The Survey of Health, Ageing and Retirement in Europe (SHARE) is a multidisciplinary and cross-national panel database of micro data on health, socio-economic status and social and family networks of more than 85,000 individuals (approximately 150,000 interviews) from 19 European countries (+Israel) aged 50 or over.

This book has two parts: “Innovations” and “Methodology”. The first six chapters of this book are devoted to the innovations in SHARE Wave 4, starting with an account of the experiences of the new countries in joining SHARE in wave four. Furthermore a detailed account of the most important new content in wave 4 is given, namely the social networks module. In the following chapter the experiences with taking and analysing dried blood spots in Germany is described. Furthermore, another chapter focuses on the linking of respondents’ retirement accounts at the German pension system with SHARE interview data. The third sub-project within the German SHARE test studies concerns response behaviour where the process and the first results of an experiment of the effects of unconditional respondent cash incentives on response rates are laid out. The section on innovations is completed by an introduction to the improvements made to key software tools such as the sample distributor and the online translation tool.

The section “Methodology” starts with a detailed outline of the sampling process and outcomes that were used to enhance the panel samples of SHARE with refreshment samples. An overview of managing a multinational survey infrastructure from an operational point of view, focusing on fieldwork monitoring and quality control in times of profound organizational changes is given as well. The book finishes with a summary of key indicators of survey response, such as contact and cooperation rates, response rates of the refreshment samples and retention rates of the panel samples.

ISBN 978-3-00-040802-1

SHARE Wave 4: Innovations & Methodology

SHARE

Survey of Health, Ageing and Retirement in Europe

SHARE Wave 4: Innovations & Methodology

Edited by
Frederic Malter
Axel Börsch-Supan

SHARE Wave 4
Innovations & Methodology

Edited by:
Frederic Malter
Axel Börsch-Supan

Authors:

Liili Abduladze
Eric Balster
Axel Börsch-Supan
Christin Czaplicki
Marcel Das
Giuseppe De Luca
Alice Delerue Matos
Róbert Gál
Matthias Ganninger
Sabine Häder
Gábor Kézdi
Thorsten Kneip
Julie Korbmacher
Markus Kotte
Ulrich Krieger
Vladimir Lavrač
Howard Litwin
Peter Lynn
Boris Majcen
Frederic Malter
Maurice Martens
Saša Mašič
Pedro Pita Barros
Anat Roll
Luule Sakkaeus
Barbara Schaan
Sharon Shiovitz-Ezra
Kim Stoeckel
Arnaud Wijnant

January 2013

The SHARE data collection has been primarily funded by the European Commission through the 5th Framework Programme (project QLK6-CT-2001-00360 in the thematic programme Quality of Life), through the 6th Framework Programme (projects SHARE-I3, RII-CT-2006-062193, COMPARE, CIT5- CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th Framework Programme (SHARE-PREP, N° 211909, SHARE-LEAP, N° 227822 and SHARE M4, N° 261982). Additional funding from the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, R21 AG025169, Y1-AG-4553-01, IAG BSR06-11 and OGHA 04-064) and the German Ministry of Education and Research as well as from various national sources is gratefully acknowledged (see www.share-project.org for a full list of funding institutions).

Published by
Munich Center for the Economics of Aging (MEA)
Amalienstrasse 33
80799 München
Tel: +49-89-38602-0
Fax: +49-89-38602-390
www.mea.mpisoc.mpg.de

Cover Design by
Markus Berger

Printed by
CS Werbezentrum
Goethestr. 28
80336 München

Suggested Citation:
Malter, F., Börsch-Supan, A. (Eds.) (2013). SHARE Wave 4: Innovations & Methodology. Munich: MEA, Max Planck Institute for Social Law and Social Policy.

ISBN 978-3-00-040802-1

Contents

1	SHARE Wave Four: New Countries, New Content, New Legal and Financial Framework *Axel Börsch-Supan, Max Planck Institute*	5
2	Becoming a New SHARE Country *Luule Sakkeus, Liili Abuladze, Estonian Institute for Population Studies, Tallinn University* *Gábor Kézdi, Central European University* *Róbert Gál, TARKI* *Pedro Pita Barros, Universidade Nova de Lisboa* *Alice Delerue Matos, Universidade do Minho* *Boris Majcen, Vladimir Lavrač, Saša Mašič, Institute for Economic Research, Ljubljana*	11
3	Social Network Measurement in SHARE Wave Four *Howard Litwin, Kim Stoeckel, Anat Roll, Sharon Shiovitz-Ezra, The Hebrew University of Jerusalem* *Markus Kotte, Max Planck Institute*	18
4	Collection of Biomarkers in the Survey of Health, Ageing and Retirement in Europe (SHARE) *Barbara Schaan, GESIS*	38
5	Linking SHARE Survey Data with Administrative Records: First Experiences from SHARE-Germany *Julie Korbmacher, Christin Czaplicki, Max Planck Institute*	47
6	Investigating Response Behavior *Axel Börsch-Supan, Max Planck Institute* *Ulrich Krieger, University of Mannheim*	53
7	Software Innovation in SHARE Wave Four *Arnaud Wijnant, Maurice Martens, Eric Balster, Marcel Das CentERdata, Tilburg University*	62
8	Sample Design in SHARE Wave Four *Peter Lynn, University of Essex* *Giuseppe De Luca, University of Palermo* *Matthias Ganninger, Sabine Häder, GESIS*	74
9	Fieldwork Management and Monitoring in SHARE Wave Four *Frederic Malter, Max Planck Institute*	124
10	Survey Participation in the Fourth Wave of SHARE *Thorsten Kneip, Max Planck Institute*	140

1 SHARE Wave Four: New Countries, New Content, New Legal and Financial Framework

Axel Börsch-Supan, Max Planck Institute

This book summarizes innovations and key methodological advancements achieved in the fourth wave of the Survey of Health, Ageing and Retirement in Europe (SHARE). SHARE's main aim is to provide data on individuals as they age and their environment in order to analyse the process of individual and population ageing in depth. SHARE is a distributed European research infrastructure which provides data for social scientists, including demographers, economists, psychologists, sociologists, biologists, epidemiologists, public health and health policy experts who are interested in population aging.

Covering the key areas of life, namely health, socio-economics and social networks, SHARE includes a great variety of information: health variables (e.g. self-reported health, health conditions, physical and cognitive functioning, health behavior, use of health care facilities), bio-markers (e.g. grip strength, body-mass index, peak flow; and piloting dried blood spots, waist circumference, blood pressure), psychological variables (e.g. mental health, well-being, life satisfaction), economic variables (current work activity, job characteristics, opportunities to work past retirement age, sources and composition of current income, wealth and consumption, housing, education), and social support variables (e.g. assistance within families, transfers of income and assets, volunteer activities) as well as social network information (e.g. contacts, proximity, satisfaction with network). Researchers may download the SHARE data free of charge from the project's website at www.share-project.org.

SHARE is unique in that it is not only multidisciplinary, but also multi-national. Wave four contains data from about 65,000 individuals aged 50 or over from 19 countries. Moreover, SHARE is harmonized with the U.S. Health and Retirement Study (HRS) and the English Longitudinal Study of Ageing (ELSA). Studies in Korea, Japan, China, India, and Brazil follow these models. Rigorous procedural guidelines, electronic tools, and instruments are designed to ensure an ex-ante harmonized cross-national design.

1.1 Design of the fourth wave

After the third wave collected retrospective life history data (SHARELIFE), SHARE returned in wave four to the "classical" longitudinal design employing the same questionnaire as in wave two in order to assess the changes in life circumstances of the respondents as they age. Hence, part of the of questionnaire design process was business as usual by modestly revising and remodelling the instrument of the second wave, thereby balancing research interests from our multi-disciplinary team of scientists, cleaning up errors found in wave two, and reducing respondent burden by cutting items with little variation or no documented scientific usage.

At the same time, however, the fourth wave introduced several areas of innovation which proved much more challenging than anticipated. First, four new countries joined SHARE to conduct their baseline interviews and become new members of the longitudinal infrastructure (Estonia, Hungary, Portugal, and Slovenia). A fifth new country, Luxembourg, joined for pilot and pre-test only.

Second, SHARE is the first cross-national multidisciplinary survey to introduce a social networks module using a name generator. Third, a set of innovative endeavors were conducted in Germany as test runs for a SHARE-wide implementation in waves five and six. These included taking dried blood spots to ascertain blood sugar level and other markers of health, a record linkage with the employment and earnings histories of the German Pension Fund (Deutsche Rentenversicherung Bund), and an experiment to examine the effect of monetary interviewer incentives on response rates and response bias.

