development of the CaseCTRL and SampleCTRL tools, we decided to further integrate the survey software tooling into one suite: the CTRL suite. The TMT was adopted into the CTRL suite and renamed TranslationCTRL. We decided to stick to TMT in this chapter since that name was used during the translation phase of Wave 8. We are confident that the developed architecture and tools provide a solid basis for the future of SHARE.

CHAPTER 5

Monitoring and Managing Fieldwork in SHARE Wave 8

05

5 MONITORING AND MANAGING FIELDWORK IN SHARE WAVE 8

5.1 Fieldwork Monitoring and Survey Participation in the Regular SHARE Wave 8

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

5.1.1 Introduction

As in the previous editions, the following is a continuation of the chapter about fieldwork monitoring in SHARE Waves 6 and 7 (Malter & Sand, 2017; Sand, 2019) with all numbers and statistics adapted to the countries of the regular Wave 8 until the suspension of fieldwork due to COVID-19 in March 2020 (see Section 10.1 of the CATI part regarding the SHARE Corona Survey after resuming fieldwork). Its conceptual basis was developed in the run-up to Wave 5 and is outlined in Kneip et al. (2015). As usual, all indicators were conceptualised strictly in accordance with the ninth 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[7] would be if fieldwork was terminated at that given moment. We are convinced that ensuring data quality has to be a key concern of any population-level survey study while putting the emphasis on all the major components of the Total Sampling Error, as described in Kneip et al. (2015).

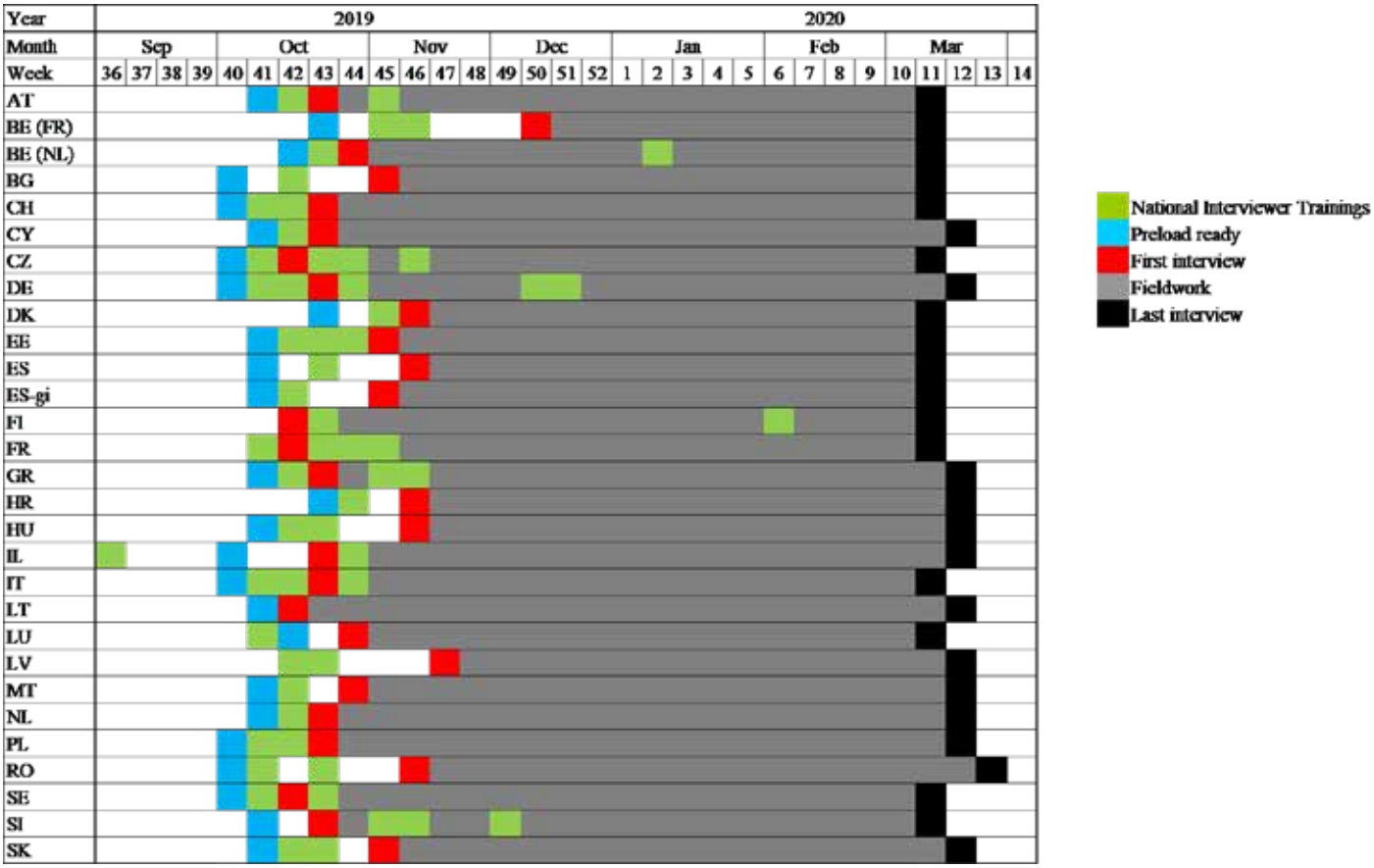
5.1.2 Fieldwork Periods of the Regular Wave 8 and Survey Agencies

Due to the fast spread of COVID-19 across Europe, SHARE fieldwork of Wave 8 had to be suspended country by country between March 10 and March 23, 2020. At this point in time, about 70 per cent of all expected longitudinal and 50 per cent of all expected refreshment interviews across countries had been done (see Scherpenzeel et al., 2020). In this respect, it is important to emphasise that compared to previous waves, the final rates of the regular Wave 8 are not as meaningful because fieldwork had to be stopped due to *force majeure*.

Figure 5.1 shows that almost all countries of Wave 8 were able to put the originally planned schedule into action. The fieldwork start of Wave 8 happened largely synchronously across countries between the end of October and the beginning of November 2019. In some countries, additional interviewer training was necessary (green squares). Notable exceptions were the French part of Belgium and Latvia, which show substantial delays between national interviewer training and delivering the first interview. The suspension of fieldwork due to the outbreak of COVID-19 happened between weeks 11 and 13 of 2020 across all countries. Since Portugal had issues with securing funding on time, there was no longer any chance of joining the regular fieldwork of Wave 8 in spring 2020.

7 In the following, we differentiate between the two 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 5.1: Fieldwork Periods of Regular SHARE Wave 8



The following organisations in Table 5.1 below conducted the fieldwork in each listed wave. There has been high stability of contracted survey agencies over time in most countries.

Table 5.1: Survey Agencies from Wave 1 to 8 of Countries Participating in Wave 8

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7	Wave 8/SCS
AT	IMAS	IMAS	IFES	IFES	IFES	IFES	IFES	IFES
BE-FR	PSBH, Liège Univ	PSBH, Liège Univ	PSBH, Liège Univ	PSBH, Liège Univ	CELLO - Antwerp Univ.	CELLO - Antwerp Univ.	CELLO - Antwerp Univ.	CELLO - Antwerp Univ.
BE-NL	PSBH Antwerp Univ.	PSBH Antwerp Univ.	CELLO - Antwerp Univ.	CELLO - Antwerp Univ.	CELLO - Antwerp Univ.	CELLO - Antwerp Univ.	CELLO - Antwerp Univ.	CELLO - Antwerp Univ.
BG	-	-	-	-	-	-	GfK Bulgaria	GfK Bulgaria
CH	MIS Trend	LINK	LINK	LINK	LINK	LINK	LINK	LINK
CY	-	-	-	-	-	-	RAI Consultants	RAI Consultants
CZ	-	SC&C	SC&C	SC&C	SC&C	SC&C	SC&C	SC&C

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7	Wave 8/SCS
DE	infas GmbH	infas GmbH	infas GmbH	infas GmbH	TNS Infratest	TNS Infratest	TNS Infratest	Kantar Public
DK	SFI Survey	SFI Survey	SFI Survey	SFI Survey	SFI Survey	SFI Survey	DST Survey	DST Survey
EE	-	-	-	Statistics Estonia	GfK	Statistics Estonia	Statistics Estonia	Statistics Estonia
EG	TNS Demoscopia	TNS Demoscopia	TNS Demoscopia	TNS Demoscopia	TNS Demoscopia	TNS Demoscopia	TNS Demoscopia	Ipsos Iberia
ES	TNS Demoscopia	TNS Demoscopia	TNS Demoscopia	TNS Demoscopia	TNS Demoscopia	TNS Demoscopia	TNS Demoscopia	Kantar TNS
FI	-	-	-	-	-	-	Taloustutki-mus	Taloustutki-mus
FR	INSEE	INSEE	INSEE	INSEE (panel)/ GFK-ISL (refresh.)	GFK-ISL	TNS SOFRES	TNS SOFRES	TNS SOFRES
GR	Kapa Research	Kapa Research	Kapa Research	-	-	Kapa Research	Kapa Research	Kapa Research
HR	-	-	-	-	-	GfK	GfK	IPSOS
HU	-	-	-	TÁRKI Social Research Institute	-	-	TÁRKI Social Research Institute	TÁRKI Social Research Institute
IL	Cohen Institute, Tel Aviv Univ.	Cohen Institute, Tel Aviv Univ.	-	-	Cohen Institute, Tel Aviv Univ.	Cohen Institute, Tel Aviv Univ.	Cohen Institute, Tel Aviv Univ.	Cohen Institute, Tel Aviv Univ.
IT	DOXA S.p.A.	DOXA S.p.A.	DOXA S.p.A.	DOXA S.p.A.	IPSOS	IPSOS	IPSOS	IPSOS
MT	-	-	-	-	-	-	Grant Thorn-ton (EMCS)	Grant Thorn-ton (EMCS)
NL								I&O Research
LU	-	-	-	-	CEPS	LISER	LISER	LISER
LT	-	-	-	-	-	-	TNS	TNS
LV	-	-	-	-	-	-	ISR	ISR
PL		TNS-OBOP	TNS-OBOP	TNS-OBOP	TNS Polska	TNS Polska	TNS Polska	Kantar Polska
PT				GfK Metris	CECS, Univer-sity of Minho	CECS, Univer-sity of Minho	CECS, Univer-sity of Minho	CECS, Univer-sity of Minho
RO	-	-	-	-	-	-	GfK Romania	GfK Romania
SE	Intervjubola-get IMRI	Intervjubola-get IMRI	Intervjubola-get IMRI	Intervjubola-get IMRI	Intervjubola-get IMRI	IPSOS Observer	IPSOS Observer	IPSOS Observer
SI	-	-	-	CJMMK	CJMMK	IPSOS	IPSOS	IPSOS
SK	-	-	-	-	-	-	GfK Slovakia	GO4insight & ACRC

5.1.3 Monitoring Fieldwork

This section includes information about the classification and computation of survey outcomes and all final rates and figures of Wave 8 based on the last data export at the end of 2020. All numbers and figures reported during fieldwork are based on information from the CaseCTRL (*read:* case control), formerly known as the SHARE Sample Management System (SMS), which is the interviewer software used to document contact attempts and conduct the interviews. As of Wave 8, all CaseCTRL 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 fieldwork progress over time. Final rates and interview numbers are then provided again in a final summary graph without trajectories to allow for easier comparison between countries.

5.1.4 Classification of Survey Outcomes

Identically to the previous waves, most representational indicators (i.e. those on unit non-response) were set out as quality targets in the specifications of the model contract of SHARE Wave 8. As usual, we follow the newest edition of AAPOR guidelines and use data from the CaseCTRL to classify the baseline/refreshment and longitudinal gross samples[8] 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 CaseCTRL is continuously converted into a so-called "household state". The algorithm that creates the household state divides the sample into three mutually exclusive categories: (i) ineligible households; (ii) eligible households; and (iii) households of unknown eligibility[9]. This is done 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 laid out 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 applies to sorting households into subcategories of the household state. A new contact only results in a change of the household state if it involves new information that conceptually trumps the previous information. For example, a household formerly classified as "non-contact" (NC) will switch 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 5.2.

Table 5.2: Detailed List of CaseCTRL Entries and Fieldwork Outcomes at the Household Level

CaseCTRL Contact Protocol Entry	Household State
Deceased[3]	
In hospital[3]	
In old-age 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]	

8 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.

9 For details on SHARE's target population and eligibility criteria, see Kneip (2013) and Bergmann et al. (2017).

CaseCTRL Contact Protocol Entry	Household State
Completed interview (incl. End-of-Life interview)	CI
Partial interview	PI
Interrupted interview	II
Refusal[1]	R
Too busy, no time	
Too old, bad 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 old-age home[4]	
Moved, new address known	
Moved, new address unknown[3]	
Address non-existent, house vacant[3]	
Household screened as eligible	
Non-contact[2]	NC
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, the latter one calling upon intervention from the agency. Neither of the refusal codes set by the interviewer closed a case.
[2] Non-contact for the eligible part of the sample does not apply to the baseline/refreshment sample in the countries in which age is not available from the sampling frame.
[3] This led to ineligibility only in the baseline/refreshment sample, and not in the longitudinal sample.
[4] Whether this 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, language barriers.

5.1.5 Formulas to Compute Survey Outcomes

Apart from eligibility, the household state variable provides information about a household's contact and cooperation status. Table 5.3 reports which fieldwork indicators are used and how they are computed based on the household state. As the current state can be determined by the CaseCTRL for every household at any given point in time, we are able to report the state of fieldwork at any time as if it was over.

In terms of household cooperation, households are considered as participating if at least one eligible household member is successfully interviewed. When looking at individual cooperation, several definitions of individual response rates are possible depending on how households of unknown eligibility are treated and how the number of eligible households of 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 and estimate this fraction by $\frac{E}{E+NE}$. Over the course of fieldwork, this estimate improves in terms of precision as the non-attempted part of the sample declines.

The number of eligible persons is only known for households with a completed Coverscreen (CV) interview. 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 preloaded information on the household composition to assess the number of eligible respondents. Here, the assumption is that this number does not change 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 5.3: 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]	$\frac{(CI_r+PI_r)}{\bar{n}(E+p\cdot UE)}$

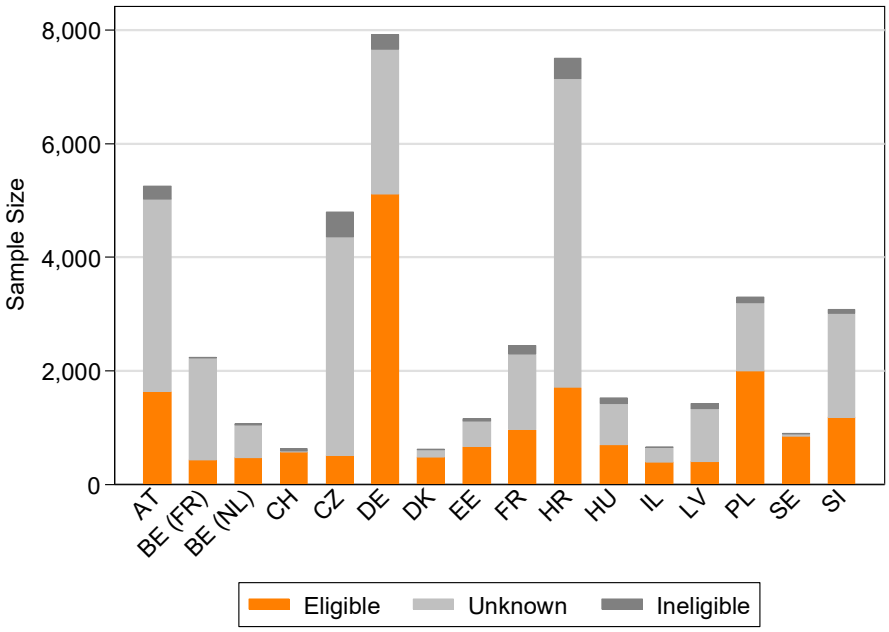
Individual response rate in subsample i[3]	$\frac{(CI_i+PI_i)}{\bar{n}_i(E+p\cdot UE)}$

Notes:
[1] $p(UE_R+UE_O)$ is not part of the denominator in AAPOR COOP2. The calculation method was adapted for the 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 household 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}.

5.1.6 Refreshment Samples

In Wave 8, 15 countries drew a refreshment sample: Austria, Belgium, Croatia, the Czech Republic, Denmark, Estonia, France, Germany, Hungary, Israel, Latvia, Poland, Slovenia, Sweden and Switzerland[10]. Figure 5.2 shows the size of the refreshment samples. Apart from the reasons leading to ineligibility in the longitudinal sample (i.e. incarceration, moving abroad, language barriers), refreshment households are also considered ineligible in the following cases: death of the drawn respondent, in-patient treatment throughout the entire field time, unknown or invalid addresses and if the Coverscreen (CV) interview yields no eligible persons in the household. While Austria's authorities only provide information on age categories for people aged 50 and above, population registers in the Czech Republic, France, Israel and Latvia do not contain any information on age. The samples in these countries 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 highest in the Czech Republic, which reflects a lack of available sample frame information. Households are classified as having "unknown eligibility" after any form of screening non-response (non-contact, refusal, other non-response). This fraction is largest in Croatia, followed by the Czech Republic, Austria and Germany. However, the refreshment samples in all four countries could not be worked off sufficiently before the suspension of fieldwork to explicitly determine household eligibility, although (with the exception of the Czech Republic) it is guaranteed that all selected units (e.g. addresses) contain at least one person aged over 50. This also holds for most countries that need screening.

Figure 5.2: Refreshment Samples by Classification of Sample Units

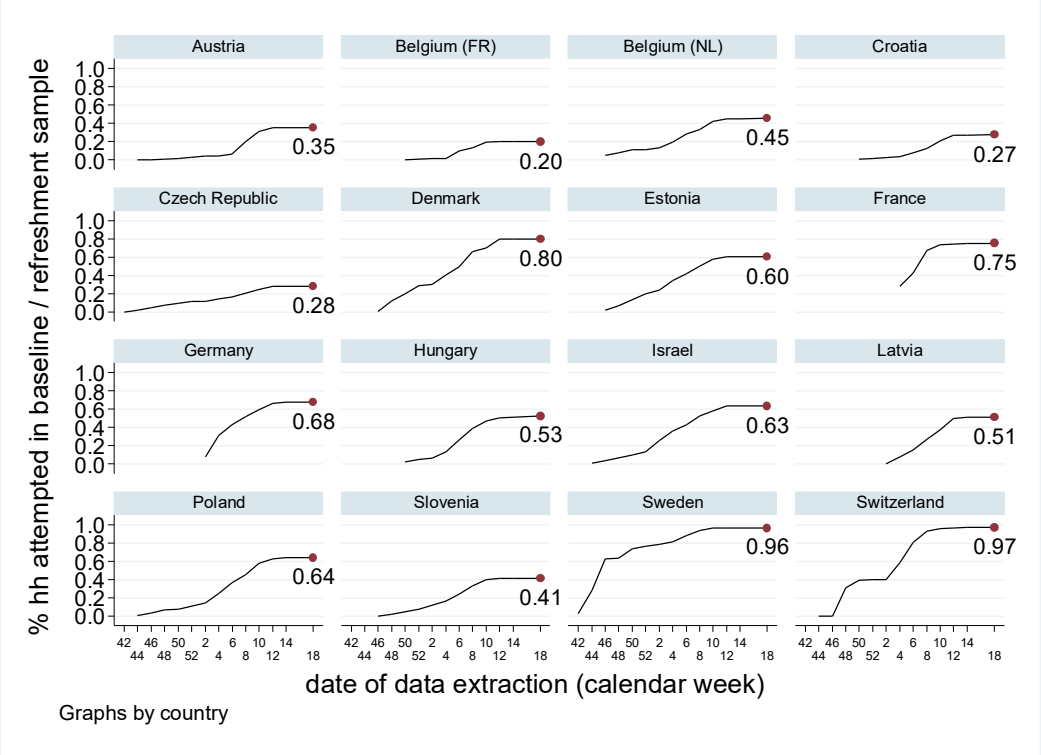


10 Furthermore, Finland, Spain and Portugal drew refreshment samples but were not able to field them due to the suspension of fieldwork in March 2020.

5.1.6.1 Contacting Households

Figure 5.3 shows the fraction of households of refreshment samples where a contact was attempted (i.e. all households where either an interviewer reports a contact attempt but was unable to actually contact anybody or where a contact is successful). By definition, this includes households with one or more conducted interviews. The data extraction from week 14 in 2020 was the last one before deciding to close regular fieldwork and switch to telephone mode. The last data point (i.e. week 18) hence provides the final rates according to completed CaseCTRL data extractions for each country.

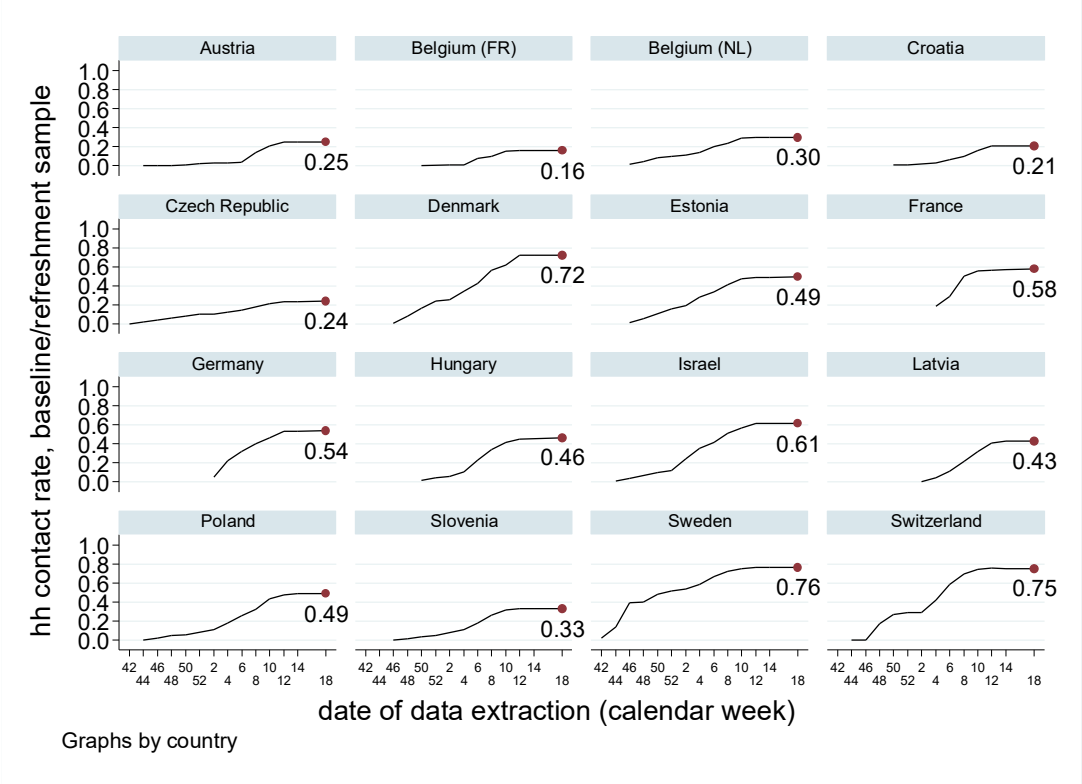
Figure 5.3: Fraction of Refreshment Sample Households with Contact Attempts by Country over Time



The suspension of fieldwork of Wave 8 took effect between weeks 11 and 13, which is why the trajectories of all countries flatten out eventually. In the majority of countries, the refreshment samples were started after the winter break in week 2 of 2020. Therefore, most countries were not able to work off their refreshment samples. With contact attempt rates of 96 and 97 per cent, the only exceptions are Sweden and Switzerland, where interviewers approached all new households immediately after starting fieldwork in autumn 2019. While Sweden's interviewers reached about 60 per cent of the sample within the first weeks of fieldwork, Switzerland's trajectory can be interpreted as two big efforts to approach the sample (one immediately after starting fieldwork in week 46, the other one after the winter break in week 2).

Figure 5.4 shows household contact rates broken down by countries. This contains contact attempts that resulted in an actual contact. By definition, this may also include households with at least one completed interview.

Figure 5.4: Contact Rate of Refreshment Sample Households by Country over Time

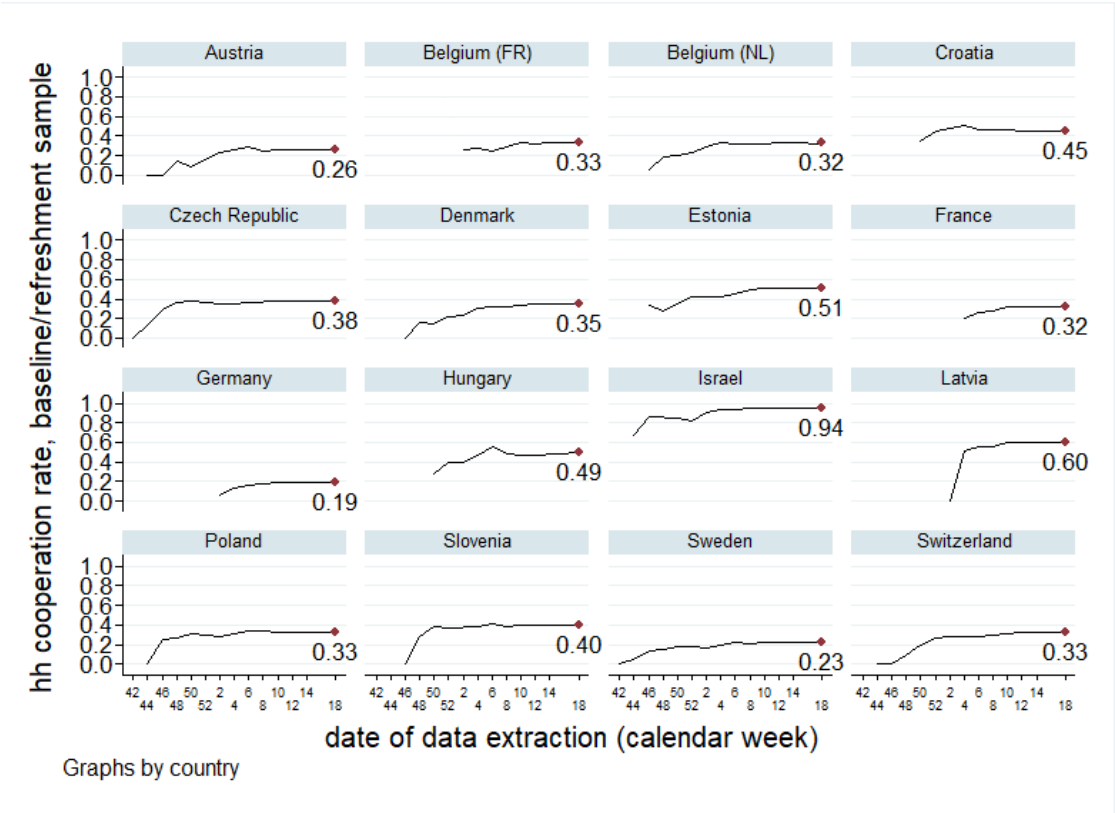


Sweden and Switzerland have the highest contact rates (76 and 75 per cent, respectively). However, even though they almost managed to exhaust their samples in terms of contact attempts before the discontinuation of fieldwork, Denmark and Israel made the best progress in establishing successful contacts relative to the share of attempted households (72 and 61 per cent, respectively).

5.1.6.2 Household Cooperation and Response Rate

Figure 5.5 shows the cooperation rate of refreshment samples by country (i.e. the rate of all contacted households that have at least one completed interview).

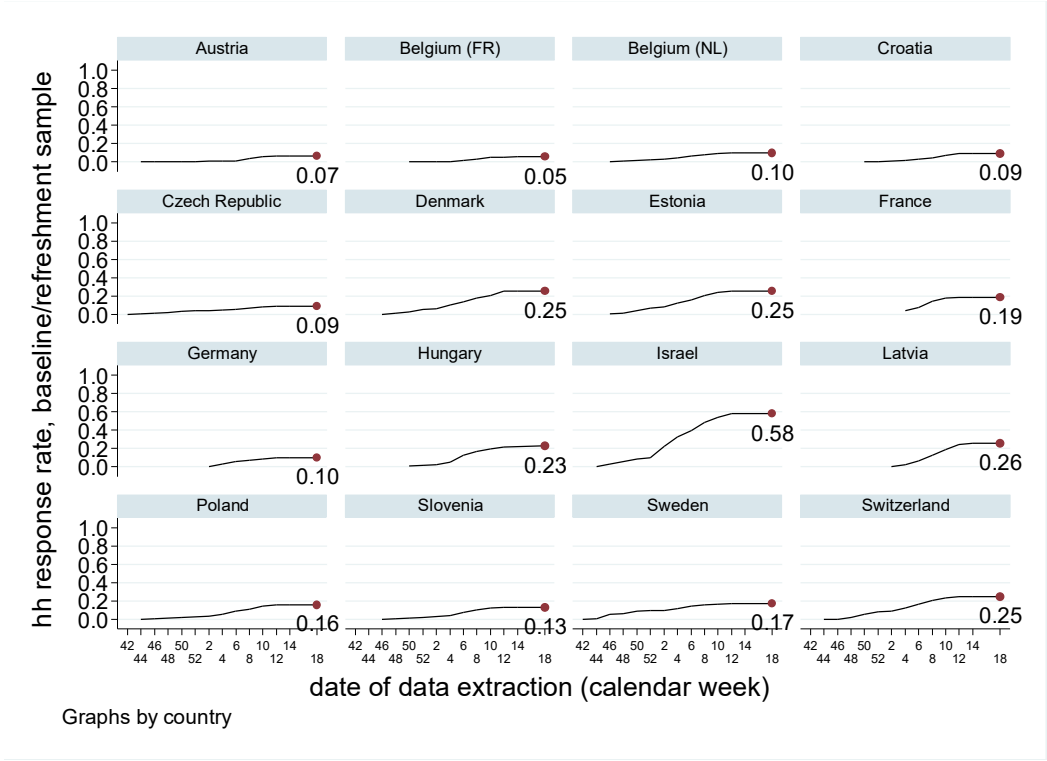
Figure 5.5: Cooperation Rate of Refreshment Sample Households by Country over Time



The cooperation rate is based only on sample units with a previous contact. For this reason, it can fluctuate up or downwards. The more contacts have been established, the less it fluctuates and the better it can be interpreted. Therefore, only the cooperation rates in Sweden and Switzerland are meaningful. It can be assumed that the rates of 23 per cent in Sweden and 33 per cent in Switzerland would not have increased much more, which implies that it became more and more difficult for interviewers to recruit new households in these countries due to the uncertainty regarding COVID-19.

Figure 5.6 shows the household response rate (i.e. the number of refreshment households with at least one complete interview divided by the total number of (estimated) eligible refreshment households).

Figure 5.6: Response Rate of Refreshment Sample Households by Country over Time

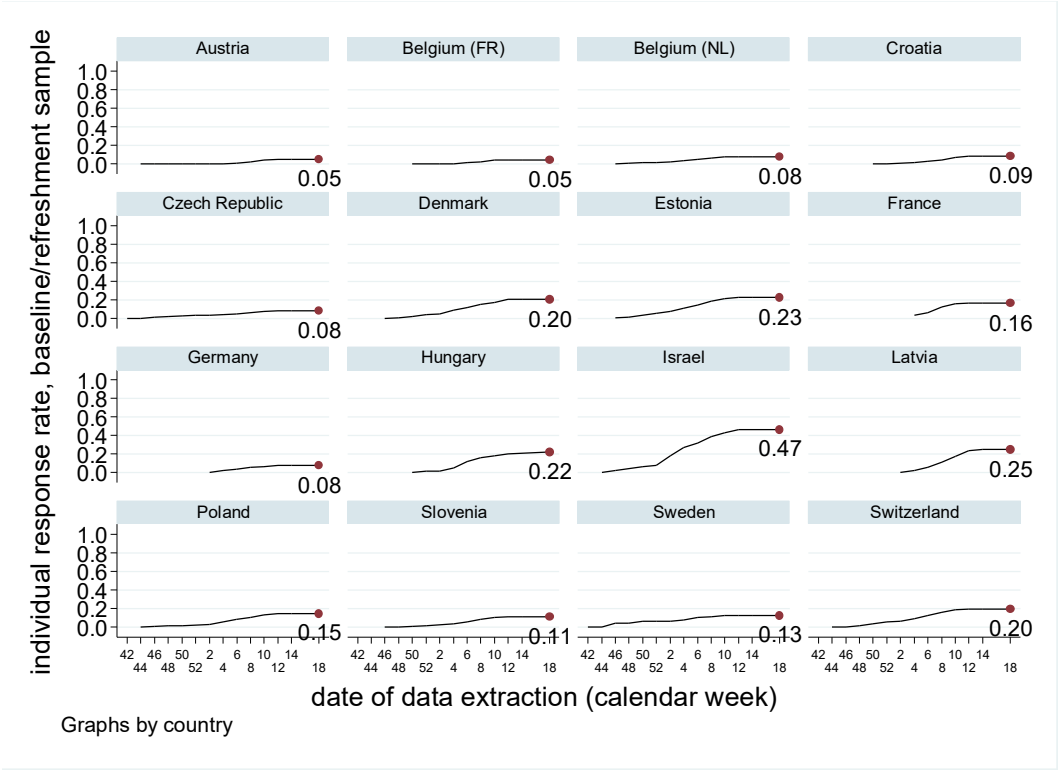


Except for Israel, all countries show slowly increasing trajectories reaching up to about 25 per cent of household response. Considering that Israel worked off 63 per cent of its sample and that refreshment samples are always harder to approach than longitudinal samples, the household response rate of 58 per cent is remarkable.

5.1.6.3 Individual Participation

Figure 5.7 shows the individual response rate of refreshment samples in Wave 8.

Figure 5.7: Individual Response Rate of Refreshment Respondents by Country over Time

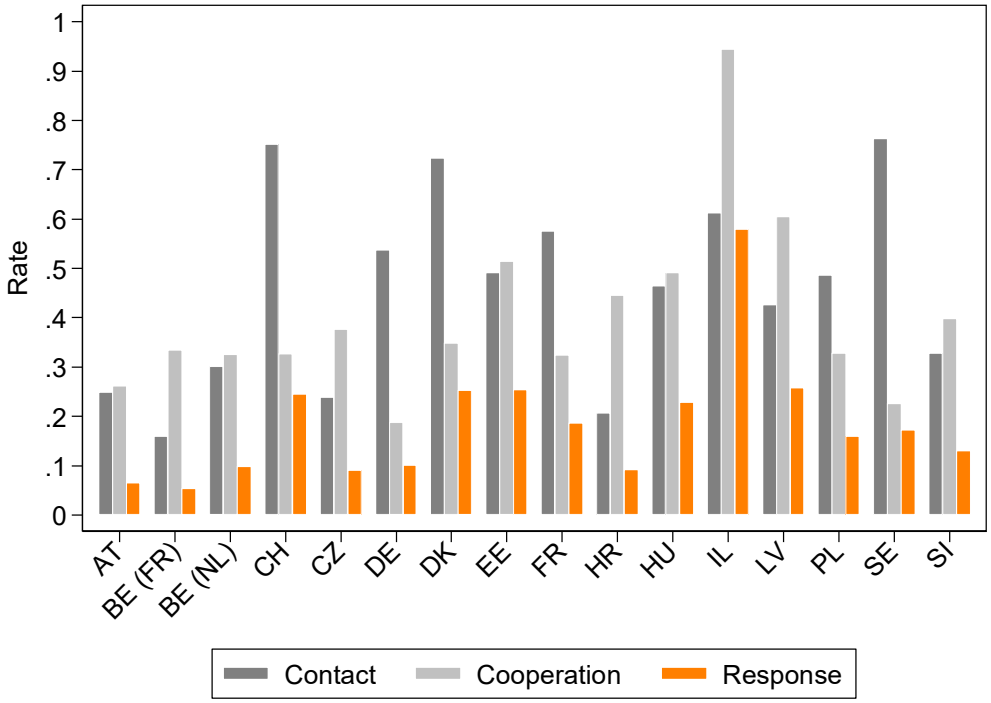


In most countries, individual response is close to household response, meaning that in the majority of cases interviewers were able to conduct an interview with all household members. Despite the suspension of fieldwork, Israel managed to surpass the target of 43 per cent (i.e. the reference response rate). However, compared to its household response rate (58 per cent), Israel's individual response rate of 47 per cent shows that it was difficult to convince spouses or partners to participate.

5.1.6.4 Summary

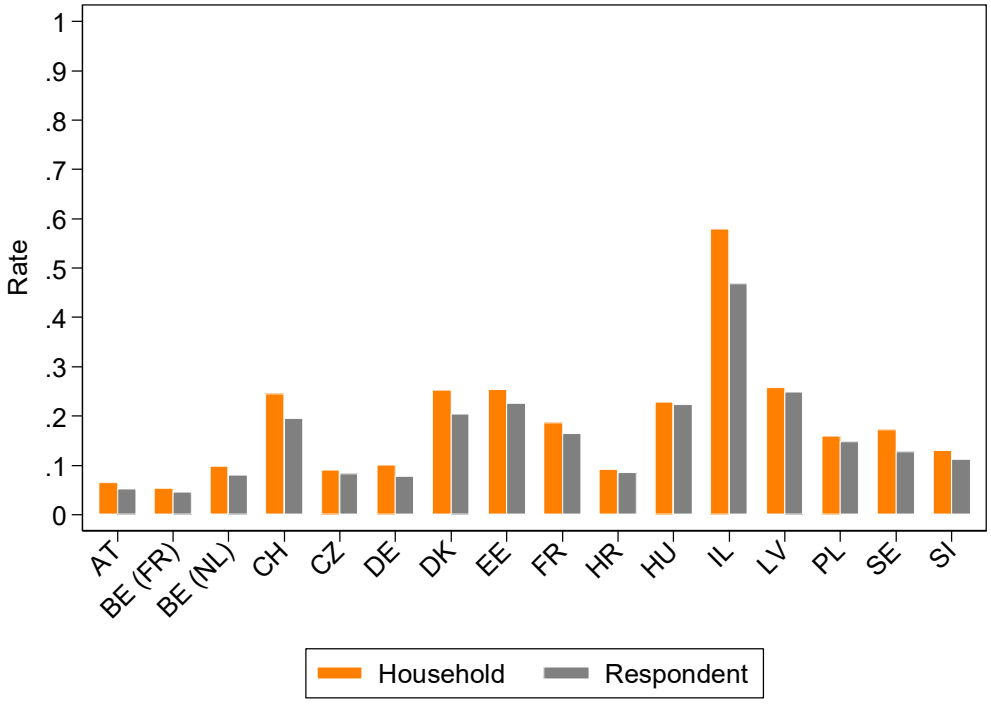
Figure 5.8 shows the final household contact, cooperation and response rates at the abrupt end of face-to-face fieldwork of Wave 8.

Figure 5.8: Contact, Cooperation and Response Rates for Refreshment Samples



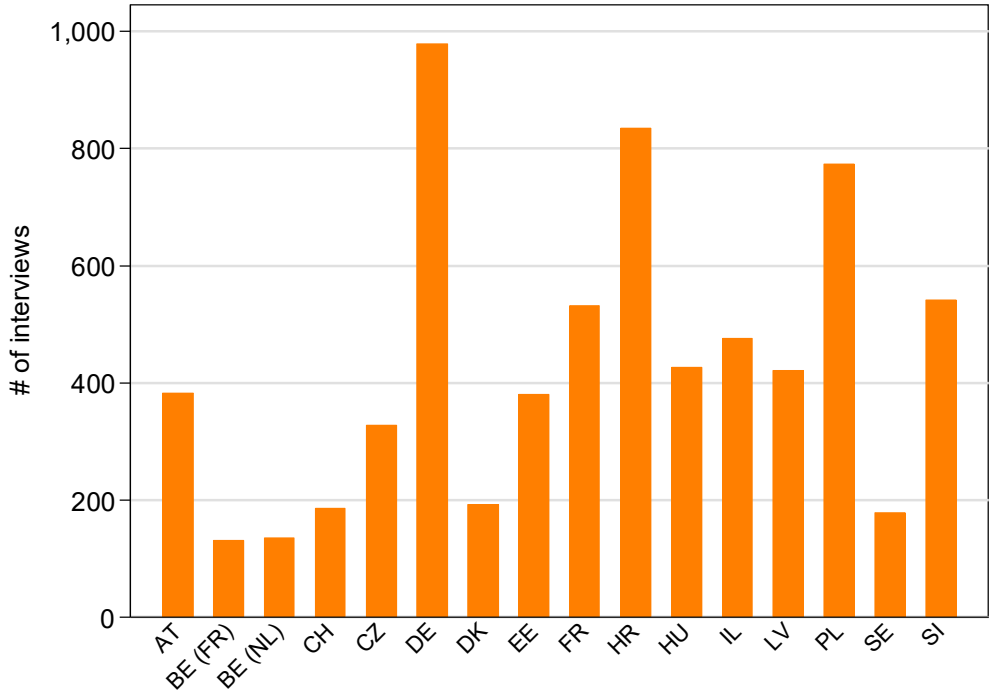
Owing to the suspension of fieldwork, the contact rates are fairly low, with the highest values in Denmark, Sweden and Switzerland (around 75 per cent). Israel is the country with the highest cooperation rate (about 94 per cent, but only 63 per cent of all households attempted) and household response rate (about 58 per cent).
Figure 5.9 shows the final household- and respondent-level response rates.