SHARE is an ex-ante harmonized survey with a great emphasis on minimizing artefacts due to methodological differences among countries. Electronic tools are essential to monitor and then eliminate cross-national deviations during the translation and survey process. As a fourth area of innovations, SHARE employed completely revised online translation and sample distributor tools.

1.2 First ERIC

The legal, governance and financial framework changed dramatically when SHARE became officially the first European Research Infrastructure Consortium (ERIC) in March 2011 as part of the ESFRI process (European Strategy Forum on Research Infrastructures). Details of becoming an ERIC can be found elsewhere[1]. SHARE has started as a pre-dominantly centrally financed enterprise. This was crucial for the harmonization across all member states. Data collection for waves one to three has been primarily funded by the European Commission through the 5th and 6th framework programmes. Substantial additional funding came from the U.S. National Institute on Aging. With becoming an ERIC, national funding is now dominant, with substantial support of the European Commission’s DG Employment, Social Affairs and Equal Opportunities to the four new countries.

Central tasks and coordination are financed by the German Ministry for Science and Education, the European Commission, and the US National Institute of Aging. Because several legal issues could not be solved in time in Germany, the official seat of the SHARE-ERIC has been set up temporarily in the Netherlands with the great help of the Dutch Ministry for Science. It will move to Munich, Germany, in 2013.

The new financial organization had substantial implications for running the survey. The transition was complicated by the coordinating institution’s move from Mannheim to Munich during wave four, becoming the Munich Center for the Economics of Aging (MEA) as part of the Max Planck Institute for Social Law and Social Policy.

1.3 Chapter overview

This book has two parts: “Innovations” and “Methodology”. After this introduction, the first six chapters of this book are devoted to the innovations in wave four, starting with an account of the experiences of the new countries in joining SHARE in wave

[1] Kleiner, Brian, Renschler, Isabelle, Wernli, Boris, Farago, Peter, and Joye, Dominique (eds). *Research Infrastructures in the Social Sciences.* Seismo: Zurich, 2013.

four. Liili Abuladze, Alice Delerue Matos, Róbert Gál, Gábor Kézdi, Vladimir Lavrač, Boris Majcen, Saša Mašič, Pedro Pita Barros and Luule Sakkaeus describe what motivated them to join with their countries the SHARE infrastructure, how they secured funding, assembled their national teams and what obstacles they overcame when implementing the survey operation in their countries. In chapter three, Howard Litwin, Kim Stoeckel, Anat Roll and Sharon Shiovitz-Ezra give a detailed account of the most important new content in wave 4, namely the social networks module. Especially users of this exciting new data will find this chapter a useful manual. In chapter four, Barbara Schaan summarizes the experiences with taking and analysing dried blood spots in Germany. In chapter five, Julie Korbmacher and Christin Czaplicki describe the linking of respondents' retirement accounts at the German pension system with their SHARE interview data. The third sub-project within the German SHARE test studies concerns response behaviour, chapter six. Here, Axel Börsch-Supan and Ulrich Krieger describe the process and the first results of an experiment of the effects of unconditional respondent cash incentives on response rates. The section on innovations is completed by chapter seven, where *Arnaud Wijnant, Maurice Martens, Eric Balster, Marcel Das* introduce the improvements made to key software tools such as the sample distributor and the online translation tool.

The section "Methodology" starts with chapter eight, in which Peter Lynn, Giuseppe De Luca, Matthias Ganninger and Sabine Häder give a detailed outline of the sampling process and outcomes that were used to enhance the panel samples of SHARE with refreshment samples. Frederic Malter gives an overview of managing a multinational survey infrastructure from an operational point of view, focusing on fieldwork monitoring and quality control in times of profound organizational changes (ERIC, see above). The book is closed out by chapter ten in which Thorsten Kneip summarizes key indicators of survey response, such as contact and cooperation rates, response rates of the refreshment samples and retention rates of the panel samples.

Acknowledgements

As in previous waves, 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 which 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 fieldwork of SHARE relied in most countries on professional survey agencies – IFES (AT), PSBH, Univ. de Liège (BE), Link (CH), SC&C (CZ), Infas (DE), SFI Survey (DK), Statistics Estonia (EE), TNS Demoscopia (ES), INSEE and GfK-ISL (FR), TARKI (HU), DOXA (IT), TNS NIPO (NL), TNS OBOP (PL), GfK Metris (PT), Intervjubolaget (SE), CJMMK (SI) – and we thank their representatives for a fruitful cooperation. Especially the work of the almost 2000 interviewers across Europe was essential to this project.

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. In wave four, member states' support accounted for 59 percent of the budget, while 33 percent came from the European Commission and eight percent from US-NIA. This is distinctively different from the first three waves, in which the

European Commission contributed 62 percent, member states 26 percent, and US-NIA 12 percent.

The EU Commission's contribution to SHARE through the 7th framework programme (SHARE-M4, No 261982) is gratefully acknowledged. The SHARE-M4 project financed all coordination and networking activities outside of Germany. We thank, in alphabetical order, Hervé Pero, Robert-Jan Smits, Maria Theofilatou, and Octavio Quintana Trias in DG Research for their continuing support of SHARE. We are also grateful for the support by DG Employment, Social Affairs, and Equal Opportunities through Georg Fischer, Fritz von Nordheim, and Ruth Paserman which made the introduction of the four new countries possible.

Substantial co-funding for add-ons such as the physical performance measures, the train-the-trainer program for the SHARE interviewers, and the respondent incentives, among others, came from the US National Institute on Ageing (P30 AG12815, R03 AG041397, R21 AG025169, R21 AG32578, R21 AG040387, Y1-AG-4553-01, IAG BSR06-11 and OGHA 04-064). We thank John Phillips and Richard Suzman for their enduring support and intellectual input.

The German Ministry of Science and Education (BMBF) financed all coordination activities at MEA, the coordinating institution. We owe special thanks to Angelika Willms-Herget, who also serves as chair of the SHARE-ERIC Council, and, in alphabetical order, Hans Nerlich, Andrea Oepen, Brunhild Spannhake and Beatrix Vierkorn-Rudolph who helped us with determination and patience to set up SHARE as a research infrastructure in Germany. Finally, we thank Richard Derksen from the Dutch Ministry of Science for his help and advice: he was instrumental for setting up the ERIC in our first host country, the Netherlands.

All SHARE countries had national co-funding which was important to carry out the study. Belgium (BE) was funded through the Belgian Federal Science Policy Administration. Switzerland (CH) received funding from the Swiss national science foundation (SNSF), grant number 10FI13_126791/1. The Czech Republic (CZ) received funding from the Ministry of Education, Youth and Sports. Germany (DE) received funding from the Bundesministerium Für Bildung und Forschung (BMBF), Deutsche Forschungsgemeinschaft (DFG), Volkswagen Stiftung and the Forschungsnetzwerk Alterssicherung (FNA) of the Deutsche Rentenversicherung (DRV). Denmark (DK) received support from The Danish Council for Independent Research: Social Sciences (FSE) (ref.no: 09-068995). Estonia (EE) received national funding from the Estonian Scientific Council, grant number SF0130018s11 and ETF 8325, grants No. 3.2.0601.11-0001 and 3.2.0301.11-0350 and the Ministry of Social Affairs. Spain (ES) acknowledges gratefully the support from MICINN (Ministerio de Ciencia e Innovación, Subprograma de Actuaciones Relativas a Infraestructuras Científicas Internacionales, AIC10-A-000457) and special thanks to Instituto Nacional de Estadística (INE) and IMSERSO. In France (FR), wave four has been realized and financed jointly by Institut national de la statistique et des études économiques (INSEE) and Institut de recherche et de documentation en économie de la santé (IRDES). Other national wave four funders were: Institut de recherche en santé publique (IReSP), Direction générale pour la recherche et l'innovation du ministère de l'enseignement supérieur et de la recherche (DGRI), Direction de la recherche, des études, de l'évaluation et des statistiques du ministère de la santé (DREES), Direction de l'animation de la recherche, des études et des statistiques du ministère du travail (DARES), Caisse nationale de solidarité pour l'autonomie (CNSA), Caisse nationale d'assurance vieillesse (CNAV), Conseil d'orientation des retraites (COR), Institut national de prévention et d'éducation pour la santé (INPES). In Italy (IT), funding for the fourth wave of SHARE was provided by the Ministry of University and Research (MIUR) and by the following foundations: Compagnia di San Paolo, Fondazione Cassa di Risparmio di Padova e Rovigo and Forum ANIA Consumatori. Generous support was also given by Bank of Italy and the National Research Council (Consiglio Nazionale delle Ricerche - CNR). Data collection in the Netherlands (NL) was nationally funded by The Netherlands Organisation for Scientific Research (NWO), The Dutch Ministry of Education, Culture and Science, by the Province of Noord-Brabant, and by Netspar and Tilburg University. Poland (PL) got funding from Educational Research Institute (Instytut Badań Edukacyjnych, IBE). Portugal acknowledges the support of the Alto-Comissariado da Saúde (High Commissioner for Health). Sweden (SE) was supported by the Swedish Council for Working Life and Social Research (FAS), the Swedish Social Insurance Inspectorate, the Swedish Pensions Agency, Riksbankens Jubileumsfond, the Ministry of Social Affairs, AFA Insurance and the Swedish Social Insurance Agency.