Figure 5.9: Household- and Respondent-Level Survey Participation in Refreshment Samples



Consequently, the household and individual response rates are low as well. With the exception of Israel, which has a household response rate of 58 per cent and an individual response rate of 47 per cent, all other countries are below the 30 per cent mark. The individual response rates are always slightly lower than the household response rates because of non-cooperation among some household members. The gap between both rates is smallest in Belgium (FR), Croatia, the Czech Republic, Hungary and Latvia, which means that in these countries, interviewers managed best to convince all household members. Figure 5.10 shows the absolute number of interviews per country in the refreshment samples at the end of fieldwork.

Figure 5.10: Absolute Number of Interviews in Refreshment Samples



The number of interviews varies with sample size and factors such as interviewers' progress until the break-off of fieldwork. Germany had the largest refreshment sample fielded until the suspension of fieldwork. It also has the highest number of interviews, followed by Croatia and Poland. It is planned to continue interviewing the refreshment samples in Wave 9 and release them together with the Wave 9 data.

5.1.7 Panel Samples

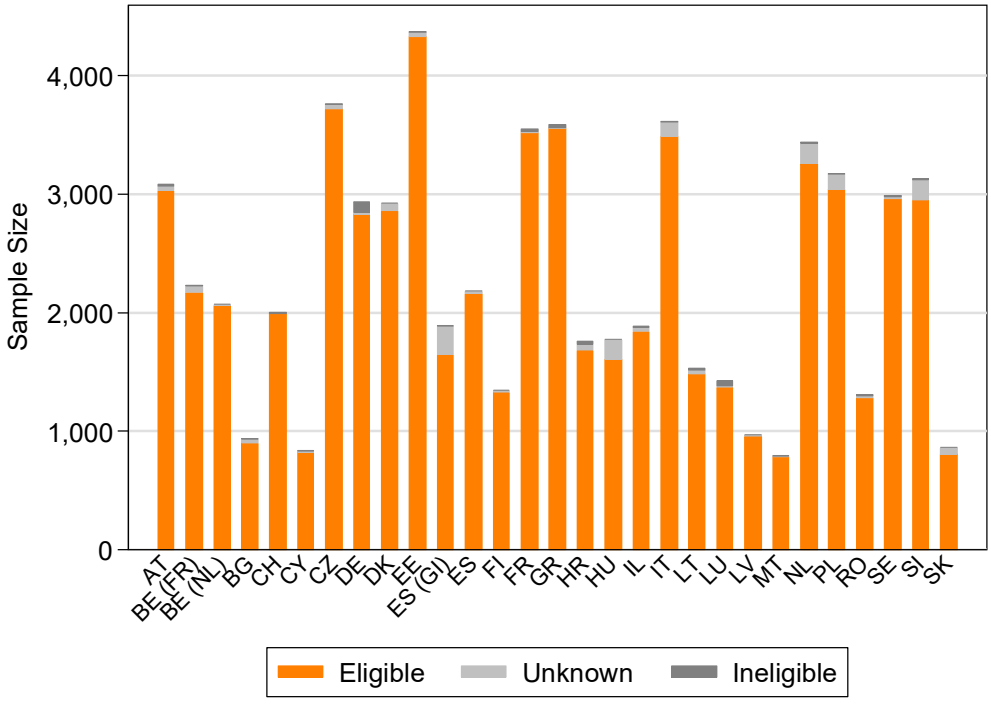
In general, longitudinal samples can be divided into five subsamples at the individual level according to SHARE's eligibility rules. While subsample A1 includes all respondents who participated in the previous wave and any other wave of the SHARE survey, subsample A2 consists of respondents who live 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 where 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 where 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 participated in SHARE so far.

Response rates are reported separately for these subsamples during fieldwork. 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 never interviewed before (i.e. either new sample members or eligible sample members who finally participated after refusals in previous waves).

Figure 5.11 shows the size and composition of the longitudinal sample per country in Wave 8. 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 preloaded into the CaseCTRL. Households that must not be approached again for legal

reasons are dropped. Overall, the longitudinal gross samples of all countries contain almost exclusively eligible cases (96.8 per cent), with the lowest share being in the Spanish region of Girona (86.8 per cent) and the highest in Switzerland (99.5 per cent). While Estonia had the largest panel sample with more than 4,000 eligible households, most of the baseline countries of Wave 7 had samples of around 1,000 eligible households (e.g. Bulgaria, Cyprus, Latvia, Malta, Slovakia).

Figure 5.11: Panel Samples by Classification of Sample Units

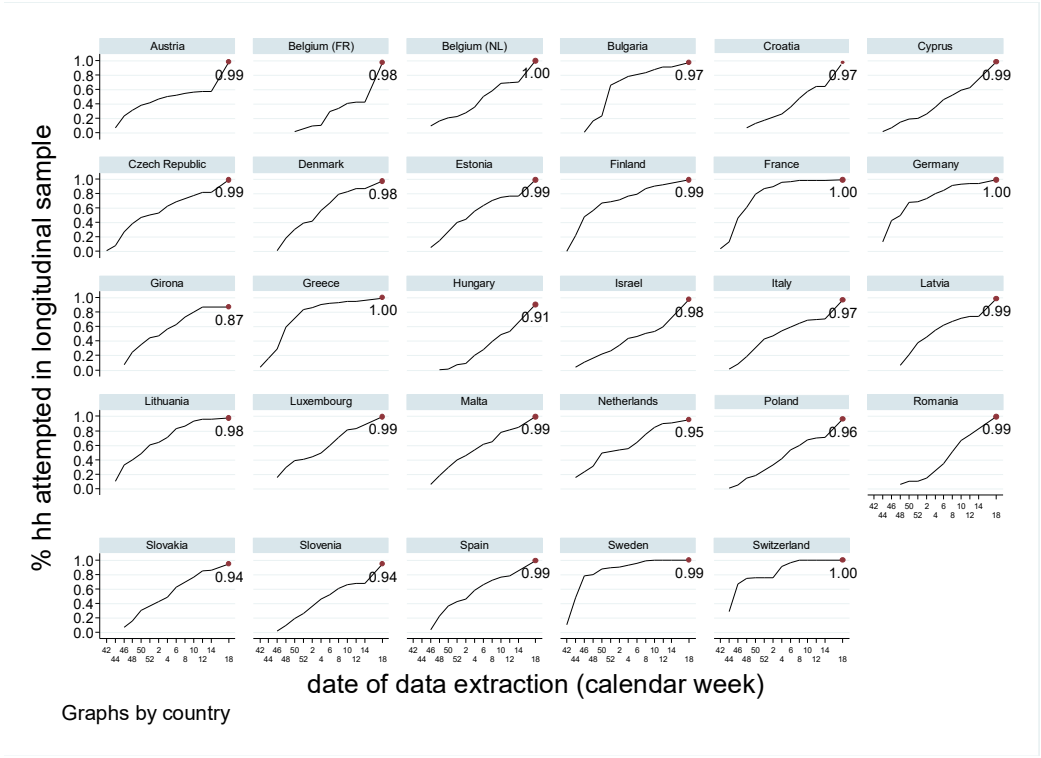


Households in the longitudinal sample can only become ineligible for the following reasons: incarceration, moving abroad and language barriers. Ineligibility applies to, on average, 0.7 per cent of all households in the longitudinal samples. Death does not lead to ineligibility. Instead, a proxy respondent 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 CaseCTRL, the eligibility of 2.5 per cent of all longitudinal households was unknown in Wave 8.

5.1.7.1 Contacting Households

Figure 5.12 shows the fraction of households in the longitudinal gross sample with a contact attempt (i.e. all households either where an interviewer reports a contact attempt but was unable to actually contact anybody or where a contact is successful). By definition, this includes households with one or more conducted interviews. It should be repeated that the suspension of fieldwork took place after the data extraction from week 14 in 2020. The data point labelled "18" provides the final rates for each country.

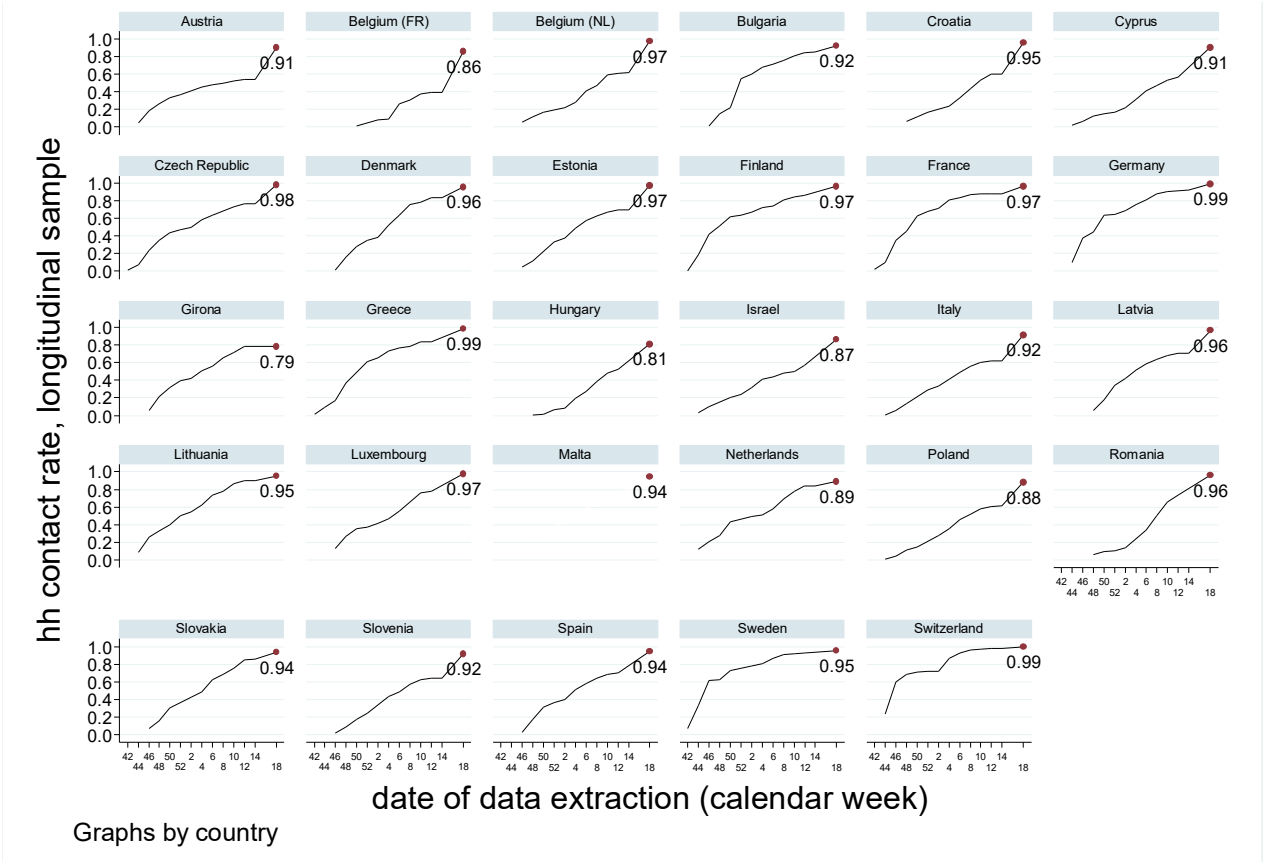
Figure 5.12: Fraction of Panel Households with Contact Attempts by Country over Time



Despite closing fieldwork earlier as planned, most countries achieved contact attempt rates of about 95 per cent or more. The only exceptions are the Spanish region of Girona (87 per cent) and Hungary (91 per cent). Belgium (NL), France, Germany, Greece and Switzerland exhausted their samples. It can be seen that some countries (e.g. France, Greece, Sweden, Switzerland) have a steep increase that levels out over time (i.e. interviewers were quick at approaching the majority of all households for contact at the beginning of fieldwork), while others show a linear trend, possibly due to a different contact strategy. The sharp change from week 14 to 18 in some countries (visible in this and the following graphs) is due to slowly incoming additional data that took some time to be completely processed.

Figure 5.13 shows country breakdowns of household contact rates over time. This contains contact attempts that resulted in an actual contact (i.e. at least one household member was reached). By definition, this may also include households with at least one complete interview.

Figure 5.13: Contact Rate of Panel Households by Country over Time

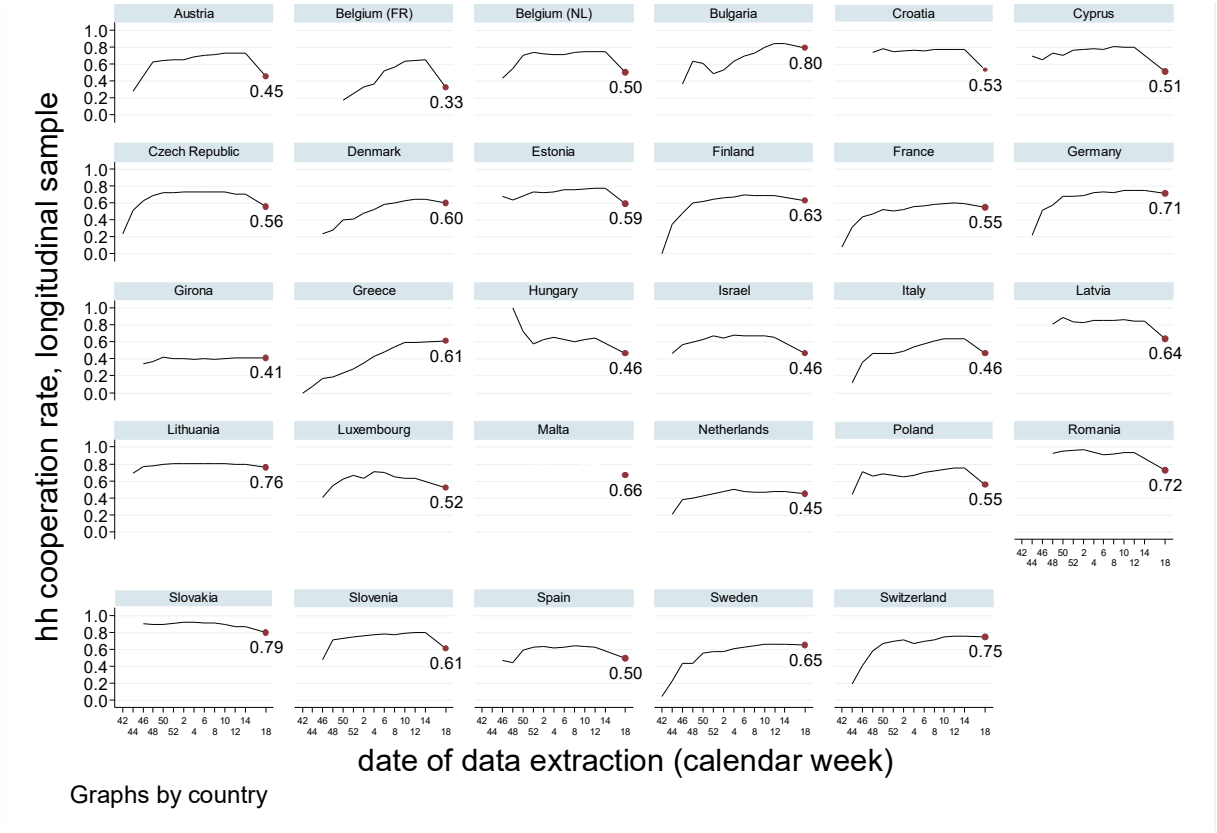


The trajectories of the contact rates are similar to the rates for approached households reported above. With contact attempt rates being the logical ceiling for contact rates, the Spanish region of Girona and Hungary also have the lowest contact rates. The highest contact rate of 99 per cent was achieved in Germany, Greece and Switzerland.

5.1.7.2 Household Cooperation and Response Rate

Figure 5.14 shows the cooperation rate of panel samples by country (i.e. the rate of all contacted households that have at least one completed interview).

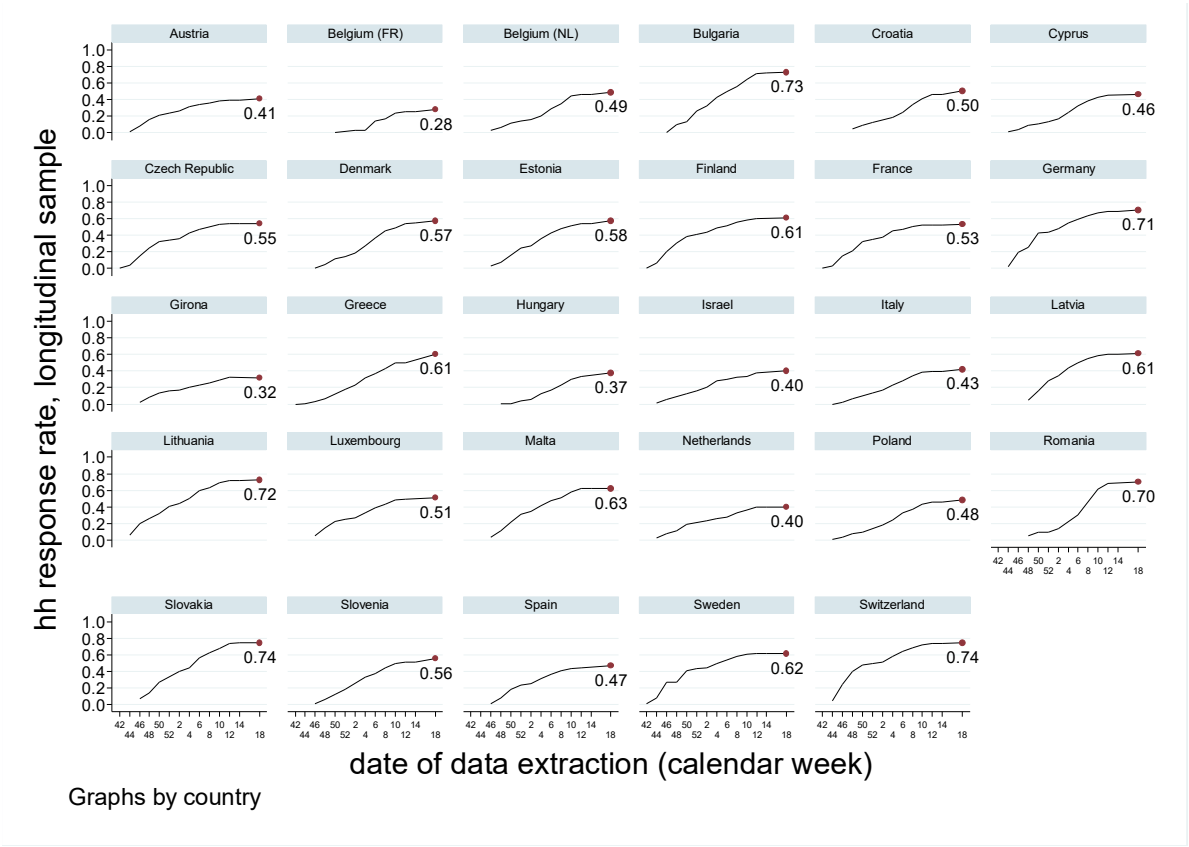
Figure 5.14: Cooperation Rate of Panel Households by Country over Time



Since the cooperation rate is based only on sample units with a previous contact, it is the only rate that can fluctuate up or downwards. Over time and with an increasing number of contacts, it stabilises and becomes more meaningful. The drop between week 14 and 18 in many countries can be attributed to the data updates done for the last data point (especially regarding the number of successfully contacted households). While Bulgaria, Slovakia and Switzerland have the highest cooperation rates ranging from 75 to 80 per cent, Belgium (FR) ended up with the lowest cooperation rate (33 per cent).

Figure 5.15 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).

Figure 5.15: Retention Rate of Panel Households by Country over Time

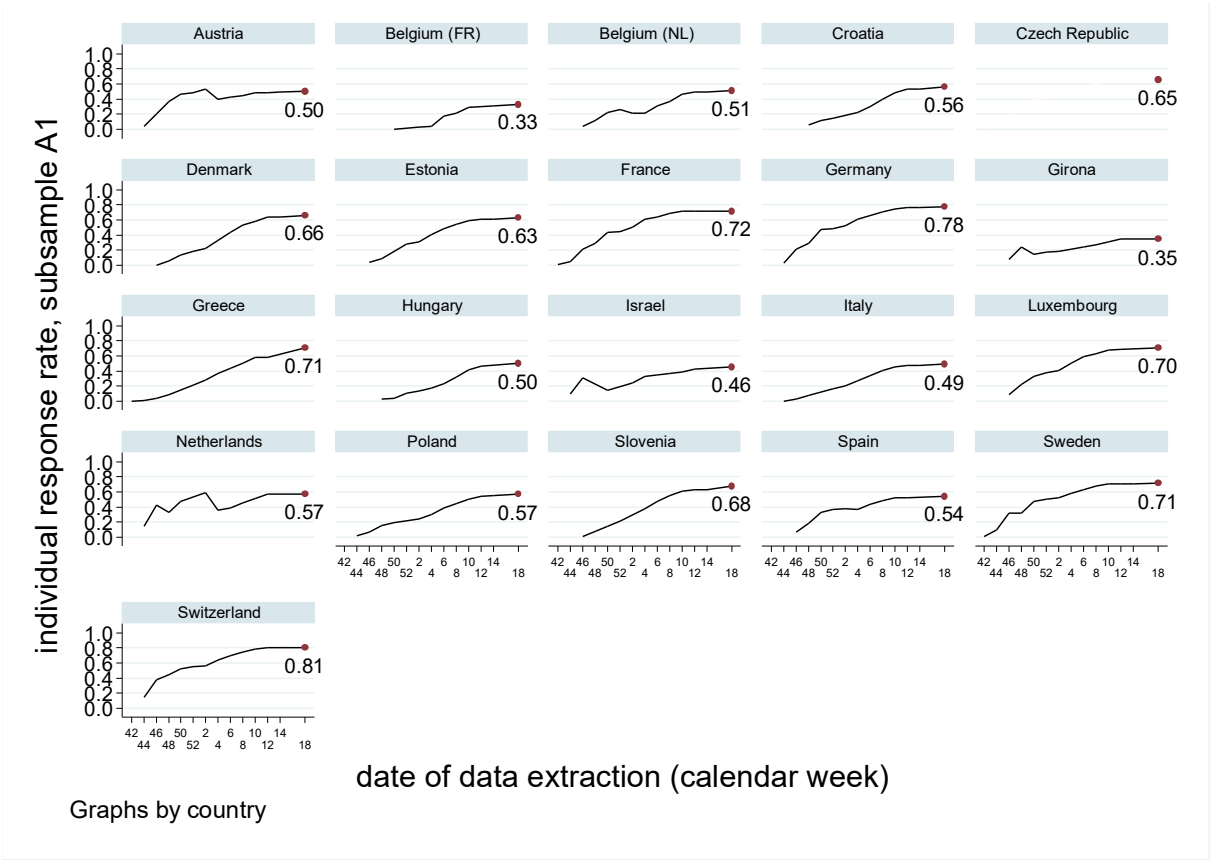


Almost all countries have steadily increasing trajectories that level out over time. Because of the suspension of fieldwork and the fact that cooperation represents the logical ceiling for the final retention rate, no single country could achieve a sufficiently high household retention rate. The highest ones can be observed in Bulgaria, Germany, Lithuania, Romania, Slovakia and Switzerland (all between 71 and 74 per cent). Belgium (FR) and Girona are at the very end with 28 and 32 per cent, mainly due to problems at the beginning that could not be resolved until the suspension of fieldwork.

5.1.7.3 Individual Participation

Figure 5.16 shows the individual retention rate of subsamples A1 and A2. As pointed out before, subsample A1 includes all respondents who participated in Wave 7 and at least one other previous wave; subsample A2 includes all respondents who were part of a baseline or refreshment sample in the previous wave. The SHARE contract stipulates bringing back at least 85 per cent of respondents in subsample A1 and 75 per cent in subsample A2 in the current wave. Survey agencies were incentivised for rates exceeding these thresholds. If the fieldwork of Wave 8 could have been continued, several countries might have reached these target rates.

Figure 5.16: Individual Retention Rates in Subsample A1 by Country over Time



With an 81 per cent individual response, Switzerland is the only country that came close to the 85 per cent mark, followed by Germany with 78 per cent. Please note that the drop-in response rates in some countries can be attributed to a bug in the household state variable. After correction, this increased the number of eligible respondents (as part of the denominator) in some countries and therefore led to an adjustment of the response rate.

Subsample A2 is comprised of respondents that participated for the first time in the previous wave. Croatia, Israel and Poland had a refreshment sample in Wave 7. All other countries in Figure 5.17 joined SHARE in Wave 7 with a baseline sample. In addition, Croatia, Israel, Latvia and Poland also had another refreshment sample in Wave 8. Bulgaria, Lithuania, Romania and Slovakia are close to the 75 per cent mark. It is remarkable that A2 retention in the countries with a refreshment sample in Wave 8 remains far below A2 retention in countries without a refreshment sample in Wave 8. This may point to specific difficulties in conducting interviews in both the panel and the refreshment sample simultaneously. The break-off of fieldwork then implied that there was not enough time to contact all panel respondents.

Figure 5.17: Individual Retention Rates in Subsample A2 by Country over Time

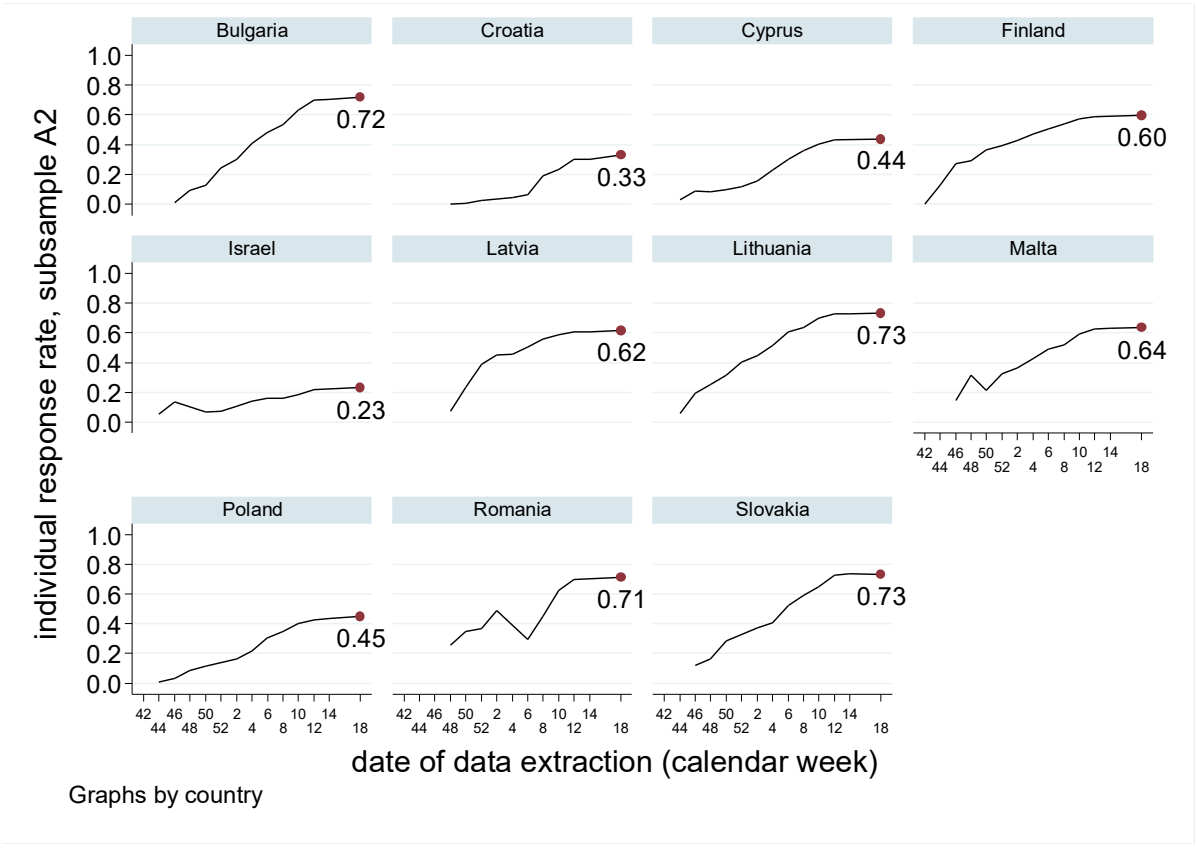
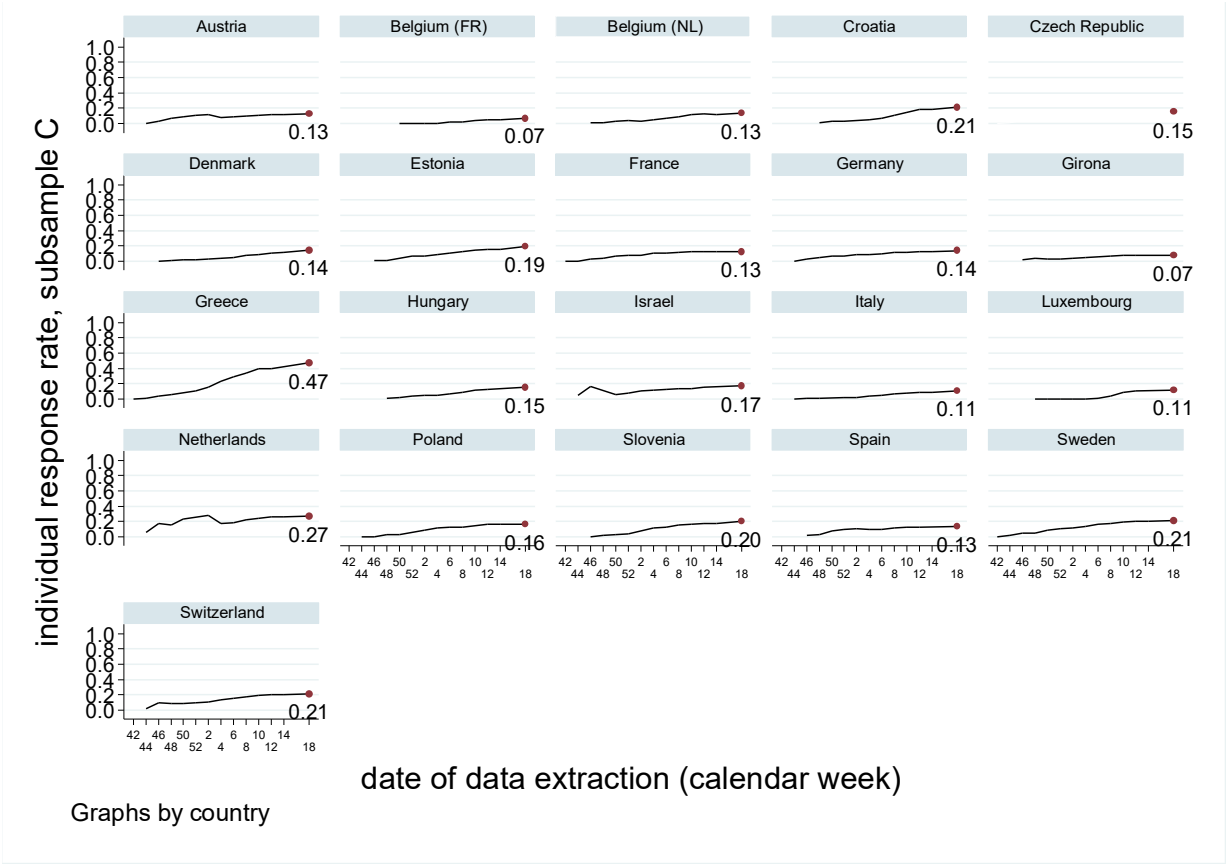


Figure 5.18 shows the individual retention rate (or, more precisely, the "recovery rate") of subsample C (i.e. the percentage of panel respondents that did not participate in Wave 7 and any combination of (non-)participation in previous waves, but that were brought back into Wave 8).

Figure 5.18: Individual Retention (Recovery) Rates in Subsample C by Country over Time

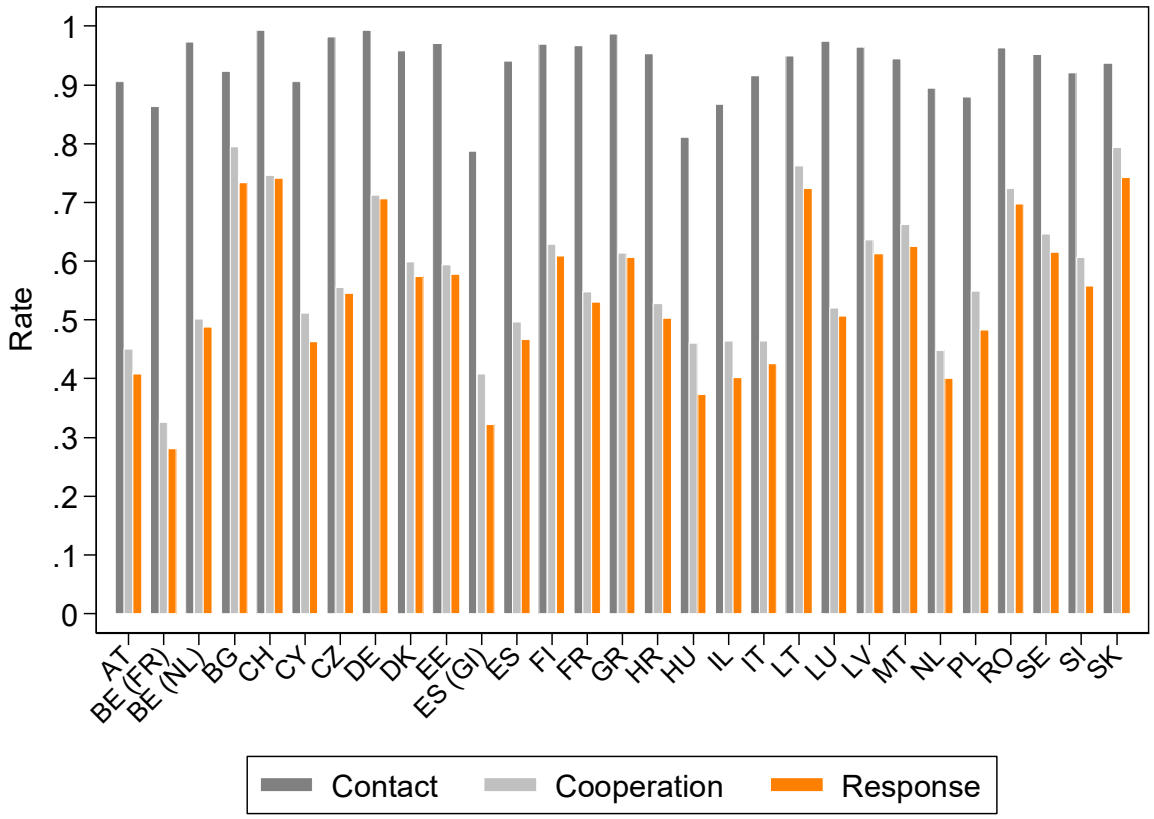


The Netherlands had the largest subsample in Wave 8 (85 per cent), followed by Greece (63 per cent) and Hungary (60 per cent). In comparison, Greece showed the best performance, almost recuperating half of its subsample C (47 per cent). The Netherlands, which did not participate in Wave 7 via face-to-face interviews, brought back about a quarter of its "lost" respondents (27 per cent). Interviewers in Hungary had more difficulty in establishing successful cooperation among this specific subsample (15 per cent).

5.1.7.4 Summary

Figure 5.19 shows the final household-level contact, cooperation and retention rates of the panel samples after the suspension of fieldwork of Wave 8.

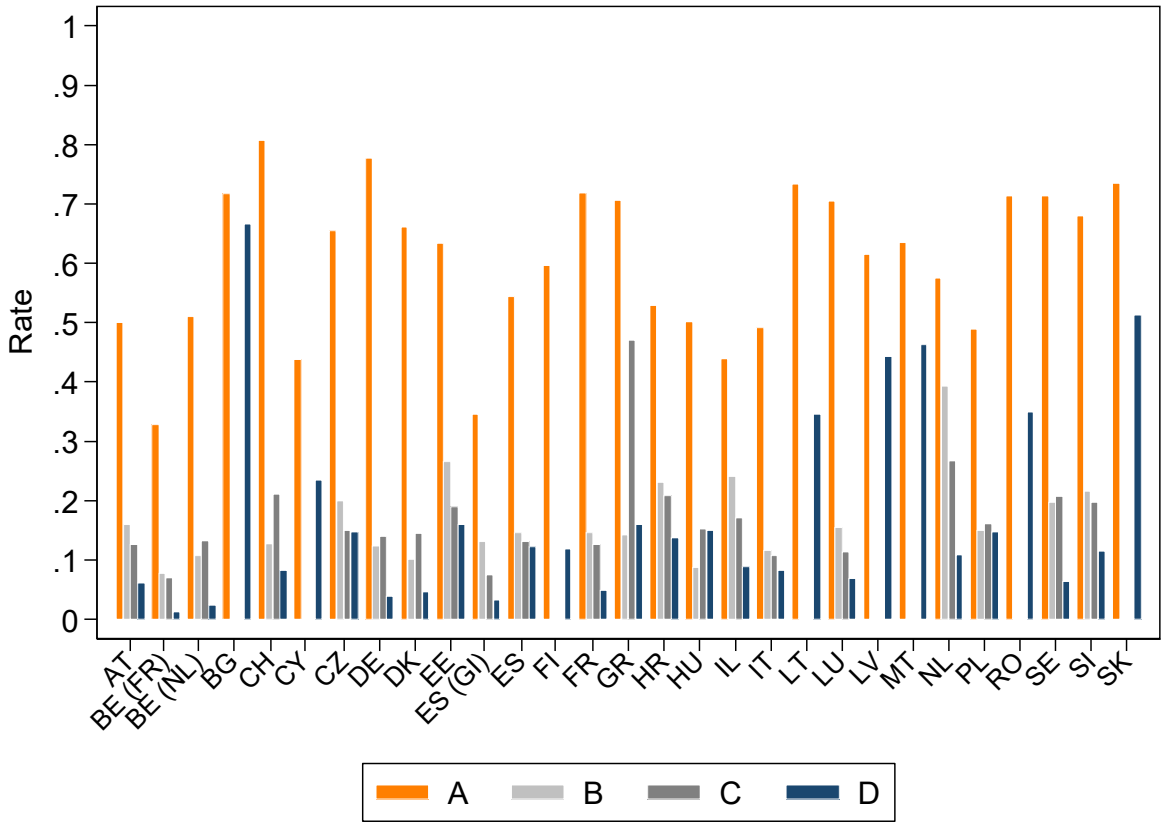
Figure 5.19: Contact, Cooperation and Retention Rates for Panel Households



The varying cooperation rates represent the upper limit for the final response/retention rates. In all countries, the contact rates are significantly higher than the cooperation and household response/retention rates, which can be attributed to the suspension of fieldwork. The highest retention rates can be observed in four countries that participated for the second time (Bulgaria, Lithuania, Romania and Slovakia) and two old ones (Germany and Switzerland). They all range between 71 and 74 per cent.

Figure 5.20 shows the final individual retention rates by subsample. Apart from the above-defined subsamples A, B and C, subsample D includes all non-responding spouses or partners and new spouses or partners that have not participated in any previous SHARE wave so far. For the new countries, i.e. Bulgaria, Cyprus, Finland, Latvia, Lithuania, Malta, Romania and Slovakia, that had only participated in one previous wave (Wave 7), the classification into subsamples B and C is not applicable yet.

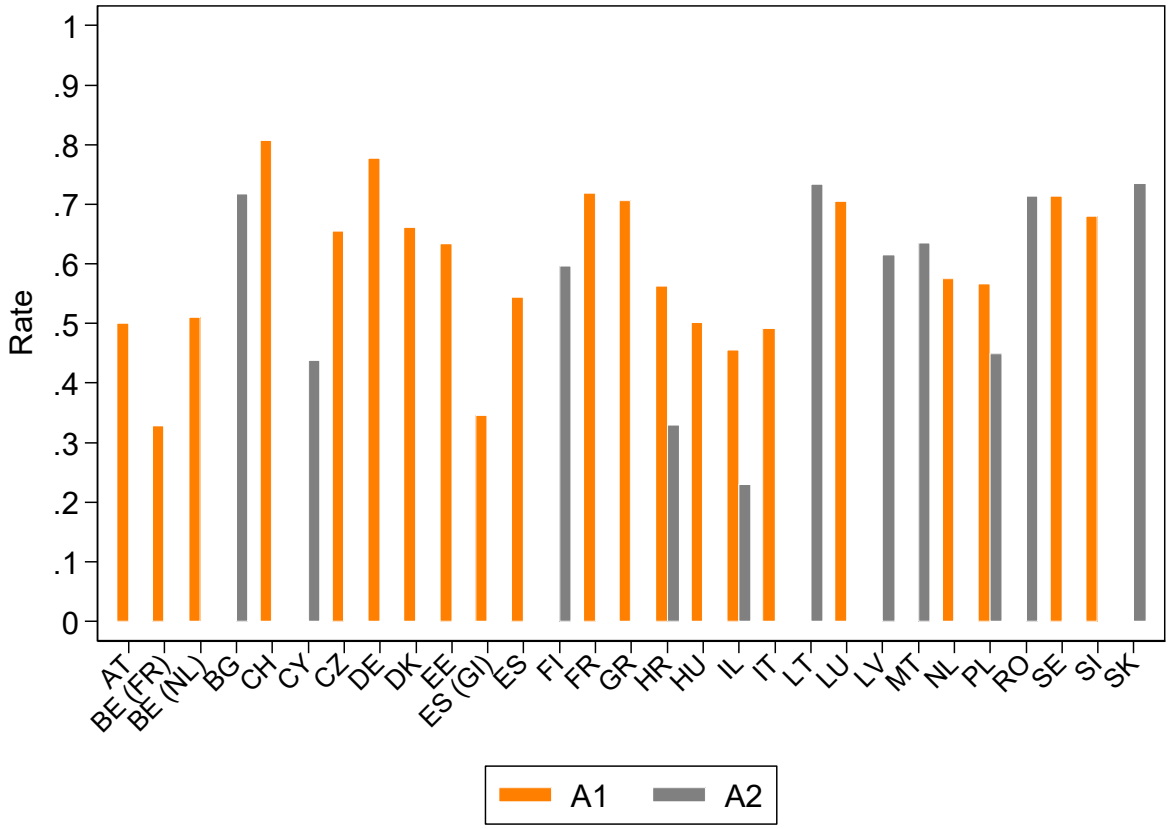
Figure 5.20: Respondent-Level Retention and Recovery for Panel Households



Across all countries, the average retention rate in subsample A (i.e. those who participated in Wave 7 and any other previous wave) is 60 per cent. Due to the discontinuation of fieldwork, the variation is high, ranging from 33 per cent in Belgium (FR) to 81 per cent in Switzerland. The Netherlands has the highest retention in subsample B (39 per cent) and a comparably high retention in subsample C (27 per cent). With regard to the latter, the Netherlands is only surpassed by Greece (47 per cent). The new countries have the highest share of interviews among spouses and partners who had never taken part in SHARE (i.e. subsample D). Of these, Bulgaria is the front-runner with 67 per cent retention.

Figure 5.21 displays individual retention by subsamples A1 and A2. 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 with a grey bar for subsample A2 had a baseline or refreshment sample in Wave 7.

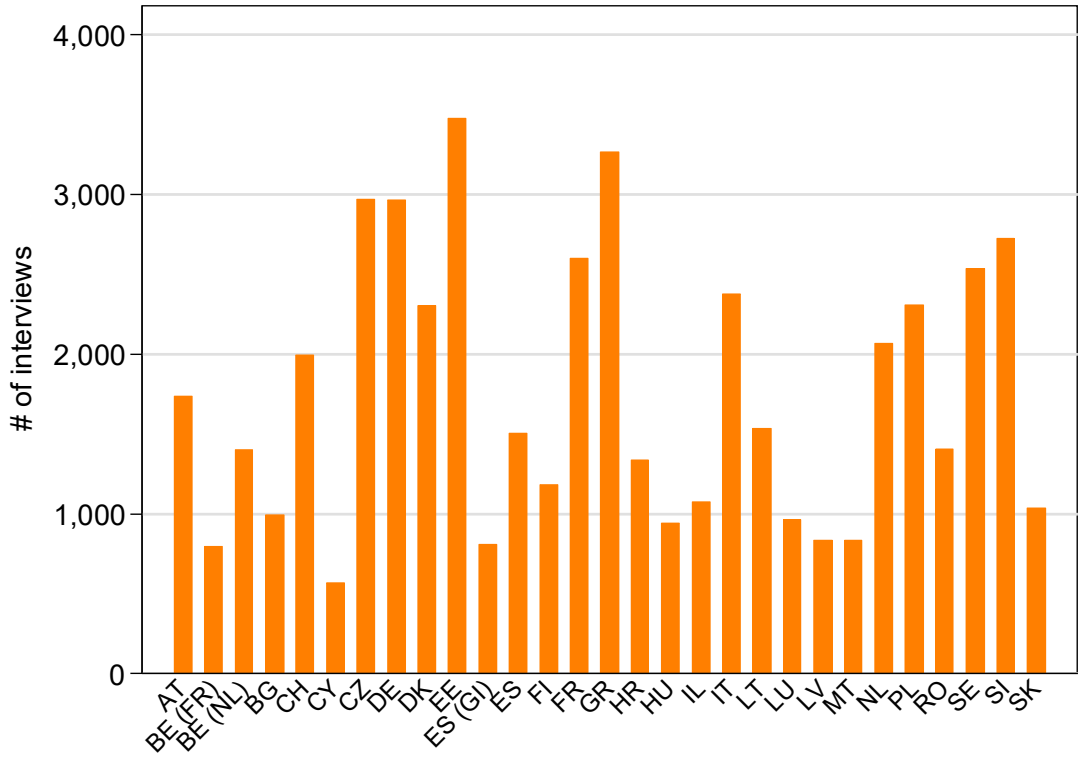
Figure 5.21: Respondent-Level Retention and Recovery for Subsamples A1 and A2



Owing to the exceptional circumstances, the final rates are not as meaningful as in previous waves. With 81 per cent, Switzerland almost reached the 85 per cent mark in subsample A1. Baseline respondents in the new Wave 7 countries are all part of subsample A2. By definition, there is no subsample A1 for them. Respondents that were part of a refreshment sample in Wave 7 are also part of subsample A2 (Croatia, Israel and Poland). As mentioned above, A2 retention is significantly higher in the countries that joined SHARE in Wave 7 and had no additional refreshment sample in Wave 8 (between 71 and 74 per cent).

Figure 5.22 shows the absolute number of panel interviews per country at the end of the fieldwork of Wave 7. Detailed breakdowns can be found in the appendix of this chapter.

Figure 5.22: Absolute Numbers of Interviews in Panel Samples



The number of completed interviews also varies with sample size in the longitudinal sample. While the Czech Republic, Estonia, Germany and Greece conducted about 3,000 or more interviews, most other countries finished fieldwork around the mark of 1,000 interviews.

5.1.8 Conclusion

Despite the exceptional circumstances, fieldwork performance was remarkable in many countries. All survey agencies managed to collect about 57,000 interviews in the regular fieldwork of Wave 8 with the help of roughly 2,000 interviewers across 27 European countries and Israel, pushing the overall numbers to about 140,000 respondents and 430,000 interviews. These numbers were augmented with almost 60,000 telephone interviews from the SHARE Corona Survey (see Section 10.1).

In Wave 8, SHARE Central benefited more than ever from large gains in efficiency by building on the conceptual framework established before Wave 5, its established software infrastructure and an effective international cooperation. As usual, we adapted the fieldwork monitoring procedures from previous waves and made some improvements. All numbers and rates are calculated bi-weekly based on formulas set by AAPOR. This standardised way of computing fieldwork outcomes allows transparency for survey agencies and comparability with other studies.

In Wave 8, we also compiled the SHARE Compliance Profiles (i.e. one of SHARE's Key Performance Indicators (KPIs)). This document, available on the SHARE website (Schuller et al., 2021), is a short evaluation report of all operative tasks in the participating countries. It usually contains a set of quality control indicators regarding the development of fieldwork, interviewer training, data transfers, and the final response and retention rates. All participating countries are evaluated on these indicators uniformly. While the known circumstances did not allow for any evaluation of contact attempts and final response and retention rates, we were able to produce the following indicators as usual.

Overall, all countries participating submitted the required input documentation and deliverables for the regular Wave 8. These deliverables include the refreshment sample (planned to be released together with the Wave 9 data) and panel gross sample data, National Training Session (NTS) dates, NTS observation protocol, NTS slides, interviewer roster, advance letters and completed Survey Agency Feedback Form (SAFF). All survey agencies submitted documentation about some kind of back-checks to validate the properness of conducted interviews. Nevertheless, the completeness of the reports and the timely submission of documentation on request could be improved in some countries. The attendance of survey agency trainers at the TTTs was very satisfactory. Data collection of both surveys was achieved in a largely synchronous fashion across all participating countries. However, there is one notable exception. Since Portugal had issues with securing funding, there was no chance of joining the regular fieldwork of Wave 8 on time. Finally, it should be noted that the envisaged end of the regular fieldwork of Wave 8 was the end of May 2020. In spite of this, some countries had almost managed to finish fieldwork regarding their panel sample before the break-off in March 2020 (e.g. Germany and Switzerland).

References

American Association for Public Opinion Research (AAPOR). (2016). Standard definitions: Final dispositions of case codes and outcome rates for surveys (9th ed). Available at: http://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 Center for the Economics of Aging (MEA).

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). MEA, Max Planck Institute for Social Law and Social Policy.

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. 101-157). MEA, Max Planck Institute for Social Law and Social Policy.

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. 94–150). MEA, Max Planck Institute for Social Law and Social Policy.

Sand, G. (2019). Fieldwork monitoring and survey participation in the seventh wave of SHARE. In M. Bergmann, A. Scherpenzeel & A. Börsch-Supan (Eds.), SHARE Wave 7 Methodology: Panel innovations and life histories (pp. 109–153). MEA, Max Planck Institute for Social Law and Social Policy.

Scherpenzeel, A., Axt, K., Bergmann, M., Douhou, S., Oepen, A., Sand, G., Schuller, K., Stuck, S., Wagner, M., & Börsch-Supan, A. (2020). Collecting survey data among the 50+ population during the COVID-19 outbreak: The Survey of Health, Ageing and Retirement in Europe (SHARE). Survey Research Methods, 14(2), 217–221. https://doi.org/10.18148/srm/2020.v14i2.7738

Schuller, K., Lasson, S., Sand, G., Bergmann, M., Pflüger, S., Hannemann, T., & Börsch-Supan, A. (2021). SHARE compliance profiles – Wave 8. MEA, Max Planck Institute for Social Law and Social Policy.

Appendix: Final Outcomes by Country, SHARE Wave 8

Austria	week 18
Baseline / refreshment sample	
Gross sample:	5248
Households attempted:	1857
Households contacted:	1373
Households estimated to be eligible:	4626.27
Households with completed coverscreen interview:	324
Households with at least one complete interview:	301
Percentage of households attempted:	35.38 %
Contact rate:	24.92 %
Cooperation rate:	26.11 %
Household response rate:	6.51 %
Refusal rate:	11.63 %
Other non-interview rate:	6.79 %
Individual interviews:	383
Estimated average number of eligibles in hh:	1.60
Individual response rate:	5.17 %
Median number of attempts for not successfully contacted hh:	2
Number of interviewers in SampleCTRL:	74
Rate of active interviewers:	71.62 %
Longitudinal sample	
Gross sample:	3086
Households attempted:	3044
Households contacted:	2796
Households estimated to be eligible:	3070.79
Households with completed coverscreen interview:	2223
Households with at least one complete interview:	1254
Percentage of households attempted:	98.64 %
Contact rate:	90.56 %
Cooperation rate:	45.09 %
Household response rate:	40.84 %
Refusal rate:	16.02 %
Other non-interview rate:	33.70 %
Individual interviews:	1738
Sample A:	1575
Sample A1:	1575
Sample A2:	0
Sample B:	13
Sample C:	142
Sample D:	8
Estimated average number of eligibles in hh:	1.46
Individual response rate:	38.72 %
Sample A:	49.98 %
Sample A1:	49.98 %
Sample A2:	. %
Sample B:	16.05 %
Sample C:	12.60 %
Sample D:	6.15 %
Median number of attempts for not successfully contacted hh:	2
Number of interviewers in SampleCTRL:	74
Rate of active interviewers:	71.62 %

Belgium (FR)	week 18
Baseline / refreshment sample	
Gross sample:	2234
Households attempted:	450
Households contacted:	360
Households estimated to be eligible:	2154.57
Households with completed coverscreen interview:	117
Households with at least one complete interview:	115
Percentage of households attempted:	20.14 %
Contact rate:	15.97 %
Cooperation rate:	33.43 %
Household response rate:	5.34 %
Refusal rate:	6.59 %
Other non-interview rate:	4.04 %
Individual interviews:	131
Estimated average number of eligibles in hh:	1.34
Individual response rate:	4.55 %
Median number of attempts for not successfully contacted hh:	1
Number of interviewers in SampleCTRL:	76
Rate of active interviewers:	63.16 %
Longitudinal sample	
Gross sample:	2229
Households attempted:	2175
Households contacted:	1926
Households estimated to be eligible:	2222.85
Households with completed coverscreen interview:	1485
Households with at least one complete interview:	625
Percentage of households attempted:	97.58 %
Contact rate:	86.38 %
Cooperation rate:	32.55 %
Household response rate:	28.12 %
Refusal rate:	17.50 %
Other non-interview rate:	40.76 %
Individual interviews:	798
Sample A:	736
Sample A1:	736
Sample A2:	0
Sample B:	13
Sample C:	45
Sample D:	4
Estimated average number of eligibles in hh:	1.51
Individual response rate:	23.78 %
Sample A:	32.89 %
Sample A1:	32.89 %
Sample A2:	. %
Sample B:	7.78 %
Sample C:	7.03 %
Sample D:	1.29 %
Median number of attempts for not successfully contacted hh:	2
Number of interviewers in SampleCTRL:	76
Rate of active interviewers:	63.16 %

Belgium (NL)	week 18
Baseline / refreshment sample	
Gross sample:	1067
Households attempted:	484
Households contacted:	326
Households estimated to be eligible:	1033.93
Households with completed coverscreen interview:	103
Households with at least one complete interview:	101
Percentage of households attempted:	45.36 %
Contact rate:	30.08 %
Cooperation rate:	32.48 %
Household response rate:	9.77 %
Refusal rate:	14.31 %
Other non-interview rate:	6.00 %
Individual interviews:	136
Estimated average number of eligibles in hh:	1.63
Individual response rate:	8.06 %
Median number of attempts for not successfully contacted hh:	1
Number of interviewers in SampleCTRL:	50
Rate of active interviewers:	96.00 %
Longitudinal sample	
Gross sample:	2073
Households attempted:	2066
Households contacted:	2018
Households estimated to be eligible:	2065.98
Households with completed coverscreen interview:	1744
Households with at least one complete interview:	1008
Percentage of households attempted:	99.66 %
Contact rate:	97.34 %
Cooperation rate:	50.12 %
Household response rate:	48.79 %
Refusal rate:	12.92 %
Other non-interview rate:	35.62 %
Individual interviews:	1401
Sample A:	1341
Sample A1:	1341
Sample A2:	0
Sample B:	15
Sample C:	40
Sample D:	5
Estimated average number of eligibles in hh:	1.59
Individual response rate:	42.76 %
Sample A:	51.09 %
Sample A1:	51.09 %
Sample A2:	. %
Sample B:	10.79 %
Sample C:	13.29 %
Sample D:	2.38 %
Median number of attempts for not successfully contacted hh:	6
Number of interviewers in SampleCTRL:	50
Rate of active interviewers:	96.00 %

Bulgaria	week 18
Longitudinal sample	
Gross sample:	937
Households attempted:	909
Households contacted:	865
Households estimated to be eligible:	925.66
Households with completed coverscreen interview:	799
Households with at least one complete interview:	679
Percentage of households attempted:	97.01 %
Contact rate:	92.26 %
Cooperation rate:	79.51 %
Household response rate:	73.35 %
Refusal rate:	5.08 %
Other non-interview rate:	13.83 %
Individual interviews:	993
Sample A:	975
Sample A1:	0
Sample A2:	975
Sample B:	0
Sample C:	0
Sample D:	18
Estimated average number of eligibles in hh:	1.50
Individual response rate:	71.59 %
Sample A:	71.74 %
Sample A1:	. %
Sample A2:	71.74 %
Sample B:	.
Sample C:	.
Sample D:	66.67 %
Median number of attempts for not successfully contacted hh:	2
Number of interviewers in SampleCTRL:	90
Rate of active interviewers:	43.33 %

Switzerland	week 18
Baseline / refreshment sample	
Gross sample:	637
Households attempted:	617
Households contacted:	489
Households estimated to be eligible:	590.54
Households with completed coverscreen interview:	149
Households with at least one complete interview:	145
Percentage of households attempted:	96.86 %
Contact rate:	75.19 %
Cooperation rate:	32.66 %
Household response rate:	24.55 %
Refusal rate:	35.90 %
Other non-interview rate:	14.73 %
Individual interviews:	186
Estimated average number of eligibles in hh:	1.61
Individual response rate:	19.50 %
Median number of attempts for not successfully contacted hh:	3
Number of interviewers in SampleCTRL:	90
Rate of active interviewers:	73.33 %
Longitudinal sample	
Gross sample:	2005
Households attempted:	2004
Households contacted:	1992
Households estimated to be eligible:	1995.00
Households with completed coverscreen interview:	1643
Households with at least one complete interview:	1479
Percentage of households attempted:	99.95 %
Contact rate:	99.35 %
Cooperation rate:	74.62 %
Household response rate:	74.14 %
Refusal rate:	19.20 %
Other non-interview rate:	6.02 %
Individual interviews:	1995
Sample A:	1863
Sample A1:	1863
Sample A2:	0
Sample B:	23
Sample C:	88
Sample D:	21
Estimated average number of eligibles in hh:	1.59
Individual response rate:	63.06 %
Sample A:	80.65 %
Sample A1:	80.65 %
Sample A2:	. %
Sample B:	12.71 %
Sample C:	21.15 %
Sample D:	8.24 %
Median number of attempts for not successfully contacted hh:	16
Number of interviewers in SampleCTRL:	90
Rate of active interviewers:	73.33 %

Cyprus	week 18
Longitudinal sample	
Gross sample:	835
Households attempted:	823
Households contacted:	756
Households estimated to be eligible:	832.97
Households with completed coverscreen interview:	660
Households with at least one complete interview:	386
Percentage of households attempted:	98.56 %
Contact rate:	90.52 %
Cooperation rate:	51.19 %
Household response rate:	46.34 %
Refusal rate:	9.48 %
Other non-interview rate:	34.70 %
Individual interviews:	570
Sample A:	534
Sample A1:	0
Sample A2:	534
Sample B:	0
Sample C:	0
Sample D:	36
Estimated average number of eligibles in hh:	1.65
Individual response rate:	41.58 %
Sample A:	43.84 %
Sample A1:	. %
Sample A2:	43.84 %
Sample B:	.
Sample C:	.
Sample D:	23.53 %
Median number of attempts for not successfully contacted hh:	4
Number of interviewers in SampleCTRL:	18
Rate of active interviewers:	83.33 %

Czech Republic	week 18
Baseline / refreshment sample	
Gross sample:	4801
Households attempted:	1361
Households contacted:	1148
Households estimated to be eligible:	2577.85
Households with completed coverscreen interview:	245
Households with at least one complete interview:	232
Percentage of households attempted:	28.35 %
Contact rate:	23.88 %
Cooperation rate:	37.69 %
Household response rate:	9.00 %
Refusal rate:	9.22 %
Other non-interview rate:	5.66 %
Individual interviews:	328
Estimated average number of eligibles in hh:	1.53
Individual response rate:	8.33 %
Median number of attempts for not successfully contacted hh:	0
Number of interviewers in SampleCTRL:	119
Rate of active interviewers:	82.35 %
Longitudinal sample	
Gross sample:	3762
Households attempted:	3723
Households contacted:	3695
Households estimated to be eligible:	3753.92
Households with completed coverscreen interview:	2505
Households with at least one complete interview:	2048
Percentage of households attempted:	98.96 %
Contact rate:	98.22 %
Cooperation rate:	55.55 %
Household response rate:	54.56 %
Refusal rate:	22.78 %
Other non-interview rate:	20.88 %
Individual interviews:	2968
Sample A:	2740
Sample A1:	2740
Sample A2:	0
Sample B:	17
Sample C:	193
Sample D:	18
Estimated average number of eligibles in hh:	1.51
Individual response rate:	52.27 %
Sample A:	65.49 %
Sample A1:	65.49 %
Sample A2:	. %
Sample B:	20.00 %
Sample C:	15.04 %
Sample D:	14.75 %
Median number of attempts for not successfully contacted hh:	8
Number of interviewers in SampleCTRL:	119
Rate of active interviewers:	82.35 %

Germany	week 18
Baseline / refreshment sample	
Gross sample:	7920
Households attempted:	5378
Households contacted:	4313
Households estimated to be eligible:	7523.85
Households with completed coverscreen interview:	773
Households with at least one complete interview:	757
Percentage of households attempted:	67.90 %
Contact rate:	53.75 %
Cooperation rate:	18.72 %
Household response rate:	10.06 %
Refusal rate:	33.43 %
Other non-interview rate:	10.26 %
Individual interviews:	979
Estimated average number of eligibles in hh:	1.67
Individual response rate:	7.77 %
Median number of attempts for not successfully contacted hh:	1
Number of interviewers in SampleCTRL:	215
Rate of active interviewers:	48.84 %
Longitudinal sample	
Gross sample:	2938
Households attempted:	2928
Households contacted:	2916
Households estimated to be eligible:	2839.67
Households with completed coverscreen interview:	2311
Households with at least one complete interview:	2006
Percentage of households attempted:	99.66 %
Contact rate:	99.24 %
Cooperation rate:	71.19 %
Household response rate:	70.64 %
Refusal rate:	17.29 %
Other non-interview rate:	11.30 %
Individual interviews:	2963
Sample A:	2865
Sample A1:	2865
Sample A2:	0
Sample B:	18
Sample C:	69
Sample D:	11
Estimated average number of eligibles in hh:	1.62
Individual response rate:	64.32 %
Sample A:	77.73 %
Sample A1:	77.73 %
Sample A2:	. %
Sample B:	12.33 %
Sample C:	13.94 %
Sample D:	3.93 %
Median number of attempts for not successfully contacted hh:	2
Number of interviewers in SampleCTRL:	215
Rate of active interviewers:	48.84 %

Denmark	week 18
Baseline / refreshment sample	
Gross sample:	622
Households attempted:	498
Households contacted:	451
Households estimated to be eligible:	607.01
Households with completed coverscreen interview:	161
Households with at least one complete interview:	153
Percentage of households attempted:	80.06 %
Contact rate:	72.32 %
Cooperation rate:	34.85 %
Household response rate:	25.21 %
Refusal rate:	30.97 %
Other non-interview rate:	16.14 %
Individual interviews:	193
Estimated average number of eligibles in hh:	1.56
Individual response rate:	20.43 %
Median number of attempts for not successfully contacted hh:	2
Number of interviewers in SampleCTRL:	66
Rate of active interviewers:	95.45 %
Longitudinal sample	
Gross sample:	2922
Households attempted:	2861
Households contacted:	2799
Households estimated to be eligible:	2919.96
Households with completed coverscreen interview:	2109
Households with at least one complete interview:	1676
Percentage of households attempted:	97.91 %
Contact rate:	95.79 %
Cooperation rate:	59.92 %
Household response rate:	57.40 %
Refusal rate:	22.50 %
Other non-interview rate:	15.89 %
Individual interviews:	2305
Sample A:	2135
Sample A1:	2135
Sample A2:	0
Sample B:	27
Sample C:	125
Sample D:	18
Estimated average number of eligibles in hh:	1.63
Individual response rate:	48.57 %
Sample A:	66.12 %
Sample A1:	66.12 %
Sample A2:	. %
Sample B:	10.11 %
Sample C:	14.47 %
Sample D:	4.66 %
Median number of attempts for not successfully contacted hh:	2
Number of interviewers in SampleCTRL:	66
Rate of active interviewers:	95.45 %

Estonia	week 18
Baseline / refreshment sample	
Gross sample:	1155
Households attempted:	697
Households contacted:	574
Households estimated to be eligible:	1101.97
Households with completed coverscreen interview:	280
Households with at least one complete interview:	279
Percentage of households attempted:	60.35 %
Contact rate:	49.18 %
Cooperation rate:	51.48 %
Household response rate:	25.32 %
Refusal rate:	12.34 %
Other non-interview rate:	11.52 %
Individual interviews:	380
Estimated average number of eligibles in hh:	1.53
Individual response rate:	22.61 %
Median number of attempts for not successfully contacted hh:	1
Number of interviewers in SampleCTRL:	97
Rate of active interviewers:	75.26 %
Longitudinal sample	
Gross sample:	4373
Households attempted:	4337
Households contacted:	4247
Households estimated to be eligible:	4364.93
Households with completed coverscreen interview:	3776
Households with at least one complete interview:	2519
Percentage of households attempted:	99.18 %
Contact rate:	97.11 %
Cooperation rate:	59.42 %
Household response rate:	57.71 %
Refusal rate:	10.40 %
Other non-interview rate:	29.00 %
Individual interviews:	3475
Sample A:	3229
Sample A1:	3229
Sample A2:	0
Sample B:	29
Sample C:	205
Sample D:	12
Estimated average number of eligibles in hh:	1.46
Individual response rate:	54.64 %
Sample A:	63.35 %
Sample A1:	63.35 %
Sample A2:	. %
Sample B:	26.61 %
Sample C:	19.00 %
Sample D:	16.00 %
Median number of attempts for not successfully contacted hh:	4
Number of interviewers in SampleCTRL:	97
Rate of active interviewers:	75.26 %

Girona	week 18
Longitudinal sample	
Gross sample:	1892
Households attempted:	1649
Households contacted:	1489
Households estimated to be eligible:	1885.12
Households with completed coverscreen interview:	662
Households with at least one complete interview:	606
Percentage of households attempted:	87.16 %
Contact rate:	78.67 %
Cooperation rate:	40.86 %
Household response rate:	32.15 %
Refusal rate:	24.88 %
Other non-interview rate:	21.64 %
Individual interviews:	809
Sample A:	732
Sample A1:	732
Sample A2:	0
Sample B:	15
Sample C:	60
Sample D:	2
Estimated average number of eligibles in hh:	1.64
Individual response rate:	26.12 %
Sample A:	34.53 %
Sample A1:	34.53 %
Sample A2:	. %
Sample B:	13.16 %
Sample C:	7.45 %
Sample D:	3.33 %
Median number of attempts for not successfully contacted hh:	2
Number of interviewers in SampleCTRL:	36
Rate of active interviewers:	41.67 %

Spain	week 18
Longitudinal sample	
Gross sample:	2182
Households attempted:	2160
Households contacted:	2053
Households estimated to be eligible:	2180.99
Households with completed coverscreen interview:	1719
Households with at least one complete interview:	1018
Percentage of households attempted:	98.99 %
Contact rate:	94.09 %
Cooperation rate:	49.61 %
Household response rate:	46.68 %
Refusal rate:	19.62 %
Other non-interview rate:	27.79 %
Individual interviews:	1504
Sample A:	1391
Sample A1:	1391
Sample A2:	0
Sample B:	15
Sample C:	93
Sample D:	5
Estimated average number of eligibles in hh:	1.56
Individual response rate:	44.07 %
Sample A:	54.36 %
Sample A1:	54.36 %
Sample A2:	. %
Sample B:	14.56 %
Sample C:	13.10 %
Sample D:	12.20 %
Median number of attempts for not successfully contacted hh:	4
Number of interviewers in SampleCTRL:	68
Rate of active interviewers:	80.88 %

Finland	week 18
Longitudinal sample	
Gross sample:	1344
Households attempted:	1330
Households contacted:	1303
Households estimated to be eligible:	1340.97
Households with completed coverscreen interview:	1094
Households with at least one complete interview:	817
Percentage of households attempted:	98.96 %
Contact rate:	96.94 %
Cooperation rate:	62.85 %
Household response rate:	60.93 %
Refusal rate:	17.75 %
Other non-interview rate:	18.27 %
Individual interviews:	1182
Sample A:	1148
Sample A1:	0
Sample A2:	1148
Sample B:	0
Sample C:	0
Sample D:	34
Estimated average number of eligibles in hh:	1.65
Individual response rate:	53.46 %
Sample A:	59.67 %
Sample A1:	. %
Sample A2:	59.67 %
Sample B:	.
Sample C:	.
Sample D:	11.85 %
Median number of attempts for not successfully contacted hh:	10
Number of interviewers in SampleCTRL:	64
Rate of active interviewers:	45.31 %

France	week 18
Baseline / refreshment sample	
Gross sample:	2448
Households attempted:	1844
Households contacted:	1412
Households estimated to be eligible:	2097.03
Households with completed coverscreen interview:	430
Households with at least one complete interview:	391
Percentage of households attempted:	75.33 %
Contact rate:	57.60 %
Cooperation rate:	32.37 %
Household response rate:	18.65 %
Refusal rate:	21.70 %
Other non-interview rate:	17.25 %
Individual interviews:	532
Estimated average number of eligibles in hh:	1.54
Individual response rate:	16.50 %
Median number of attempts for not successfully contacted hh:	0
Number of interviewers in SampleCTRL:	141
Rate of active interviewers:	87.23 %
Longitudinal sample	
Gross sample:	3549
Households attempted:	3537
Households contacted:	3432
Households estimated to be eligible:	3531.94
Households with completed coverscreen interview:	2169
Households with at least one complete interview:	1871
Percentage of households attempted:	99.66 %
Contact rate:	96.69 %
Cooperation rate:	54.79 %
Household response rate:	52.97 %
Refusal rate:	23.41 %
Other non-interview rate:	20.30 %
Individual interviews:	2601
Sample A:	2347
Sample A1:	2347
Sample A2:	0
Sample B:	30
Sample C:	215
Sample D:	9
Estimated average number of eligibles in hh:	1.52
Individual response rate:	48.56 %
Sample A:	71.80 %
Sample A1:	71.80 %
Sample A2:	. %
Sample B:	14.56 %
Sample C:	12.67 %
Sample D:	4.89 %
Median number of attempts for not successfully contacted hh:	5
Number of interviewers in SampleCTRL:	141
Rate of active interviewers:	87.23 %

Greece	week 18
Longitudinal sample	
Gross sample:	3587
Households attempted:	3582
Households contacted:	3541
Households estimated to be eligible:	3563.97
Households with completed coverscreen interview:	2864
Households with at least one complete interview:	2160
Percentage of households attempted:	99.86 %
Contact rate:	98.71 %
Cooperation rate:	61.40 %
Household response rate:	60.61 %
Refusal rate:	22.22 %
Other non-interview rate:	15.88 %
Individual interviews:	3267
Sample A:	2155
Sample A1:	2155
Sample A2:	0
Sample B:	10
Sample C:	1068
Sample D:	34
Estimated average number of eligibles in hh:	1.57
Individual response rate:	58.28 %
Sample A:	70.61 %
Sample A1:	70.61 %
Sample A2:	. %
Sample B:	14.29 %
Sample C:	47.03 %
Sample D:	15.96 %
Median number of attempts for not successfully contacted hh:	4
Number of interviewers in SampleCTRL:	123
Rate of active interviewers:	98.37 %

Croatia	week 18
Baseline / refreshment sample	
Gross sample:	7500
Households attempted:	2060
Households contacted:	1639
Households estimated to be eligible:	6218.45
Households with completed coverscreen interview:	606
Households with at least one complete interview:	573
Percentage of households attempted:	27.47 %
Contact rate:	20.70 %
Cooperation rate:	44.52 %
Household response rate:	9.21 %
Refusal rate:	8.04 %
Other non-interview rate:	3.44 %
Individual interviews:	835
Estimated average number of eligibles in hh:	1.57
Individual response rate:	8.56 %
Median number of attempts for not successfully contacted hh:	1
Number of interviewers in SampleCTRL:	114
Rate of active interviewers:	67.54 %
Longitudinal sample	
Gross sample:	1761
Households attempted:	1710
Households contacted:	1678
Households estimated to be eligible:	1728.05
Households with completed coverscreen interview:	1494
Households with at least one complete interview:	869
Percentage of households attempted:	97.10 %
Contact rate:	95.25 %
Cooperation rate:	52.79 %
Household response rate:	50.29 %
Refusal rate:	9.90 %
Other non-interview rate:	35.07 %
Individual interviews:	1337
Sample A:	1256
Sample A1:	1144
Sample A2:	112
Sample B:	3
Sample C:	66
Sample D:	12
Estimated average number of eligibles in hh:	1.62
Individual response rate:	47.90 %
Sample A:	52.88 %
Sample A1:	56.22 %
Sample A2:	32.94 %
Sample B:	23.08 %
Sample C:	20.89 %
Sample D:	13.79 %
Median number of attempts for not successfully contacted hh:	1
Number of interviewers in SampleCTRL:	114
Rate of active interviewers:	67.54 %

Hungary	week 18
Baseline / refreshment sample	
Gross sample:	1529
Households attempted:	805
Households contacted:	723
Households estimated to be eligible:	1316.27
Households with completed coverscreen interview:	340
Households with at least one complete interview:	300
Percentage of households attempted:	52.65 %
Contact rate:	46.42 %
Cooperation rate:	49.10 %
Household response rate:	22.79 %
Refusal rate:	17.55 %
Other non-interview rate:	6.08 %
Individual interviews:	427
Estimated average number of eligibles in hh:	1.46
Individual response rate:	22.26 %
Median number of attempts for not successfully contacted hh:	1
Number of interviewers in SampleCTRL:	90
Rate of active interviewers:	66.67 %
Longitudinal sample	
Gross sample:	1777
Households attempted:	1613
Households contacted:	1442
Households estimated to be eligible:	1769.29
Households with completed coverscreen interview:	1070
Households with at least one complete interview:	661
Percentage of households attempted:	90.77 %
Contact rate:	81.11 %
Cooperation rate:	46.06 %
Household response rate:	37.36 %
Refusal rate:	15.94 %
Other non-interview rate:	27.81 %
Individual interviews:	941
Sample A:	763
Sample A1:	763
Sample A2:	0
Sample B:	3
Sample C:	163
Sample D:	12
Estimated average number of eligibles in hh:	1.53
Individual response rate:	34.74 %
Sample A:	50.10 %
Sample A1:	50.10 %
Sample A2:	. %
Sample B:	8.82 %
Sample C:	15.19 %
Sample D:	15.00 %
Median number of attempts for not successfully contacted hh:	1
Number of interviewers in SampleCTRL:	90
Rate of active interviewers:	66.67 %

Israel	week 18
Baseline / refreshment sample	
Gross sample:	651
Households attempted:	412
Households contacted:	399
Households estimated to be eligible:	647.71
Households with completed coverscreen interview:	384
Households with at least one complete interview:	375
Percentage of households attempted:	63.29 %
Contact rate:	61.29 %
Cooperation rate:	94.46 %
Household response rate:	57.90 %
Refusal rate:	2.62 %
Other non-interview rate:	0.77 %
Individual interviews:	476
Estimated average number of eligibles in hh:	1.57
Individual response rate:	46.83 %
Median number of attempts for not successfully contacted hh:	0
Number of interviewers in SampleCTRL:	24
Rate of active interviewers:	37.50 %
Longitudinal sample	
Gross sample:	1886
Households attempted:	1846
Households contacted:	1635
Households estimated to be eligible:	1878.85
Households with completed coverscreen interview:	1296
Households with at least one complete interview:	756
Percentage of households attempted:	97.88 %
Contact rate:	86.65 %
Cooperation rate:	46.44 %
Household response rate:	40.24 %
Refusal rate:	20.28 %
Other non-interview rate:	26.13 %
Individual interviews:	1077
Sample A:	929
Sample A1:	894
Sample A2:	35
Sample B:	27
Sample C:	110
Sample D:	11
Estimated average number of eligibles in hh:	1.59
Individual response rate:	35.99 %
Sample A:	43.92 %
Sample A1:	45.54 %
Sample A2:	23.03 %
Sample B:	24.11 %
Sample C:	17.16 %
Sample D:	8.94 %
Median number of attempts for not successfully contacted hh:	2
Number of interviewers in SampleCTRL:	24
Rate of active interviewers:	37.50 %

Italy	week 18
Longitudinal sample	
Gross sample:	3617
Households attempted:	3497
Households contacted:	3313
Households estimated to be eligible:	3607.69
Households with completed coverscreen interview:	2719
Households with at least one complete interview:	1534
Percentage of households attempted:	96.68 %
Contact rate:	91.58 %
Cooperation rate:	46.43 %
Household response rate:	42.52 %
Refusal rate:	15.30 %
Other non-interview rate:	33.76 %
Individual interviews:	2376
Sample A:	2229
Sample A1:	2229
Sample A2:	0
Sample B:	12
Sample C:	120
Sample D:	15
Estimated average number of eligibles in hh:	1.65
Individual response rate:	40.01 %
Sample A:	49.17 %
Sample A1:	49.17 %
Sample A2:	. %
Sample B:	11.65 %
Sample C:	10.70 %
Sample D:	8.29 %
Median number of attempts for not successfully contacted hh:	3
Number of interviewers in SampleCTRL:	157
Rate of active interviewers:	78.34 %

Lithuania	week 18
Longitudinal sample	
Gross sample:	1533
Households attempted:	1498
Households contacted:	1456
Households estimated to be eligible:	1516.63
Households with completed coverscreen interview:	1238
Households with at least one complete interview:	1098
Percentage of households attempted:	97.72 %
Contact rate:	94.95 %
Cooperation rate:	76.25 %
Household response rate:	72.40 %
Refusal rate:	8.97 %
Other non-interview rate:	13.58 %
Individual interviews:	1536
Sample A:	1461
Sample A1:	0
Sample A2:	1461
Sample B:	0
Sample C:	0
Sample D:	75
Estimated average number of eligibles in hh:	1.46
Individual response rate:	69.55 %
Sample A:	73.34 %
Sample A1:	. %
Sample A2:	73.34 %
Sample B:	.
Sample C:	.
Sample D:	34.56 %
Median number of attempts for not successfully contacted hh:	3
Number of interviewers in SampleCTRL:	50
Rate of active interviewers:	74.00 %

Luxembourg	week 18
Longitudinal sample	
Gross sample:	1425
Households attempted:	1414
Households contacted:	1389
Households estimated to be eligible:	1377.63
Households with completed coverscreen interview:	864
Households with at least one complete interview:	698
Percentage of households attempted:	99.23 %
Contact rate:	97.41 %
Cooperation rate:	52.01 %
Household response rate:	50.67 %
Refusal rate:	28.53 %
Other non-interview rate:	18.22 %
Individual interviews:	963
Sample A:	856
Sample A1:	856
Sample A2:	0
Sample B:	7
Sample C:	73
Sample D:	27
Estimated average number of eligibles in hh:	1.66
Individual response rate:	42.03 %
Sample A:	70.45 %
Sample A1:	70.45 %
Sample A2:	. %
Sample B:	15.56 %
Sample C:	11.41 %
Sample D:	6.92 %
Median number of attempts for not successfully contacted hh:	3
Number of interviewers in SampleCTRL:	29
Rate of active interviewers:	79.31 %