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

We are very grateful to the contributions of the four area coordination teams involved in the design process. Guglielmo Weber (University of Padua) led the economic area with Agar Brugiavini, Anne Laferrère, Viola Angelini and Giacomo Pasini. The health area was led by Karen Andersen-Ranberg and assisted by Mette Lindholm Eriksen (South Denmark University) with support from Johan Mackenbach and Mauricio Avendano, Simone Croezen at Erasmus University. Health care and health services utilisation fell into the realm of Hendrik Jürges (University of Wuppertal). The fourth area, family and social networks, was led by Howie Litwin from Hebrew University with assistance from Kim Stoeckel, Anat Roll and Marina Motsenok.

Many countries drew refreshment samples in the fourth wave. We gratefully thank our external sampling experts Matthias Ganninger, Sabine Häder, Guiseppe de Luca and Peter Lynn for providing their expertise to the country teams.

The coordination of SHARE entails a large amount of day-to-day work which is easily understated. We would like to thank Yvonne Berrens, Kathrin Axt, and Marie-Louise Kemperman for their management coordination, Stephanie Lasson, and Eva Schneider, Sabine Massoth, and Hannelore Henning at MEA in Mannheim and Munich for their administrative support throughout various phases of the project. Martina Brandt, and Frederic Malter provided as assistant coordinators the backbone work in coordinating, developing, and organizing wave four of SHARE. Preparing the data files for the fieldwork, monitoring the survey agencies, testing the data for errors and consistency are all tasks which are essential to this project. The authors and editors are grateful to Annelies Blom, Johanna Bristle, Christine Czaplicki, Christian Hunkler, Markus Kotte, Thorsten Kneip, Julie Korbmacher, Gregor Sand, Barbara Schaan, Morten Schuth, Stephanie Stuck, and Sabrina Zuber for data cleaning and monitoring services at MEA in Mannheim and Munich. We owe thanks to Guiseppe de Luca and Claudio Rosetti for weight calculations and imputations in Palermo and Rome. Markus Berger and Lisa Schug were responsible for the design work around the book and we greatly appreciate their work.

Programming and software development for the SHARE survey was done by CentERdata at Tilburg. We want to thank Alerk Amin, Marcel Das, Maurice Martens, Corrie Vis, Iggy van der Wielen and Arnaud Wijnant for their support, patience and time.

Last but by no means least, the country teams are the flesh to the body of SHARE and provided invaluable support: Rudolf Winter-Ebmer, Nicole Halmdienst, Michael Radhuber and Mario Schnalzenberger (Austria); Marie-Thérèse Casman, Xavier Flawinne, Stephanie Linchet, Dimitri Mortelmans, Laurent Nisen, Sergio Perelman, Jean-François Reynaerts, Jérôme Schoenmaeckers, Greet Sleurs, Karel Van den Bosch and Aaron Van den Heede (Belgium); Radim Bohacek, Michal Kejak and Jan Kroupa (Czech Republic); Karen Andersen Ranberg and Henriette Engberg (Denmark); Luule Sakkeus, Enn Laansoo Jr., Heidi Pellmas Silja Karu, Kati Karelson, Tiina Linno, Anne Tihaste and Lena Rõbakova (Estonia); Anne Laferrère, Nicolas Briant, Pascal Godefroy, Marie-Camille Lenormand and Nicolas Sirven (France); Annelies Blom, Christine Diemand and Ulrich Krieger (Germany); Anikó Biró, Róbert Iván Gál, Gábor Kézdi, Lili Vargha (Hungary), Liam Delaney and Colm Harmon (Ireland), Guglielmo Weber, Danilo Cavapozzi, Elisabetta Trevisan, Chiara Dal Bianco, Alessio Fiume and Omar Paccagnella (Italy); Frank van der Duyn Schouten, Johannes Binswanger, Adriaan Kalwij and Irina Suanet (Netherlands); Michał Myck, Malgorzata Kalbardczyk,Anna Nicinska, Monika Oczkowska and Michał Kundera (Poland); Pedro Pita Barros and Alice Delerue A. Matos (PT), Pedro Mira and Laura Crespo (Spain); Per Johansson and Daniel Hallberg (Sweden); Carmen Borrat-Besson (FORS), Alberto Holly (IEMS), Peter Farago (FORS), Thomas Lufkin (IEMS), Pierre Stadelmann (IEMS) Boris Wernli (FORS) (Switzerland), Boris Majcen, Vladimir Lavrač and Saša Mašič (Slovenia).

2 Becoming a New SHARE Country

This chapter reflects the experience of four countries which entered SHARE as new countries in wave four. We asked the authors to give an account of what motivated them to become part of SHARE, what obstacles and challenges they encountered, how these were overcome and to give a brief outlook of future directions. Countries appear in alphabetical order.

2.1 Estonia

Luule Sakkeus, Liili Abuladze, Estonian Institute for Population Studies, Tallinn University

2.1.1 Introduction

As is true for all European countries, population ageing affects Estonia as well. However, Estonia experienced a slower change in age distribution than is typical for the ageing process in the other European countries. Three main determinants of population ageing withheld the rapid increase of population ageing during the second half of the 20th century in Estonia compared to Western and Northern Europe: a) post-war fertility trend in Estonia lacked the baby-boom effect, b) intensive immigration of younger cohorts into Estonia during five post-war decades and c) the same post-war period was characterised by mortality stagnation. Combined, these three determinants withheld the ageing process (Katus et al. 2003). However, immigrant populations that previously slowed down population ageing due to their younger age structure became one of the main determinants of rapid ageing starting in the 1990s, when the numerous inflows of immigrants from countries of former Soviet Union stopped. Thus, Estonia, together with Slovenia, has been among the European countries with the biggest annual average growth of the elderly during the last two decades, which amounts to around 4-5 percentage points per year. Adding sharp fertility declines due to postponement of childbirths into later ages since the early 1990s, the demographic trends present new challenges for the future. Additionally, long-term accumulation of bad health conditions during the period of mortality stagnation pose extra challenges for Estonia, in particular achieving the targets set in Europe 2020.

2.1.2 Funding and assembling a national working group

The Social Agenda of the EU and the Open Method of Coordination programme of the PROGRESS call in 2009 came very timely. The negotiations between the Estonian Institute for Population Studies, Tallinn University (EDI) and the Ministry of Social Affairs (SOM) resulted in the decision of the ministry to give the mandate for implementing SHARE in Estonia to Statistics Estonia (SE), which applied for V2009/009 funds together with EDI, the Institute of International and Social Research (RASI) and the Institute for Social Work (STI), Tallinn University and the National Institute of Public Health (TAI). SHARE Estonia was funded by EC grant (VS/2009/0561), a grant from SOM, grants from Ministry of Education and Research

(HTM No 586/2011 and 146/2012, SF0130018s11, SF0130018s11AP12) and a grant from Estonian Science Fund (ETF No. 8325).

A SHARE Estonia Working Group was formed with these institutes (SE, EDI, RASI, STI and TAI). This consortium monitored the translation of the generic SHARE questionnaire into Estonian and Russian, prepared the manuals for the interviewers and together with SE carried out the training sessions. The consortium also succeeded to respond to the call of the SHARE Management Board to enlarge the sample size to about 6000 respondents with extra funding from SOM. In 2011 under the lead of Tallinn University, the SHARE Estonia Steering Committee was constituted which comprises of representatives of four ministries - aside SOM also the Ministry of Education and Research (HTM), the Ministry of Economics and Communication (MKM) and the Ministry of Finance (RM)) and research institutes from Tallinn University (EDI, RASI, STI) and Tartu University (Institute for Health Care (TI), the Department of Economics (MTK), the Institute of Psychology (PI), the Institute of Internal Medicine (SK) and the Estonian Genome Center (GV). The Steering Committee was responsible for helping to find resources to maintain SHARE in the long run and make the most important decisions related to survey management and implementation. The representatives of the scientific institutions of the SHARE Estonia Steering Committee formed the core of the Scientific Board which decided on methodological questions of the survey. The country team leader and scientific coordinator of SHARE Estonia were located at EDI. SOM was the representative of Estonia in negotiations with the SHARE-ERIC.