Latvia	week 18
Baseline / refreshment sample	
Gross sample:	1424
Households attempted:	732
Households contacted:	607
Households estimated to be eligible:	1176.60
Households with completed coverscreen interview:	306
Households with at least one complete interview:	303
Percentage of households attempted:	51.40 %
Contact rate:	42.61 %
Cooperation rate:	60.43 %
Household response rate:	25.75 %
Refusal rate:	10.22 %
Other non-interview rate:	6.64 %
Individual interviews:	421
Estimated average number of eligibles in hh:	1.44
Individual response rate:	24.92 %
Median number of attempts for not successfully contacted hh:	0
Number of interviewers in SampleCTRL:	39
Rate of active interviewers:	56.41 %
Longitudinal sample	
Gross sample:	969
Households attempted:	957
Households contacted:	934
Households estimated to be eligible:	965.96
Households with completed coverscreen interview:	839
Households with at least one complete interview:	592
Percentage of households attempted:	98.76 %
Contact rate:	96.38 %
Cooperation rate:	63.59 %
Household response rate:	61.29 %
Refusal rate:	8.59 %
Other non-interview rate:	26.50 %
Individual interviews:	836
Sample A:	793
Sample A1:	0
Sample A2:	793
Sample B:	0
Sample C:	0
Sample D:	43
Estimated average number of eligibles in hh:	1.43
Individual response rate:	60.36 %
Sample A:	61.57 %
Sample A1:	. %
Sample A2:	61.57 %
Sample B:	.
Sample C:	.
Sample D:	44.33 %
Median number of attempts for not successfully contacted hh:	4
Number of interviewers in SampleCTRL:	39
Rate of active interviewers:	56.41 %

Malta	week 18
Longitudinal sample	
Gross sample:	792
Households attempted:	786
Households contacted:	748
Households estimated to be eligible:	789.98
Households with completed coverscreen interview:	652
Households with at least one complete interview:	494
Percentage of households attempted:	99.24 %
Contact rate:	94.43 %
Cooperation rate:	66.22 %
Household response rate:	62.53 %
Refusal rate:	11.52 %
Other non-interview rate:	20.38 %
Individual interviews:	834
Sample A:	796
Sample A1:	0
Sample A2:	796
Sample B:	0
Sample C:	0
Sample D:	38
Estimated average number of eligibles in hh:	1.69
Individual response rate:	62.47 %
Sample A:	63.53 %
Sample A1:	. %
Sample A2:	63.53 %
Sample B:	.
Sample C:	.
Sample D:	46.34 %
Median number of attempts for not successfully contacted hh:	2
Number of interviewers in SampleCTRL:	16
Rate of active interviewers:	56.25 %

Netherlands	week 18
Longitudinal sample	
Gross sample:	3439
Households attempted:	3267
Households contacted:	3077
Households estimated to be eligible:	3425.32
Households with completed coverscreen interview:	1581
Households with at least one complete interview:	1373
Percentage of households attempted:	95.00 %
Contact rate:	89.45 %
Cooperation rate:	44.81 %
Household response rate:	40.08 %
Refusal rate:	28.46 %
Other non-interview rate:	20.90 %
Individual interviews:	2067
Sample A:	1138
Sample A1:	1138
Sample A2:	0
Sample B:	100
Sample C:	780
Sample D:	49
Estimated average number of eligibles in hh:	1.64
Individual response rate:	36.82 %
Sample A:	57.47 %
Sample A1:	57.47 %
Sample A2:	. %
Sample B:	39.37 %
Sample C:	26.66 %
Sample D:	10.86 %
Median number of attempts for not successfully contacted hh:	2
Number of interviewers in SampleCTRL:	95
Rate of active interviewers:	29.47 %

Poland	week 18
Baseline / refreshment sample	
Gross sample:	3300
Households attempted:	2109
Households contacted:	1629
Households estimated to be eligible:	3131.01
Households with completed coverscreen interview:	513
Households with at least one complete interview:	498
Percentage of households attempted:	63.91 %
Contact rate:	48.58 %
Cooperation rate:	32.74 %
Household response rate:	15.91 %
Refusal rate:	15.87 %
Other non-interview rate:	16.80 %
Individual interviews:	773
Estimated average number of eligibles in hh:	1.66
Individual response rate:	14.83 %
Median number of attempts for not successfully contacted hh:	1
Number of interviewers in SampleCTRL:	114
Rate of active interviewers:	86.84 %
Longitudinal sample	
Gross sample:	3175
Households attempted:	3051
Households contacted:	2792
Households estimated to be eligible:	3168.76
Households with completed coverscreen interview:	2403
Households with at least one complete interview:	1531
Percentage of households attempted:	96.09 %
Contact rate:	87.92 %
Cooperation rate:	54.95 %
Household response rate:	48.32 %
Refusal rate:	8.96 %
Other non-interview rate:	30.64 %
Individual interviews:	2307
Sample A:	2217
Sample A1:	860
Sample A2:	1357
Sample B:	3
Sample C:	27
Sample D:	60
Estimated average number of eligibles in hh:	1.62
Individual response rate:	44.95 %
Sample A:	48.86 %
Sample A1:	56.69 %
Sample A2:	44.93 %
Sample B:	15.00 %
Sample C:	16.07 %
Sample D:	14.71 %
Median number of attempts for not successfully contacted hh:	3
Number of interviewers in SampleCTRL:	114
Rate of active interviewers:	86.84 %

Romania	week 18
Longitudinal sample	
Gross sample:	1308
Households attempted:	1297
Households contacted:	1260
Households estimated to be eligible:	1290.86
Households with completed coverscreen interview:	1176
Households with at least one complete interview:	900
Percentage of households attempted:	99.16 %
Contact rate:	96.29 %
Cooperation rate:	72.41 %
Household response rate:	69.72 %
Refusal rate:	6.20 %
Other non-interview rate:	20.37 %
Individual interviews:	1407
Sample A:	1364
Sample A1:	0
Sample A2:	1364
Sample B:	0
Sample C:	0
Sample D:	43
Estimated average number of eligibles in hh:	1.58
Individual response rate:	69.18 %
Sample A:	71.38 %
Sample A1:	. %
Sample A2:	71.38 %
Sample B:	.
Sample C:	.
Sample D:	34.96 %
Median number of attempts for not successfully contacted hh:	2
Number of interviewers in SampleCTRL:	51
Rate of active interviewers:	78.43 %

Sweden	week 18
Baseline / refreshment sample	
Gross sample:	900
Households attempted:	868
Households contacted:	692
Households estimated to be eligible:	875.12
Households with completed coverscreen interview:	152
Households with at least one complete interview:	151
Percentage of households attempted:	96.44 %
Contact rate:	76.33 %
Cooperation rate:	22.60 %
Household response rate:	17.25 %
Refusal rate:	45.82 %
Other non-interview rate:	13.26 %
Individual interviews:	179
Estimated average number of eligibles in hh:	1.61
Individual response rate:	12.69 %
Median number of attempts for not successfully contacted hh:	3
Number of interviewers in SampleCTRL:	130
Rate of active interviewers:	89.23 %
Longitudinal sample	
Gross sample:	2986
Households attempted:	2969
Households contacted:	2842
Households estimated to be eligible:	2975.94
Households with completed coverscreen interview:	2004
Households with at least one complete interview:	1831
Percentage of households attempted:	99.43 %
Contact rate:	95.13 %
Cooperation rate:	64.68 %
Household response rate:	61.53 %
Refusal rate:	25.37 %
Other non-interview rate:	8.23 %
Individual interviews:	2535
Sample A:	2271
Sample A1:	2271
Sample A2:	0
Sample B:	46
Sample C:	194
Sample D:	24
Estimated average number of eligibles in hh:	1.59
Individual response rate:	53.58 %
Sample A:	71.33 %
Sample A1:	71.33 %
Sample A2:	. %
Sample B:	19.74 %
Sample C:	20.77 %
Sample D:	6.33 %
Median number of attempts for not successfully contacted hh:	6
Number of interviewers in SampleCTRL:	130
Rate of active interviewers:	89.23 %

Slovenia	week 18
Baseline / refreshment sample	
Gross sample:	3079
Households attempted:	1260
Households contacted:	1025
Households estimated to be eligible:	2895.73
Households with completed coverscreen interview:	397
Households with at least one complete interview:	378
Percentage of households attempted:	40.92 %
Contact rate:	32.81 %
Cooperation rate:	39.79 %
Household response rate:	13.05 %
Refusal rate:	12.67 %
Other non-interview rate:	7.08 %
Individual interviews:	542
Estimated average number of eligibles in hh:	1.66
Individual response rate:	11.25 %
Median number of attempts for not successfully contacted hh:	2
Number of interviewers in SampleCTRL:	66
Rate of active interviewers:	89.39 %
Longitudinal sample	
Gross sample:	3133
Households attempted:	2957
Households contacted:	2886
Households estimated to be eligible:	3120.29
Households with completed coverscreen interview:	2496
Households with at least one complete interview:	1742
Percentage of households attempted:	94.38 %
Contact rate:	92.11 %
Cooperation rate:	60.61 %
Household response rate:	55.83 %
Refusal rate:	14.36 %
Other non-interview rate:	21.92 %
Individual interviews:	2724
Sample A:	2490
Sample A1:	2490
Sample A2:	0
Sample B:	21
Sample C:	169
Sample D:	44
Estimated average number of eligibles in hh:	1.60
Individual response rate:	54.51 %
Sample A:	67.96 %
Sample A1:	67.96 %
Sample A2:	. %
Sample B:	21.65 %
Sample C:	19.79 %
Sample D:	11.49 %
Median number of attempts for not successfully contacted hh:	2
Number of interviewers in SampleCTRL:	66
Rate of active interviewers:	89.39 %

Slovakia	week 18
Longitudinal sample	
Gross sample:	867
Households attempted:	816
Households contacted:	812
Households estimated to be eligible:	856.38
Households with completed coverscreen interview:	708
Households with at least one complete interview:	636
Percentage of households attempted:	94.12 %
Contact rate:	93.65 %
Cooperation rate:	79.30 %
Household response rate:	74.27 %
Refusal rate:	9.46 %
Other non-interview rate:	9.93 %
Individual interviews:	1036
Sample A:	1016
Sample A1:	0
Sample A2:	1016
Sample B:	0
Sample C:	0
Sample D:	20
Estimated average number of eligibles in hh:	1.66
Individual response rate:	72.91 %
Sample A:	73.46 %
Sample A1:	. %
Sample A2:	73.46 %
Sample B:	.
Sample C:	.
Sample D:	51.28 %
Median number of attempts for not successfully contacted hh:	2
Number of interviewers in SampleCTRL:	56
	71.43 %

5.2 Data Quality Back-checks in the Regular SHARE Wave 8

Tessa-Virginia Hannemann – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC) and Michael Bergmann – 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)

5.2.1 Introduction

One central aspect of ensuring the data quality of SHARE is to check whether interviews have taken place as reported. This chapter describes the process implemented in Wave 8 to improve interview verification and the identification of interview falsification in a continuous manner, throughout the fieldwork of the regular Wave 8. In the following, we will explain the procedures set in place to systematically check interviews, in a subset of randomly selected completed interviews, i.e. random back-checks, and in interviews that had been identified as suspicious, i.e. focused back-checks. Further, we will report on the detection of suspicious interviews throughout this process as well as alternative back-checking procedures in Germany and the Netherlands.

In Wave 8 for the first time SHARE Central selected the interviews to be randomly back-checked, rather than countries' survey agencies. The frequency of the back-checks was adjusted to select interviews after each data delivery, typically every fortnight. This allowed for needed interventions during fieldwork. As mentioned above, focused back-checks were conducted in addition every four to six weeks similarly to Wave 7 in order to detect interviews suspected of falsification (Bergmann & Schuller, 2019). The following will describe in further detail the selection process of both random and focused back-checks.

5.2.2 Back-checks: Selection, Distribution and Verification of Interviews

Random Back-checks. The selection of interviews was based on the completion of an interview. Only interviews that were completed in the time since the last data delivery were included in the selection. We aimed at verifying 20 per cent of interviews that were conducted by each interviewer throughout the fieldwork. In order to compensate for possible exclusions due to the random selection of the interviews, a check was put in place to ensure all interviewers that had completed a minimum number of interviews were also verified through the back-check process.

A template was sent out that included the interviews to be back-checked along with verification questions. An overview of the questions used for the random back-checks during CAPI can be seen in the Annex of this chapter. These templates were shared via a secure server to ensure data protection according to GDPR. Once checking had been completed, the survey agencies were instructed to send back the completed back-checking template, again via the secure server.

End-of-Life interviews were excluded from the back-checking procedure. These interviews are conducted using a proxy for the deceased respondent. As such, a further back-checking would entail further strain for the proxy and possible attrition of the sample. In addition, the verification questions would not line up with the questions asked during an End-of-Life interview, which would result in incorrect conclusions from the checks.

The process was adaptable during fieldwork to accommodate unforeseen requirements. One such adaptation that was implemented was an indicator to signify proxy involvement to aid the survey agencies during back-checking. Using the variables that indicate proxy involvement during several modules, we were able to identify interviews that (partially) involved a proxy. This allowed survey agencies to evaluate the responses to verification questions better.

Focused Back-checks. Focused back-checks are checks to verify interviews that have been flagged as suspicious, i.e. it is not certain that these interviews have been conducted properly. Focused back-checks were conducted every four to six weeks.

There were 14 indicators of suspicion considered that are presented below (see Table 5.4). In order to calculate the indicators information from the CaseCTRL, the CAPI interview as well as paradata were taken into consideration. The main aspects that were looked into were answer patterns, interview timing and deviations from previous waves. Straight-lining, repetitive answer patterns, missing items, rounding of physical measurement values, few follow-up questions and few extreme answer options selected, as well as few answers selected in multiple answer options, were considered suspicious response behaviour, and could signpost a falsified interview. Similarly, the timing of the interview was considered. Thus, completing more than five households in one day, unrealistic times of day as well as unrealistic interview durations after controlling for respondent characteristics were examined in further detail. Information from previous waves was also taken into account. Unrealistic deviations,

i.e. large weight changes between waves, changing dominant hands, as well as contradicting information from the preload and the current interview, in particular year of birth and gender, were also indicators of suspicious interviews. The mere presence of an indicator did not deem an interview as suspicious but rather a threshold of indicators (7/14) created the need to verify the interview. However, contradicting information from the preload and CAPI regarding respondents' gender and year of birth was also sufficient to flag an interview as suspicious.

Table 5.4: Overview of Indicators Regarding Interview Falsifications

Indicator	Description
1	**Deviation from last wave:** interview is flagged as suspicious if a change in the respondent's dominant hand and/or an absolute weight change of more than 20 kg have been recorded.
2	**Rounding in grip strength measurement:** interview is flagged as suspicious if multiples of 5 have been recorded in all four grip strength measures (two measurements with each hand).
3	**Number of interviews on same day:** interview is flagged as suspicious if five or more interviews have been conducted by an interviewer on the same day.
4	**Interview rate at first contact:** interview is flagged as suspicious if the rate of interviews with no recorded contact attempt across all completed interviews of an interviewer is above 95%.
5	**Cooperation rate:** interview is flagged as suspicious if the cooperation rate (i.e. the number of interviews divided by the total number of interviews plus the number of non-interviews that include contact with an eligible respondent; see AAPOR, 2016: COOP1) of an interviewer with at least five households contacted is above 95%.
6	**Cooperation rate of partner:** interview is flagged as suspicious if the cooperation rate in households with two eligible respondents (i.e. the number of households with two interviews divided by the total number of all partner households) of an interviewer with at least five partner households is above 95%.
7	**Interview duration:** interview is flagged as suspicious if the residual of a linear regression, using the log normal distribution of interview length (based on keystroke data without the IV module) regressed on key respondent and interview characteristics (year of birth, self-rated health, frequency of questionnaire clarifications, number of asked items, sequence number of interview, questionnaire version and interview language) is below the fifth percentile in the respective country.
8	**Same answer pattern (duplicate):** interview is flagged as suspicious if the ratio of an identical answer pattern across all items in a module asked by an interviewer is above the 95th percentile in the respective country.
9	**Straight-lining:** interview is flagged as suspicious if the ratio of selecting the same answer category across all items in five multi-item sets (standardised by the number of items) is above the 95th percentile in the respective country.
10	**Item non-response:** interview is flagged as suspicious if the ratio of missing values across all substantial and answered items is above the 95th percentile in the respective country.
11	**"Other" answers:** interview is flagged as suspicious if the ratio of "other" answers across all questions with an "other" category is below the fifth percentile in the respective country.
12	**"Code all that apply" answers:** interview is flagged as suspicious if the ratio of selecting more than one answer option across all "code all that apply" questions is below the fifth percentile in the respective country.

Indicator	Description
13	**Follow-up questions:** interview is flagged as suspicious if the ratio of selecting "no" across all filter questions with follow-up items is below the fifth percentile in the respective country.
14	**Extreme answers:** interview is flagged as suspicious if the ratio of extreme answers across all numerical items with five, seven or eleven-point scales is below the fifth percentile in the respective country.

5.2.3 Results

During Wave 8, 13,477 households were selected for random back-checking. The first back-checks were sent out after the data synchronisation of calendar week 42 (mid-October) in 2019 and the final back-checks were sent out after data synchronisation in calendar week 12 (mid-March) in 2020 that marks the suspension of fieldwork due to the outbreak of COVID-19. A total of 12,115 back-checks were received at SHARE Central. Table 5.5 gives an overview of the back-checks sent out to each country and the number of checked interviews received, as well as the number of households successfully reached. Focused back-checks were sent out after data synchronisation in calendar week 48 in 2019, and in calendar weeks 2 and 10 in 2020. In total, 363 interviews were identified as suspicious during the regular Wave 8. Of the checked interviews none warranted reinterview after consultation with the respondent.

Table 5.5: Results of Random Back-checks

Country	households selected	households checked	households reached
AT	541	541	304
BE_nl	418	437	303
BE_fr	257	230	133
BG	273	148	100
CH	544	459	291
CY	155	127	96
CZ	821	655	390
DK*	734	732	552
EE	1,011	975	746
ES	400	546	371
EG	215	215	160
FI	300	300	237
FR	843	835	540
GR	761	5	4
HR	527	527	477
HU	354	117	25
IL	312	312	274
IT*	654	1122	1033
LT	396	372	303
LU	257	257	226
LV	362	275	191
MT	226	147	143
PL	824	633	535
RO	393	382	277
SE	668	650	497
SI	855	855	537
SK	376	261	195
Total	**13,477**	**12,115**	**8,940**

Note: * checked and reached households include those checked during the first SHARE Corona Survey.

Back-checking made us aware of potential problems in three cases during fieldwork. In one instance, the respondent reported that no partner interview had taken place in the household, though a completed interview had been logged. After consultation with the interviewer, it was decided to re-interview the partner and the information was relayed to database management. A further interview had been flagged as suspicious. After consultation with the respondent and the interviewer, it was discovered that the preload information had been incorrect. Correcting the preload information was made possible through the back-checking procedure. In a final instance, through back-checking it was brought to our attention that not all information that is assessed during the verification process was available for the interview, as there had been a proxy involved. This was the impulse

to include a proxy indicator in the templates, as described above. The cases described here prove the flexibility the back-checking process provides for fieldwork.

5.2.4 Alternative Back-checking Procedures in Germany and the Netherlands

Not all participating countries participated in the random back-checking procedure as described above. In Germany and the Netherlands, 100 per cent verification was done on completed interviews. That is, all households that had participated in the survey were contacted. In both countries, this was done via postal mail. In Germany, the verification questions were sent out in conjunction with the thank you letter sent to all respondents that took part in the survey. In the Netherlands, the questions were sent out in conjunction with the thank you letter that contained the incentive.

Harmonisation is of great importance in a survey with as many participating countries as SHARE. It was therefore vital that although the procedure differed, the information gained from the verification questions was comparable to that of other countries. Table A 5.1 in the Annex shows all verification questions that had been implemented by the two countries that conducted the verifications by mail. We received 1,800 returned pre-pandemic fieldwork verification questionnaires from Germany, 26 of which had been undeliverable. The completed verification questionnaires had been checked again at the survey agency, as was evident in the feedback added by Kantar. If there were any uncertainties about the conduct of the interview, the survey agency checked either again with the interviewer or directly with the respondent via telephone. We received 1,612 pre-pandemic fieldwork verification questionnaires from the Netherlands. One interview caused concern in the survey agency but was deemed fine after consulting with the interviewer.

Back-checking all completed interviews, as done in Germany and the Netherlands, has a number of advantages. They tend to require less time and personnel resources for the survey agencies if done via mail. There is no need to randomly select interviews that need to be back-checked, which may increase the risk of oversampling interviews conducted by very productive interviewers.

As postal communication is not as instantaneous as communication via telephone, there is a lag in timing. The opportunity to intervene closely during fieldwork was the great advantage of the back-checking procedure described in this chapter. Additional checks by the survey agencies need to be in place in order to compensate for the timing disadvantage. Moreover, although all households were addressed, the scope of verification questionnaires that were returned was similar to that of the randomly selected 20 per cent of completed interviews per interviewer.

5.2.5 Discussion and Lessons Learned

Overall there was good acceptance of the procedures. Although the verification questions were viewed as helpful, there seemed to be some concern over older respondents forgetting about the interview and therefore reflecting badly on the interviewer. Timing is therefore of great importance and back-checks should not be sent out less frequently than every two weeks, to ensure the interview is still fresh in the respondents' memory. In addition, it can be considered to leave each respondent with an individual code, they can report during the back-check, or prompting their memory by asking specific questions about the interviewer.

Alternative back-checking procedures were presented in this chapter. This highlights the different ways in which interview verification can be done. Although procedures differed, great care was taken to harmonise the information gleaned from the verification interviews.

During fieldwork, the process was quick to adapt to challenges that emerged. For instance, we were able to implement an indicator for proxy involvement in the interview as an additional check of the information relayed from the respondent. In the instances described above in which an intervention was warranted, the process in place allowed for quick reaction, such as reinterviewing as well as relaying information on invalid interviews internally to database management. The focused back-checks were also a way for survey agencies to receive further continuous information on the course of the interviews, again allowing for possible adjustments to be made during fieldwork.

Because timing played such an important role in the effectiveness of the back-checks, it is important, for future waves, to ensure that sufficient back-checking procedures are in place prior to the start of fieldwork. This also means early check-ins to the survey agencies and timely assistance if needed.

References

American Association for Public Opinion Research (AAPOR). (2016). Standard definitions: Final dispositions of case codes and outcome rates for surveys (9th ed). Available at: http://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf

Bergmann, M., & Schuller, K. (2019). Improving the efficiency of data quality back checks: A new procedure to prevent curbstoning. In M. Bergmann, A. Scherpenzeel & A. Börsch-Supan (Eds.), SHARE Wave 7 Methodology: Panel innovations and life histories (pp. 154-159). MEA, Max Planck Institute for Social Law and Social Policy.

Table A 5.1: Verification Questions of Random Back-checks

Verification questions	Included in back-checks...	
	DE	NL
Household ID (hhidcom)		
Person ID (pidcom)		
Laptop ID (laptop_s8)		
Interviewer ID (interviewerid_s8)		
Mode and sequence of contact attempts	x	
Successful contact in the end?	x	x
If no: Reason for non-contact	x	
1) Has one of our interviewers recently interviewed you for the SHARE study?	x	x
2) If yes: With whom was the interview conducted?	x	
3) How was the interview conducted?	x	
4) Where was the interview conducted?	x	
5a) How long was the interview with you?	x	x
5b) If partner interview: How long was the interview with your partner?	x	
6) Did the interviewer use a device to measure the strength of your hands?	x	x
7) Did the interviewer use showcards during the interview?	x	x
8) Did you receive a monetary incentive for your participation in the SHARE study?	x	
9) What is your year of birth?	x	x
10) Only if not sure: What is your gender?		x
Other comments by respondent	x	x
Evaluation of back-check by survey agency	x	x
If interview not ok: Consultation with interviewer regarding suspicion	x	x
Final decision by survey agency after consultation with interviewer	x	x
Only if interview not ok/fake: Reinterview	x	x
Only if reinterviewed: Information on new interview	x	x

CHAPTER 6

Weights and Imputations in SHARE Wave 8

06

6 WEIGHTS AND IMPUTATIONS IN SHARE WAVE 8

Giuseppe De Luca, Paolo Li Donni and Moslem Rashidi – University of Palermo

6.1 Introduction

This chapter provides a description of the weighting and imputation strategies used for dealing with problems of unit non-response, sample attrition and item non-response in the eighth wave of the Survey of Health, Ageing and Retirement in Europe (SHARE). As discussed in the previous chapters, the data collection process of Wave 8 was suddenly interrupted in March 2020 by the COVID-19 outbreak and the subsequent lockdowns enforced by the national governments of the various countries. SHARE reacted promptly to this deep pandemic shock through the design of a special COVID-19 questionnaire, which was fielded between June and July 2020. We expect that the data collected in the regular Wave 8 will become an extraordinary source of information for studying health and socio-economic implications of the shock for the elderly population. To best exploit the available data, it is important for the user to have a basic understanding of the fieldwork rules adopted for the standard interview and the special COVID-19 interview of Wave 8, the different types of non-response errors that occurred in the implementation of these two interview instruments and the basic strategies adopted to cope with these errors. In the following, we first use the different patterns of participation to define three subsamples of primary interest for the analysis of the data collected in Wave 8: CAPI, CATI and CAPI & CATI. We then describe the procedure used to construct calibrated cross-sectional and longitudinal weights for handling, respectively, problems of unit non-response and attrition in the CAPI subsample. Afterwards, we describe the model used to obtain multiple imputations of the missing values due to item non-response in the CAPI data. The construction of calibrated weights and multiple imputations for the CATI data is discussed in Chapter 11.

6.2 Composition of the Sample in Wave 8

The data collection process of Wave 8 started regularly in October 2019 by means of a face-to-face Computer-Assisted Personal Interview (CAPI) administered in 28 countries. As usual, the sample in Wave 8 consisted of a longitudinal subsample and a refreshment subsample. The longitudinal subsample includes all respondents already interviewed in any previous wave of the study. The refreshment subsample, on the other hand, includes the new sample units drawn in Wave 8 to maintain the representation of the younger cohorts of the target population that were not age-eligible in the previous waves (i.e. people born between 1967 and 1969) and to compensate for the reduction of sample size due to attrition across waves of the SHARE panel.

The fieldwork activities of Wave 8 were suddenly interrupted in March 2020 due to the COVID-19 outbreak. To study the impact of the pandemic on the health and socio-economic conditions of SHARE respondents, a new COVID-19 questionnaire was promptly fielded between June and July 2020 by a Computer-Assisted Telephone Interview (CATI). By design, this new survey instrument was administrated to the longitudinal part of the sample only (not to the refreshment sample). Table 6.1 and Table 6.2 provide, respectively, a breakdown of the number of individual interviews and the number of household interviews by country and type of interview (CAPI and/or CATI) based on SHARE Wave 8, Release 0 as well as SHARE Wave 8, Release 0.0.1 beta (Börsch-Supan, 2020a, 2020b). In total, 23 per cent of respondents answered the CAPI only, 28 per cent answered the CATI only, and 49 per cent answered both the CAPI and CATI instrument. For the type of data collected in Wave 8 one can then distinguish three subsamples of primary interest: CAPI, CATI and CAPI & CATI. The CAPI subsample consists of 51,018 respondents in 35,914 households who have answered the CAPI questionnaire irrespective of whether they have also answered the CATI questionnaire. The CATI subsample consists of 54,600 respondents in 37,222 households who have answered the CATI irrespective of whether they have also answered the CAPI. The CAPI & CATI subsample consists of 34,916 respondents in 24,191 households who have answered both interviews.

Table 6.1: Number of Individual Interviews of Wave 8 by Country and Type of Interview

Country	CAPI only	CATI only	CAPI & CATI	Total CAPI	Total CATI
AT	607	1,204	1,265	1,872	2,469
BE	518	2,095	1,687	2,205	3,782
BG	170	171	640	810	811
CH	364	246	1,640	2,004	1,886
CY	123	416	374	497	790
CZ	884	579	2,040	2,924	2,619
DE	1,406	378	2,278	3,684	2,656
DK	854	530	1,453	2,307	1,983
EE	577	1,836	2,706	3,283	4,542
ES	961	1,037	1,011	1,972	2,048
FI	119	457	1,006	1,125	1,463
FR	1,189	316	1,727	2,916	2,043
GR	184	1,039	2,595	2,779	3,634
HR	862	961	1,048	1,910	2,009
HU	666	513	483	1,149	996
IL	640	763	687	1,327	1,450
IT	171	1,860	1,846	2,017	3,706
LT	269	179	1,086	1,355	1,265
LU	193	202	726	919	928
LV	490	322	656	1,146	978
MT	104	200	628	732	828
NL	1,400	276	504	1,904	780
PL	1,055	1,300	1,628	2,683	2,928
PT	0	1,118	0	0	1,118
RO	73	378	1,101	1,174	1,479
SE	1,367	238	1,121	2,488	1,359
SI	819	983	2,129	2,948	3,112
SK	37	87	851	888	938
Total	**16,102**	**19,684**	**34,916**	**51,018**	**54,600**

Data: SHARE Wave 8, Release version: 0.

Table 6.2: Number of Household Interviews of Wave 8 by Country and Type of Interview

Country	CAPI only	CATI only	CAPI & CATI	Total CAPI	Total CATI
AT	467	843	913	1,380	1,756
BE	396	1,543	1,258	1,654	2,801
BG	130	111	437	567	548
CH	275	135	1,236	1,511	1,371
CY	72	249	270	342	519
CZ	602	414	1,442	2,044	1,856
DE	1,120	241	1,515	2,635	1,756
DK	614	379	1,073	1,687	1,452
EE	423	1,260	1,981	2,404	3,241
ES	723	665	679	1402	1344
FI	95	268	684	779	952
FR	881	230	1,254	2,135	1,484
GR	152	666	1,680	1,832	2,346
HR	600	598	670	1,270	1,268
HU	475	336	335	810	671
IL	482	506	490	972	996
IT	125	1149	1,160	1,285	2,309
LT	193	121	786	979	907
LU	152	123	513	665	636
LV	355	226	471	826	697
MT	52	124	388	440	512
NL	950	182	357	1,307	539
PL	696	859	1,090	1,786	1,949
PT	0	725	0	0	725
RO	52	236	719	771	955
SE	1,021	161	819	1,840	980
SI	601	629	1,430	2,031	2,059
SK	19	52	541	560	593
Total	**11,723**	**13,031**	**24,191**	**35,914**	**37,222**

Data: SHARE Wave 8, Release version: 0.

The distinction between these three subsamples has important implications for the information available in the analysis of Wave 8 data. Specifically, the CAPI subsample contains the data collected before the COVID-19 outbreak by the regular SHARE questionnaire of Wave 8, and its longitudinal part (about 86 per cent) can be merged with the data collected in one or more previous waves. The CATI subsample contains the data collected after the COVID-19 outbreak by the SHARE Corona Survey and can be fully merged with some of the previous waves of SHARE as it consists of longitudinal respondents only. The CAPI & CATI subsample exploits the full force of the survey instruments implemented in Wave 8 as it contains the data collected before and after the outbreak and can be fully merged with previous waves. As discussed in the next section, the SHARE weights database provides different sets of calibrated cross-sectional weights for the three subsamples. SHARE also provides different sets of imputations for the missing values due to item non-response in the CAPI and CATI data. In this chapter, we shall focus attention on calibrated weights and imputations for the standard CAPI data of Wave 8.

6.3 Calibrated Weights

In the ideal situation of complete response, the availability of design weights allows the users to account for the randomness of the sampling process by compensating for unequal selection probabilities of the various sampling units. Unfortunately, properties of inferential procedures based on the sampling design weights depend 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 of each wave suffer from problems of unit non-response (Groves & Peytcheva, 2008). Moreover, the longitudinal part of the sample is subject to attrition problems (Lynn, 2009). Because of these non-sampling errors, we discourage the 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 weights database only to favour the implementation and comparison of alternative statistical procedures for handling non-response and attrition errors.

The baseline strategy adopted by SHARE to handle problems of unit non-response and attrition relies on the calibration approach proposed by Deville and Särndal (1992). This approach allows the sample and population distributions of some benchmark variables to be aligned without the need for specifying an explicit model for the non-response mechanism. Under the assumption that the missing data mechanism is missing at random (Rubin, 1987), calibrated weights may help reduce the potential selection bias generated by non-response errors. Thus, unless these sources of non-sampling errors are controlled for in other ways, these are the types of weights that we generally recommend using 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. Then, we describe the implementation of the calibration procedure for constructing the various types of calibrated cross-sectional and longitudinal weights available in the public release of SHARE Wave 8 data.

6.4 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 a probability-based sampling design. Unless otherwise specified, we shall assume that the inclusion probability $\pi_i=Pr(i \in s)$ is known and strictly positive for all population units. To describe the basic ideas and the key properties of the calibration approach, we consider first the ideal situation of complete response where all units in the sample s agree to participate in the survey. Then, we relax this ideal set-up to describe the key implications of non-response errors for the properties of this weighting method.

The sampling design weights $w_i=\pi_i^{-1}$ are typically used to account for the randomness of the sampling process and the variability of the inclusion probabilities across sample units due to stratification and clustering strategies (additional details can be found in Chapter 2). For example, one can estimate the population total $t_y=\sum_{i\in U}y_i$ of a variable of interest y using the Horvitz-Thompson estimator:

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

Under the ideal set-up of complete response, this estimator is known to be design unbiased, that is $E_p(\hat{t}_y)=t_y$, where $E_p(\cdot)$ denotes the expectation with respect to the sampling design.

Let us assume now that the sampling frame or other external sources such as census data and administrative archives provide supplementary data on a q-vector of categorical auxiliary variables $x_i=(x_{i1},\ldots,x_{iq})^T$ with known population totals $t_x=\sum_{i\in U}x_i$. We shall refer to the auxiliary variables x_i as calibration variables and to their population totals t_x as calibration margins. The basic idea of the calibration approach is to determine a set of *calibrated weights* w_i^* that are as close as possible to the design weights w_i and that satisfy the constraints

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

Thus, given a distance function $G(w_i^*,w_i)$ and the availability of survey data on $(w_i,x_i:i=1,\ldots,n)$ and population data on

the calibration margins t_x, the aim of the procedure is to determine the calibrated weights w_i^* by minimizing 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). Under some regularity conditions on the distance function $G(w_i^*, w_i)$ (see Deville & Särndal, 1992), the solution of this constrained optimisation problem exists, is unique and can be written as

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

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 the q-vector of Lagrangian multipliers associated with the constraints (2), and $F(\cdot)$ is a calibration function, which is uniquely determined by the distance function $G(w_i^*, w_i)$.

A key feature of the calibration approach is that many traditional reweighting methods such as post-stratification, raking and generalised linear regression (GREG) correspond to special cases of the calibration estimator

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

for particular choices of the calibration function $F(\cdot)$ (or, equivalently, of the distance function $G(\cdot,\cdot)$). Deville and Särndal (1992) present various functional forms for $G(w_i^*, w_i)$ and $F(\eta_i)$. 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^*. However, this distance function is unbounded, and depending on the chosen set of calibration variables it may also lead to negative weights. Different specifications of the calibration function may avoid these issues, but the underlying optimisation problems may not admit a solution and the Lagrange multipliers must be computed numerically. In SHARE, we rely on the logit specification of the distance function

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

which leads to a calibrated function of the form

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

where $a = (u - l)/[(1 - l)(u - 1)]$. Unlike other distance functions, these functional forms restrict in advance 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 it must satisfy the restriction $w_i l \le w_i^* \le w_i u$.

As discussed in 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 when y can be expressed as a linear combination of x, it is clear that the calibrated estimator $\hat{t}_y^*$ gives an exact estimate of t_y for every realised sample s. Under suitable regularity conditions, the class of calibration estimators $\hat{t}_y^*$ satisfies other desirable asymptotic properties. For example, the estimators obtained by alternative specifications of the distance function are asymptotically equivalent to the GREG estimator based on a chi-squared distance function. Thus, in large samples, calibrated weights are robust to arbitrary choices of the calibration function $F(\cdot)$.

Unfortunately, this property does not necessarily extend to the more realistic cases where survey data are affected by non-response errors. Previous studies by Lundström and Särndal (1999) and Haziza and Lesage (2016) suggest that in these cases alternative specifications of the calibration function $F(\cdot)$ correspond in practice to imposing different parameterisation of the relationship between response and calibration variables. Moreover, statistical properties of calibration estimators depend as usual on the validity of the missing-at-random assumption. Brick (2013), Molenberghs et al. (2015), Vermeulen and Vansteelandt (2015) and Haziza and Lesage (2016), among others, discuss a variety of robust weighting methods based on a propensity-score approach. One key issue in the implementation of these methods for SHARE is that selection probabilities and auxiliary variables are usually known for the subsample of respondents only.

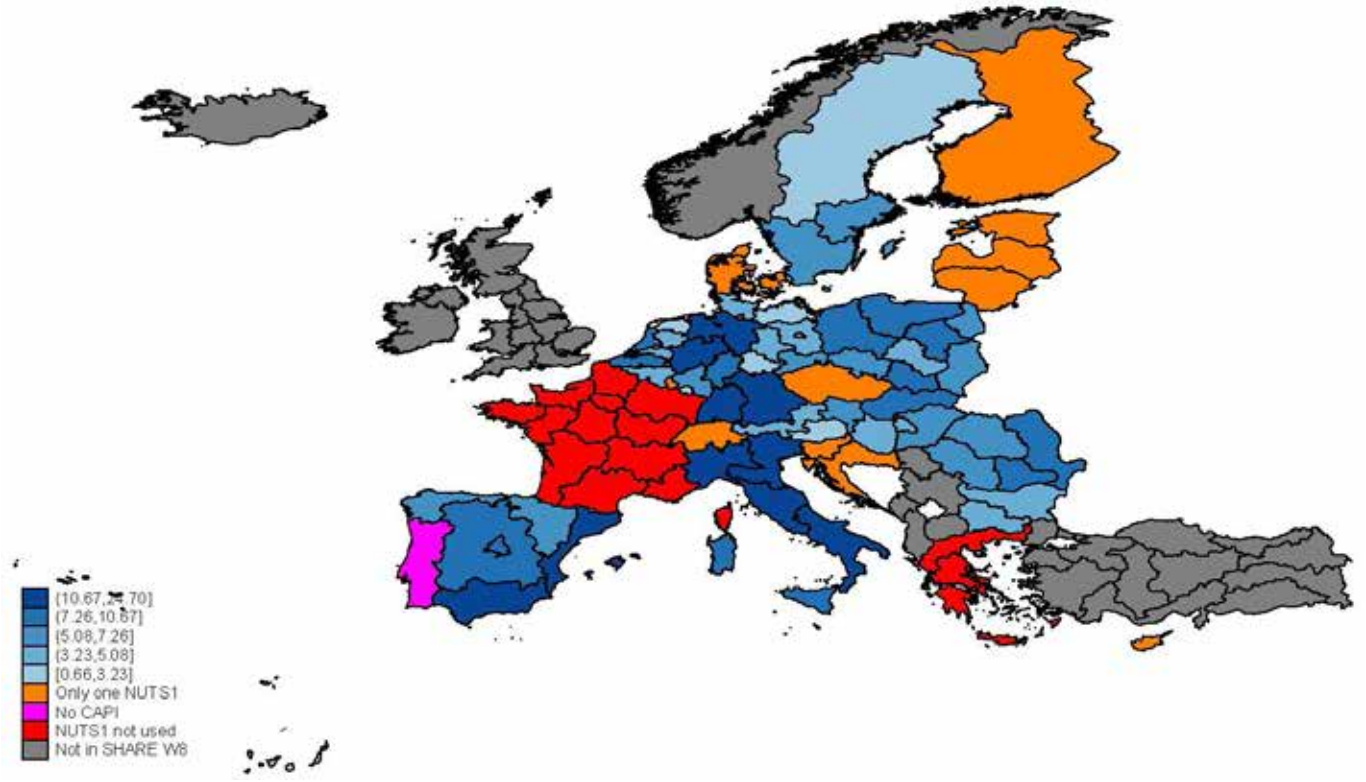
6.5 Calibrated Cross-sectional Weights for the CAPI Subsample

The calibrated cross-sectional weights of the CAPI subsample of Wave 8 were computed separately by country to match the size of national 50+ populations of individuals in 2019. In each country, we used a logit specification of the calibration function $F(\cdot)$ and a set of calibration margins for the size of the target population across the eight gender-age groups (i.e. males and females in the age classes ([50 – 59], [60 – 69], [70 – 79], [80+]) reported in Table A 6.1 in the Annex.

In 11 countries (Austria, Belgium, Bulgaria, Germany, Hungary, Italy, the Netherlands, Poland, Romania, Spain and Sweden), we also included an additional set of calibration margins for the size of the 50+ population across 2016 NUTS1 regional areas. In Figure 6.1, we illustrate the distribution of the 50+ population across 2016 NUTS1 regional

areas (Israel is excluded from the figure). Notice that this additional set of calibration margins were ineffective in all countries containing only one NUTS1 region[11]. In France and Greece, NUTS1 calibration margins were excluded because of inconsistency between sample and population data. In Israel, where no NUTS nomenclature is available, we used an additional set of calibration margins for the Jewish Israeli and Arab Israeli population groups, and immigrants from the former USSR. Population data about the calibration margins come from the Central Bureau of Statistics for Israel and from the EUROSTAT regional database for all other countries.

Figure 6.1: NUTS1 Population Margins for the Calibrated Cross-sectional Weights of Wave 8 (Millions of People)



Data: SHARE Wave 8, Release version: 0.

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+ respondent that depends on the household design weight and the respondents' set of calibration variables (namely, gender, age class and NUTS1 code). 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 set of calibration variables for all 50+ respondents in that household. By construction, calibrated cross-sectional weights are missing for respondents younger than 50 (i.e. age-ineligible partners of an age-eligible respondent), for those with missing information on the calibration variables and for those with missing sampling design weights (i.e. respondents from households for which we do not have sampling frame information).

6.6 Calibrated Longitudinal Weights for the CAPI Subsample

In addition to calibrated cross-sectional weights, SHARE Wave 8 Release 8.0.0 also includes calibrated longitudinal weights for the purposes of panel data analyses. Although calibration relies on the same procedure, calibrated longitudinal weights differ from calibrated cross-sectional weights in two important respects. First, the calibrated longitudinal weights are defined only for the balanced subsample of respondents who have participated in at least two waves of the study. Second, since

11 That is the case in Croatia, Cyprus, the Czech Republic, Denmark, Estonia, Finland, Latvia, Lithuania, Luxembourg, Malta, Slovakia, Slovenia and Switzerland.

mortality is a source of attrition that affects both the sample and the population, calibrated longitudinal weights account for the mortality of the target population across waves. In other words, the target population for panel data analysis is defined as the target population at the beginning of a reference time period that survives up to the end of the period considered (see, for example, Lynn, 2009).

To simplify the structure of the public release of the data, we provide calibrated longitudinal weights only for selected wave combinations of the SHARE panel. Those available in Release 8.0.0 are the seven possible couples of any two adjacent waves (namely, the wave combinations 1 – 2, 2 – 3, 3 – 4, 4 – 5, 5 – 6, 6 – 7 and 7 – 8) and the fully balanced panel (i.e. the wave combination 1 – 2 – 3 – 4 – 5 – 6 – 7 – 8). The weights of the generic wave combination t – … – s are computed separately by country to represent the national 50+ population of Wave t that survives up to the interview year of Wave s. For example, the wave combination 1 – 2 allows the population of people aged 50+ in 2004 that survived up to 2006 to be represented, while the fully balanced panel allows the population of people aged 50+ in 2004 that survived up to 2019 to be represented.

For the calibrated longitudinal weights of two adjacent waves, 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 whose ages at the time of the starting wave were in the four classes [50 – 59], [60 – 69], [70 – 79] and [80+]). Compared to calibrated cross-sectional weights, we do not control for NUTS1 calibration margins due to the smaller number of observations available in the national longitudinal subsamples. Moreover, we account for the mortality of the target population by subtracting from each calibration margin the corresponding number of deaths that occurred between the interview years of Wave t and Wave s. Table A 6.2 in the Annex provides the population margins used to compute the calibrated longitudinal weights of the wave combination 7 – 8. Population margins for the calibrated longitudinal weights of the other wave combinations can be found in De Luca and Rossetti (2019a, Tables A.3 – A.8).

For the calibrated longitudinal weights of the fully balanced panel, we further restricted the set of calibration margins to six gender-age groups (i.e. males and females whose ages in 2004 were in the three classes [50 – 59],[60 – 69] and [70+]). Table A 6.3 in the Annex shows the population margins used to construct the longitudinal weights of the fully balanced panel.

As with calibrated cross-sectional weights, calibrated longitudinal weights are available both at the individual level and at the household level. For the individual weights, the balanced sample consists of respondents interviewed in each wave of the selected 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, according to these definitions, the balanced sample of households is larger than the balanced sample of individuals. For example, couples with one partner participating in Wave 7 and the other partner participating in Wave 8 belong to the balanced sample of households for the wave combination 7-8, even if neither of the two partners belongs to the corresponding balanced panel of individuals.

6.7 Supplementary Material and User Guide on Calibrated Weights

Since the SHARE panel now consists of eight waves, one can compute many different types of calibrated longitudinal weights depending on the selected combination of waves and the selected 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 collected 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 8). 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 non-response and attrition by computing their own calibrated weights or by implementing some alternative correction method.

To support users in this non-trivial 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. In addition, we provide one data set with updated information on population size and number of deaths by year, gender, age and NUTS1 regions. Registered users can download this supplementary material on calibrated weights from the SHARE Research Data Center dissemination website (https://releases.sharedataportal.eu/releases), under the link "Generate Calibrated Weights Using Stata (2018)". A discussion of the step-by-step operations can also be found in the SHARE Technical Report "Computing Calibrated Weights in Stata" (De Luca & Rossetti, 2019b).

6.8 Imputations of Missing Values in the CAPI Data

Imputations of the missing values due to item non-response errors in the regular face-to-face interview of Wave 8 were

constructed using the same procedure adopted in the previous regular waves of SHARE (see, for example, De Luca et al., 2015). Of course, we adapted the imputation model to the specific features of the regular Wave 8 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 also attempted to preserve as much as possible the comparability of the imputations across different waves of the SHARE panel. The imputation procedure is essentially based on either the hot-deck method or the fully conditional specification (FCS) method depending on the prevalence of missing values for the variables collected in the regular interview of Wave 8.

6.9 Hot-deck Imputations

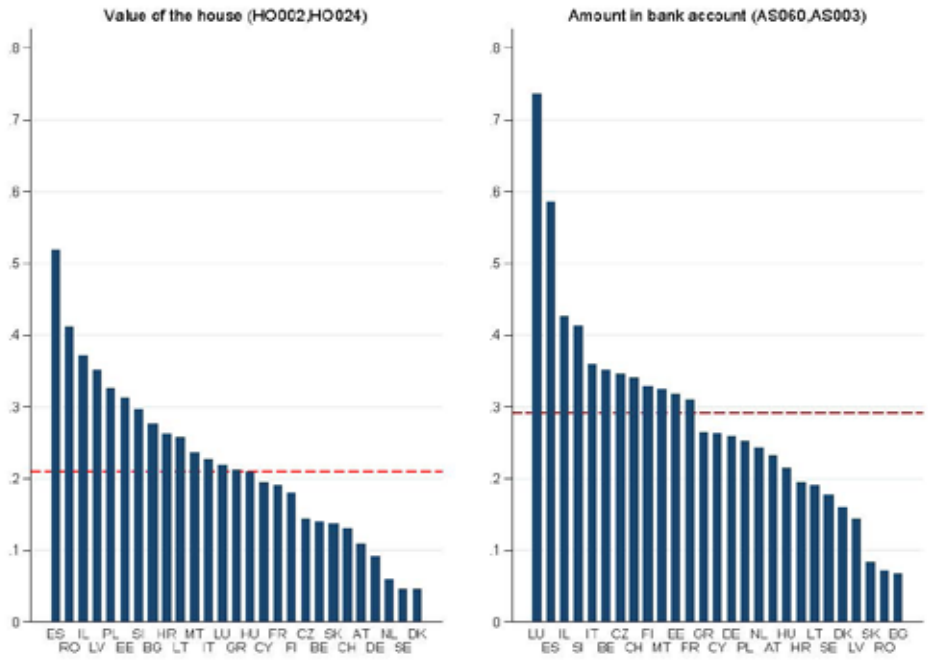
In SHARE, we always used the hot-deck method for variables affected by negligible fractions of missing values (usually, much less than 5 per cent of the respondents eligible to answer a specific item on the CAPI questionnaire). The hot-deck method consists of replacing the missing values in one or more variables for a non-respondent (called the recipient) with the observed values in the same variables obtained from a respondent (called the donor) who is "similar" to the recipient according to some metric (see, for example, Andridge & Little, 2010).

In Wave 8, we computed hot-deck imputations in an early stage, separately by country, and according to a convenient order that accounts for branching and skip patterns included in the various modules of the CAPI questionnaire. For each variable imputed through this method, we select the donors randomly from imputation classes determined by auxiliary variables that are observed for both donors and recipients. We imputed first basic socio-demographic characteristics such as age and education, which contained very small fractions of missing values. These characteristics were then used as auxiliary variables to impute the missing values in the other variables. Our baseline set of auxiliary variables consisted of 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 good/bad health. For some variables, we exploited a larger set of auxiliary variables. For example, we also used the number of children to impute the number of grandchildren and an indicator for being hospitalised overnight during the last year to impute other health-related variables. Variables that are known to be logically related, such as respondent's weight, height and body mass index, were imputed jointly.

6.10 FCS Imputations

In the second stage of the imputation procedure, we dealt with the more worrisome issue of item non-response in monetary variables, such as income from various sources, assets and consumption expenditures, which were typically collected by retrospective and open-ended questions that are sensitive and difficult to answer precisely (see Figure 6.2).

Figure 6.2: Item Non-response Rates for Value of the House and Amount in Bank Account by Country



Data: SHARE Wave 8, Release version: 0.

Figure 6.2 shows the item non-response rates of two monetary variables: value of the house (*HO002, HO024*) and amount in bank account (*AS060, AS003*). For the first variable, the percentage of missing values among the eligible respondents ranges from a minimum of 5 per cent in Denmark and Sweden to a maximum of 52 per cent in Spain (21 per cent on average). The percentage of missing values becomes even more dramatic for questions that are likely to be very sensitive for the respondents. For example, the financial respondent was asked *"Do you (or your husband/wife/partner) currently have a bank account, or transaction account, or saving account or postal account?" (AS060)* and then *"About how much do you (and your husband/wife/partner) currently have in bank accounts, transaction accounts, saving accounts or postal accounts?" (AS030)*. In 12 out of 27 countries participating in Wave 8, more than 30 per cent of the eligible respondents either refused or did not know how to answer these two questions. The unweighted cross-country average of the item non-response rate is equal to 29 per cent.

Since Wave 1, we have handled these large fractions of missing values with the fully conditional specification (FCS) method of van Buuren et al. (1999). The FCS method uses a Gibbs sampling algorithm, which imputes multiple variables jointly and iteratively through a sequence of regression models. Assume we want to impute arbitrary patterns of missing values on a set of J variables. The basic idea of the FCS method is that, at each step of the iterative process, we impute the missing values on the j-th variable (j=1,…,J) by drawing from the predictive distribution of a regression model that includes as predictors the most updated imputations of the other J – 1 variables (as well as other fully observed predictors). The process is applied sequentially to the whole set of J variables and is repeated in a cyclical manner by overwriting at each iteration the imputed values computed in the previous iteration. Despite a lack of rigorous theoretical justification (see, for example, 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).

In Wave 8, we computed FCS imputations separately by country and household type. The household types considered were singles and third respondents (sample 1), couples with both partners interviewed (sample 2), and all couples with and without a non-responding partner (sample 3). The distinction between the first two samples was primarily motivated by the fact of using socio-demographic characteristics of the partner of the designed respondent as additional predictors to impute the missing monetary amounts within couples. The overlapping partition of the last two samples was instead motivated by the need to impute properly total household income in the couples with a non-responding partner.

The set of monetary variables imputed jointly in the Gibbs sampling algorithm was country- and sample-specific as we required a minimum number of donor observations for estimating the regression model associated with each variable[12]. Variables that did not satisfy this requirement were imputed first (either by hot-deck or by regression imputations) and then used as fully observed predictors for computing the FCS imputations of missing values in the other monetary variables.

The imputation of each monetary variable was typically based on a two-part model that involved a probit model for ownership and a linear regression model for the amount conditional on ownership[13]. Depending on eligibility and ownership, we converted (if needed) non-zero values of monetary variables in annual euro amounts to avoid modelling differences in the time reference periods of the various variables and the national currencies of non-euro countries. In an early stage of the imputation process, we also symmetrically trimmed 2 per cent of the complete cases from the country-specific distribution of annual euro amounts to exclude (and then impute) outliers that may have a large influence on survey statistics. Moreover, we applied logarithm or inverse hyperbolic sine transformations to reduce skewness in the right tails of the conditional distribution of each monetary variable[14].

The set of fully observed predictors was also sample-specific. For singles and third respondents *(sample 1)*, it 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 (as perceived by the interviewer in the IV module of the CAPI instrument). For couples with both partners interviewed *(sample 2)*, we added a similar set of predictors for the partner of the designed respondent. For couples with a non-responding partner (those remaining in *sample 3* after excluding the couples in *sample 2*), we restricted the

12 The minimum number of observations was equal to 100 in sample 1 and 150 in samples 2 and 3.

13 For the few variables without an ownership question, such as food at home expenditure *(CO002)* and total household income *(HH017)*, we used a simple linear regression model.

14 We apply the log transformation to variables with a positive support and the inverse hyperbolic sine transformation to variables that may take negative values (e.g. income from self-employment, bank account and value of own business).

additional set of predictors referring to the non-responding partner to age and years of education only[15].

Imputations of the monetary amounts were always constrained to fall within individual-level bounds that incorporated the partial information available on the missing observations (e.g. country-specific thresholds used to trim outliers in the tails of the observed distribution of each monetary variable, bounds obtained from the sequence of unfolding bracket questions asked by design to non-respondents of open-ended monetary variables and lower bounds based on the observed components of aggregated monetary variables).

As usual, the imputation of total household income received particular attention because the CAPI questionnaire provides two alternative measures of this variable. The first measure (*thinc*) can be obtained by a suitable aggregation at the household level of all individual income components, while the second (*thinc2*) can be obtained via the one-shot question on monthly household income (*HH017*). As discussed in De Luca et al. (2015), it is not easy to find strong arguments to prefer one measure over the other. Moreover, the availability of two alternative measures may greatly improve the imputation process because each measure could contribute relevant information on the missing values of the other measure. Specifically, to avoid understating the first measure of total household income in couples with a non-responding partner, we adopted the following three-stage algorithm:

Stage 1. For singles and third respondents (*sample 1*), we imputed all monetary variables by the FCS method discussed before. At the end of each iteration of the Gibbs sampling algorithm, we also computed total household income (thinc), household net worth (*hnetw*) and total household expenditure (*thexp*) by suitable aggregations of the imputed income, wealth and expenditure items. Next, we imputed the second measure of total household income (*thinc2*) using the first measure of total household income (*thinc*), household net worth (*hnetw*), total household expenditure (*thexp*) and socio-demographic characteristics of the household respondent as predictors. The imputed values of *thinc2* were constrained to fall in the bounds derived from the sequence of unfolding bracket questions for the variable *HH017*.

Stage 2. For couples with both partners interviewed (*sample 2*), the imputation strategy is similar to that adopted in stage 1 for the sample of singles and third respondents (*sample 1*). The main difference is that in each iteration of the Gibbs sampling algorithm we employed a larger set of predictors that also included socio-demographic characteristics and the most updated imputations of the monetary variables of the partner of the designed respondent.

Stage 3. Imputed values of all monetary variables for the subsample of couples with both partners interviewed were obtained in stage 2. In stage 3, these couples were included in the imputation sample only as donor observations to impute the missing values in monetary variables for the remaining subsample of couples with a non-responding partner. In this case we imputed first all monetary variables for the responding partners using the FCS method. Unlike stage 2, the predictors referring to the non-responding partner now consisted, however, of age and years of education only. At the end of each iteration of the Gibbs sampling algorithm, we also imputed the second measure of total household income (*thinc2*) using household net worth (*hnetw*), total household expenditure (*thexp*) and socio-demographic characteristics of the responding partner as predictors and bound information obtained from the sequence of unfolding bracket questions for the variable *HH017*. Finally, we imputed the first measure of total household income (*thinc*) using the second measure of total household income (*thinc2*), household net worth (*hnetw*), total household expenditure (*thexp*) and socio-demographic characteristics of the responding partner as predictors, couples with two partners interviewed as donor observations and the imputed sum of individual income sources of the responding partner as a lower bound.

To account for the additional variability generated by the imputation process, we always provide five different imputations of the missing values. Multiple imputations were constructed through five independent replicates of the hot-deck/FCS imputation method. Notice that neglecting this additional source of uncertainty by selecting only one of the five available replicates in the generated imputations module (gv_imputations) may result in misleadingly precise estimates. Convergence of the Gibbs sampling algorithm for FCS imputations 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.

15 In the few cases where the number of donor observations available in the estimation step was lower than 30, we employed a smaller subset of predictors, namely gender, age, years of education and self-reported health.

References

Andridge, R. R., & Little, R. J. A. (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. Springer.

Arnold, B. C., Castillo, E., & Sarabia, J. M. (2001). Conditionally specified distributions: An introduction. Statistical Science, 16(3), 249–274.

Börsch-Supan, A. (2020a). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8. COVID-19 Survey 1. Release version: 0.0.1. beta. SHARE-ERIC. Data set. https://doi.org/10.6103/SHARE.w8cabeta.001

Börsch-Supan, A. (2020b). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8. Release version: 0. SHARE-ERIC. Preliminary data set.

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 (pp. 85-100). MEA, Max Planck Institute for Social Law and Social Policy.

De Luca, G., & Rossetti, C. (2019a). Weights and imputations. In M. Bergmann, A. Scherpenzeel & A. Börsch-Supan (Eds.), SHARE Wave 7 methodology: Panel innovations and life histories (pp. 167–189). MEA, Max Planck Institute for Social Law and Social Policy.

De Luca, G., & Rossetti, C. (2019b). Computing calibrated weights in Stata. SHARE Working Paper Series 43–2019. Munich Center for the Economics of Aging (MEA).

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/10.2307/2290268

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. Statistical Science, 7(4), 457–511.

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

Groves, R. M., & Peytcheva, E. (2008). The impact of nonresponse rates on self-selection bias: A meta-analysis. Public Opinion Quarterly, 72(2), 167–189. https://doi.org/10.1093/poq/nfn011

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.

Lynn, P. (2009). Methods for longitudinal surveys. In P. Lynn (Ed.), Methodology of longitudinal surveys (pp. 1–19). Wiley.

Molenberghs, G., Fitzmaurice, G., Kenward, M. G., Tsiatis, A., & Verbeke, G. (2015). Handbook of missing data methodology. 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.

Rubin, D. B. (1987). Multiple imputation for nonresponse in surveys. Wiley.

Van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. Statistical Methods in Medical Research, 16, 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.

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

Table A 6.1: Gender-Age National Calibration Margins for the Calibrated Cross-sectional Weights of Wave 8

Country	Men				Women				Total
	[50-59]	[60-69]	[70-79]	[80+]	[50-59]	[60-69]	[70-79]	[80+]	
AT	692,191	474,962	350,075	161,732	690,277	514,324	429,922	280,785	3,594,268
BE	802,163	648,581	415,223	238,790	792,093	677,454	487,635	408,179	4,470,118
BG	470,525	433,913	273,241	117,795	476,004	520,171	407,681	220,815	2,920,145
CH	646,594	457,734	328,274	168,216	635,284	473,791	377,379	275,436	3,362,708
CY	53,572	45,812	29,742	13,348	54,845	47,747	34,164	18,665	297,895
CZ	666,784	643,683	415,820	145,542	653,390	715,219	557,948	287,365	408,5751
DE	6,767,896	4,987,359	3,503,497	2,025,017	6,706,270	5,315,052	4,182,432	3,364,089	36,851,612
DK	400,964	325,943	262,429	103,820	396,811	336,986	289,629	159,926	2,276,508
EE	82,421	69,650	39,223	19,117	89,403	92,800	70,564	55,600	518,778
ES	3,427,300	2,509,740	1,734,281	1,068,505	3,486,303	2,703,891	2,084,157	1,812,379	18,826,556
FI	366,458	351,333	246,439	108,055	366,324	373,436	293,480	194,655	2,300,180
FR	4,292,095	3,793,351	2,481,629	1,464,385	4,496,175	4,205,880	2,964,334	2,642,280	26,340,129
GR	714,255	602,074	450,866	312,655	786,535	676,684	545,121	447,779	4,535,969
HR	284,323	264,078	147,496	71,412	297,231	296,359	210,691	146,221	1,717,811
HU	590,879	579,975	324,657	127,425	627,392	732,233	514,932	305,608	3,803,101
IL	406,588	347,274	210,114	109,198	426,100	388,107	252,035	163,941	2,303,357
IT	4,578,610	3,511,037	2,727,000	1,605,281	4,773,621	3,826,173	3,235,533	2,724,793	26,982,048
LT	197,840	143,592	80,935	43,367	227,204	199,804	154,548	118,172	1,165,462
LU	45,729	30,197	17,583	9,021	42,096	29,946	19,751	15,261	209,584
LV	125,437	100,102	57,730	26,854	145,230	140,106	113,267	80,659	789,385
MT	30,110	29,526	19,955	7,912	28,852	29,778	22,730	12,934	181,797
NL	1,258,588	1,038,005	730,336	307,968	1,249,800	1,051,908	791,774	490,852	6,919,231
PL	2,314,260	2,362,249	1,070,816	514,817	2,406,732	2,790,943	1,574,555	1,145,559	14,179,931
PT	696,521	595,393	415,892	236,885	782,400	691,534	548,704	424,571	4,391,900
RO	1,242,261	1,157,669	602,331	316,528	1,233,703	1,389,591	880,666	589,870	7,412,619
SE	651,921	554,220	466,408	207,684	634,895	560,157	497,859	314,449	3,887,593
SI	153,747	136,366	75,445	36,422	149,902	140,115	95,536	74,611	787,533
SK	349,422	315,265	147,993	54,719	358,875	369,921	227,013	124,794	1,948,002

Table A 6.2: Gender-Age National Calibration Margins for the Longitudinal Weights of Waves 7 and 8

Country	Men				Women				Total
	[50-59]	[60-69]	[70-79]	[80+]	[50-59]	[60-69]	[70-79]	[80+]	
AT	641,916	425,171	293,227	101,159	649,968	474,602	377,426	197,514	3,160,983
BE	780,772	600,343	338,619	150,758	781,348	641,318	426,507	287,540	4,007,205
BG	462,860	404,824	210,042	69,239	486,490	518,959	338,629	136,263	2,627,306
CH	609,106	427,735	273,061	107,327	597,391	452,618	330,130	194,427	3,123,558
CY	51,799	42,136	24,034	7,671	53,876	44,947	28,440	11,674	264,577
CZ	655,340	623,733	296,595	87,898	660,149	726,857	435,224	191,820	3,677,616
DE	6,398,988	4,387,725	3,330,805	1,132,381	6,402,957	4,796,985	4,204,401	2,149,719	32,803,961
DK	377,632	320,726	205,745	60,700	376,853	336,264	238,206	105,935	2,022,061
EE	80,083	61,275	32,113	10,799	91,955	88,242	65,943	35,666	466,076
ES	3,186,720	2,273,951	1,459,554	717,612	3,269,944	2,508,630	1,861,721	1,316,017	16,594,149
FI	364,057	348,678	184,520	65,441	368,773	379,871	237,100	134,048	2,082,488
FR	4,179,699	3,649,049	1,936,799	965,970	4,442,788	4,105,991	2,467,897	1,913,064	23,661,257
GR	676,197	567,672	396,558	198,600	755,403	647,373	506,928	292,844	4,041,575
HR	288,285	234,875	122,872	39,536	307,264	277,605	192,420	90,979	1,590,340
HU	569,112	523,511	251,723	76,040	635,270	696,793	445,832	196,556	3,394,837
IL	380,842	323,108	161,092	69,022	406,329	367,274	202,439	109,958	2,020,064
IT	4,271,615	3,392,051	2,313,368	1,000,711	4,497,087	3,747,290	2,915,342	1,902,851	24,040,315
LT	193,340	118,862	71,173	24,963	232,518	180,405	149,832	75,514	1,046,607
LU	41,157	26,281	14,261	5,696	38,514	26,498	17,380	10,754	180,541
LV	125,019	85,017	49,187	14,585	150,639	130,578	108,960	49,985	713,970
MT	29,833	28,068	14,697	4,486	29,610	29,283	17,890	8,377	162,244
PL	2,442,865	2,053,353	784,862	311,989	2,620,141	2,556,668	1,302,070	766,896	12,838,844
PT	678,140	550,935	358,266	143,870	761,523	656,475	500,876	285,679	3,935,764
RO	1,128,197	1,014,797	494,872	181,635	1,196,540	1,285,956	798,850	354,295	6,455,142
SE	612,168	550,185	375,389	128,547	600,384	565,018	417,055	217,611	3,466,357
SI	150,661	121,342	61,937	21,233	148,776	129,297	85,917	50,960	770,123
SK	351,918	271,024	108,079	32,221	372,424	337,740	186,921	79,537	1,739,864

Table A 6.3: Gender-Age National Calibration Margins for the Longitudinal Weights of the Fully Balanced Panel (Waves 1-8)

Country	Men			Women			Total
	[50-59]	[60-69]	[70+]	[50-59]	[60-69]	[70+]	
AT	395,994	265,326	67,998	444,475	351,754	148,713	1,674,260
BE	557,019	304,036	97,811	602,188	399,757	212,144	2,172,955
CH	423,805	245,054	73,879	446,730	303,050	150,123	1,642,641
DE	4,133,958	3,273,280	748,171	4,516,076	4,205,946	1,525,452	18,402,883
DK	316,423	165,945	36,175	334,147	201,451	74,751	1,128,892
ES	2,046,682	1,238,257	437,780	2,333,482	1,701,382	914,894	8,672,477
FR	3,320,951	1,756,371	694,005	3,775,566	2,342,234	1,519,876	13,409,003
IT	3,112,610	2,106,999	608,107	3,455,659	2,784,107	1,329,301	13,396,783
SE	547,372	306,847	83,886	563,213	356,514	162,161	2,019,993

II

SHARE CORONA SURVEY (CATI)

CHAPTER 7

Sampling for the First SHARE Corona Survey after the Suspension of Fieldwork in Wave 8

07

7 SAMPLING FOR THE FIRST SHARE CORONA SURVEY AFTER THE SUSPENSION OF FIELDWORK IN WAVE 8

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) and Giuseppe De Luca – University of Palermo

7.1 Introduction

After fieldwork was suspended in mid-March due to the pandemic, it soon became clear that a quick return to the regular face-to-face Computer-Assisted Personal Interview (CAPI) was unlikely. Therefore, it was decided to complement the regular SHARE data with measurements of the current situation, especially during the lockdown. This led to the development of the SHARE Corona Survey, which resumed interviewing with a Computer-Assisted Telephone Interview (CATI), collecting data on the same topics as in the regular SHARE questionnaire but shortened and targeted at the COVID-19 living situation of people aged 50 or older (see Chapter 8; Scherpenzeel et al., 2020).

7.2 Sample for the SHARE Corona Survey

For the new CATI instrument on the COVID-19 outbreak, a sample was selected in each country that included: 1) panel members who had not been interviewed before the suspension of fieldwork; and 2) panel members who had already been interviewed face-to-face in Wave 8. In this respect, refusals were excluded before to be in accordance with GDPR rules. While in most countries the complete eligible panel sample could be fielded again, only in two countries, the Netherlands and Sweden, did a stratified sample based on all panel households have to be selected for cost reasons. While stratification was built on regions (i.e. NUTS-3 level), the size of the subsample took into account information from previous waves to calculate an estimated retention rate in order to reach the affordable number of interviews. Overall, about 84,000 eligible respondents were fielded for an interview (see Table 7.1).

Table 7.1: Overview of Gross Sample Sizes for the SHARE Corona Survey

Country	Size of CATI sample (households)	Size of CATI sample (respondents)	CATI sample: not yet interviewed respondents	CATI sample: already interviewed respondents
Austria	2805	4060	2491	1569
Bulgaria	848	1255	344	911
Belgium (fr.)	2087	3145	2394	751
Belgium (nl.)	1927	3043	1652	1391
Switzerland	1814	2859	678	2181
Cyprus	817	1339	769	570
Czech Republic	3175	4722	1982	2740
Germany	2302	3569	896	2673
Denmark	2395	3828	1513	2315
Estonia	3914	5586	2569	3017

Country	Size of CATI sample (households)	Size of CATI sample (respondents)	CATI sample: not yet interviewed respondents	CATI sample: already interviewed respondents
Spain	1918	2965	1592	1373
Finland	1147	1864	666	1198
France	2841	4193	1729	2464
Greece	2942	4614	1673	2941
Croatia	1621	2604	1390	1214
Hungary	1557	2354	1563	791
Israel	1632	2556	1603	953
Italy	3353	5493	3264	2229
Lithuania	1347	1947	497	1450
Luxembourg	1097	1794	794	1000
Latvia	895	1277	467	810
Malta	697	1171	361	810
Netherlands	1063	1645	940	705
Poland	2976	4773	2627	2146
Portugal	1200	1948	1948	0
Romania	1230	1911	591	1320
Sweden	1281	1999	694	1305
Slovenia	2719	4273	1842	2431
Slovakia	804	1302	293	1009
Total	**54,404**	**84,089**	**39,822**	**44,267**

Note: Due to funding issues a subsample was drawn in the Netherlands and in Sweden.
Data: SHARE Wave 8 COVID-19 Survey 1, Release version: 0.0.1 beta.

Other than the panel sample, the recruitment of the Wave 8 refreshment samples was not continued after the suspension of the regular face-to-face fieldwork, nor were any of the already recruited refreshment sample members reinterviewed. The reason for this choice was that telephone numbers are unavailable for most refreshment sample households. Further, the benefit of merging the COVID-19 data with the wealth of the SHARE panel data from respondents in previous waves was valued higher than refreshment respondents. However, it was decided to continue the already drawn and so far not recruited refreshment samples in Wave 9, provided face-to-face interviewing will be possible again.

In contrast to many other cross-national studies, SHARE includes persons living in nursing homes. It was decided that these panel members should also be asked to participate in the SHARE Corona Survey. However, interviewers were instructed to avoid pressing refusal conversion attempts among nursing home respondents, or on the caretaker or staff members of the nursing home if they are hesitant to allow the interview. This was considered ethically undesirable given the burden that the COVID-19 outbreak puts on nursing home staff and inhabitants. In addition, End-of-Life interviews were also conducted by phone whenever possible.

7.3 The Future of SHARE and the Second SHARE Corona Survey

The continuation of the regular SHARE Wave 8 by conducting a special SHARE Corona questionnaire over the phone was carried out in 27 European countries and Israel from June until early September 2020. Whether a normal CAPI data collection will be possible again in Wave 9 or even whether returning to CAPI data collection among older or nursing home respondents will ever be feasible again is an open question. Therefore, previous plans to move SHARE gradually towards new and multiple modes of data collection will now be accelerated. In addition, SHARE will conduct a second SHARE Corona Survey in 2021 in order to study the long-term impact of COVID-19[16]. For this, all panel households that participated in the first SHARE Corona Survey and are still eligible will be part of the gross sample for the second SHARE Corona Survey. This will, for example, allow a comparison of how the high-risk group of older respondents coped with the crisis, how the national healthcare and social systems responded to the pandemic and which lessons for the future should be drawn from the very different political reactions of the SHARE countries (e.g. in Sweden) towards the pandemic.

References

Scherpenzeel, A., Axt, K., Bergmann, M., Douhou, S., Oepen, A., Sand, G., Schuller, K., Stuck, S., Wagner, M., & Börsch-Supan, A. (2020). Collecting survey data among the 50+ population during the COVID-19 pandemic: The Survey of Health, Ageing and Retirement in Europe (SHARE). Survey Research Methods, 14(2), 217–221. https://doi.org/10.18148/srm/2020.v14i2.7738

16 At the time of publishing this volume, the second SHARE Corona Survey was already successfully fielded.

CHAPTER 8

SHARE COVID-19 Questionnaire

08

8 SHARE COVID-19 QUESTIONNAIRE

Yasemin Yilmaz, Melanie Wagner and Axel Börsch-Supan – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC)

8.1 Motivation

While SHARE was still collecting data for Wave 8, the pandemic was unfolding at great speed across the globe. Our first priority was to secure the health of our interviewers and respondents. Hence, in March 2020, we had to stop the fieldwork of Wave 8 because face-to-face interviewing would carry an irresponsible risk of infection. On the other hand, however, SHARE was the ideal data set to start research on the effects of the pandemic and the related epidemic control measures. While SHARE is not an epidemiological survey, its interdisciplinarity with a wealth of data on the socio-economic circumstances and the prior information that we have collected thanks to previous panel waves (since 2004) and life histories make it a perfect foundation for COVID-specific additional questions, now elicited via telephone interview to avoid any in-person contacts.

The main design decision was to set up a before-during-after scheme of data collection with four elements. The first element is Wave 8, which took place immediately before the pandemic. The second element is a *first SHARE COVID-19* questionnaire described in this chapter, which is dedicated to straightforward questions with maximum relevance for health, economic, work and family events during the first wave of the pandemic. Third, a *follow-up SHARE COVID-19* questionnaire, again via telephone, will take place around June and July 2021 and is dedicated to the long-run implications of the pandemic and the epidemic control measures, such as long-term health effects, economic hardship due to company closures and unemployment, and social disruptions due to the contact limitations. Finally, the fourth element in this design is *Wave 9* in the winter of 2021/2022.

Our main design principles were *comparability* to the standard Waves 8 and 9, *relevance* to measure the immediate impact of pandemic and control measures on health and everyday life, and *simplicity* to be easily understood over the phone in a relatively short time (max. 30 minutes).

In order to balance the trade-offs between these three principles, we differentiated between three age groups, classified roughly as follows:

1. Professionally active (approx. ages 50–65)
2. Young retirees (approx. ages 65–80)
3. Elderly aged (approx. ages 80–100)

For the first group, those who are still *professionally* active, we survey the economic consequences of the COVID-19 crisis: short-time work, loss of income, unemployment and business closures for the self-employed. In this respect, the shock from the pandemic is comparable to that of the 2008 financial crisis. From earlier waves, we know who was already in a difficult labour market situation and who only fell into one through COVID-19. Again, we can use the new data to distinguish from previous waves, whether such difficulties take place against the backdrop of a functioning or an already precarious socio-economic situation. People in this age group often still have living parents. How good is the contact with them? How often do they help them? On the other hand, some in this age group still have schoolchildren and thus belong to the sandwich generation. How are they dealing with the double burden and to what extent has it increased due to the pandemic?

The *second group of young retirees* appears at first glance to face the fewest problems during the pandemic, since they no longer have a job and therefore no more earnings risks. Their public and occupational pensions are generally unaffected by the pandemic. The questionnaire also focuses on their social contacts. To what extent do they look after their children or help them financially? This is an age group that is still very sociable. So how do they experience restrictions on contact with their children, their grandchildren and their friends? They still face economic consequences from the pandemic, since this is the age group that suffers the most from the loss of wealth in the wake of the economic crisis. Their savings are typically intended for long-term care, accommodation in a comfortable nursing home or the education of their grandchildren. If their assets lose value, they will have a difficult time offsetting such losses in their remaining lifetime. How do they cope with this situation?

With regard to the *third age group of the elderly aged*, an important focus is on their existential fears. Their health is very much threatened by the virus: they are the least likely to survive an infection. How many of them are now losing the already few but vital contacts? Are they isolated from their

children, friends and those who are supposed to care for them? Within Europe, we know that migrant labour is the backbone of care workers for countries in the West (Genet et al., 2012), and many care workers have had to rush back to their home countries in the East. How many among the elderly aged have been left alone as a result? How do these people manage now? Who steps in? There are increasing access and admission restrictions in nursing homes: who provides care now?

The questionnaire is set up to differentiate the answers for all these questions not only by these three age groups, but also by income, education, previous social embedding, health history, etc., since the new data can be linked to all previous waves of SHARE, including the life histories.

8.2 Contents

The questionnaire is structured around five main sections: health (physical and mental) and health behaviour; COVID-19 infections for respondents and their social network; quality of healthcare; work and economic situation; and social relationships.

The time frame of interest was determined by natural events, as a comparison of pre- and post-COVID-19 outbreak experiences. Throughout the survey, the reference point of "since the outbreak of Corona" is used interchangeably to imply this pre-/post comparison as well as encompassing the experiences over the course of the pandemic. Interviewers were instructed to let respondents interpret the "start" of the pandemic relatively freely, as when media attention in their respective countries started intensifying, when the first lockdown measures were introduced, etc.

In addition to special pandemic-related questions, a handful of regular panel items were also included in the questionnaire, as data collection for Wave 8 had come to an abrupt halt. These items were limited to those with immediate relevance for the COVID-19 outbreak, such as existing health conditions, medication intake, household income. Furthermore, the questionnaire included mental health questions from the regular SHARE interview and respondents who reported experiencing mental health troubles were then asked to compare their recent situation to before the outbreak of the pandemic.

Under health behaviour, data collection focused on frequency of contact and select activities, abiding by public health measures and recommendations, such as mask wearing, distancing and hygiene measures. Policies pertaining to mask wearing in public spaces such as supermarkets, public squares and restaurants were mostly public health recommendations or guidelines at the time when the questionnaire was developed. Only later, closer to the fieldwork of the COVID-19 questionnaire, did they become mandatory policies.

Of major interest was the personal exposure of respondents to COVID-19 illness. Respondents were asked if they themselves or any relations had had COVID-like symptoms, positive tests, negative tests or hospitalisations, and if so, the number of people. A further question asked if respondents knew anyone who had died due to/with COVID-19. This set of questions are meant to provide a measure of how prevalent exposure to COVID-19 was for the 50+ population in Europe and Israel.

The quality of healthcare section aims to measure the extent and nature of disruptions in the access to care that the pandemic caused, whether it was due to respondents voluntarily cancelling medical appointments or cancellations or denials of appointments by healthcare providers. As the long-term health effects of the pandemic are still unknown, the SHARE COVID-19 questionnaire provides a first look into questions such as access to the kinds of care that were disrupted, respondents' evaluation of their satisfaction, or lack thereof, with the healthcare they received, including the reasons for it, e.g. lack of attention or long waiting times.

One of the immediate consequences of the pandemic was a widespread rise in unemployment and partial work arrangements. The questions in the work section aim to first determine whether respondents in the SHARE sample have lost their employment due to the pandemic, and if so, for how long. Additional information on the pandemic's effects is also collected in the number of hours worked before the pandemic and if they have increased/decreased since the outbreak of the pandemic. Another aspect of work life in the face of the pandemic was the different work arrangements, the availability of a home office, and, in its absence, the extent to which safety measures were provided, and how safe it felt to work. Questions on new technological skills demanded by these changed work conditions or the provision of safety measures at work are also asked in this section.

Furthermore, the COVID-19 shock meant important changes in incomes for many households. In keeping with existing SHARE conventions, income questions were only asked to the first respondent in couple households, to lower respondent burden. In addition to collecting data on disposable household income, the section also asks questions on difficulty in getting by and financial support received due to the COVID-19 crisis.

In the final questionnaire section, we turn to the impact of the pandemic on social relationships. Respondents were asked about different kinds of contact they had with their

social networks, including their children, parents, family and friends, and the frequency of these contacts. Receiving and providing care and volunteering work are also part of this section. Furthermore, respondents who regularly received home care before the outbreak were also asked about the difficulties they experienced in receiving care during pandemic times.

References

Genet, N., Boerma, W., Kroneman, M., Hutchinson, A., & Saltman, R. B. (2012). Home care across Europe: Current structure and future challenges. Observatory Studies Series. WHO. Available at: https://intranet.euro.who.int/__data/assets/pdf_file/0008/181799/e96757.pdf

CHAPTER 9

Software Innovations: Quest

09

9 SOFTWARE INNOVATIONS: QUEST

Marika de Bruijne, Sebastiaan Pennings and Iggy van der Wielen – Centerdata

9.1 Background

In March 2020, the development of the SHARE Wave 9 CAPI survey was well underway. The programming proceeded smoothly as we could still profit from the larger software improvements that had taken place for Wave 8 (Martens, 2020). Then, the COVID-19 outbreak brought the ongoing Wave 8 fieldwork to an abrupt halt. It was clear that the face-to-face interviews could not be continued. As the world was figuring out how to react to the new circumstances, SHARE rapidly made a plan to capture the effects of the crisis among the European 50+ population.