2.1.3 Survey implementation

Estonia implemented SHARE wave four as its first round, using only the baseline version of the questionnaire. The sample frame of SHARE was based on a population register which allowed drawing age-eligible target individuals from each household. The data of the population register also helped to review the household composition. Our main challenges in implementing SHARE were mostly related to a shortage of information of the scale, organisation and management of the survey before submitting our grant applications. That put us into a very tight economic situation. Another challenge for newcomers like Estonia was the translation of the generic CAPI questionnaire. In Estonia, we needed double efforts as the questionnaire was implemented in two languages – Estonian and Russian - to accommodate the 30 percent of foreign-born target persons.

Another challenge was the late decision to increase the sample size up to 6000 respondents. There was, however, common understanding among members of the SHARE Estonia Working Group that this large increase was the only suitable way to obtain data that would allow any meaningful country-specific results over time.

2.1.4 Summary

The main challenge for SHARE Estonia appears to be securing sustainable funding for future longitudinal waves. Despite SHARE being on ESFRI and now (Sept. 2012) being the first ERIC, SHARE is quite costly from the social sciences perspective. The national funders, especially in new countries, where no country-specific results from SHARE are yet available, might be hesitant to include SHARE on national ESFRI roadmap or commit themselves to ERIC. In our opinion, SHARE would benefit substantially from centralised funding for national data collection to be sustainable in the long run.

2.2 Hungary

Gábor Kézdi, Central European University

Róbert Gál, TARKI

2.2.1 Introduction

Many factors motivated Hungary's joining the SHARE project. Challenges arising from population ageing in the areas of health, employment and retirement are as important in Hungary as in other European countries. Yet, appropriate data were scarce in Hungary. One important advantage of SHARE is the possibility to analyse cross-country variation in institutions and policies of the welfare state using highly comparable data. We were convinced that joining SHARE would be beneficial from both perspectives: on the one hand, Hungary would add especially useful variation to SHARE by being a new Member State of the European Union where welfare regimes have undergone a very different trajectory than in the 15 older Member States. On the other hand, providing country-specific findings framed in a pan-European context would give Hungarian policy makers a better foundation for evidenced-based decision making around challenges of population ageing.

2.2.2 Funding and assembling a national working group

As other new countries in wave four, Hungary was funded by DG Employment, under grant VS/2009/0560. The grant was supplemented by contributions from the TARKI Foundation. No other funding was obtained for wave four. Putting together a research team for Hungary was accomplished without major problems. The two country team leaders, Róbert Gál (TARKI) and Gábor Kézdi (Central European University, CEU) have been long-time research collaborators in economic demography. We have been thinking of getting Hungary into SHARE for quite some time. Anikó Biró, an operative staff in the Hungarian country team, was a doctoral student at Central European University during wave four. Lili Vargha, another key operative person in SHARE Hungary, was already affiliated with TARKI. The fact that the survey agency TARKI was a research institute at once, hosting half of the research team, proved to be very helpful as well.

2.2.3 Survey implementation

Translating the generic questionnaire to Hungarian went relatively smoothly. The most important challenges in the translation process were fitting the Hungarian health and retirement institutions into the structure and response options of the generic questionnaire. The Hungarian pension system went through fundamental changes extremely quickly right before the fieldwork began, causing some extra difficulties. These challenges were overcome by the dedicated, hard-working members of the team and the fact that one member had been an expert in pension systems. A novelty of SHARE, compared to other survey operations, was that the CAPI questionnaire and the electronic contact protocol allowed for real-time, very quick feedback on some aspects of fieldwork quality. All members of CEU and TARKI gave feedback to the interviewers and used indicators generated by the electronic survey infrastructure to check on their performance with respect to an incentive scheme. For example, the quality of interviewers' performance was checked through keystroke data right after interviews were completed, and outlier cases were double-checked with the interviewers. This was one of the many instances where having half of our team work at TARKI proved to be very helpful.

2.2.4 Summary

The most important lessons from wave four include the importance of having an excellent team. The need to design more detailed incentives for interviewers that include feedback on data quality is another important lesson. We believe that our experience will help further improve the quality of the SHARE data not only in Hungary but in other countries, too. While internal funding for SHARE is still not secure in Hungary, we believe that an appropriate solution will be found so that the exceptionally rich SHARE data can help policy analysis in Hungary and appropriate data from Hungary can add to the scientific value of SHARE.

2.3 Portugal

Pedro Pita Barros, Universidade Nova de Lisboa

Alice Delerue Matos, Universidade do Minho

2.3.1 Introduction

Like most other European countries, Portugal is undergoing rapid population ageing. Not only has longevity of the population increased, fertility rates have declined sharply. This implies a fast increasing ratio of the elderly population to the young and to the working-age population. Society and public policies at all levels – from central to local – will have to prepare and adjust for a different age structure of the population. SHARE fills a gap in scientific knowledge by providing a multidisciplinary and longitudinal approach of the ageing process in Portugal and a comparative analysis at the European level.

2.3.2 Funding and assembling a national working group

Funding for wave four was provided by a DG Employment grant and by national funding from Alto Comissariado da Saúde (High Commissioner for Health). The national funding partner was very enthusiastic about the project from the start. After grants were approved, the main issue has been the rules applying to the use of funds, which had to follow the stricter rules of civil service in Portugal. The team was formed with an economist and a sociologist, which allowed for a suitable division of work. Several junior research assistants were also hired to participate in the project. The main lesson to take is that managing SHARE at the country level is a full time occupation.

2.3.3 Survey implementation

The translation process posed no particular problems, but was quite time consuming. It required expertise and knowledge of the technical terminology used in some items. A translator's work was reviewed, especially for translation of terms that are usual among experts but not commonly used outside a certain context. More problematic was to adapt concepts that are adjusted for one set of countries but do not correspond to a general situation in another country. One example of this were the questions regarding asset holdings and investments, which in Portugal are not widespread (due to general poverty of the old age population, a public pension system and lack of widespread knowledge on how to buy and sell financial assets, either through banks' funds or directly at the stock exchange).

There were two main challenges worth mentioning. The first challenge was the definition of the sample. The source of information ended up being the national registry of National Health Service's beneficiaries. The National Health Service covers the total population that resides in Portugal. The registry had, however, inaccuracies. These necessitated two different types of procedures. First, before selection of the sample, addresses without zip codes were dropped (about 6 percent of the total sample). Checks on representativeness yielded no statistically significant differences in the age and sex distributions of the units included and excluded from the sampling frame. Second, after sample selection took place, a time-consuming check of each incomplete address had to be performed (by telephone). The second main challenge was to have the survey agency comply with all quality requirements, and permanent communication was required.

2.3.4 Summary

The main lessons learned were the importance of securing sufficient funding, and setting up a very clear contract with the survey agency. We have high hopes of the possibility of using ERIC to facilitate the availability of national funds. The expected gain will not only be financial but also add flexibility. For example, to achieve a long-term goal of 6,000 interviews, the expected costs would force us to launch an international tender procedure, which according to current Portuguese rules can take six months or more. A full-time executive manager would improve the management of the SHARE process considerably.

2.4. Slovenia

Boris Majcen, Vladimir Lavrač, Saša Mašič, Institute for Economic Research, Ljubljana

2.4.1 Introduction

Slovenia is among those EU countries where the process of population ageing is most pronounced. Projected trends of demographic change present big challenges: population ageing burdens public finance, health and pension system and affects the labour market negatively if no corrective action is taken. These demographic trends are taking place in the context of preparations for serious structural reforms (welfare state, labour market, pension, health, long-term care reforms) and severe austerity measures, adopted recently (as of July. 2012).

Research on public policies and consequences of population ageing at the Institute for Economic Research (IER), Ljubljana has shown a severe lack of consistent and reliable data, which would enable researchers to assess the overall situation of this segment of the population, carry out scientific analysis and suggest measures and reforms to the policy makers in relevant areas. Joining SHARE with its multidisciplinary, internationally comparable and longitudinal dimension based on an ex-ante harmonised questionnaire was therefore most welcome. This was finally also recognised by all five relevant ministries (labour, health, science, finance, development) in Slovenia, which crucially contributed to securing funding for the inclusion of Slovenia in SHARE by signing letters of support.