9.2 Challenge

Since face-to-face interviews were clearly not an option, an alternative mode for a new SHARE Corona Survey was called for. Telephone interviews were soon established as the best way to proceed (Scherpenzeel et al., 2020). The technical challenge was: how to develop a new questionnaire for telephone interviews, have it translated by all participating SHARE countries and keep managing the longitudinal sample. The time frame until the start of the main fieldwork: two months, including a pretest round.

9.3 Solution

A survey mode can be defined by several dimensions: for example, some think of it as the means of communication, others as the technology that is used (Couper, 2011). To meet the given timeline, the new questionnaire needed to be robust and quick to implement. The most feasible option was to develop a web-based solution. The regular approach for the CAPI interviews, programming an offline questionnaire in Blaise, would have been more time-consuming and complicated to install on interviewers' laptops due to a lockdown or similar situations in many countries. The strategy for the SHARE Corona Survey was therefore to conduct telephone interviews that would be running an online questionnaire – in other words, using the telephone as the communication medium and a web survey as the underlying technology. In the following, we will refer to this survey mode simply as CATI (Computer-Assisted Telephone Interviewing).

After an investigation of the available products, SHARE Central decided to use Quest, the software package developed by Centerdata, to handle (online) questionnaires with large numbers of respondents and multiple languages. Quest is a PHP-based web application that is split into two parts. The first and primary part is the "Engine", which holds the core functionalities for running questionnaires and handling respondent interaction. The second part is the "Designer", which connects to the engine and holds functionalities for questionnaire programming and theming, as well as importing and exporting of questionnaire structure, translations, respondent data and time stamp information. The application is split up in this way to provide extra security: the part that respondents use (the engine) and the part that developers and managers use (the designer) are separated, meaning that respondents can never gain access to any of the management and import/export functionality.

A multilingual Quest survey typically runs on one master version, with one database in which the data for all language versions are stored. This central approach significantly increases the efficiency of the many steps of data collection and data delivery. Via the "Designer", Quest also makes the data and metadata available to administrators in real time and in various formats, including CSV exports and imports, as well as labelling syntaxes for SPSS. An automated html codebook ("paper version") of the questionnaire can be exported in each language. Furthermore, Quest supports metadata exports of the questionnaire in different formats, such as XML, .json and DDI 3. Figure 9.1 shows an example of the Corona Survey that was developed for SHARE as it appears to respondents in the Quest "Engine" user interface.

Figure 9.1: An Example of the Quest "Engine" User Interface of the SHARE Corona Survey

One of the advantages of Quest was that it could communicate with the other SHARE tools: TranslationCTRL, the tool to manage translations and formerly known as Translation Management Tool (TMT; see Martens et al., 2015), and CaseCTRL (see Chapter 4), the application to manage the respondents' contact and household information and eligibility. The use of TranslationCTRL enabled a controlled translation process. Continuing to use CaseCTRL, which was already in the field and installed on interviewers' laptops, prevented a disruption of the longitudinal household data and data structure. Furthermore, the separate End-of-Life interview that had been programmed in Blaise and that could be conducted via telephone without any technical changes was already available via CaseCTRL. Altogether, the SHARE Corona Survey was made possible through the integration of three applications: Quest, TranslationCTRL and CaseCTRL.

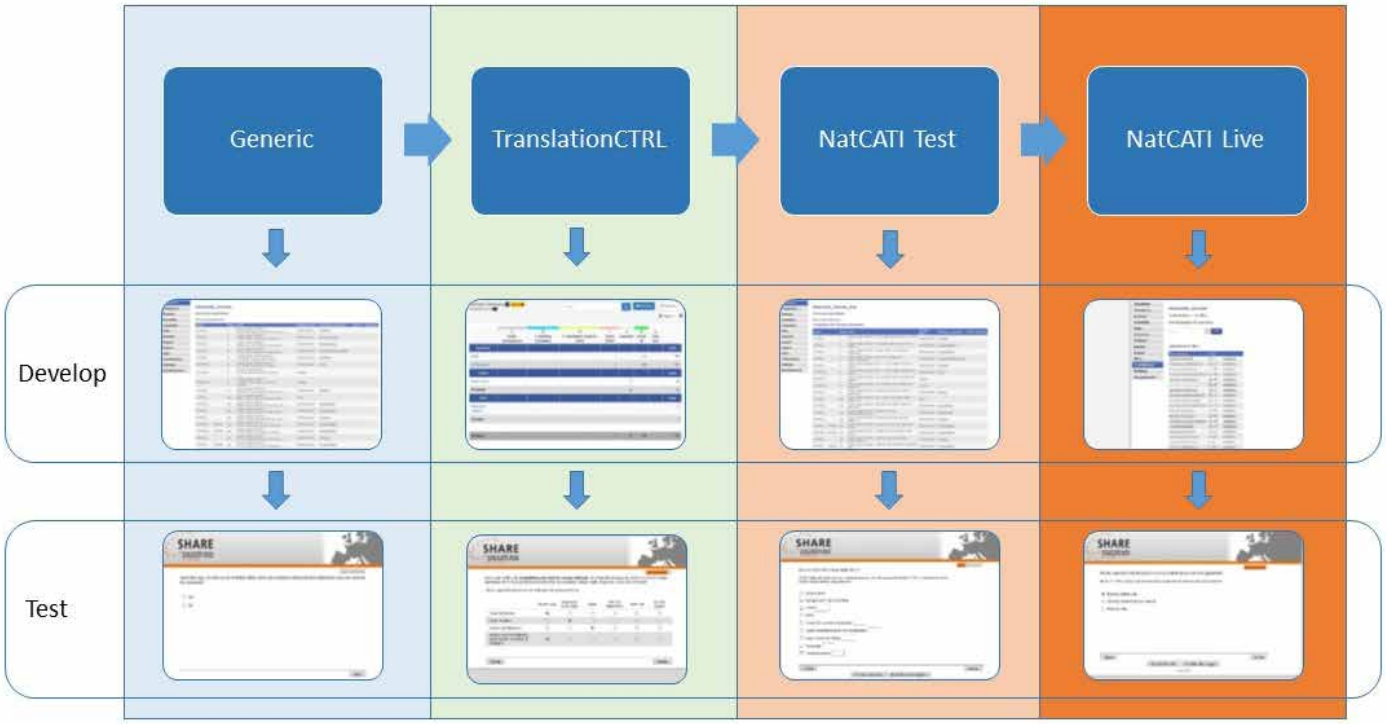
9.4 Survey Development

The preparation of the Quest survey included four main phases: developing the generic version, translating it, creating the national CATIs and making them ready for the fieldwork. Each phase included preparation and testing. The process is illustrated in Figure 9.2.

First, the generic, English questionnaire was programmed in Quest. The online mode enabled quick iterations between the developers and SHARE Central, as there was no need to install a local instrument on the user's device. The developers could provide test links that led directly to the web survey or even showcase the changes in real time. In this phase, the questionnaire was developed and tested as a stand-alone web survey.

When this generic source version of the questionnaire was ready in Quest, the definitions were exported and subsequently imported into TranslationCTRL. There, Wave 8 translations already existing within the system were copied into the CATI when sufficient text matches could be found based on the item names and source texts. Once this was done, the translators could use their personal login to view and translate the questions and the related instructions that SHARE Central had added per question. They could also ask questions about the translations and view examples of other translations via the tool. What was new was the possibility of using a preview function in which TranslationCTRL uploaded translations to Quest, allowing the translators to see how their translation appeared in the Quest "Engine" user interface. This way the translators could already run the first tests on the national version and immediately fix any errors in the translation. This reduced the need for corrections at a later stage.

Figure 9.2: The Development Process of the Quest Survey



When the translators were ready, the translations were exported from TranslationCTRL as PO files and uploaded to the master Quest survey on a test server. Via SHARE Central, test links to the online survey were provided to each Country Team, so that those Country Team members who do not use TranslationCTRL could also test the national CATI survey. If any changes to the translation were still needed, another round of test links was provided after the translation had been improved.

When the Country Team gave the green light to their national version, it was set ready on the SHARE Live Server. Using dedicated test links, the Country Team could start their CATI to confirm the final version. The agency also tested the survey via CaseCTRL using country-specific test cases and confirmed that the survey was ready for the fieldwork. CaseCTRL had been modified for the SHARE Corona Survey so that it showed a new COVID-19 start button in the household members' overview that initiated the respondent's individual CATI interview. No new installation of CaseCTRL was needed on interviewers' laptops to enable this feature, but the interviewers did need to run a patch update. When the fieldwork started, interviewers could click on a button in their local CaseCTRL application that directed them to the online Quest survey.

9.5 Changes in Implementation

The implementation of the SHARE Corona Survey as a web-based application brought about several changes to the usual workflows in SHARE. In Table 9.1, we give an overview of the differences between using an offline survey that is locally installed on interviewers' laptops, as is the case with the SHARE CAPI, and a centrally running web-based survey, as was the case with the CATI.

To deliver a generic or national version of the SHARE CAPI survey, we normally build an executable program that contains both the CaseCTRL and the CAPI survey. We deliver this executable installer to the tester, either at SHARE Central or the Country Team, who then installs it locally on his or her computer. Prior to the installation, the tester needs to uninstall any old versions of the SHARE instrument if present. Later, when the survey is ready for the field, the agency installs the instrument on the laptops of all interviewers. The main benefit of the local installation is of course that no Internet connection is needed to conduct the interviews, allowing the interviewers to visit the respondents at home even at locations with inferior Internet coverage. During the interviews, the data are stored locally on the interviewers' laptops. Thanks to this, all users can use the same test cases.

At frequent intervals, for example at the end of the day, the interviewers connect their laptops to the Internet to synchronise the data to the agency's server. The agencies then synchronise the data to Centerdata, usually every two weeks. After this, all the interviewers' databases are unpacked and merged, and delivered to SHARE Central.

Table 9.1: Comparison of Technical Aspects of Using Local and Web-based SHARE Surveys

	Local survey	Central web-based survey
Delivering generic and national surveys	Survey is built into an executable installer and delivered to the user	Generic version is directly ready, translations are uploaded to a central server
Testing	Each survey version is installed locally, old versions need to be uninstalled, testers can share test cases	Possible via test links, no installation needed, separate test cases needed for each tester
Installation of final version	Survey is installed on interviewers' laptops	No installation needed for the survey itself
Internet	No Internet connection needed	Stable Internet connection is required
Database	Separate database on each laptop	One database for all countries
Data extraction	Synchronisation of all interviewers' laptops required, databases need to be merged	Survey data directly available on central server

In the first phase of developing the generic version of the web-based CATI, we did not integrate it with CaseCTRL but enabled direct access to the stand-alone online survey. Since the survey could be accessed via a web browser, no user installations or uninstallations were needed for testing. The generic version as well as the translated national versions were tested using unique test links. All testers needed their own set of unique test cases, because all responses were stored in the same central database. Prior to the fieldwork, the interviewers did not need to install anything for the Quest survey itself. However, as noted earlier, they still needed an updated CaseCTRL to initiate the survey. The main limitation of the web-based survey was that the interviewer needed to have a good and stable Internet connection throughout the interview. During the interview, the answers to the questions were stored on the central server after completion of each survey page. This continuous interim storage guarantees that there is no data loss if the questionnaire is prematurely terminated, whether or not this is done intentionally. After completing the interview, the synchronisation was still needed for the CaseCTRL data, but not for the Quest data. All Quest survey data were stored in one central database and could be automatically extracted to SHARE Central.

9.6 Concluding Remarks

The switch to a web-based questionnaire in Quest to support the telephone interviews of the SHARE Corona Survey introduced many benefits in terms of efficiency. The programming of the generic version, the translation using a preview function, preparing and testing national versions as well as the data extraction were all conducted in record time. With Quest, SHARE took a first step towards a more flexible data collection infrastructure that can rapidly respond to the changing world and be deployed when needed.

References

Couper, M. P. (2011). The future of modes of data collection. Public Opinion Quarterly, 75(5), 889–908. https://doi.org/10.1093/poq/nfr046

Martens, M. (2020). Blaise 5 in SHARE. Proceedings of the 19th International Blaise Users Conference, Limassol, Cyprus, (pp. 127–136). Available at: https://www.blaiseusers.org/2020/papers/IBUC2020_S5_5.pdf

Martens, M., van der Wielen, I., Wijnant, A., & Sand, G. (2015). Software innovations in SHARE Wave 5. In F. Malter & A. Börsch-Supan (Eds.), SHARE Wave 5: Innovations & methodology (pp. 51-66). MEA, Max Planck Institute for Social Law and Social Policy.

Scherpenzeel, A., Axt, K., Bergmann, M., Douhou, S., Oepen, A. Sand, G., Schuller, K., Stuck, S., Wagner, M., & Börsch-Supan, A. (2020). Collecting survey data among the 50+ population during the COVID-19 outbreak: The Survey of Health, Ageing and Retirement in Europe (SHARE). Survey Research Methods, 14(2), 217–221. https://doi.org/10.18148/srm/2020.v14i2.7738

CHAPTER 10

Monitoring and Managing Fieldwork in the First SHARE Corona Survey

10

10 MONITORING AND MANAGING SHARE FIELDWORK IN THE FIRST SHARE CORONA SURVEY

10.1 Fieldwork Monitoring and Survey Participation the First SHARE Corona Telephone Survey

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

10.1.1 Introduction

SHARE Wave 8 was special because fieldwork had to be suspended in March 2020 due to the outbreak of COVID-19. Since face-to-face interviewing was not possible anymore, all SHARE stakeholders agreed to pick up the survey again via Computer-Assisted Telephone Interviewing (CATI) as quickly as possible. Evidently, data about the health and living situation of the 50+ population in Europe were needed more than ever to shed light on the short- and long-term health, economic and social implications of the epidemic-control decisions. SHARE data should be complemented by measurements of the current situation by conducting a specifically developed SHARE Corona Survey (SCS) covering topics such as health and health behaviour, mental health, infections and healthcare, changes in the work and economic situation, and social networks (for details, see Chapter 8 on survey content of the SCS). By switching the interview mode to CATI, some of the existing SHARE software tools could be more easily adapted since they were already installed on the interviewers' laptops (for details, see Scherpenzeel et al., 2020).

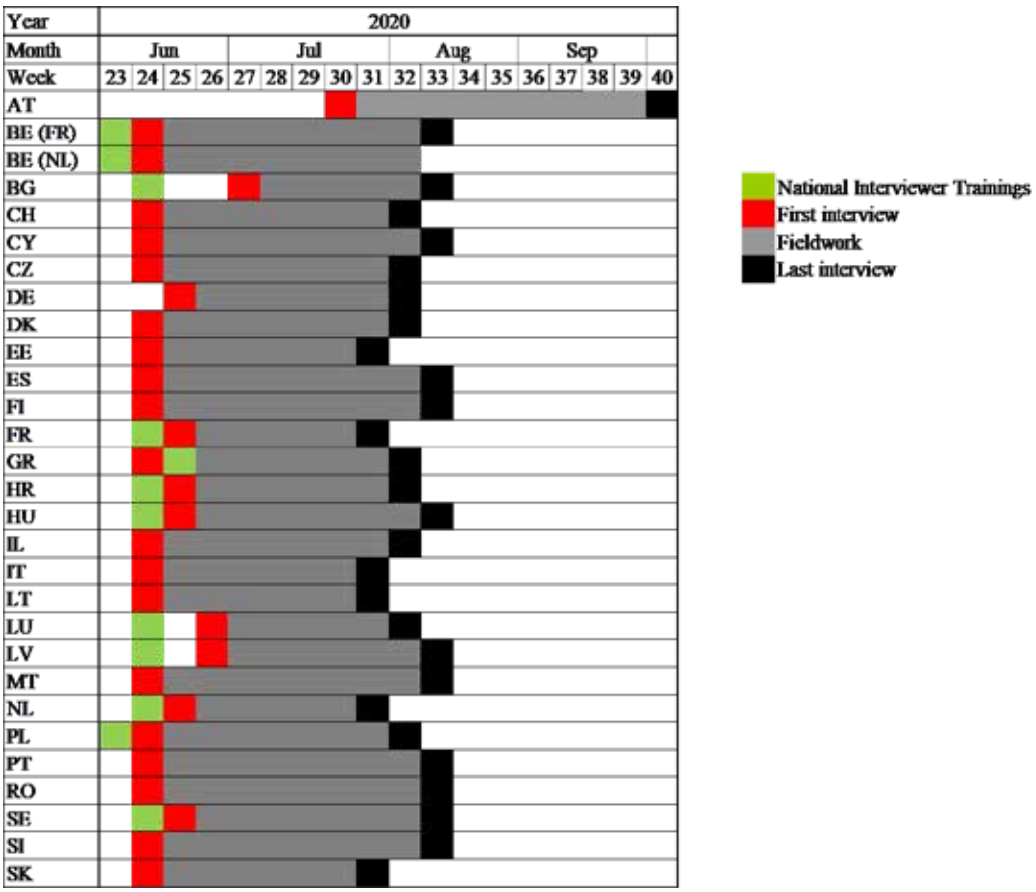
In April and May 2020, the software and fieldwork monitoring procedures were adapted; the SCS was developed, programmed and translated into the 39 SHARE languages; a new preload was prepared containing all panel respondents with or without an interview from Wave 8, plus persons living in nursing homes; the instrument was tested and distributed to the interviewers; the survey agencies and their interviewers received virtual training sessions; and advance letters were sent out to all eligible households. The change from CAPI to CATI mode, questionnaire content and fieldwork design required amendments to the existing contracts. Nevertheless, SHARE maintained its principle of providing the same software tools and programmed questionnaire to all survey agencies in order to harmonise and standardise fieldwork and monitoring.

From June to August 2020, the fieldwork of Wave 8 was resumed via the newly developed SCS in 27 European countries and Israel. Only, the Spanish region of Girona was out of funding sources and was unable to participate in the CATI. Overall, the telephone survey yielded almost 60,000 interviews. In autumn, all preliminary retention rates for the CATI data collection were calculated, an internal release was provided and the SHARE Compliance Profiles (Schuller et al., 2021) were compiled. A follow-up of the SCS took place in 2021.

10.1.2 Fieldwork Periods of the First SHARE Corona Survey

As mentioned above, the fieldwork of Wave 8 was resumed in the form of a shorter telephone survey with questions tailored to the current situation of SHARE respondents. The Train-the-Trainer sessions (TTT) carried out as a webinar by SHARE Central took part at the end of May 2020, followed by virtual national training sessions (NTS) at the beginning of June 2020. In most countries, the first CATIs were conducted in the second week of June (see Figure 10.1). With the exception of Austria, all participating countries managed to complete the SCS within up to two months by mid-August 2020. Austria's delay was for country-specific administrative reasons. Furthermore, no procurement was done for the SCS. Instead, the same agencies that conducted the regular face-to-face fieldwork in Wave 8 carried out the telephone survey.

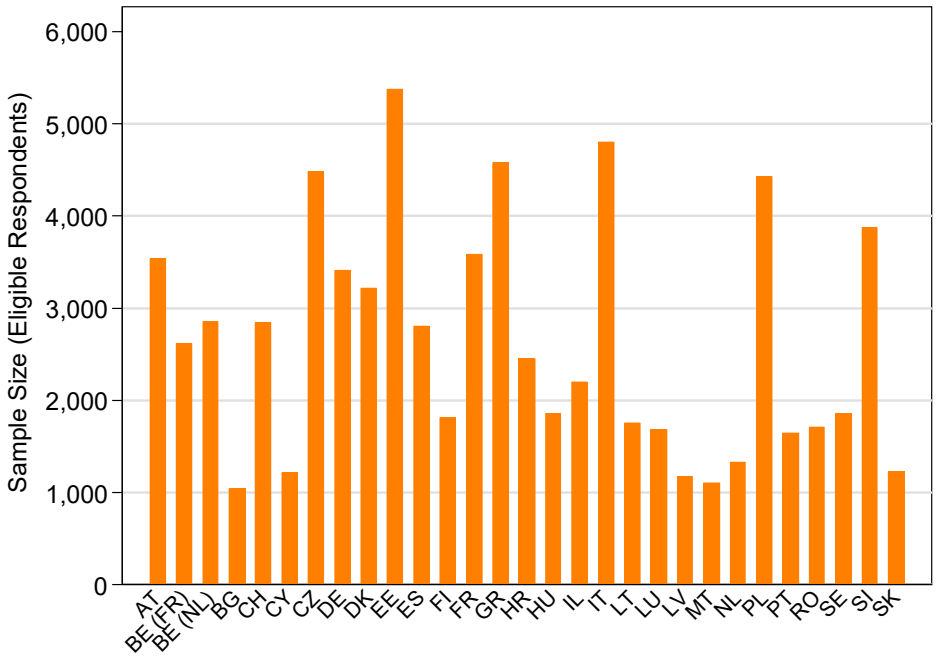
Figure 10.1: Fieldwork Periods of SHARE Corona Survey Wave 8



10.1.3 Monitoring Fieldwork

As in Section 5.1, this section includes information about survey outcomes of the SCS based on the last data export at the end of 2020. All numbers and figures reported during fieldwork are based on information from the CaseCTRL (*read:* case control), formerly known as the SHARE Sample Management System (SMS), which is the interviewer software used to document contact attempts and conduct the interviews (see Sections 5.1.4 and 5.1.5 for detailed information on the classification of survey outcomes and the formulas to compute them). The main data source used to calculate the numbers for the CATI sample was drawn from the Quest software, which ran the SCS from July to August 2020. Due to time constraints, the CaseCTRL software could only be adapted marginally for the switch to telephone interviewing. While retrieving interview numbers from Quest was possible at any time and allowed for weekly updates on the number of collected telephone interviews, the number of End-of-Life interviews was still obtained from CaseCTRL data. The technical quick fix did not enable the extensive fieldwork monitoring that comprises the usually reported AAPOR indicators. We had to resort to providing short weekly updates on the number of completed telephone interviews and the corresponding individual response rates.
Figure 10.2 shows the size of the CATI sample per country. The samples consists of all panel respondents from Wave 8, both with and without already completed CAPI. It can be seen that the sample sizes vary significantly between about 1,000 eligible respondents in countries such as Bulgaria, Cyprus, Latvia and Malta and up to more than 5,000 respondents in Estonia.

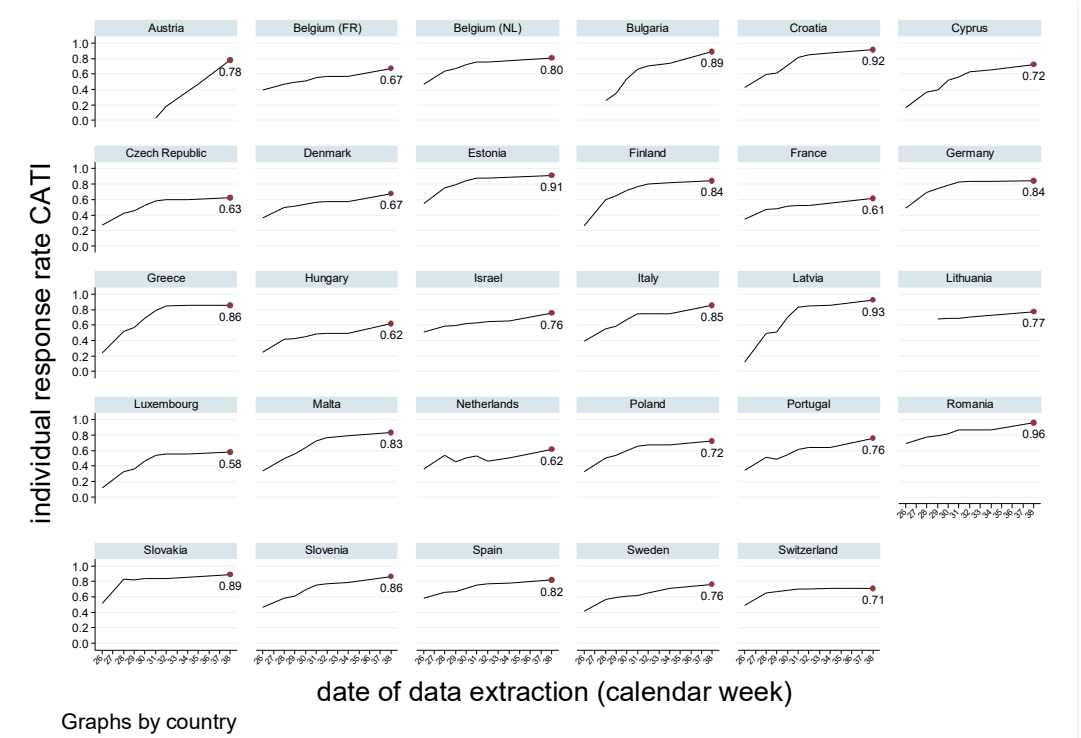
Figure 10.2: CATI Sample Size



10.1.4 Individual Participation in the SHARE Corona Survey

Figure 10.3 shows the individual retention rates of the CATI sample. The rates were adjusted by removing all households for which no telephone numbers could be obtained.

Figure 10.3: Individual Retention Rates in CATI Sample by Country over Time

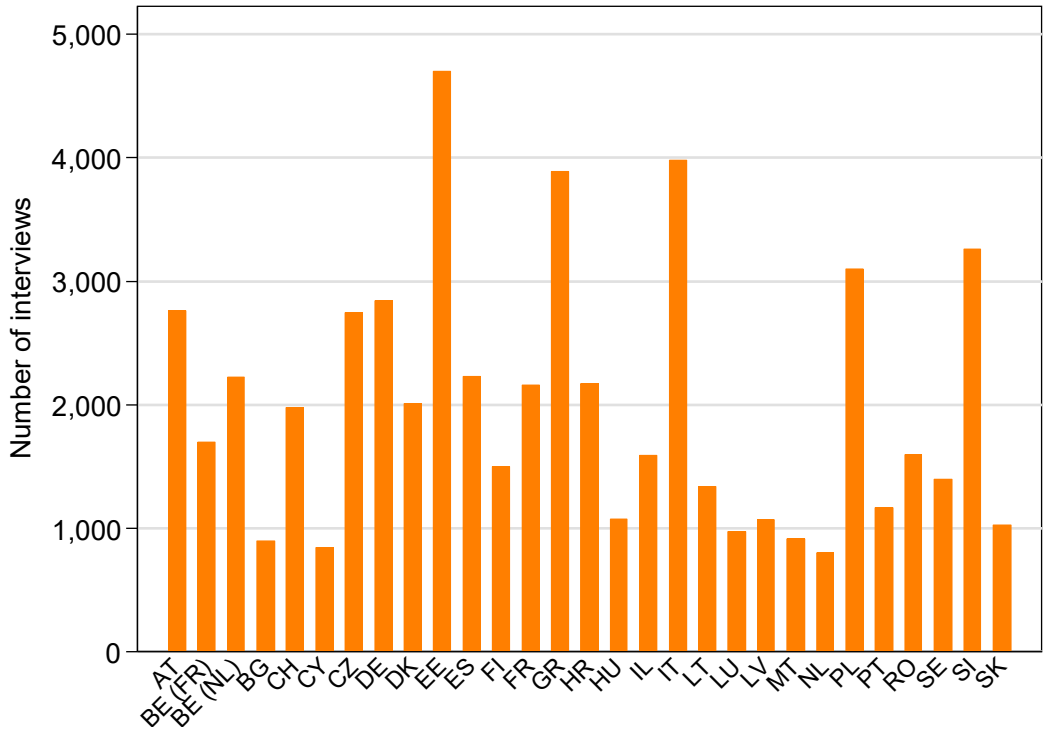


While four countries move around the 60 per cent mark (Belgium (FR), Hungary, Luxembourg and the Netherlands), about half of all countries achieved a retention rate of 80 per cent or more. Among these, the best performance of 96 per cent was achieved by Romania.

10.1.5 Absolute Number of Telephone Interviews

Figure 10.4 shows the absolute number of telephone interviews per country at the end of fieldwork.

Figure 10.4: Absolute Numbers of Interviews in Panel Samples



Owing to the large differences in sample size, the number of completed interviews also varies widely across countries. While Estonia, Greece and Italy conducted about 4,000 or more interviews, a considerable number of countries finished fieldwork around the mark of 1,000 interviews (Bulgaria, Cyprus, Hungary, Luxembourg, Latvia, Malta, the Netherlands, Portugal and Slovakia).

10.1.6 Conclusion

The quick mode switch from CAPI to CATI because of the outbreak of COVID-19 required some sacrifices on the technical side and forced SHARE Central to reduce fieldwork monitoring procedures to the basics. Despite this, fieldwork performance was remarkable in many countries. All survey agencies managed to collect almost 60,000 telephone interviews from the SHARE Corona Survey (see Table A 10.1 in the Annex). Similarly to the regular Wave 8, all participating countries are evaluated on a set of quality control indicators uniformly (see SHARE Compliance Profiles; Schuller et al., 2021).With the exception of Greece, where one file is missing for the SCS, all countries participating submitted the required input documentation and deliverables, including the refreshment sample and panel gross sample data, National Training Session (NTS) dates, NTS observation protocol, NTS slides, interviewer roster, advance letters and completed Survey Agency Feedback Form (SAFF). Attendance of survey agency trainers at the virtual TTTs was extremely satisfactory. Data collection for both surveys was achieved in a largely synchronous fashion across all participating countries, with one exception. Austria had a substantial delay (more than one month) between obtaining their sample software and delivering the first interview in the SCS. However, this has to be evaluated against the difficult circumstances and the extremely tight schedule. Therefore, we conclude that despite some challenges, the first SHARE Corona Survey was a great success.

References

Scherpenzeel, A., Axt, K., Bergmann, M., Douhou, S., Oepen, A., Sand, G., Schuller, K., Stuck, S., Wagner, M., & Börsch-Supan, A. (2020). Collecting survey data among the 50+ population during the COVID-19 outbreak: The Survey of Health, Ageing and Retirement in Europe (SHARE). Survey Research Methods, 14(2), 217–221. https://doi.org/10.18148/srm/2020.v14i2.7738

Schuller, K., Lasson, S., Sand, G., Bergmann, M., Pflüger, S., Hannemann, T., & Börsch-Supan, A. (2021). SHARE compliance profiles – Wave 8. MEA, Max Planck Institute for Social Law and Social Policy.

Table A 10.1: Final Outcomes by Country

Country	Gross sample (HH without telephone numbers excluded)	Number of CATI inter-views	Number of End-of-Life Interviews	Total	Response rate (HH without telephone numbers excluded)	Percentage of HH without telephone numbers
Austria	3541	2773	1	2774	78	12
Belgium (FR)	2621	1715	51	1766	67	16
Belgium (NL)	2859	2242	55	2297	80	5
Bulgaria	1040	914	10	924	89	17
Switzerland	2843	1995	10	2005	71	0
Cyprus	1213	867	7	874	72	9
Czech Republic	4478	2782	18	2800	63	5
Germany	3413	2855	20	2875	84	0
Denmark	3215	2083	73	2156	67	15
Estonia	5375	4738	161	4899	91	4
Spain	2807	2251	54	2305	82	5
Finland	1813	1509	15	1524	84	2
France	3584	2187	7	2194	61	14
Greece	4574	3901	30	3931	86	1
Croatia	2452	2188	65	2253	92	5
Hungary	1857	1099	55	1154	62	21
Israel	2201	1626	36	1662	76	13
Italy	4802	4003	102	4105	85	12
Lithuania	1758	1349	9	1358	77	9
Luxembourg	1687	982	3	985	58	4
Latvia	1174	1079	9	1088	93	8
Malta	1106	920	3	923	83	5
Netherlands	1328	813	8	821	62	19
Poland	4427	3128	67	3195	72	7
Portugal	1643	1180	63	1243	76	15
Romania	1709	1620	15	1635	96	9
Sweden	1859	1414	0	1414	76	7
Slovenia	3873	3271	67	3338	86	9
Slovakia	1225	1087	0	1087	89	4
Total	**76477**	**58571**	**1014**	**59585**	**78**	**9**

Data: SHARE Wave 8 COVID-19 Survey 1, end of fieldwork.

10.2 Data Quality Back-checks in the First SHARE Corona Survey

Tessa-Virginia Hannemann – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC) and Michael Bergmann – 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)

10.2.1 Introduction

As with all aspects of fieldwork, the process of interview verification needed adjustements in response to the pandemic and change of course in Wave 8. The intent of the procedures is presented in Section 5.2 (see also Bergmann & Schuller, 2019); this chapter will only outline the adaptations that were implemented as well as the results of the back-checking done within the frame of the First SHARE Corona Survey (SCS1).

10.2.2 Back-checks: Accommodations Made for CATI

Random back-checking resumed in calendar week 24 (beginning of June 2020). As data from the CATI were synchronised continuously, interviews considered for the selection of the back-checks were those that had been completed between the previous back-check and the current one. Templates including the households to be checked were relayed to the survey agencies using a secure server, as they had been during the main fieldwork.

One adaptation that was undertaken was the exclusion of previously selected households. As some households had been interviewed during the main fieldwork as well as during the CATI, an effort was made to avoid rechecking households. This was done to avoid further stress on the sample households.

Furthermore, a number of verification questions had to be disregarded, as they did not reflect the CATI interview. These included the use of the hand-strength measure and the use of showcards by the interviewer. Table A 10.2 in the Annex of this chapter shows the verification questions asked during the CATI back-checks.

One round of focused back-checks was sent out during the CATI fieldwork. The focused back-checks not only take information on the completion of the interview into account, but also incorporate information from the interview (see Section 5.2 for details). As the CATI interviews differed, not all indicators used in the focused back-checks of the regular Wave 8 could be calculated. The number of indicators was therefore reduced, focusing on interview timing and cooperation rates, as well as deviations in preload information. The indicators used to flag suspicious interviews are presented below in Table 10.1.

Table 10.1: Overview of Indicators Regarding Interview Falsifications

Indicator	Description
1	**Number of interviews on same day:** interview is flagged as suspicious if five or more interviews have been conducted by an interviewer on the same day.
2	**Cooperation rate of partner:** interview is flagged as suspicious if the cooperation rate in households with two eligible respondents (i.e. the number of households with two interviews divided by the total number of all partner households) of an interviewer with at least five partner households is above 95%.
3	**Interview duration:** interview is flagged as suspicious if the residual of a linear regression, using the log normal distribution of interview length (based on keystroke data without the IV module) regressed on key respondent and interview characteristics (year of birth, self-rated health, frequency of questionnaire clarifications, number of asked items, sequence number of interview, questionnaire version and interview language) is below the fifth percentile in the respective country.

Indicator	Description
4	**Unrealistic time of interview:** interview is flagged as suspicious if the starting time of an interview is between 11 pm and 6 am.
5	**Different preload information:** interview is flagged as suspicious if the preload information on gender and year of birth of the respondent does not match the given coverscreen information.

10.2.3 Results

During the SHARE Corona Survey, random back-checks were selected in calendar weeks 24, 26, 29, 31 and 34. A total of 12,245 interviews were selected to be verified. SHARE Central received 9,296 completed back-checks. Of the countries verifying 100% of their interviews, we received 1,326 successful back-checks from DE after the SHARE Corona Survey fieldwork. Table 10.2 describes further the distribution of sent and received back-checks across countries. Focused back-checks were sent out during calendar week 34 (mid-August 2020) after completion of fieldwork. Fifty-five interviews were identified as suspicious. Among the checked interviews, none of the suspicions could be verified, thus interviews needed not be repeated.

Table 10.2: Results of Random Back-checks (CATI)

Country	Households selected	Households checked	Households reached
AT	462	838	656
BE_nl	553	540	398
BE_fr	475	254	191
BG	192	129	116
CH	413	413	239
CY	201	201	172
CZ	640	559	426
DK*	560	732	552
EE	1,051	342	325
ES	533	525	251
FI	326	240	185
FR	441	441	290
GR	857	-	-
HR	515	515	475
HU	249	249	202
IL	385	375	305
IT*	947	1122	1033
LT	262	262	255

Country	Households selected	Households checked	Households reached
LU	198	198	172
LV	235	-	-
MT	180	179	167
PL	706	-	-
PT	350		
RO	331	-	-
SE	302	302	111
SI	704	703	447
SK	177	177	176
Total	**12245**	**9296**	**7144**

Note: * checked and reached households include those checked during Wave 8 CAPI.

10.2.4 Discussion and Lessons Learned

The process of random and focused back-checking implemented at the start of Wave 8 proved itself to be very adaptable to mitigating circumstances. This reflects further the capacity to intervene, which was gained with the currently implemented random back-checking process. Furthermore, we received the feedback in the initial weeks of the CATI fieldwork that households had been selected to be verified, which had previously been selected during the CAPI fieldwork. Again, due to the adaptability of the process we were able to implement a further check of our selection to avoid overstraining the sample. It needs to be taken into account that the number of CATI interviews conducted in a week surpasses the number of CAPI interviews in a week. Thus, once fieldwork restarted, the number of interviews to be checked was high in comparison to the weeks pre-pandemic. There is potential to overwhelm the survey agencies. However, given the extenuating circumstances all faced in the situation of a global pandemic, the reuptake of back-checking and the adaptation of the questionnaire were very successful.

References

Bergmann, M., & Schuller, K. (2019). Improving the efficiency of data quality back checks: A new procedure to prevent curbstoning. In M. Bergmann, A. Scherpenzeel & A. Börsch-Supan (Eds.), SHARE Wave 7 Methodology: Panel innovations and life histories (pp. 154-159). MEA, Max Planck Institute for Social Law and Social Policy.

Table A 10.2: Verification Questions of Random Back-checks

Verification questions	Included in back-checks…		
	DE	NL	SCS1
Household ID (hhidcom)			x
Person ID (pidcom)			x
Laptop ID (laptop_s8)			x
Interviewer ID (interviewerid_s8)			x
Mode and sequence of contact attempts	x		x
Successful contact in the end?	x	x	x
If no: Reason for non-contact	x		x
1) Has one of our interviewers recently interviewed you for the SHARE study?	x	x	x
2) If yes: With whom was the interview conducted?	x		x
3) How was the interview conducted?	x		x
4) Where was the interview conducted?	x		
5a) How long was the interview with you?	x	x	x
5b) If partner interview: How long was the interview with your partner?	x		x
6) Did the interviewer use a device to measure the strength of your hands?	x	x	
7) Did the interviewer use showcards during the interview?	x	x	
8) Did you receive a monetary incentive for your participation in the SHARE study?	x		x
9) What is your year of birth?	x	x	x
10) Only if not sure: What is your gender?		x	x
Other comments by respondent	x	x	x
Evaluation of back-check by survey agency	x	x	x
If interview not ok: Consultation with interviewer regarding suspicion	x	x	x
Final decision by survey agency after consultation with interviewer	x	x	x
Only if interview not ok/fake: Reinterview	x	x	x
Only if reinterviewed: Information on new interview	x	x	x

CHAPTER 11

Weights and Imputations in the First SHARE Corona Survey

11

11 WEIGHTS AND IMPUTATIONS IN THE FIRST SHARE CORONA SURVEY

Giuseppe De Luca, Paolo Li Donni and Moslem Rashidi – University of Palermo

11.1 Introduction

As described in Chapter 1, the regular Computer-Assisted Personal Interview (CAPI) of Wave 8 was suspended in March 2020 due to the COVID-19 outbreak and a new SHARE Corona questionnaire was fielded in June and July 2020 by Computer-Assisted Telephone Interview (CATI) to study the impact of the pandemic on the socio-economic and health conditions of SHARE respondents. Based on the participation in these two survey instruments, one can distinguish three subsamples of primary interest: CAPI, CATI and CAPI & CATI. Calibrated weights and imputations for the CAPI data are discussed in Chapter 6. In this chapter, we describe the construction of calibrated weights for the CATI and CAPI & CATI subsamples and the imputation of missing values due to item non-response in the CATI data.

11.2 Calibrated Weights for the CATI and CAPI & CATI Subsamples

Based on the publicly available beta data release of the SHARE Corona Survey (Börsch-Supan, 2020), the CATI subsample includes 54,600 respondents in 37,222 households who have participated in the telephone survey irrespective of whether they have also participated in the regular CAPI of Wave 8. The CAPI & CATI subsample includes instead the smaller subset of 34,916 respondents in 24,191 households who have participated in both interviews. Notice that, even though the CATI was fielded in June and July 2020, the target population of these two subsamples is the same as the CAPI subsample because they consist of longitudinal respondents only (i.e. people born in 1969 or earlier). The representativeness of the 50+ population in the middle of 2020 is therefore complicated by two issues: (i) the lack of refreshment samples of younger cohorts who were not age-eligible in Wave 7; and (ii) the sizeable effects of mortality due to the COVID-19 pandemic in the first six months of 2020.