2.4.2 Funding and assembling a national working group

Securing long-term funding of SHARE was a complicated and demanding process which took us 3 years and finally resulted in the inclusion of SHARE among SISFRI projects (Slovenian version of ESFRI road map), which in principle should be financed until 2020. Unfortunately, the ability to finance each consecutive wave of SHARE still depends on each year's available budget. As a consequence, it was uncertain until the end whether funding for SHARE wave four would be available or not. As all the activities (translating and preparing questionnaire, pilot and pre-test etc.) had to be undertaken in time anyway, the country team had to invest their time and own finance in the interim period, hoping that financing of SHARE in Slovenia would finally be approved.

The Slovenian SHARE country team was assembled at the Institute for Economic Research, founded in 1965 and led by Dr. Boris Majcen, director of the institute. Fieldwork coordination, data cleaning and questionnaire development were led by Vladimir Lavrač and Saša Mašič. As part of the University of Ljubljana, the Public Opinion and Mass Communication Research Centre (CJMMK) at the Faculty of Social Sciences was chosen to carry out the fieldwork.

2.4.3 Survey implementation

CJMMK, founded in 1965, has a long tradition in national and cross-national projects and well-established fieldwork procedures, including monitoring and incentives for interviewers. The translation of the questionnaire was done by the SHARE country team at the University of Ljubljana. Our ample experience in translating other social surveys carried out in Slovenia made the process of translating the SHARE questionnaire easier. The sampling procedures took less effort due to the existence of the Central Register of Population (CRP), where all residents with current address (citizens and non-citizens) are included and which is regularly updated. However, strict regulations apply to protection of personal data, which prolonged the planning phase that led to obtaining the sample.

2.4.4 Summary

The most important lesson learned was the need to secure financing for the entire wave, not just for one budget year, and the need to sign a contract with the survey agency that contains specified deliverables for both parties (the country team and the survey agency).

References

Katus, K., Puur, A., Põldma, A., Sakkeus, L. (2003): *Population Ageing and Socio-Economic Status of Older Persons in Estonia.* New York, Geneva: United Nations.

3 Social Network Measurement in SHARE Wave Four

Howard Litwin, Kim Stoeckel, Anat Roll, Sharon Shiovitz-Ezra, The Hebrew University of Jerusalem

Markus Kotte, Max Planck Institute

3.1 Introduction

The personal social networks of older people are linked to a wide range of outcomes, behaviors and perceptions (Litwin 2011; Litwin 2010). Consequently, the study of personal social networks, their antecedents, correlates and effects, has become an enterprise of increasing importance in social surveys, particularly those geared to the older segment of the population. This unique realm of inquiry enables researchers to clarify the contribution of one's social entourage to key dynamics that shape the nature and the quality of late life.

Although the importance of social network analysis is well-established in the literature, there is less agreement as to how personal social networks should best be measured for analytical purposes. Two main opposing thrusts can be discerned in this regard—a direct approach and an indirect approach. Indirect measurement of personal social networks is exemplified by the role-relational orientation which records the collection of social ties that one has, by category. Such indicators have also been termed socio-demographic proxies (Pescosolido 2011). Researchers engaged in this line of inquiry adopt an inferred interpretation of the social network (Litwin 1996). That is, the very existence of a social relationship is assumed to constitute sufficient evidence for comprising part of one's network. Thus, for example, if a given individual has an adult child, the child in question is considered to be a member of that person's network. Proponents of the indirect approach contend that this method provides an objective delineation of the social network phenomenon. This approach has been the principal basis for the collection of social network data in such major surveys as the American Health and Retirement Study, the English Longitudinal Study of Ageing and the first two waves of SHARE.

In comparison, advocates of the direct approach maintain that a social network is essentially a subjective phenomenon and that social ties function mainly if they are perceived to be meaningful or important to a given individual. This implies that one cannot infer the existence of a personal social network simply on the basis of the numeration of existing role relations. Rather, one must directly derive the network by querying specifically as to who it is that is important to a given respondent. This direct or derived approach to social network identification most usually entails the use of name generators through which network members are nominated. The use of name generators for network identification has been applied in the American General Social Survey (Burt 1986; Burt & Guilarte 1986), the Longitudinal Aging Study Amsterdam (van Tilburg 1995) and the National Social life, Health and Ageing Project (NSHAP) (Cornwell, Schumm, Laumann, & Graber 2009).

In order to widen and to diversify the measurement of social networks in SHARE, the fourth wave of the survey introduced a new social network module (SN) that employed the direct approach for social network derivation. It was based largely upon the instrument employed in NSHAP, along with important new additions. The key feature of the SN module in SHARE is the compilation of a list of meaningful people in the life of the respondent. The basis of the list is subjective, that is, it is based upon the respondent's own appraisal of who is important to him or her.

The interviewer asks a general probe and the respondent supplies a list of names, up to six in the first round, in response to the probe: "Over the last 12 months, who are the people with whom you most often discussed important things?" This question focuses the respondents to consider their confidants, persons with whom they interact, discuss things of relative importance, and maintain a degree of trust. In order to represent only the true confidents in a social network, respondents are limited to listing only six persons. As respondents may have a tendency to focus exclusively on family relationships, the question is also worded in a way to encourage consideration of other important people in their lives, such as friends, neighbors or other acquaintances.

An additional probe is asked in the SN name generating mechanism in SHARE by which respondents are allowed to mention one additional person who is important to them "for any other reason." This question was instituted so that respondents could cite a meaningful relationship they maintain that may not be otherwise classified in the main name-generating probe. As a result, the number of names on the social network roster can reach a maximum of seven; that is, up to six cited confidants and one additional person of importance. For each of the individuals listed in the network roster, additional information about the person is solicited from the respondent. These probes are applied to obtain information that describes the persons cited and the nature of the tie with them, respectively. In SHARE wave four, these include role relation categories, gender, residential proximity, frequency of contact and emotional closeness. The full list of SN role relation categories appears in Figure 3.1.1.

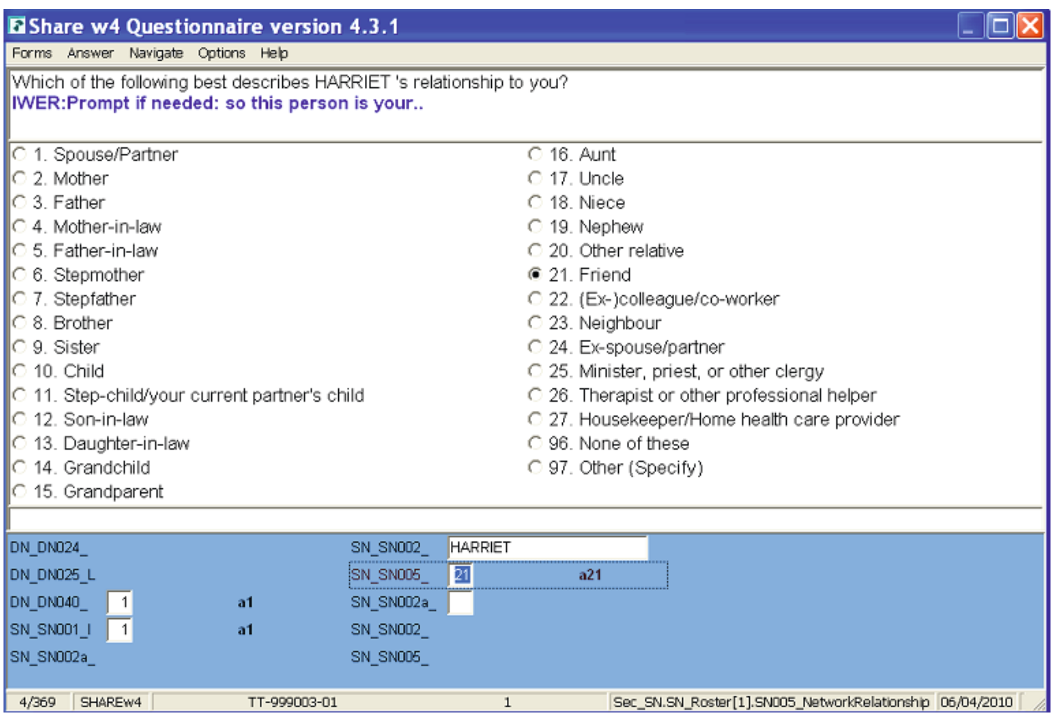


Figure 3.1.1 SN role relation categories (CAPI screenshot)

An innovative feature of the social network module fourth wave of SHARE calls up the names recorded on the SN name roster in two subsequent survey modules; social support (SP) and financial transfers (FT). No other major survey currently solicits this kind of information. The linkage between the SN module and the FT and SP modules consider the nature and extent of exchange within the context of personal social networks. The data from the fourth wave of the SHARE survey thus allows, for the first time, to distinguish between the exchange of money and support within one's personal social network and the exchange of money and support with other persons. In addition, characteristics of network members with whom financial or social support exchanges occur can be identified in the corresponding SN module.