The baseline strategy adopted by SHARE to deal with problems of unit non-response and panel attrition relies again on the calibration approach of Deville and Särndal (1992).

For each subsample we always provide calibrated cross-sectional weights and calibrated longitudinal weights, as well as calibrated weights at individual level for inference to the target population of individuals and calibrated weights at the household level for inference to the target population of households. The calibration procedure and the population calibration margins coincide with those described in Chapter 6 for the CAPI subsample of Wave 8. The only difference is that the weights are now calculated on two different subsamples of respondents. For example, since the CAPI & CATI subsample is nested in the CAPI (or CATI) subsample, the calibrated weights of the first subsample must be greater than or equal to the calibrated weights of the second subsample.

11.3 Imputations of Missing Values in the CATI Data

Missing values due to item non-response in the CATI data of Wave 8 were imputed separately from the missing values in the CAPI data. Since the fraction of missing values in the CATI data was generally much less than 3 per cent, the imputation procedure for this data set draws mainly on the hot-deck method. Two exceptions worth noting are the monetary variables on overall monthly household income before the outbreak (*CAHH017*) and the lowest overall monthly household income since the outbreak (*CAE005*). As for the monetary variables collected in the standard CAPI questionnaire, these two open-ended questions are very sensitive and difficult to answer precisely, especially in a CATI mode. In addition, unlike the CAPI questionnaire, the CATI questionnaire does not include sequences of unfolding bracket questions that may allow valuable interval data on the missing monetary amounts to be recovered.

In addition to missing data due to "Don't know" and "Refusal" answers, for some variables we also imputed other types of data inconsistencies due to routing errors in Section E (Economic situation) and measurement errors in Section W (Work) of the CATI questionnaire. Specifically, Section E depends on a filter variable, *CAE001*, which controls in turn all other questions included in this section (i.e. all variables starting with *CAE*). The problem is that, by design, this section would have to be asked only to the first respondent in the household. However, the filter variable was not automatically assigned by the interviewing

software, but rather it was left open for selection by the respondent/interviewer. Thus, instead of having only one respondent per household, the data contain a set of 1,902 individuals in 864 households who have skipped all questions in the economic section and another set of 2,112 individuals in 1,056 households with two respondents per household. For the specific purposes of imputations, we adopted the following strategy: in households that skipped the economic section, we imputed the missing values on all variables of this section by selecting the household member with the minimum person identifier (*mergeid*). In households that have two respondents per household, we selected first the respondent with the largest number of valid answers to the remaining variables of the economic section. Depending on the number of valid answers, we then selected the respondent with the minimum person identifier within each household. The CATI imputation database (gv_imputations) contains an indicator variable (*RESP_E*) that allows the household member who was selected as the eligible respondent for the economic section to be identified.

In Section W, respondents who reported being employed or self-employed at the time when COVID-19 broke out were first asked about their usual working hours per week before the outbreak (*CAW020*). Next, respondents were asked whether they reduced/increased the number of working hours and finally the lowest/highest number of hours of work since the outbreak (*CAW021*, *CAW022*, *CAW024*, *CAW025*). These variables were affected by two types of measurement errors (236 observations in total): (i) people who reported a reduction in the number of working hours, but the lowest number of hours of work since the outbreak was in fact greater than or equal to the number of hours of work before the outbreak; (ii) people who reported an increase in the number of working hours, but the highest number of hours of work since the outbreak was in fact smaller than or equal to the number of hours of work before the outbreak. To handle these types of measurement errors, we imputed in this case all data inconsistencies due to misreporting on the sequence of variables *CAW020*, *CAW021*, *CAW022*, *CAW024* and *CAW025*.

The imputation procedure for these five variables of Section W and other eleven variables of Section E is based on the Fully Conditional Specification (FCS) method. One important difference with respect to the FCS method used to impute the monetary variables of the CAPI questionnaire (see Chapter 6) is that we account for the continuous, binary or categorical nature of the sixteen variables that need to be imputed jointly.

11.4 Hot-deck Imputations

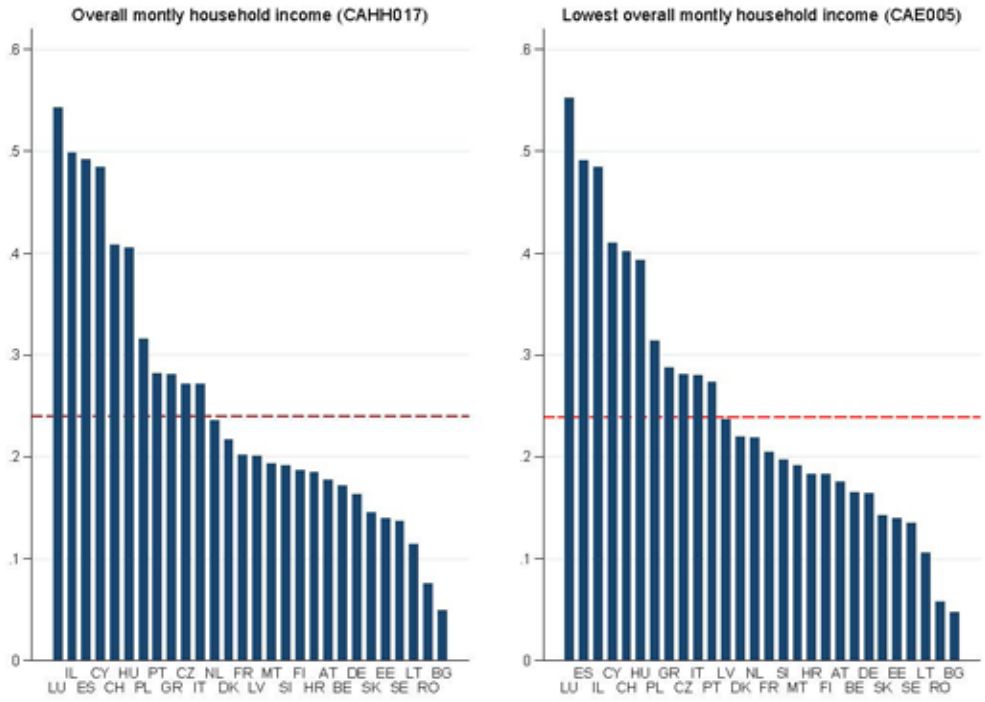
Similarly to Chapter 6, we performed hot-deck imputations separately by country and according to a convenient order of the variables that accounts for branching and skip patterns in the CATI questionnaire. Imputation classes for the implementation of this method were based on the following set of auxiliary variables: country, gender, five age classes ([– 49], [50 – 59], [60 – 69], [70 – 79], [80+]), a binary indicator for respondents living with a spouse/partner, five groups for years of education, a binary indicator for good self-perceived health before the outbreak, and a binary indicator for changes in the self-perceived health status during the outbreak. The first four auxiliary variables are fully observed, while the last three auxiliary variables contain very small fractions of missing values that were imputed first using only the first four variables. Since the CATI data consist of longitudinal respondents only, the information on years of education was obtained by the most recent CAPI data available for each respondent. For some variables we exploited a larger set of auxiliary variables. For example, we used two additional binary indicators for wearing a face mask in public spaces and keeping distance from others in public when imputing several variables included in Section H (Health and health behavior), Section C (Corona-related infection) and Section Q (Quality of healthcare) of the CATI questionnaire. Furthermore, we jointly imputed missing values on the variables that are logically related. Specifically, in Section H, we jointly imputed variables regarding illness or health conditions since the last interview (CAH003 and CAH004). In Section C, we imputed jointly variables regarding the COVID-19 symptoms (*CAC002*, *CAC003*), whether testing positive (*CAC004*, *CAC005*) or negative (*CAC007* and *CAC008*) for coronavirus, and whether being hospitalised (CAC010 and *CAC011*) or having died (*CAC013* and *CAC014*) due to an infection from coronavirus. In Section Q, we imputed jointly variables regarding forwent medical treatment since the outbreak (*CAQ005* and *CAQ006*), postponed medical appointment (*CAQ010* and *CAQ011*), denied medical appointment (*CAQ015* and *CAQ016*), whether treated in hospital and the associated level of satisfaction (*CA025* and *CA027*), and whether visited by a doctor and the associated level of satisfaction (*CAQ020* and *CAQ022*).

In total, we imputed sequentially by the hot-deck method about 200 variables of the CATI questionnaire. As for the hot-deck imputations of the CAPI data, the CATI imputation database contains five multiple imputations of the missing values and a flag variable associated to each imputed variable which allows the users to identify the imputed observations.

11.5 FCS Imputations

After the hot-deck imputations, we also constructed FCS imputations for the missing values on the five hours of work variables collected in Section W (*CAW020*, *CAW021*, *CAW022*, *CAW024* and *CAW025*) and other eleven variables collected in Section E: overall monthly household income before the outbreak (*CAHH017*), lowest overall monthly household income since the outbreak (*CAE005*), a set of six binary indicators for received financial support (*CAE003* and *CAE004*), household's ability to make ends meet (*CACO007*), a binary indicator for the postponement of regular payments (*CAE011*) and a binary indicator for dipping into savings to cover the necessary day-to-day expenses (*CAE012*). In addition to the increased set of missing values generated by routing and measurement errors, we imputed all these variables jointly by the FCS method to preserve their correlations. As explained in Section 11.3, the most worrisome variables are the two monthly household income variables before and after the outbreak (*CAHH017* and *CAE005*). Figure 11.1 shows the item non-response rates of these two monetary variables by country. For both variables, the unweighted cross-country average of the item non-response rate is equal to 24 per cent. However, in six countries (LU, IL, ES, CY, CH and HU), the item non-response rate was considerably greater than the cross-country average (about 40 per cent or more) and it reaches a maximum of 55 per cent in Luxembourg.

Figure 11.1: Item Non-response Rates for Overall Monthly Household Income before the Outbreak (left) and Lowest Overall Monthly Household Income since the Outbreak (right) by Country



Data: SHARE Wave 8 COVID-19 Survey 1, Release version: 0.0.1 beta.

FCS imputations of these sixteen variables of the CATI questionnaire were always constructed separately by country. We do not use separate imputation models for different household types, but we always include a binary indicator for living with a spouse/partner in our set of observed predictors.

At each iteration of the Gibb sampling algorithm, we used a linear regression model for the continuous variables (*CAW020*, CAW022, *CAW025*, *CAHH017*, *CAE005*), a logit model for four binary variables (*CAW021*, *CAW024*, *CAE011*,

and *CAE012*), a multinomial logit model for the categorical variable *CACO007*, and the multivariate hot-deck method for the six binary indicators related to financial support received since the outbreak (*CAE003* and *CAE004*).

For the *CAHH017* and *CAE005* variables, we symmetrically trimmed 2 per cent of the complete cases from the country-specific distribution of each variable to exclude (and then impute) outliers that may have a large influence on survey statistics. Furthermore, we transformed all continuous variables using either the logarithm (*CAW020*, *CAW025* and *CAHH017*) or the inverse hyperbolic sine (*CAW022* and *CAE005*) transformations to reduce skewness in the left tails of their distributions.[17]

In addition to the variables imputed jointly within the Gibb sampling (see Chapter 6), we used as observed predictors a binary indicator for female respondents, a quadratic polynomial in age, years of education, a binary indicator for living with a spouse/partner and its interaction with age of the spouse/partner, a binary indicator for good self-perceived health and a binary indicator for changes in the self-perceived health status during the outbreak. In the linear regression models for *CAHH017* and *CAE005* and in the multinomial model for *CACO007*, we also included as observed predictors a binary indicator for being employed before the outbreak. For the multivariate hot-deck imputations of the six binary indicators related to financial support received since the outbreak we used instead a similar set of observed predictors (properly discretized to form the imputation classes) plus the quantiles of *CAE005* computed at each iteration of the Gibb sampling algorithm.

The final FCS imputation model adopted in each country was subject to an accurate fine-tuning for the choice of the predictors. Specifically, we had to impose a set of country and item-specific exclusion restrictions to avoid possible problems of collinearity, very imprecise estimates, as well as problems of convergence and perfect prediction in the context of non-linear models. As for the other types of imputations in SHARE, we always provide five multiple imputations of the missing values. After an initial set of 15 burn-in iterations, convergence of the Gibbs sampling algorithm was assessed by the Gelman – Rubin criterion applied to the mean, median and 90th percentile of the distribution of each continuous variable and the mean of the distribution of each discrete variable. In all countries, convergence was achieved before the 50th iteration.

References

Börsch-Supan, A. (2020). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8. COVID-19 Survey 1. Release version: 0.0.1. beta. SHARE-ERIC. Data set. https://doi.org/10.6103/SHARE.w8cabeta.001

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/10.2307/2290268

17 For *CAW022* and *CAE005*, we used the inverse hyperbolic sine transformation because these variables may take value zero.

III

ENHANCE SHARE DATA COLLECTION

CHAPTER 12

Measuring Physical Activity in SHARE: The SHARE Accelerometer Study

12

12 MEASURING PHYSICAL ACTIVITY IN SHARE: THE SHARE ACCELEROMETER STUDY

Annette Scherpenzeel – Netherlands Institute for Health Services Research (Nivel), Nora Angleys, Fabio Franzese and Luzia Weiss – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC)

12.1 Introduction

SHARE puts major efforts into harmonisation and standardisation of survey questionnaires across countries. However, the consistency in the set of questions is not necessarily sufficient for measuring the same concepts. For many variables, items suffer from "differential item functioning" (DIF), i.e. interpersonal and intercultural variation in interpreting and using the response categories for the same question (Teresi & Fleishman, 2007). SHARE therefore employs "objective" measures that minimise the DIF of measures based on self-report, such as the measurement of grip strength, peak flow and chair stand measures. Such measures facilitate the comparison across countries and permit adjustments of self-reported measures of health.

The self-reported assessment of physical activity suffers from DIF in several dimensions that are especially relevant in the SHARE sample: old age and (limitations in) mobility make it even more difficult to compare the respondents' answers concerning their physical activites, because the interpretation of "moderate" and "vigorous" activities might differ due to individual health conditions and living situations. Therefore, SHARE conducted the accelerometer project to measure the level of activity and sedentary behaviour of the elderly with a sensor to gather data that are comparable across countries.

An accelerometer is a sensor that records acceleration, i.e. the rate of change of velocity. Accelerometers are integrated in wearable devices such as smart watches and activity trackers as well as smartphones. Most modern accelerometers – like the one used in SHARE – measure the acceleration along all axes of the three-dimensional space, which allows the position, i.e. the inclination of the device, to be detected. The acceleration is recorded typically several times per second to catch even short periods of acceleration.

The SHARE accelerometer study was implemented in Wave 8. SHARE Wave 8 fieldwork started in October 2019, with the first accelerometers being sent out one month later. In March 2020, the entire fieldwork was stopped due to the COVID-19 pandemic and ensuing regulations that prevented the continuation of face-to-face interviews (see Scherpenzeel et al., 2020). Furthermore, the delivery of accelerometers was stopped as the devices have to be worn on the body and then cleaned by the survey agency, which may have posed a risk of infection with coronavirus. Another reason for stopping the accelerometer study was the very likely distortion of usual routines and behaviour patterns of participants. This chapter describes the study design in detail and presents the fieldwork results that were obtained until the suspension of fieldwork due to COVID-19.

12.2 Study Design

The accelerometer study design can be summarised as follows. A subsample of the panel respondents was drawn up before fieldwork started. During the SHARE face-to-face interviews, the sampled respondents were asked for their consent to participate in the accelerometer study (see Section 3.6). Next, the survey agencies sent the devices and related material to the participants who had consented via postal mail. The respondents put on the device by themselves, wore it for eight consecutive days (day and night) and sent it back to the survey agency. In the following, we will describe this study design in more detail.

12.3 Used Device and Supplementary Materials

SHARE aims to measure the level of activity and sedentary behaviour of the elderly. Therefore, it is vital to detect the posture of the respondents. Posture recognition performs best when using multiple accelerometers in different body positions (Stewart et al., 2018; Tang et al., 2020). However, in the SHARE setting in which respondents wear the device for several days and put it on and take it off themselves, it was not feasible to use more than one accelerometer. The optimal solution for the goal of the measurement and relatively low burden for respondents is the use of one accelerometer worn on the upper thigh.

The SHARE accelerometer study used an Axivity AX3 (Axivity Ltd, Newcastle upon Tyne, United Kingdom), a small (24.2 x 35.4 x 8.9 mm) and lightweight (11 g) triaxial accelerometer. The device is waterproof and therefore suitable

for long wear times. It is possible to wear the accelerometer whilst showering and swimming. It should be removed only when having a sauna or scuba diving. The Open Movement software (Open Movement, Newcastle University, United Kingdom) enabled the configuration of the devices before sending them out and downloading the data after the measurement. The accelerometers were set up to a sampling frequency of 50 Hz[18] (with a range of ± 8 *g*).

The device has no button to turn it on and off. Recording can either be programmed to start on a certain date and at a particular time or it starts immediately after the device has been configured, i.e. as soon as the device is disconnected from the computer. As devices were shipped via mail, which means that the exact day of arrival was unknown, the survey agencies were instructed to start the recording immediately after the configuration (and this recording mode was also factored into the decision regarding the recording frequency). Consequently, the devices recorded shipping before and after wear time as well. While this procedure resulted in large raw data files, it was still considered the best approach to ensure that all the respondents' wear times were recorded. The Axivity AX3 stores the raw sensor data in a binary packed format called the "Continuous Wave Accelerometer (CWA)" format. The raw sensor data enable comparability with measurements based on other devices and therefore allow replicable post-processing methods to be used (Crowley et al., 2019; Rowlands, 2018; Welk et al., 2012).

The commonly used wrist-worn accelerometers might be easier to handle for the respondents, however there is evidence from a study of older persons that suggests that wrist-worn devices underestimate sedentary time (Suorsa et al., 2020). Another study of older adults shows that a thigh-mounted triaxial accelerometer can be used to confidently measure sedentary behaviour and moderate-to-vigorous physical activity (Wullems et al., 2017). Although taping the device to the thigh might be challenging for elderly respondents, this procedure can be simplified to a large extent by providing them with clear instructions and an easy-to-handle device. The two test runs that preceded the SHARE Wave 8 main study proved that it was feasible to do this for respondents, using instructions provided by SHARE Central. No major problems in attaching the device were reported. In addition, respondents could always contact the interviewer or survey agency to receive help if needed. Respondents participating in the test rounds reported back that once the device was properly attached, they tended to forget it was there since it is very lightweight and barely noticeable under any type of clothing. This might lead to more continuous wear time and contribute to compliance (Montoye et al., 2016; Schneller et al., 2017).

By attaching the device to the thigh of the respondent, it is possible to detect postures, which is essential to reliably define sedentary behaviour (Byrom et al., 2016). Algorithms use the information on the inclination of the thigh for posture recognition, which results in highly accurate detection of sitting and lying time (Kozey-Keadle et al., 2011; Lyden et al., 2012; Skotte et al., 2014; Stemland et al., 2015).

With the accelerometer, participating respondents also received a so-called "accelerometer kit", which contained all the material needed for wearing the accelerometer, including an information letter and an instruction brochure on how to put on the device, as well as a reply card and a prepaid return envelope to send it back after the wear time. The information letter gave a date when to start the wearing period at the latest to avoid the device memory and battery capacity running out before completion of the eight-day wearing time. The brochure used easy-to-understand pictures to ensure clear understanding of the instructions (see Figure 12.1). The information letter and instruction brochure gave a telephone number for respondents to contact the survey agency if they had any questions or problems with the device. Medical adhesive tape and gauze pads (including spares) were included for attaching the device to the upper thigh. The gauze pad was placed between the skin and the device. A piece of medical adhesive tape was used to cover and attach the device and the gauze pad to the thigh. Different adhesive tapes were tested in the field rehearsal. Since the reports of reactions to the tape (rashes, allergic reactions, etc.) were similar for any type of tape, we selected the tape that had the longest wear time and which respondents reported to be the least noticable. The reply card covered a few questions about wearing time (start and end), wearing position (right or left thigh), possible wearing breaks (time) and occuring problems (if any). The filled-in card was sent back to the survey agency together with the accelerometer itself. The aforementioned prepaid envelope was provided for this pupose.

18 A sampling frequency of 50 Hz is high enough to apply a low pass filter (to filter out non-human movement), if desired. The sampling frequency must be at least twice the bound of the filter (Medical Research Council, 2020), e.g. 30 Hz sampling frequency for a 15 Hz filter. Using a higher sampling frequency, e.g. 100 Hz, means more battery use and also comes at the risk of not capturing the full wear time of respondents because the storage of the device is full before wear time ends (as the device also records shipping, etc.).

Figure 12.1: Cover and First Page of Instruction Brochure

12.4 Sampling

The SHARE accelerometer study was conducted in ten countries to ensure geographic variation: two northern (Denmark, Sweden), two southern (Italy, Spain), three eastern (Czech Republic, Poland, Slovenia) and three central (Belgium, France, Germany) European countries.

For each of the countries, the defined target was a net sample of 200 participants. The expectations for calculating the size of the gross sample were:

- 75 per cent average response rate of long-term panel members in the Wave 8 CAPI interview;
- 50 per cent consent rate of Wave 8 CAPI respondents to participate in the accelerometer study[19];
- 85 per cent of consenting participants will wear and return the device[20];
- 75 per cent of consenting participants will have completed eight days of capture;

To obtain complete data from 200 respondents, a gross sample of about 200/(0.75*0.50*0.75) = 710 individuals per country was needed. The gross sample drawn up in reality is somewhat larger as a result of the stratified sampling design, which is described below. In the case of higher response and/or consent rates, subselections of respondents who consented were drawn up by SHARE Central during the fieldwork.

The gross sample was a stratified sample selected from each country's longitudinal sample before the fieldwork started. It included only the panel sample, i.e. respondents who participated in SHARE before to allow the use of information from previous interviews, and excluded younger partners aged under 50. Strata were defined by age group and self-reported activity level in previous waves[21] (see Table 12.1). There was no stratification by region or clustering within region.

19 Based on experiences in the LISS panel accelerometer study (Scherpenzeel, 2017).
20 Based on experiences in the LISS panel accelerometer study (Scherpenzeel, 2017).
21 For most respondents the information from Wave 6 was used as this was the latest information due to the SHARELIFE questionnaire in Wave 7 (which did not include questions on physical activity).

Self-reported activity level is based on the CAPI question *br015* (sports or activities that are vigorous) and question *br016* (activities that require a low or moderate level of energy). It is defined as:

Activity level = (2*(4-*br015*) + (4-*br016*))

Activity level categorised = Activity level (0–1=very low) (2–3 = moderately low) (4–7 = moderately high) (8–9 = high).

In principle, sample members were selected in proportion to the size of the strata in the Wave 7 panel sample by country but with a fixed minimum number of 50 respondents per stratum. The minimum size for strata was chosen because some strata were very small, especially those of repondents with low self-reported physical activity. The fixed minimum resulted in an intended oversample of this group, as well as a somewhat larger gross sample size of about 760 individuals on average (see Table A 12.1 in the Annex for detailed information on the gross sample sizes for each country).

Table 12.1 illustrates the stratification design. It can be seen that the five smallest strata (mainly the youngest age group) are oversampled in the gross sample, by implementing the minimum of 50 cases. In one stratum, 50 cases are not even available, hence the maximum number of available cases is used (n = 26). The oversampling of the smallest strata in this example increases the gross sample size from 710 to 735. As the oversampling of smaller strata is so prevalent, we reduce the calculatory gross sample size to 600 and still arrive at the desired sample sizes easily. The gross samples for each country are shown in Table A 12.1 in the Annex of this chapter.

Table 12.1: Calculation of Strata Sample Sizes (Example)

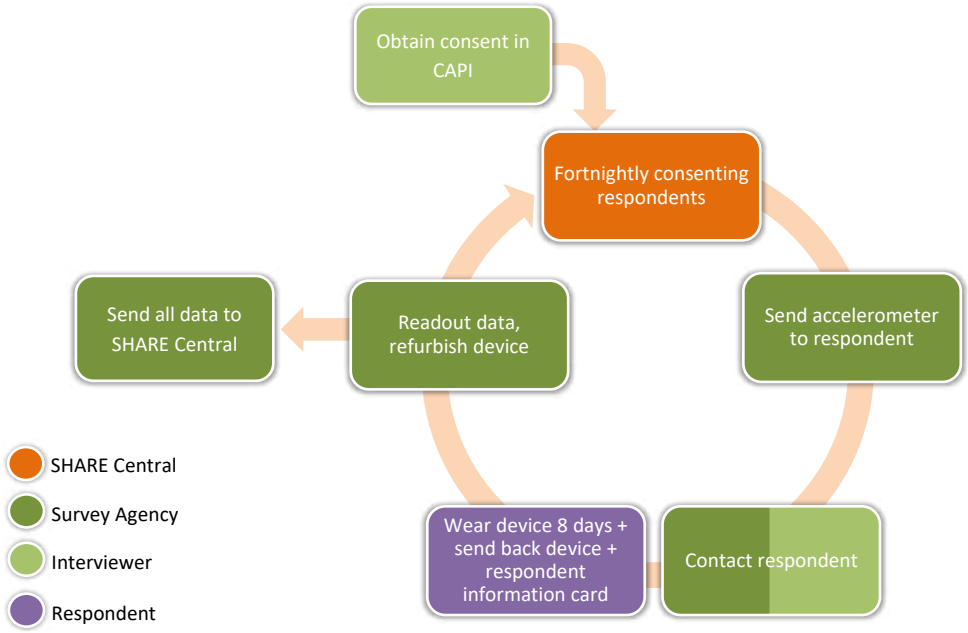
Age	**Physical activity level**	**N** *(in panel sample)*	**Proportion** *(in panel sample, weighted with field rehearsal consent rate)*	**N** *(in accelerometer gross subsample; if (Pi x 600) < 50 then n = 50 or all available)*
50-59	Very Low	26	0.8%	~~5~~ 26
	Moderately Low	105	2.6%	~~16~~ 50
	Moderately High	139	3.1%	~~19~~ 50
	Very High	313	6.3%	~~38~~ 50
60-69	Very Low	73	2.4%	~~14~~ 50
	Moderately Low	307	8.7%	52
	Moderately High	328	8.3%	50
	Very High	549	12.4%	74
70+	Very Low	230	9.0%	54
	Moderately Low	562	19.0%	114
	Moderately High	399	11.9%	72
	Very High	579	15.6%	93
Total		**3610**	**100.0%**	**735**

12.5 Fieldwork Design

Respondents were asked to participate in the accelerometer study during the SHARE face-to-face interview. A short extra CAPI module was programmed that guided the interviewer through the consent-gaining process (see Section 0).

The accelerometers available to each participating country (50 devices per country) were distributed among respondents who were a priori sampled in the stratified gross sample and consented to participate in the CAPI interview by the survey agency. To achieve the planned number of 200 respondents with accelerometry data per country, each device had to be used several times (four times on average). The cyclic procedure is illustrated in Figure 12.2 and described below.

Figure 12.2: Cyclic Fieldwork Design



Four weeks into fieldwork, the first CAPI data were processed and lists of all (so far) consenting respondents (compiled by SHARE Central) were delivered to the survey agencies, which then sent accelerometers to those respondents. In the case of high consent rates that resulted in a sample that was too large (either in terms of total number of respondents or available devices), SHARE Central randomly drew up a subsample.

The survey agency used an online platform to register which accelerometer was sent to which respondent on which date. The interviewers or the survey agency contacted the respondents to check whether they had received the accelerometer kit, to give further explanations and to offer assistence (with another home visit) if necessary or wished.

Once the accelerometers had been received, respondents were asked to start wearing the device as described in the instructions mentioned above. Respondents were not supposed to change their usual activity pattern for the time of the accelerometer measurement. After having been worn for eight days, the device was removed and sent back to the survey agency together with the reply card. Respondents received an incentive for participating in the accelerometer study, e.g. 20 euros in Germany.

As soon as the device was retrieved, the survey agency downloaded the respondent's activity data from the returned accelerometer, refurbished it and updated their register of accelerometers. They then uploaded the data as well as the information from the short questionnaire card (wear time, etc.) to the online platform where it was downloaded by SHARE Central for processing.

Two weeks after the first batch of accelerometers had been sent out, the agency received the next list of respondents who had given consent and sent out the next batch. From then on, new consents were selected every two weeks and accelerometers (including those that had been returned in the meantime) were sent out accordingly. A maximum number of 20 accelerometers per country were sent out every two weeks, regardless of the actual number of consents obtained in that two-week period. This was meant to distribute the accelerometer measurements evenly over the entire

fieldwork period, which was scheduled from November 2019 until June 2020. There are a number of reasons why this is important. First, due to the variation in seasons, experienced temperatures and exposure to daylight will differ, which in turn may affect physical behaviour (Turrisi et al., 2021). Second, spreading the time of measurements reduces the probability of oversampling motivated and easy-to-reach SHARE respondents who participate at the beginning of the SHARE fieldwork, and allows inclusion of SHARE respondents that are more difficult to reach and participate late in the survey (cf. Heffetz & Reeves, 2019). Finally, a logistical reason for spreading the measurements is the limited number of devices that were available per country (50 per country).

Based on these considerations, not all respondents who agreed to participate in the accelerometer study received a device immediately. Some respondents received the device a few weeks later (median: 31 days), some did not receive it at all, because of the unexpected early fieldwork termination. As a consequence, the achieved number of participants in the accelerometer study is relatively low, as less than half of the targeted number could be realised (see fieldwork result section).

12.6 Agency Training: Train-the-Trainer Programme

In line with the regular SHARE interview, the accelerometer study included two test phases – the pretest and the field rehearsal – preceding the main survey. For each of the three phases, a Train-the-Trainer (TTT) session was held to instruct survey agencies from the participating countries how to train their interviewers, handle the devices, communicate with respondents and transfer data to SHARE Central (for general information on the SHARE TTT programme, see Sand et al. (2019).

Particular focus was placed on how to put on the device to ensure harmonious data recordings, i.e. the axes along which the device measures acceleration are always arranged the same way for every data recording. Another training focus was downloading the data from the device and uploading them to the online platform specifically built for the accelerometer study. The training preceding the second test phase focused specifically on obtaining consent to participate. The first pretest showed that many respondents who did not consent to participate in the accelerometer study had concerns that expressed a specific information need or a lack of understanding of why their participation was valuable – concerns that would be relatively easy to relieve with additional, tailored information (see Section 0).

12.7 Fieldwork Results

The overall willingness to participate was slightly higher than expected: approximately 54 per cent of the SHARE respondents asked to participate[22] gave their consent in the face-to-face interview (see Table 12.2). However, there are substantial differences in the consent rate among countries, ranging from 33.7 per cent in the Czech Republic to 70.2 per cent in Poland. The actual participation rate was lower, as some respondents who initially consented withdrew from the accelerometer study later.

Table 12.2: Consent Rates in CAPI

Country	Consent rate
Belgium	61.9
Czech Republic	33.7
Denmark	68.6
France	55.1
Germany	50.7
Italy	36.0
Poland	70.2
Slovenia	61.7
Spain	35.8
Sweden	68.4
Total	**54.4**

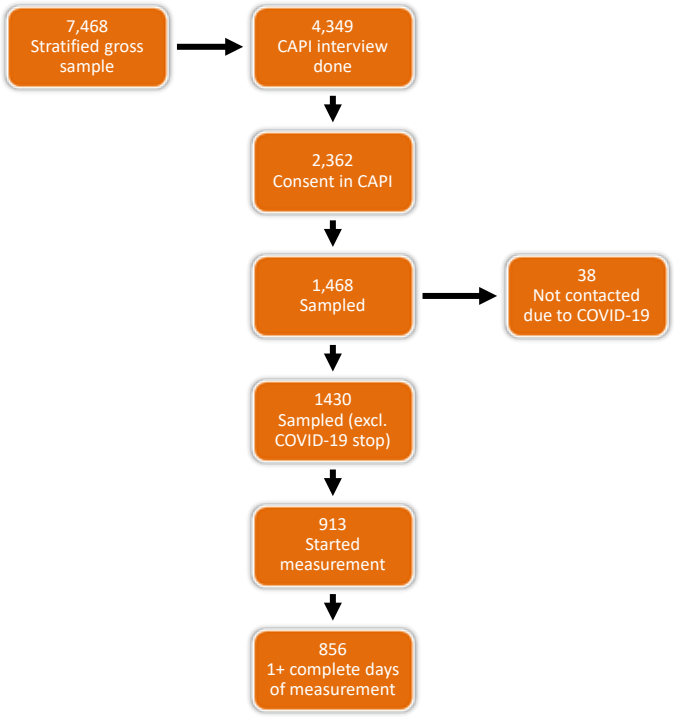
Data: SHARE Wave 8, Release version: 1.0.0

Among respondents who put on the device and started the accelerometer measurement, approximately 79 per cent wore the device for the requested eight days. Fifteen per cent of respondents who started the measurement had some – but fewer than requested – valid days of wear time. For 6 per cent of the respondents who started to wear the device, no or less than one day of wear time is available. In two-thirds of the cases with less than one day of wear time, the respondent started to wear the device too late and the device's memory was full before wear time started. Other reasons for missing data are short wear time (for example, respondent had a problem with the tape), corrupt devices (very short battery power, no data recorded) and wrong use by respondent (e.g. wearing device on arm instead of thigh).

22 This was a stratified sample that included an oversampling of inactive respondents that have generally lower participation rates. This should be considered when interpreting the reported consent rates.

The study was successful in terms of consent and compliance, but the targeted number of 2,000 valid measurements was not achieved due to the early termination of fieldwork as a result of the COVID-19 outbreak. Measurements (at least one complete day) were collected from 856 respondents (see Figure 12.3). Due to the data cleaning procedures, the number of respondents with accelerometer measurements is slightly lower in the release data. The composition of the net sample in regard to the stratification is similar to the gross sample that was drawn up prior to Wave 8 (see Table A 12.2 in the Annex).

Figure 12.3: Sample



12.8 Data Processing and Release of Data

Raw accelerometer data are processed at SHARE Central to provide users with aggregated measures for activity on different levels in generated modules "gv_accelerometer_total", "gv_accelerometer_day" and "gv_accelerometer_hour". The data are processed with GGIR (Migueles et al., 2019), an open-source package for the statistical computing software R (R Core Team, 2020). GGIR includes several steps of raw data processing to maximise the quality and comparability of generated measures. Raw accelerometer data are calibrated with respect to local gravity and deviations between devices (van Hees et al., 2014). Moreover, non-wear time is detected and imputation of non-wear time is performed. Some measures based on GGIR are provided in the generated modules, e.g. the vector magnitude (van Hees et al., 2013) and intensity gradient (Rowlands et al., 2018). For details on the data processing and generated variables, see SHARE Release Guide.

12.9 Acknowledgements

The SHARE accelerometer study is financed by the NIA grant "Enhancing the Comparability of SHARE with HRS and ELSA", submitted in June 2015 (5R01AG052527). Data processing and release of the accelerometer data were supported by the European Union within the Social Science & Humanities Open Cloud (SSHOC: GA N°823782).

References

Börsch-Supan, A. (2021). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8. Release version: 1.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w8.100

Byrom, B., Stratton, G., Mc Carthy, M., & Muehlhausen, W. (2016). Objective measurement of sedentary behaviour using accelerometers. International Journal of Obesity, 40(11), 1809–1812. https://doi.org/10.1038/ijo.2016.136

Crowley, P., Skotte, J., Stamatakis, E., Hamer, M., Aadahl, M., Stevens, M. L., Rangul, V., Mork, P. J., & Holtermann, A. (2019). Comparison of physical behavior estimates from three different thigh-worn accelerometers brands: A proof-of-concept for the Prospective Physical Activity, Sitting, and Sleep consortium (ProPASS). The International Journal of Behavioral Nutrition and Physical Activity, 16(65). https://doi.org/10.1186/s12966-019-0835-0

Heffetz, O., & Reeves, D. B. (2019). Difficulty of reaching respondents and nonresponse bias: Evidence from large government surveys. The Review of Economics and Statistics, 101(1), 176–191. https://doi.org/10.1162/REST_a_00748

Kozey-Keadle, S., Libertine, A., Lyden, K., Staudenmayer, J., & Freedson, P. S. (2011). Validation of wearable monitors for assessing sedentary behavior. Medicine and Science in Sports and Exercise, 43(8), 1561–1567. https://doi.org/10.1249/MSS.0b013e31820ce174

Lyden, K., Kozey Keadle, S. L., Staudenmayer, J. W., & Freedson, P. S. (2012). Validity of two wearable monitors to estimate breaks from sedentary time. Medicine and Science in Sports and Exercise, 44(11), 2243–2252. https://doi.org/10.1249/MSS.0b013e318260c477

Medical Research Council (2020). DAPA Measurement Toolkit: Accelerometers. Available at: https://dapa-toolkit.mrc.ac.uk/physical-activity/objective-methods/accelerometers

Migueles, J. H., Rowlands, A. V., Huber, F., Sabia, S., & van Hees, V. T. (2019). GGIR: A research community–driven open source R Package for generating physical activity and sleep outcomes from multi-day raw accelerometer data. Journal for the Measurement of Physical Behaviour, 2(3), 188–196. https://doi.org/10.1123/jmpb.2018-0063

Montoye, A. H. K., Pivarnik, J. M., Mudd, L. M., Biswas, S., & Pfeiffer, K. A. (2016). Validation and comparison of accelerometers worn on the hip, thigh, and wrists for measuring physical activity and sedentary behavior. AIMS Public Health, 3(2), 298–312. https://doi.org/10.3934/publichealth.2016.2.298

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Available at: https://www.R-project.org/

Rowlands, A. V. (2018). Moving forward with accelerometer-assessed physical activity: Two strategies to ensure meaningful, interpretable, and comparable measures. Pediatric Exercise Science, 30(4), 450–456. https://doi.org/10.1123/pes.2018-0201

Rowlands, A. V., Edwardson, C. L., Davies, M. J., Khunti, K., Harrington, D. M., & Yates, T. (2018). Beyond cut points: Accelerometer metrics that capture the physical activity profile. Medicine and Science in Sports and Exercise, 50(6), 1323–1332. https://doi.org/10.1249/MSS.0000000000001561

Sand, G., Philip, J. T., & Pettinicchi, Y. (2019). Interviewer training. In M. Bergmann, A. Scherpenzeel & A. Börsch-Supan (Eds.), SHARE Wave 7 Methodology: Panel innovations and life histories (pp. 89–99). MEA, Max Planck Institute of Social Law and Social Policy.

Scherpenzeel, A. (2017). Mixing online panel data collection with innovative methods. In S. Eifler & F. Faulbaum (Eds.), Methodische Probleme von Mixed-Mode-Ansätzen in der Umfrageforschung (pp. 27–49). Springer Fachmedien. https://doi.org/10.1007/978-3-658-15834-7_2

Scherpenzeel, A., Axt, K., Bergmann, M., Douhou, S., Oepen, A., Sand, G., Schuller, K., Stuck, S., Wagner, M., & Börsch-Supan, A. (2020). Collecting survey data among the 50+ population during the COVID-19 outbreak: The Survey of Health, Ageing and Retirement in Europe (SHARE). Survey Research Methods, 14(2), 217–221. https://doi.org/10.18148/srm/2020.v14i2.7738

Schneller, M. B., Bentsen, P., Nielsen, G., Brønd, J. C., Ried-Larsen, M., Mygind, E., & Schipperijn, J. (2017). Measuring children's physical activity: Compliance using skin-taped accelerometers. Medicine and Science in Sports and Exercise, 49(6), 1261–1269. https://doi.org/10.1249/MSS.0000000000001222

Skotte, J., Korshøj, M., Kristiansen, J., Hanisch, C., & Holtermann, A. (2014). Detection of physical activity types using triaxial accelerometers. Journal of Physical Activity & Health, 11(1), 76–84. https://doi.org/10.1123/jpah.2011-0347

Stemland, I., Ingebrigtsen, J., Christiansen, C. S., Jensen, B. R., Hanisch, C., Skotte, J., & Holtermann, A. (2015). Validity of the Acti4 method for detection of physical activity types in free-living settings: Comparison with video analysis. Ergonomics, 58(6), 953–965. https://doi.org/10.1080/00140139.2014.998724

Stewart, T., Narayanan, A., Hedayatrad, L., Neville, J., Mackay, L., & Duncan, S. (2018). A dual-accelerometer system for classifying physical activity in children and adults. Medicine and Science in Sports and Exercise, 50(12), 2595–2602. https://doi.org/10.1249/MSS.0000000000001717

Suorsa, K., Pulakka, A., Leskinen, T., Pentti, J., Holtermann, A., Heinonen, O. J., Sunikka, J., Vahtera, J., & Stenholm, S. (2020). Comparison of sedentary time between thigh-worn and wrist-worn accelerometers. Journal for the Measurement of Physical Behaviour, 3(3), 234–243. https://doi.org/10.1123/jmpb.2019-0052

Tang, Q. U., John, D., Thapa-Chhetry, B., Arguello, D. J., & Intille, S. (2020). Posture and physical activity detection: Impact of number of sensors and feature type. Medicine and Science in Sports and Exercise, 52(8), 1834–1845. https://doi.org/10.1249/mss.0000000000002306

Teresi, J. A., & Fleishman, J. A. (2007). Differential item functioning and health assessment. Quality of Life Research, 16, 33–42. https://doi.org/10.1007/s11136-007-9184-6

Turrisi, T. B., Bittel, K. M., West, A. B., Hojjatinia, S., Hojjatinia, S., Mama, S. K., Lagoa, C. M., & Conroy, D. E. (2021). Seasons, weather, and device-measured movement behaviors: A scoping review from 2006 to 2020. International Journal of Behavioral Nutrition and Physical Activity, 18(24). https://doi.org/10.1186/s12966-021-01091-1

Van Hees, V. T., Fang, Z., Langford, J., Assah, F., Mohammad, A., da Silva, Inacio C. M., Trenell, M. I., White, T., Wareham, N. J., Brage, S. (2014). Autocalibration of accelerometer data for free-living physical activity assessment using local gravity and temperature: An evaluation on four continents. Journal of Applied Physiology, 117(7), 738–744. https://doi.org/10.1152/japplphysiol.00421.2014

Van Hees, V. T., Gorzelniak, L., Dean León, E. C., Eder, M., Pias, M., Taherian, S., Ekelund, U., Renström, F., Franks, P. W., Horsch, A., & Brage, S. (2013). Separating movement and gravity components in an acceleration signal and implications for the assessment of human daily physical activity. PloS One, 8(4), e61691. https://doi.org/10.1371/journal.pone.0061691

Welk, G. J., McClain, J., & Ainsworth, B. E. (2012). Protocols for evaluating equivalency of accelerometry-based activity monitors. Medicine and Science in Sports and Exercise, 44(1S), 39–49. https://doi.org/10.1249/MSS.0b013e3182399d8f

Wullems, J. A., Verschueren, S. M. P., Degens, H., Morse, C. I., & Onambélé, G. L. (2017). Performance of thigh-mounted triaxial accelerometer algorithms in objective quantification of sedentary behaviour and physical activity in older adults. PloS One, 12(11), e0188215. https://doi.org/10.1371/journal.pone.0188215

Table A 12.1: Size of Stratified Gross Sample by Country

Age	Physical activity level	Belgium (FR)	Belgium (NL)	Czech Rep.	Germany	Denmark	France	Spain	Italy	Poland	Sweden	Slovenia
50–59	Very Low	25	24	14	26	21	33	18	50	9	3	9
	Moderately Low	25	25	50	50	50	50	25	50	31	22	50
	Moderately High	25	25	50	50	50	50	22	50	34	32	50
	Very High	25	25	50	50	50	50	47	50	50	50	50
60–69	Very Low	25	25	50	50	46	50	50	50	50	22	49
	Moderately Low	33	33	64	53	50	73	52	50	78	50	56
	Moderately High	25	25	50	50	50	50	50	50	50	50	53
	Very High	29	30	51	75	99	56	50	50	61	64	85
70+	Very Low	43	42	78	53	50	84	145	147	125	50	61
	Moderately Low	64	68	173	114	101	168	171	108	119	151	124
	Moderately High	29	27	80	72	60	53	53	63	50	98	77
	Very High	26	32	64	93	105	50	58	50	50	145	87
Total		**374**	**381**	**774**	**736**	**732**	**767**	**741**	**768**	**707**	**737**	**751**

Table A 12.2: Gross Sample and Net Sample by Strata (in %)

Age	Physical activity level	Belgium (FR)		Belgium (NL)		Czech Rep.		Germany		Denmark		France		Spain		Italy		Poland		Sweden		Slovenia	
		gross sample	AX done	gross sample	AX done	gross sample	AX done	gross sample	AX done	gross sample	AX done	gross sample	AX done	gross sample	AX done	gross sample	AX done	gross sample	AX done	gross sample	AX done	gross sample	AX done
50–59	Very Low	6.7	0	6.3	7.5	1.8	1	3.5	4.2	2.9	2.6	4.3	2.5	2.4	0	6.5	9	1.3	0.8	0.4	1.4	1.2	1
	Moderately Low	6.7	14.8	6.6	3.8	6.5	5.7	6.8	8.5	6.8	0	6.5	6.2	3.4	2.8	6.5	6	4.4	3.9	3	0	6.7	9
	Moderately High	6.7	7.4	6.6	7.5	6.5	4.8	6.8	8.5	6.8	2.6	6.5	6.2	3	4.2	6.5	7.5	4.8	3.9	4.3	5.4	6.7	6
	Very High	6.7	11.1	6.6	5.7	6.5	3.8	6.8	7.6	6.8	5.3	6.5	8.6	6.3	6.9	6.5	10.4	7.1	8.6	6.8	1.4	6.7	6
60–69	Very Low	6.7	0	6.6	13.2	6.5	5.7	6.8	2.5	6.3	7.9	6.5	8.6	6.7	5.6	6.5	9	7.1	6.3	3	2.7	6.5	8
	Moderately Low	8.8	18.5	8.7	5.7	8.3	9.5	7.2	5.9	6.8	2.6	9.5	9.9	7	13.9	6.5	3	11	15.6	6.8	6.8	7.5	12
	Moderately High	6.7	3.7	6.6	11.3	6.5	14.3	6.8	8.5	6.8	10.5	6.5	7.4	6.7	9.7	6.5	6	7.1	8.6	6.8	6.8	7.1	2
	Very High	7.8	7.4	7.9	9.4	6.6	8.6	10.2	12.7	13.5	21.1	7.3	11.1	6.7	18.1	6.5	13.4	8.6	10.2	8.7	10.8	11.3	11
70+	Very Low	11.5	7.4	11	5.7	10.1	1.9	7.2	5.1	6.8	7.9	11	2.5	19.6	9.7	19.1	11.9	17.7	6.3	6.8	5.4	8.1	3
	Moderately Low	17.1	18.5	17.8	15.1	22.4	28.6	15.5	13.6	13.8	21.1	21.9	23.5	23.1	12.5	14.1	6	16.8	22.7	20.5	18.9	16.5	13
	Moderately High	7.8	7.4	7.1	3.8	10.3	8.6	9.8	7.6	8.2	0	6.9	9.9	7.2	4.2	8.2	9	7.1	8.6	13.3	17.6	10.3	13
	Very High	7	3.7	8.4	11.3	8.3	7.6	12.6	15.3	14.3	18.4	6.5	3.7	7.8	12.5	6.5	9	7.1	4.7	19.7	23	11.6	16

CHAPTER 13

Harmonising Record Linkage Procedures in SHARE

13

13 HARMONISING RECORD LINKAGE PROCEDURES IN SHARE

Imke Herold, Yuri Pettinicchi and Daniel Schmidutz – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC)

13.1 Introduction

Record linkage broadens the possibilities of investigating research questions by enriching self-reported survey data with administrative data. Linking data is a complex procedure and goes along with several challenges and obstacles on different levels – legal, ethical, technical, and organisational such as reaching an agreement with data-providing institutions or dealing with privacy concerns (Künn, 2015). This especially holds true for record linkage in an international context. The intention of SHARE record linkage is to enhance survey data with very accurate administrative data, such as data from public authorities, insurances or governmental institutions, in order to optimally exploit the advantages of both data sources.

Although administrative data are produced for organisational purposes of the respective institution in the first place, it can open up new perspectives to research since it provides more precise and accurate information than self-reported data or even contains information that might be unknown to survey respondents. In particular, the chronological consistency of administrative data and the richness of detail are an advantage for longitudinal research. Administrative data on employment histories, for example, provide many more details than it would be possible to collect during a retrospective interview due to memory biases or time constraints. Surveys, on the other hand, provide the opportunity to collect certain background information as well as information on attitudes, personality or well-being that is not captured in administrative data sets. The different advantages of administrative data and survey data complement each other when the data sources are linked. Linked data sets combine both objective and subjective perspectives on respondents' lives. In addition, linked data sets allow for methodological research, such as reciprocal validation of both data sources.

In SHARE, the first record linkage project was established in 2009 in Germany. Since then, more and more SHARE Country Teams have implemented record linkage projects (see also Korbmacher & Schmidutz, 2015). The data content varies from pension and employment data to income and health data. Up to Wave 8, seven countries that participate in the SHARE study have implemented linkage projects: Austria, Germany, Denmark, Estonia, Finland, Italy and the Netherlands.

The linkage projects have to be specifically developed for each country and the procedures have to be adapted to the respective country-specific needs and requirements. SHARE Central carries out the central coordination of the linkage projects based on what has been achieved thus far, and strives to harmonise the procedures as much as possible. Our aim is to continuously advance the existing linkage projects, including the related procedures, and enhance record linkage in SHARE by exploring the possibilities of cross-country comparative analyses based on linked data. In this chapter we explain the legal and technical framework of record linkage in SHARE. Furthermore, we provide an overview of the different country-specific projects by pointing out challenges and solutions. We conclude the chapter with an outlook on the SHARE record linkage project.

13.2 Consent as the Legal Basis for Record Linkage in SHARE

One of the main responsibilities of researchers when setting up a record linkage project is to implement it in a data protection-compliant manner. On the European level, the General Data Protection Regulation (GDPR), which has directly applied in all EU member states since May 2018, regulates the protection of natural persons with regard to the processing of personal data. Accordingly, the collection and processing of personal data in SHARE as well as any subsequent linkage with administrative data is subject to the GDPR. The survey data collection and processing in SHARE is based on the consent of the respondents. This, however, does not cover the linkage of the survey data with national administrative data. Therefore, additional explicit consent for the purpose of record linkage projects must be obtained from the respondents.

According to Article 4 (11) of the GDPR, consent must be "freely given, specific, informed and unambiguous". Informed consent means that the respondents have to have full knowledge of what they are consenting to. When consenting to the record linkage the respondents have to be aware that their consent is the legal basis for the data processing and that the consent is completely voluntary. Furthermore, the respondents have to be informed about

the purpose of data processing, the processing of special categories of personal data (if applicable), the procedure of storing and processing the data (including involved parties and storage period of personal data) as well as the identity and contact details of the controller and its Data Protection Officer. In addition to this, the respondents have to be notified about their rights. Besides the right to withdraw their consent, respondents have the right to access, correct and delete their personal data, the right of restriction of processing concerning their person or to object to processing as well as the right to data portability. Furthermore, they have the right to complain to the supervisory authority.

The consent has to be obtained at the time when the personal data are obtained, i.e. during the SHARE interview. For record linkage in SHARE, this is done separately in the respective CAPI (Computer-Assisted Personal Interview) module. The consent to record linkage can be either obtained orally (documented in the CAPI instrument) or in written form (documented by signature).

Even though the requirements of the GDPR apply to the SHARE survey data collection as well as to the national record linkage projects in the same manner, there are several aspects that have to be considered when preparing the information for consent to the record linkage. While the SHARE survey data collection relies on a generic procedure, which includes a data protection statement that – apart from a few exceptions (e.g. names and contact details of the SHARE Country Teams and Survey Agencies) – contains identical information for the respondents, the information provided for the linkage consent differs between the countries for several reasons (Schmidutz, 2018). First, in each country different institutions (data controllers and processors) are involved in the linkage procedure, including the administrative data providers. Second, depending on the type of administrative data, different variables and in some cases special categories of data are linked. Accordingly, the procedure of data processing and the recipients of data vary across countries, also depending on the requirements of the respective data providing institutions. And third, the consent type can differ between oral consent and written consent. This is why each SHARE Country Team that plans a record linkage project has to develop individual consent documents (such as CAPI consent question, information leaflet and/or consent form). SHARE Central supports this by providing instructions and a checklist as well as assistance in defining and documenting procedures and finalising the respective documents.

13.3 Linkage Procedure: From Obtaining Consent to Data Access

In order to link SHARE survey data with administrative data, a procedure for the linkage has to be defined. The linkage procedure can broadly be divided into four steps:

- Step 1: obtaining consent during the interview
- Step 2: internal preparation of linkage
- Step 3: data processing and transfer by the involved institutions
- Step 4: actual linkage and granting access to the linked data

Figure 13.1 below presents the first three steps in a flow chart. The fourth step is described in detail in the following section.

Step 1: During the interview, respondents are asked for their consent to the linkage. The consent is obtained verbally and documented in the CAPI instrument or obtained in a written form by signature on a consent form. Since record linkage in SHARE is based on a direct linkage, further personal information – such as social security number (SSN) and/or date of birth – is needed to identify the respondents in the respective administrative data.

Verbal consent was obtained in Austria and in the Netherlands in Wave 7. While in Austria the SSN was collected via the CAPI instrument, in the Netherlands, the already available personal identifying information was sufficient to identify respondents in the administrative data. All other countries (Germany, Denmark, Estonia, Finland and Italy) obtained respondents' consent via a signed form. The consent form is sent back by the respondents via postal mail after the interview. In those countries the SSN was collected via the consent form, except for Denmark, where it was collected via the CAPI. In order to be able to correctly assign the consent forms to the respondents, a random number – the consent form number (*internal ID1* in Figure 13.1) – is both printed onto the consent form and typed into the CAPI by the interviewer.

Step 2: The internal preparation of the linkage only needs to be conducted for countries with written consent. This step is very important given that the respondents' consent is a necessary legal requirement for linkage. It consists of two operations: first, the consent forms need to be checked for completeness and correctness (e.g. whether the consent form is signed); second, the consent forms have to be matched to the CAPI interview data via the consent form number (internal ID1) and the personal information provided by the respondents has to be verified. The latter is especially challenging and demanding because the consent form number inserted in the CAPI by the interviewer might be incorrect due to typing mistakes or due to switching of consent forms between cohabiting partners. Sometimes manual checks and corrections are necessary in the CAPI. One approach for mitigating the risk of typing mistakes is asking the interviewers to type the number twice into the CAPI instrument. In order to cut the direct link between survey data and consent forms that contain identifying information

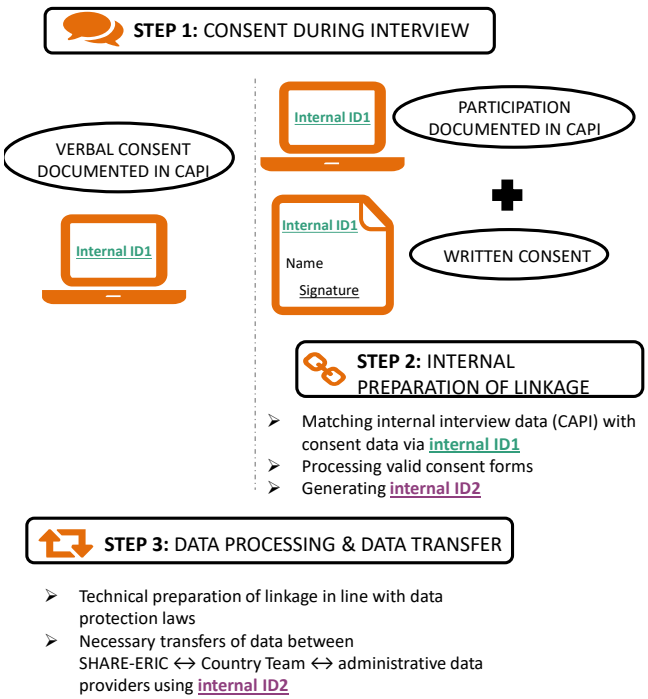
about the respondents, the consent form number is replaced by a new random number (internal ID2). Internal ID2 is then used during the data processing and data transfer processes.

Step 3: In order to enable the actual linkage of the SHARE survey data with the administrative data it has to be ensured that the same identifier will be available in both data sources. This is a necessary prerequisite to correctly link the data sets (i.e. the survey data and the administrative data on the same person). In order to achieve this, several data processing and data transfer processes are necessary. These processes are agreed upon by SHARE-ERIC, the SHARE Country Team and the administrative data-providing institution, taking into account all relevant data protection requirements. In line with the agreements between the involved actors, two scenarios can be identified.

Scenario A: A specific ID number (national *mergeid*) is generated and included in the national SHARE release data and this data set is transferred to the respective institution that holds the administrative data. In SHARE this is the case for Denmark, Estonia, Finland and the Netherlands.

Scenario B: The administrative data set is transferred to SHARE Central in a format suitable for research use in order to include the regular SHARE ID (*mergeid*) in the administrative data. SHARE implements this procedure for Germany, Austria and Italy.

Figure 13.1: Flow Chart Technical Linkage Procedure



13.4 Actual Linkage and Data Access

The final step of the linkage procedure (step 4) is to grant access to the record linkage data for researchers (see Figure 13.2 below). Depending on the scenarios described above, the actual linkage will either be carried out by the institution that holds the administrative data before access is granted (scenario A) or will be conducted by researchers themselves after having been granted access (scenario B). In both cases, however, data access has to be requested separately for SHARE data and administrative data. Besides the usual user registration for the use of SHARE data, an additional application for usage of administrative data is required at the respective national administrative data-providing institution. There are two different data access procedures that correspond with the described scenarios.

Scenario A: Access to the linked data set is granted by the respective administrative data-providing institution to registered SHARE users via secure remote access. The corresponding part of SHARE data can be accessed via this connection too.

Scenario B: The administrative data are provided via a secure data transfer as a scientific use file, which includes the regular SHARE ID (*mergeid*). This allows registered SHARE users to merge both data sources by themselves.

The data access procedures can change over time due to new regulations or requirements. Further information and documentation for Germany, Denmark and the Netherlands can be found on the SHARE Website[23]. The information on the website will be extended to additional linkage projects in other SHARE countries once the linked data sets are available for access. SHARE is continuing to improve the access procedures in order to make it as easy as possible for researchers to use the data.

Figure 13.2: Flow Chart Preparation and Provision of Data Access

23 http://www.share-eric.eu/special-data-sets/record-linkage-project.html

13.5 National Linkage Projects at a Glance

Up to Wave 8 seven Country Teams implemented record linkage projects in SHARE. This section gives a brief overview regarding the current status of the projects as well as the content of the administrative data that are used in the national record linkage projects.

In *Germany*, the cooperation between SHARE-ERIC and the German Pension Insurance (DRV), known as "SHARE-RV", has been ongoing since 2009 (see Börsch-Supan et al., 2020). In the third SHARE wave German respondents were asked for their written consent to record linkage for the first time. The linkage module was implemented again in Waves 5, 6 and 8. The administrative data provided by the DRV consist of two parts. The first is a cross-sectional data set, the RTBN ("Rentenbestand" – policy holder pension portfolio), which is available for respondents who already receive a pension. It contains information on pension claims, such as type of pension, amount of paid pension and some indicators on how these claims were achieved. The second data set, named VSKT ("Versicherungskontenstichprobe" – insurance account sample), is available for non-pensioners as well as for pensioners and includes information on the employment history. This longitudinal data set contains all activities that are relevant to the pension system on a monthly basis from the age of 14 until the age of 65. The administrative data are updated every year (together with the SHARE Release data) and available after a successful application.

Denmark implemented record linkage in Wave 5 for the first time. Due to the Danish data protection law at that time, no consent from the respondents was needed. Within the so-called "REGLINK-SHAREDK" project, the SHARE data of Waves 1–6 of all Danish respondents who took part in SHARE in Wave 5 and Wave 6 could be linked to administrative data of Statistics Denmark and the Danish Health Authority. The administrative data contain information on employment, income, education and health. As a result of the requirements of the GDPR, Denmark introduced a written consent question in Wave 8. Until now, the administrative data have been available for around 5,000 respondents. The linked data are available after a successful application for researchers with an affiliation to a Danish research institution.

In the *Netherlands* the linkage module was introduced for the first time in Wave 5. Dutch law allows for a documented verbal consent during the interview. In the first release version of the "Linkage SHARE NL" data about 90 per cent of the Dutch respondents from Wave 5 consented to the linkage. Based on their consent, their data from Waves 1 to 6 were linked to administrative data of Statistics Netherlands (CBS). Besides information on employment, income and social security, CBS also provides certain health data. Consent was asked for again in Wave 7 and will also be requested in Wave 9.

In *Austria*, data from the records of the Main Association of Austrian Social Security Institutions and the Public Employment Service Austria are prepared for linkage with the assistance of the Austrian Ministry of Social Affairs. The "SHARE HV-AMS" project combines information on demographics, employment, education, income, insurance and reception of social benefits with panel survey data of SHARE. The linkage consent question was implemented in Waves 5 and 8 and it is planned again for Wave 9.

In *Estonia*, cooperation with Statistics Estonia has been established. Since Wave 5, the linkage module has been implemented in every wave. Statistics Estonia is in charge of several registries covering information on pension, education and health that are planned to be included in the linkage project.

Finland started a record linkage project for the first time in Wave 8. Written consent of the respondents was obtained. The main cooperation partner in Finland is Statistics Finland, which provides data on employment histories and income. Furthermore, SHARE data of the Finnish respondents will be linked to the two administrative data sets of the main pension institutions: the National Institute for Pensions (Kela) and the Finnish Centre for Pensions (ETK).

Italy is the second country in which record linkage was started in Wave 8. The Italian project followed a similar structure with regard to the linkage to that applied in Germany and Austria. With the CAPI module, a written consent procedure was introduced. Respondents who agreed to the linkage will be linked to the administrative data of the Italian Social Security Institute (INPS), which provides information on earnings and social security contributions.

13.6 Conclusion and Outlook

Ten years after SHARE's first record linkage project was set up, SHARE has developed a well-grounded framework regarding relevant legal and technical challenges for the implementation of further linkage projects. Following changes to EU data protection law, a reliable and coherent basis regarding consent and data processing procedures was established, enabling country-specific implementations.

SHARE's framework for the implementation of record linkage projects consists of four steps that are cooperatively conducted by SHARE Central, the Country Teams and the administrative data-providing institutions. Until Wave 8, a linkage module allowing for both verbal and written consent has been implemented, which has become an integral part of the questionnaire now. Moreover, well-developed procedures for the internal preparation of the linkage based on written consent have been established. In addition, the necessary data processing and data transfer pro-

cesses have been harmonised to a large extent across the national linkage projects. Taking into account country-specific requirements and restrictions regarding the transfer of administrative data, two standard scenarios have been identified. These scenarios result in two corresponding data access procedures.

The priority for the near future will be to conduct the actual linkage and to set up data access procedures in those countries where consent has already been obtained but the linked data are not available for research yet. In the long run, the work on the international harmonisation of the linkage projects in SHARE will continue, both in terms of standardising the different aspects of the linkage procedures across countries and with regard to the challenging aim to facilitate comparative analyses based on linked data between two or more countries.

References

Börsch-Supan, A., Czaplicki, C., Friedel, S., Herold, I., Korbmacher, J., & Mika, T. (2020). SHARE-RV: Linked data to study aging in Germany. Jahrbücher für Nationalökonomie und Statistik. 240(1), 121-132. https://doi.org/10.1515/jbnst-2018-0034

Korbmacher, J., & Schmidutz, D. (2015). A note on record linkage. In F. Malter & A. Börsch-Supan (Eds.), SHARE Wave 5: Innovations & methodology (pp. 60-66). MEA, Max Planck Institute for Social Law and Social Policy.

Künn, S. (2015). The challenges of linking survey and administrative data. IZA World of Labor, 214. https://doi.org/10.15185/izawol.214

Schmidutz, D. (2018). Implementing consent for record linkage in a cross-national survey: A practical example from SHARE Wave 8. SERISS Survey Experts Network Workshop. Available at: https://seriss.eu/wp-content/uploads/2018/11/Schmidutz_Consent_SHARE_linkage.pdf

CHAPTER 14

SPLASH – A Social Policy Archive for SHARE

14

14 SPLASH – A SOCIAL POLICY ARCHIVE FOR SHARE

Diana López-Falcón – Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy (MPISOC)

14.1 Introduction

The complexity of demographic processes such as population aging represents challenges for all types of welfare states, leading to an increased need for policy-oriented research. In this context, to identify the best policy options, it is crucial to learn from experiences about the interactions and effects of policies on individual behaviours.

The "Social Policy Archive for SHARE" (SPLASH, www.splash-db.eu; see Figure 14.1) aims to overcome existing data limitations in order to foster comparative policy-oriented research using – but not limited to – SHARE microdata. It provides easy access to macro-indicators and policy information, as well as socio-economic contextual data organised in two substantive sections: Data and Policy. Unlike other databases, which often only have data for cross-sectional research, it supports longitudinal multilevel research by providing time-series data at the national level, but also, whenever possible, at the regional level.

Figure 14.1: SPLASH website

SPLASH represents a milestone in the evolution of multiple projects involving the support and vision of several international partner institutions. Originally developed in collaboration with the Max Planck Institute for Demographic Research, the project has been led since 2016 by the Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy in Munich.

As SHARE provides microdata for the study of aging as a lifelong process, the aim of SPLASH is to provide macro-level data on the heterogeneity of the different welfare regimes. To do so, SPLASH's Policy section incorporates the repository of social policies collected up until 2016 by the Max Planck Institute for Demographic Research. The collection covers up to 19 European countries in the fields of education, family, health and migration, as well as work and retirement policies as early as 1800. Moreover, it includes comprehensive information on the supporting legislation and additional background details on specific policy contexts, and it is envisaged to add more countries and fields in the future.

The Data section offers contextual cross-national quantitative indicators in the research domains covered by SHARE and based on official statistics and research outcomes. This section offers resources covering European countries in fields such as education, health, migration, living conditions, work and retirement, among others.

The combination of policy details and quantitative macro-indicators substantially enriches the analytical potential of the SHARE data. For instance, the data gathered in SPLASH could help us understand not only today's living conditions of the SHARE respondents, but the policies that might have affected their behaviours at earlier ages while using the retrospective Waves 3 and 7 data.

The recently collected respondent data from the SHARE Corona Survey stand to offer many insights regarding the outcomes of the aged population during the pandemic. This includes identifying the effectiveness of the different government policies implemented throughout Europe. Contextual data covering numerous aspects (i.e. political, social, healthcare) must be incorporated into the study for a comparative analysis. This paper provides an overview of SPLASH's background and the development of the SPLASH COVID-19 curated data collection.

14.2 Recent Development

To effectively support the study of societal changes while addressing pressing social policy issues, SPLASH maintains a collection of core indicators and continually evaluates the website's content to remain relevant. Aligning with these aims, SPLASH has experienced a shift in its endeavours as a result of the recent COVID-19 pandemic. The current

events have generated a demand for high-quality data covering all aspects of the crisis, ranging from epidemiologic indicators to the various policy responses. As such, the website has been restructured and expanded with the adoption of collaborative data collection methods, which will enable the site to offer specialised indicators updated on a continuous basis.

The effects of the COVID-19 pandemic remain pervasive; however, individual hardships endured as a result of the virus differ dramatically throughout the EU, as do the associated measures implemented.

This further highlights the importance of contextual indicators for policy research, particularly to understand the non-intended consequences of the containment measures. As a result, SPLASH is currently focusing on collecting data for the analysis of the SHARE Corona questionnaire with the aim of understanding the COVID-19 living situation of individuals aged 50 or older. Furthermore, the database plays an active role in the H2020 SHARE-COVID19 research project.

The SPLASH COVID-19 contextual data collection will address the cross-national differences related to the pandemic, as well as the severity and stringency of the epidemic control actions to measure their non-intended effects on health, economic and social outcome variables for all SHARE countries. In addition, the portal will also serve as a platform to exchange and publish related research outcomes. The main lines of development of the SPLASH COVID-19 contextual data can be grouped into the following categories: healthcare (access, treatment and preventive measures), employment and social life. It will document societal and cultural structures, health system preparedness and resilience, population densities and population risk groups, as well as environmental and economic factors at the national and regional level.

To enhance the data collection and maintenance processes, the first step carried out was surveying the newly available COVID-19 data sources covering the SHARE countries. This mapping exercise considered specialised quantitative and qualitative resources from different disciplines, as well as contextual sources of information. While doing so, we were also able to identify gaps in the data available and then use this knowledge to refine the database project goals.

This way we provide content that will aid in the analysis of all SHARE data sets by addressing the gaps in available data. In terms of geographical coverage, our aim is to provide data for all SHARE countries. To that end, the reports provided by SHARE Country Team Leaders at the beginning of the pandemic documenting the public policy measures affecting fieldwork allowed us to identify the key actors and sources of information at national and subnational level. In addition, the reported containment measures were later used as checkpoints to verify the completeness of the information provided by external sources. Subsequently, we mapped the COVID-19-related sources covering specific health and epidemiological indicators (e.g. number of reported cases, deaths, hospitalisations), as well as policy responses such as containment measures at different levels.

14.3 Contextual Data for COVID-19 Research

The need to include COVID-19 contextual indicators in SPLASH's regular operation required broadening the scope and adapting the project at a faster pace. As SPLASH already maintains a collection of contextual indicators for the benefit of SHARE microdata analysis, the goal of meeting the need for COVID-19-related data was a natural extension of the partnership. However, the project still required several major overhauls. In terms of data collection, this meant SPLASH's prior objectives and development strategies were redirected to fit the broad scope of COVID-19 resources, which includes specialised quantitative and qualitative information with the limited resources at hand. Ongoing communication with SHARE researchers throughout the pandemic directed much of the content's focus, reflecting the evolution of their findings. Additionally, the website required technical as well as organisational updates to accommodate the requested data.

Despite the limitations, the demand for carrying out a significantly broader data collection exercise followed the guiding principles of SPLASH for offering open access to high-quality data and sources in one site. The strategy followed for the development of the quantitative data collection seeks to prioritize the cross-country comparability of the data produced by international agencies and organisations, such as Eurostat, the OECD, the World Health Organization (WHO) and the European Centre for Disease Prevention and Control (ECDC). In addition, the published data and outcomes from research infrastructures were also revised. These include policy databases and data sets, contextual data sources and reports with information relevant for the study of the SHARE Corona microdata. The results were collected in SPLASH's standardised format containing a basic description of each indicator, keywords, the countries and years covered, as well as a hyperlink to the provider's website. Users are also alerted about the possibility of methodological breaks or corrections, particularly in epidemiological COVID-19 data.

For the main areas of research covered by the SHARE COVID-19 Project and the SHARE Corona Survey, the mapping and data collection exercises considered the periods before and during the pandemic. In addition, efforts have been targeted towards the identification of specific policy instruments and the quantitative data for its evaluation. Thus, the SPLASH COVID-19 contextual data will support the comparative study of the living conditions of the SHARE respond-

ents prior to the pandemic and the effects of the reactive policy measures on their lives: for instance, the indirect impact of containment measures on households' work-related income, social contacts or well-being disparities.

After conducting a preliminary search, the first efforts at qualitative data collection have been concentrated on labour market and particularly short-term employment and wage subsidies policies for all SHARE countries. During this exercise, it has been possible to identify and collect the main characteristics of the measures applied, such as the requirements, amount and duration of the benefits. The results are explained in more detail in the next section.

After revising the available data and associated metadata, the sources suitable for academic studies will be available in SPLASH's map of external resources, whereas collections of quantitative contextual indicators for all SHARE countries will be available in its data section. The indicators support the analysis of the SCS and address, for instance, the severity and stringency of the epidemic control actions and the employment-related measures implemented in response. All the associated resources will be identified with the keyword "SHARE-COVID19".

As many of the compiled resources and extracted data will be shared on the SPLASH website, the requirements for the terms of use for each provider are at all times reviewed. In cases where authorisation is required, the organisation or representative in charge is contacted to request permission. Several adjustments to the SPLASH website were mandatory to accommodate the COVID-19 contextual data and take advantage of the website as a platform for exchanging preliminary research results among project partners. Further, to promote and facilitate the use of the data collected, it is envisaged to include sample Stata do-files to show how to combine the contextual indicators in SPLASH with the SHARE microdata.

The mapping and first data collection exercises described above were not free from struggles. For providing data at the intersection of research and policy, one of the main challenges refers to the scattered sources. Even if the number of COVID-19 sources and research outcomes are increasing on almost a daily basis, the resources needed to maintain live data sets result in a limited number of cross-national sources of information. These mainly address national and summary measures. However, once the level of detail or geographical disaggregation is increased, the barriers to access grow exponentially. For instance, accessing adequate policy and statistical sources of information is accompanied by language barriers, which complicate the development of a detailed cross-country comparison at the subnational level. Further, methodological and timely differences in the reported indicators and retrospective corrections – such as the number of COVID-19 cases and associated deaths – can result in incompatibilities across regional, national and international sources.

14.4 Outlook

Despite the challenges faced during the data collection, the integration of COVID-19 resources has significantly enriched SPLASH in terms of contents and methodological scope. The survey of COVID-19 resources produced a compilation of supplementary contextual data accessed for the preliminary study of policy developments in relation to the individual outcomes captured by the SHARE microdata. A review of data quality and the associated metadata was conducted during the first surveying exercise of COVID-19-related data. It resulted in the inclusion of more than 30 additional external sources (such as policy databases and data sets) in SPLASH's data map. These include resources such as the OECD Country Policy Tracker[24], the Oxford COVID-19 Government Response Tracker (OxCGRT)[25], the John Hopkins Coronavirus Resource Center[26], the Human Mortality Database[27], the ACAPS COVID-19 Government Measures Dataset[28] and resources developed by stakeholder organisations. The variety allows for several avenues of inspection concerning the unintended effects of containment measures – immigration, gender, business/investments and more. Further, these are individually reported in SPLASH's standardised format, which lists information about the country, level of detail covered, main contents, methodological resources, provider details and a link to the original source.

Data from some of the resources listed were harmonised to code basic COVID-19-related indicators for gauging the severity of the pandemic in the respective SHARE countries. These include new cases, cumulative cases, new deaths and cumulative deaths, as well as information about the evolution of the containment measures. These can be analysed in combination with the health-related indicators part of SPLASH's core set of indicators developed by international organisations such as the WHO[29] covering health infrastructure, health risk-related behaviours, morbidity and preventive healthcare. As with all indicators in SPLASH, these are continuously monitored and updated. Whenever possible, the contents include information about metadata and any potential methodological changes.

24 Available at: https://www.oecd.org/coronavirus/country-policy-tracker/
25 Available at: https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker
26 Available at: https://coronavirus.jhu.edu/map.html
27 Available at: https://www.mortality.org/
28 Available at: https://www.acaps.org/projects/covid19/data
29 For instance, https://covid19.who.int/ for COVID-19-specific resources and the Global Health Observatory for healthcare infrastructure indicators: https://www.who.int/data/gho

As mentioned in the previous section, the first qualitative data collected by SHARE's Country Team Leaders provided a cross-comparative timeline of government responses with respect to COVID-19. These data have been particularly useful in analysing the relationship between the resulting severity of the pandemic in the respective SHARE countries and the varying implementation dates of lockdown measures.

Subsequently, the efforts on qualitative data analysis and collection were focused on labour market policies. For a contextual overview of the employment and the working conditions of the respondents, the contents of diverse reports and policy databases were revised and restructured. These include the OECD Job Retention Measures report (OECD, 2020a), the European Commission Reports on policy measures against the spread of COVID-19 (European Commission, 2020a, 2020b, 2020c), the European Trade Union Confederation (ETUC) Briefing Notes on alternatives to redundancy (ETUC, 2020a, 2020b), the ILO Country Policy Responses[30] and the ACAPS COVID-19 Government Measures Dataset. The main point of interest was job retention measures, namely short-time employment and wage subsidy programmes, given their relevance as instruments to combat unemployment and loss of income during the current crisis. The information was revised to see whether the measures existed prior to the pandemic, the requirements for accessing the benefit, as well as its duration and the amount offered. Once the information was compiled, the various policies were organised into the SPLASH main categories to help with cross-country comparisons. Almost 400 employment-related measures have been identified for the SHARE countries up until December 2020. A potential line of action is related to the development of synthetic, regulation-based indicators. Mainly addressed to researchers and analysts, the indicators would be used as a basis for the development of composite indicators or variables in their analyses.

The data collected showed that SHARE countries that previously had no short-time work policy implemented one during the crisis. If they already had one, they updated the original one to increase eligibility, duration and amount of benefit. Many benefits were extended to the self-employed under these adjustments. Another trend was that countries that did not have a short-time work policy or wage subsidy prior to the pandemic crisis often offered protection in the form of guaranteed wages in case of the employer's bankruptcy. However, the delayed reporting of unemployment and short-time work participants posed a challenge in observing its effect on welfare.

While we have been able to retrieve qualitative data detailing policy measures for wage subsidy and short-time work policies implemented by the SHARE countries, we are still searching for complementary quantitative indicators. There are several challenges, however, in locating these data that are also applicable to other policy fields. Firstly, it appears that the events are too recent for international organisations and statistical offices to offer harmonised cross-country or subnational data on beneficiaries from COVID-19-related instruments. Further, despite their availability at the national level, the methodological disparities and language prevail as main barriers to constructing a cross-national data set.

For instance, for labour market policy instruments, although some countries report the number of applicants received and approved, they do not report whether an applicant is an employer or employee. In addition, the instances used for reporting, as well as the periods reported, might also differ. Nevertheless, the content offered by the national providers is relevant for the SPLASH COVID-19 data collection and the potential development of indicators. In this case, to overcome language barriers we used Eurofound's European Monitoring Centre on Change search tool[31] to retrieve the official name of the short-time work policy in each country's native language. Subsequently, we have been able to retrieve a limited number of contextual employment indicators for a better understanding of the labour market conditions existing prior to and – whenever possible – during the pandemic in the SHARE countries. These include data from the OECD Employment Outlook 2020 (OECD, 2020b) such as average hours worked, participation in short-time work programmes and unemployment rates.

As with all contents in SPLASH, we will continue monitoring and collecting complementary quantitative indicators for the SPLASH COVID-19 data collection. We expect that more employment-related and policy-specific data will become available during the next months, in line with the release calendar of national and international offices.

We will actively collaborate with SHARE Country Team Leaders and project partners to validate our findings on national policies and the evaluation of high-quality sources of information at national and subnational level. To make the resources available to the public, the process of integrating the content into SPLASH has already begun. All the resources collected in the framework of the project will be identified with the "SHARE-COVID19" keyword to facilitate their identification and access.

30 Available at: https://www.ilo.org/global/topics/coronavirus/regional-country/country-responses/lang--en/index.htm
31 Available at: https://www.eurofound.europa.eu/observatories/emcc

References

European Commission. (2020a). Policy measures taken against the spread and impact of the coronavirus (28 May 2020). European Commission - Directorate General Economic and Financial Affairs. Available at: https://ec.europa.eu/info/files/policy-measures-against-spread-coronavirus_en

European Commission. (2020b). Policy measures taken against the spread and impact of the coronavirus (16 November 2020). European Commission – Directorate General Economic and Financial Affairs, Brussels. Available at: https://ec.europa.eu/info/files/policy-measures-against-spread-coronavirus_en

European Commission. (2020c). Policy measures taken against the spread and impact of the coronavirus (8 December 2020). European Commission – Directorate General Economic and Financial Affairs, Brussels. Available at: https://ec.europa.eu/info/files/policy-measures-against-spread-coronavirus_en

European Trade Union Confederation (ETUC). (2020a). Short time work measures across Europe. ETUC Briefing Note: Short Time Work 25 May 2020. Available at: https://www.etuc.org/sites/default/files/publication/file/2020-05/Covid_19%20Briefing%20Short%20Time%20Work%20Measures%2024%20May.pdf

European Trade Union Confederation (ETUC). (2020b). Short time work measures across Europe. ETUC Briefing Note: Short Time Work 27 November 2020. Available at: https://www.etuc.org/sites/default/files/publication/file/2020-11/Covid_19%20Briefing%20Short%20Time%20Work%20Measures%2027%20November.pdf

OECD. (2020a). Job retention schemes during the COVID-19 lockdown and beyond, OECD policy responses to coronavirus (COVID-19). OECD Publishing. https://doi.org/10.1787/0853ba1d-en

OECD. (2020b). OECD Employment Outlook 2020: Worker security and the COVID-19 crisis, OECD Publishing. https://doi.org/10.1787/1686c758-en

ISBN 978-3-00-069877-4