The linkage between the SN and FT or SP modules in the wave four CAPI instrument in SHARE works in the following way: in each query as to the identity of persons with whom respondents exchange time or money, the names from the social network roster (and their corresponding role relationships) appear first. That is, the first seven answer categories on the answer screen are assigned to potential SN members. Answer categories eight and above are a general list of relationships to classify persons with whom exchanges occurred but are not part of the respondent's social network. These general role relationship categories, that is, the role categories numbered 8 or higher on the answer screen (e.g. child, brother, uncle etc.), are cited only if the recipient or giver of the exchange of help (time or money) was not listed as part of the survey respondent's social network roster of meaningful persons in his or her life. Figure 1.1.2, a screenshot from the CAPI, illustrates this.

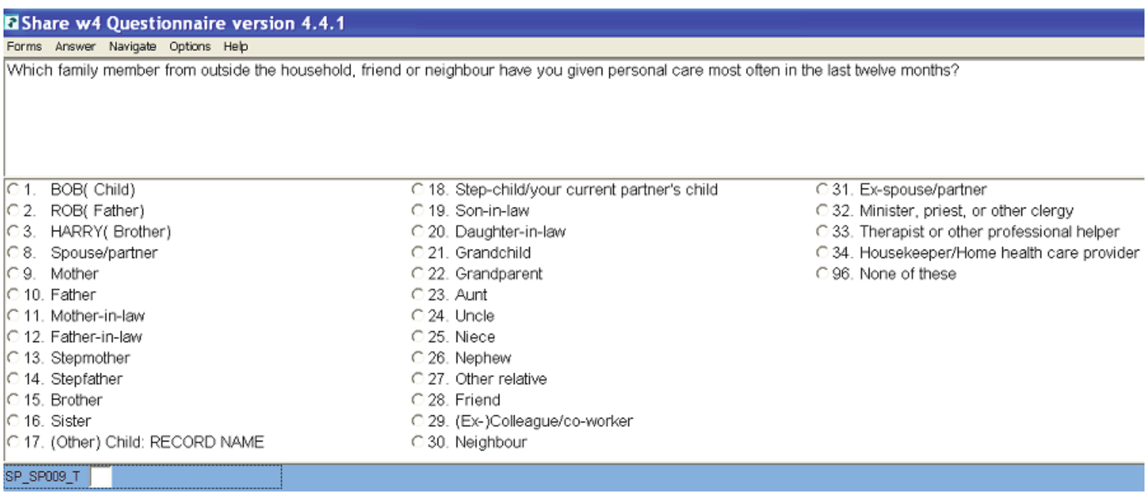


Figure 3.1.2 SP and FT role relation categories (CAPI screenshot)

Users wishing to synchronize the relationship categories in the SP and FT modules and the SN module should be aware of the differences in the relationship type variable values across the modules. For example, the value identifying a specific relationship category (e.g. neighbour, friend) is not necessarily the same in the SN module in comparison to the SP and FT modules. Identification of specific social network members in the variables referring to relationship types in the SP and FT modules is also possible. For most of the variables, the values 101 to 107 are always reserved for the up to seven members of the respondent's social network. If, for example, the variable ft017_1ft is coded 103 ("social network person 3"), users need to refer to the third social network member relationship type variable (sn005_3) to ascertain the role relation of the specific social network member. However, several variables in the SP module allow for unlimited response options for the respondents ("code all that apply"). For variables of this nature, the information is coded as dummies in the data, and the dummy variables identifying social network members are identified with the suffix "sn". For example, the dummy variable sp021d6sn is coded 1 ("respondent provided help to: social network person 6"), users should refer to the sixth social network member relationship type variable (sn005_6). Additional clarification about how these relationship categories appear in the data is outlined in the Wave 4 Release Guide 1.0.0[1].

3.2 Derived social network variables

The data gathered in the Social Network (SN) module and the corresponding sections in the Financial Transfer (FT) and Social Support (SP) modules in wave four of SHARE allow the derivation of a wide range of social network variables for analysis. Generated social network variables were derived to assist researchers with the dissemination of the social network data in the SN, FT and SP modules. Descriptions and listings of the derived social network variables follow.

3.3 Social network composition and satisfaction

The derived social network variables related to network composition and satisfaction are listed in Table 3.2.1.

Table 3.2.1A Derived SN variables: network composition and satisfaction

Variable name	Variable description	Generated variable coding description	-8 "Does not apply" description
	Global network descriptors		
sizeofsocialnetwork	Count of social network members	sn005_1 – sn005_7	
sn_satisfaction	satisfaction with social network – combined	sn012_; sn017_	
	Relationship Composition of Social Network		
famnet famnet2 famnet3	Family members in social network	sn005_x = 1-20	-8 = no social network
childnet childnet2 childnet3	Children in social network	sn005_x = 10, 11	-8 = no social network; no children

[1] http://www.share-project.org/data-access-documentation/documentation0.html

Table 3.2.1B Derived SN variables: network composition and satisfaction

gchildnet gchildnet2 gchildnet3	Grandchildren in social network	sn005_x = 14	-8 = no social network; no grandchildren
spousenet2 spousenet3	Spouse in social network	sn005_x = 1	-8 = no social network; no spouse
siblingnet siblingnet2 siblingnet3	Sibling in social network	sn005_x = 8, 9	-8 = no social network; no living siblings
parentnet* parentnet2* parentnet3*	Parent in social network	sn005_x = 2, 3	-8 = no social network; no living parents
friendnet friendnet2 friendnet3	Friends in social network	sn005_x = 21	-8 = no social network
formalnet formalnet2 formalnet3	Formal helpers in social network	sn005_x = 25 - 27	-8 = no social network
othernet othernet2 othernet3	Others in social network	sn005_x = 22-24, 96	-8 = no social network
womennet womennet2 womennet3	Women in social network	sn005a_x = 2	-8 = no social network
mennet mennet2 mennet3	Men in social network	sn005a_x = 1	-8 = no social network

*Survey limitations did not allow for identification of living status of parents for all survey respondents. These cases are coded as missing for these derived variables.

Social network size

This variable is derived from variables sn005_1 – sn005_7 and is a count of these variables which identify a social network member's relationship to the respondent.

Social network satisfaction

In the raw data, satisfaction with network was divided into two questions distinguishing respondents with (sn012_) or without (sn017_) cited social network members. The derived network satisfaction variable combines the data from these two variables into one overall measure of satisfaction with the state of one's interpersonal network.

Social network relationship composition

Generated variables identifying the relationship composition of the social network are derived from variables sn005_1 thru sn005_7. Three generated variables were created for each relationship composition type.

xxxnet – This variable is a count of the total number of social network members for each relationship category

xxxnet2 – This variable dichotomizes the count variable into two categories: (1) the relationship category is present in the social network & (0) the relationship category is not present in the social network.

xxxnet3 – This variable provides the percentage of the total social network comprised of members of the designated relationship category.

3.4 Geographic proximity of social network members

A series of variables, listed in Table 3.2.2, summarizes characteristics of the geographic proximity between survey respondents and social network members. These variables are derived from variables sn006_1 thru sn006_7.

Table 3.2.2 Derived SN variables: geographic proximity

Variable name	Variable description	Generated variable coding description	-8 "Does not apply" description
prx_mean	SN proximity – Average		-8 = no social network
most_prx	proximity of closest SN member		-8 = no social network
prx_5km	SN members within 5 km – count	sn006_x = 1 - 4	-8 = no social network
prx_5km3	SN members within 5 km - % of SN		
prx_1km prx_1km3	SN members within 1 km – count SN members within 1 km - % of SN	sn006_x = 1 - 3	-8 = no social network

3.5 Frequency of contact with social network members

A series of variables, listed in Table 3.2.3., identify the frequency of contact between survey respondents and members of their social network. Variables sn007_1 – sn007_7 are used to construct the derived variables. Frequency of contact was not asked about social network members who lived with the respondent. For these social network members, frequency of contact was coded as (1) daily contact for the derivation of these variables. Variables that identify the mean frequency of contact by a specific relationship type of social network members were also derived.

Table 3.2.3A Derived SN variables: frequency of contact[2]

Variable name	Variable description	Generated variable doding description	-8 "Does not apply" description
contact_mean	SN contact - average		-8 = no social network
most_contact	Contact with most contacted SN member		-8 = no social network
daily_contact	SN members with daily contact – count	sn007_x = 1	-8 = no social network
daily_contact3	SN members with daily contact - % of SN		
week_contact	SN members with weekly or more contact – count	sn007_x = 1 - 3	-8 = no social network
week_contact3	SN members with weekly or more contact - % of SN		
month_contact	SN members with monthly or less contact – count	sn007_x= 5-7	-8 = no social network
month_contact3	SN members with monthly or less contact - % of SN		
Average contact with social network by relationship type			
fam_contact	Average contact with family members in SN	Average of sn007_x when sn005_x = 1-20	-8 = no social network; no family members in SN
child_contact	Average contact with children in SN	Average of sn007_x if sn005_x = 10, 11	-8 = no social network; no children; no children in SN
gchild_contact	Average contact with grandchildren in SN	Average of sn007_x if sn005_x = 14	-8 = no social network; no grandchildren; no grandchildren in SN
spouse_contact	Average contact with spouse in SN	Average of sn007_x if sn005_x = 1	-8 = no social network; no spouse; no spouse in SN

[2] if missing sn007_x and sn006_x = 1 then sn007_x recoded as 1

Table 3.2.3B Derived SN variables: frequency of contact

sibling_contact	Average contact with sibling in SN	Average of sn007_x if sn005_x = 8-9	-8 = no social network; no living siblings; no sibling in SN
parent_contact	Average contact with parent in SN	Average of sn007_x if sn005_x = 2-3	-8 = no social network; no living parents; no parent in SN
friend_contact	Average contact with friends in SN	Average of sn007_x if sn005_x = 21	-8 = no social network; no friends in SN
formal_contact	Average contact with formal helpers in SN	Average of sn007_x if sn005_x = 25-27	-8 = no social network; no formal helpers in SN
other_contact	Average contact with others in SN	Average of sn007_x if sn005_x = 22-24, 96	-8 = no social network; no others in SN

3.6 Emotional closeness of social network members

A series of derived variables, listed in Table 3.2.4, identifies several characteristics about how emotionally close survey respondents feel towards members of their social network. The derived variables were calculated using the variables sn009_1 – sn009_7.

Table 3.2.4 Derived SN variables: emotional closeness

Variable name	Variable description	Generated variable coding description	-8 “Does not apply” description
close_mean	SN emotional closeness – average		-8 = no social network
most_close	Emotional closeness of closest SN member		-8 = no social network
very_close	Very to extremely close - count	sn009_x = 3 or 4	-8 = no social network
very_close3	Very to extremely close - % of SN		
not_close	Somewhat close or less – count	sn009_x = 1 or 2	-8 = no social network
not_close3	Somewhat close or less - % of SN		

3.7 Financial transfers with social network members

The list of social network members gathered in the Social Network (SN) module was linked with the Financial Transfer (FT) module in SHARE wave four. For each exchange variable in the FT module that identifies to whom or from whom financial transfers were exchanged, the first seven answer categories are reserved for the social network roster of names. As previously stated, social network members identified in the FT exchange variables can be correctly identified in the SN module via the variable value indicating the member's numbered placement in the SN roster listing. The FT module is collected from the identified financial respondent of the household. Consequently, all FT information pertaining to social network members is applicable only to the financial respondent because of the individual nature of the social network roster. It is recommended that all research utilizing the financial transfer and social network linkage be performed at the individual level of analysis, relating only to the financial respondent.

Receipt/provision of financial help

A series of variables was derived to identify if survey respondents gave or received financial help of the equivalent of 250 Euros or more from a social network member. These variables were derived using the variables ft002_, ft003_1 – ft003_3; ft009_ and ft010_1 – ft010_3 and are listed in Table 3.2.5.A.

Table 3.2.5.A Derived SN variables: transfers of financial help (250 Euros or more)

Variable name	Variable description	Generated variable coding description	-8 "Does not apply" description
fin_gave	Gave financial help – count		-8 = non-financial respondent
fin_gave2	Gave financial help – dummy		
snfin_gave	Gave financial help to SN members – count	ft003_1 – ft003_3 = 1 - 7	-8 = non-financial respondent; no social network; fin_gave=0
snfin_gave2	Gave financial help to SN members - dummy		
fin_received	Received financial help – count		-8 = non-financial respondent
fin_received2	Received financial help – dummy		
snfin_received	Received financial help from SN members – count	ft010_1 – ft010_3 = 1 -7	-8 = non-financial respondent; no social network; fin_received=0
snfin_received2	Received financial help from SN members – dummy		

Receipt/provision of financial gift

A series of variables was derived to identify if survey respondents gave or received financial gifts of the equivalent of 5000 Euros or more to or from a social network member. These variables were derived using the variables ft015_, ft017_1 – ft017_5; ft025_ and ft027_1 – ft027_5 and are listed in Table 3.2.5.B.

Table 3.2.5.B Derived SN variables: transfers of financial gifts (5.000 Euros or more)

Variable name	Variable description	Generated variable coding description	-8 "Does not apply" description
gift_gave gift_gave2	Gave financial gift – count Gave financial gift – dummy		-8 = non-financial respondent
sngift_gave sngift_gave2	Gave financial gift to SN members – count Gave financial gift to SN members - dummy	ft017_1 – ft017_5 = 1 -7	-8 = non-financial respondent; no social network; fin_gave=0
gift_received gift_received2	Received financial gift – count Received financial gift – dummy		-8 = non-financial respondent
sngift_received sngift_received2	Received financial gift from SN members – count Received financial gift from SN members – dummy	ft027_1 – ft027_5 = 1 -7	-8 = non-financial respondent; no social network; fin_received=0

3.8 Social support exchanges with social network members

The list of social network members gathered in the Social Network (SN) module was also linked with the Social Support (SP) module in SHARE wave four. Here too, the first seven answer categories are reserved in each support exchange variable for persons listed in the social network module. Thus, as was the case in the FT module, social network members identified in the SP exchange variables can also be correctly identified in the SN module via the variable value indicating the member's numbered placement in the SN roster listing.

Received or gave personal care or practical help from outside household

A series of variables was derived to identify if survey respondents provided or received personal care or practical help to or from social network members living outside the survey respondent's household. The derived variables, listed in Table 3.2.6.A were calculated using information from sp002_, sp003_1 – sp003_3, sp008_, sp009_1 – sp009_3. These data were only collected from family respondents. Because the social network roster is individually generated, it is advised that social network research utilizing these variables be performed at the individual level of analysis, relating only to the family respondent.

Table 3.2.6.A Derived SN variables: social support exchanges outside the household

Variable name	Variable description	Generated variable coding description	-8 "Does not apply" description
outhh_receive_care	Received help outside HH – count		-8 = non-family respondent
outhh_receive_care2	Received help outside HH – dummy		
outhh_snreceive_care	Received help from SN members outside HH– count	sp003_1 – sp003_3 = 1-7	-8 = non-family respondent; no social network; outhh_receive_care =0
outhh_snreceive_care2	Received help from SN members outside HH - dummy		
outhh_gave_care	Gave help outside HH – count		-8 = non-family respondent
outhh_gave_care2	Gave help outside HH - dummy		
outhh_sngave_care	Gave help to SN members outside HH – count	sp009_1 – sp009_3 = 1-7	-8 = non-family respondent; no social network; outhh_gave_care = 0
outhh_sngave_care2	Gave help to SN members outside HH - dummy		

Received personal care from inside household

A series of variables was derived to identify if survey respondents received personal care from a person in their social network living in their household. The derived variables, listed in Table 3.2.6.B were calculated using information from sp020_, sp021_1 – sp021_35. The questions were asked of survey respondents living in households greater than one person who reported having had difficulty with one or more physical functions due to health problems (ph048).

Table 3.2.6.B Derived SN variables: social support received within the household

Variable name	Variable description	Generated variable coding description	-8 "Does not apply" description
hh_receive_care	Received care inside HH – count		-8 = household size = 1; ph048_ = 96
hh_receive_care2	Received care inside HH – dummy		
hh_snreceive_care	Received care from SN member inside HH– count	sp021_1 – sp021_35 = 1-7	-8 = household size = 1; ph048_ = 96; no social network; hh_receive_care = 0
hh_snreceive_care2	Received care from SN member inside HH – dummy		

Provided personal care within household

A series of variables was derived to identify if survey respondents provided personal care to a social network member living in their household. The derived variables, listed in Table 3.2.6.C were calculated using information from sp018_, sp019_1 – sp019_35. Only survey respondents living in a household of more than one person were asked these questions.

Table 3.2.6.C Derived SN variables: social support provided within the household

Variable name	Variable description	Generated variable coding description	-8 "Does not appl description
hh_gave_care	Provided care inside HH – count		-8 = household si = 1
hh_gave_care2	Provided care inside HH – dummy		
hh_sngave_care	Provided care to SN member inside HH– count	sp019_1 – sp019_35 = 1-7	-8 = household si = 1; no social network; hh_gave_care =
hh_sngave_care2	Provided care to SN member inside HH – dummy		

3.9 Network types

Although there is evidence showing that different network variables may be variously related to a range of antecedents and outcomes, there is a growing body of research that suggests that a social network may be more than just "the sum of its parts." That is to say, social networks may be best represented by unique combinations of individual network indicators. Wenger's (1991) groundbreaking work in this domain has drawn attention to the concept of network type. This analytic construct allows for the identification of key personal social network configurations, as measured by the constellation of selected variables. The notion of network type is represented in a series of unique characterizations of sets of social ties, often referred to as a network typology.

Network types may be derived through several analytic procedures for data reduction. One such recommended procedure is K-means cluster analysis in which designated criterion variables are employed to identify relatively homogeneous groupings in a population of interest. The process uses an algorithm that can handle a large number of cases, a characteristic particularly suitable for large scale surveys such as SHARE. In the K-means cluster procedure, initial cluster centers are assigned for each of a number of selected criterion variables and are then iteratively updated until optimal groupings are achieved based upon Euclidean distance (See Milligan & Cooper, 1987 and Rapkin & Luke, 1993 for additional recommendations for running K-means cluster analyses).

It should be noted that this statistical procedure is essentially an exploratory one, insofar as the researcher selects in advance the number of clusters to be derived in each trial. In analysis of the network types of older people, cluster combinations of four, five, or six groupings are frequently tested, as is reflected in the number of cluster solutions obtained in studies in three different countries (Litwin & Landau, 2000; Melkas & Jylhä, 1996; Stone & Rosenthal, 1996). Three guiding principles must be taken into account in any network clustering procedure. First, the criterion variables employed must reflect the specific aims of the researcher. That is, different analyses may employ different sets of criterion variables. Second, the criterion variables employed in the clustering procedure should be measured on similar scales, or should be otherwise standardized before the clustering process takes place. The third guiding principle is that the ultimate preferred solution is the choice of the analyst (backed up by prior evidence, if it exists). The researcher must identify the best cluster solution, that is, the number of clusters that best reflects the field of inquiry, to be employed in the analysis. The best solution is based upon the distinctiveness of one cluster from another, the parsimony of the overall cluster set, the theoretical relevance of the derived groupings and the degree to which the solution is grounded in established knowledge.

For purpose of illustration of the derivation of network types from the SHARE wave four data, we present a network typology based upon the compositional characteristics of the personal social networks that were generated by the SN module. The typology was derived from an early internal release of the SHARE data that included some 43,000 respondents. The criterion variables, all structural/compositional in nature, were the respective proportions of the networks comprised by spouse (spousenet3), children (childnet3), grandchildren (gchildnet3), siblings (siblingnet3), parents (parentnet3), friends (friendnet3) and formal professional helpers (formalnet3). Other ties, e.g. neighbors, colleagues and other non-specified relationships were not addressed in this particular exercise. The results appear in Figure 3.3.1.

	1 Family	2 Helpers	3 Spouse	4 Diverse	5 Friends	6 Parent	7 Grandkids	8 Siblings	M/ *Total*
% Spouse	23.17	7.39	99.92	20.84	4.86	23.92	12.12	14.07	33.07
% Children	67.98	8.63	0.04	18.03	3.68	11.64	26.70	8.25	29.87
% Grandkids	0.50	0.42	0.01	0.40	0.16	0.10	42.63	0.34	1.37
% Siblings	1.55	3.80	0.00	6.44	2.40	3.69	2.98	59.50	7.44
% Parents	0.54	0.96	0.00	1.33	0.68	48.72	0.38	1.38	2.12
% Friends	1.85	4.90	0.01	24.11	81.56	5.21	3.78	6.63	15.46
% Helpers	0.12	63.31	0.00	0.43	0.22	0.25	0.39	0.16	0.61
N	14,788	279	7,467	9,854	4,400	1,248	1,051	3,552	*42,639*
%	34.7	0.7	17.5	23.1	10.3	2.9	2.5	8.3	*100.00*
SIZE	2.7	2.2	1.0	3.4	2.9	2.7	3.4	2.7	2.6
AGE	67.7	71.3	65.0	64.5	64.8	56.7	73.5	64.6	65.8
WOMEN	59.8	57.7	34.0	61.5	62.5	53.4	72.3	62.5	56.3

Figure 3.3.1 Network types based upon compositional criteria

As may be seen in the top part of the figure, some two thirds of the networks in the first cluster were children, and about a quarter was comprised by a spouse. Accordingly, this network type was named the family network. It was the most prevalent network type in the sample, accounting for more than a third of the respondents. The second cluster was the least frequent in the sample. Its unique identifying characteristic was the strong representation of formal helpers in this network grouping, accounting for almost two thirds of the network. It was termed the helpers network. The third cluster included mainly respondents who named their spouse as the sole network member. Consequently, the network size for respondents in this cluster was only one and the smallest network size across all the clusters. It accounted for the network of a bit less than a fifth of the sample.

The fourth cluster was the second most prevalent network type. It was termed the diverse network because it was comprised by similar proportions of spouse, children and friends. In comparison, the fifth cluster was comprised almost entirely by friends, and was named accordingly. This network type represented about a tenth of the sample. The remaining network types were relatively less frequent in the sample, and were named according to the majority category present in the cluster. They were termed parent, grandchildren and sibling network types, respectively.

The figure also shows that there was variation in the age and gender of respondents in the different network types. Women were much less likely to be embedded in spouse networks, and were more frequently located in grandchildren type networks. Respondents in parent network types were the youngest in the sample, on average, while those in grandchildren and helper networks were the oldest. This example illustrates how the construct of network type can distinguish between various aspects of research interest. We should also note that social network types also vary by national setting. Figure 3.3.2 shows the distribution of the network types in each of the SHARE countries. As may be seen, almost all the network types appeared in all the SHARE countries (except for the helper network). However, their relative distribution differed somewhat across the respective countries.

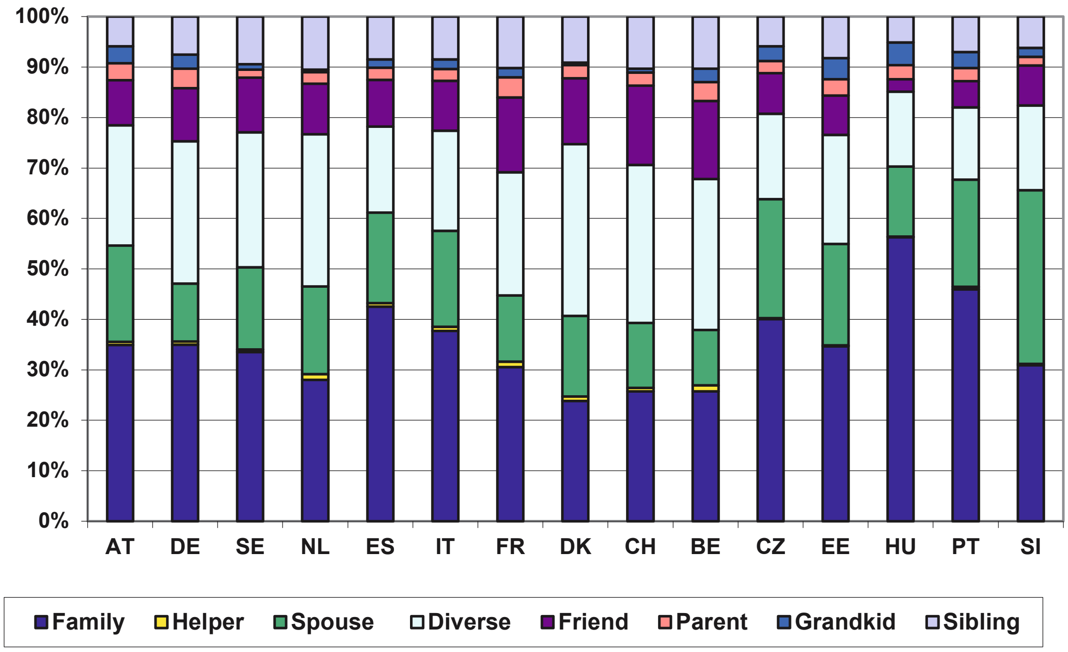


Figure 3.3.2 Network types based upon compositional criteria by country

3.10. Loneliness – the absence of social network

The notion of social network can also be considered in its absence, that is, the extent to which people perceive themselves as being lonely. Loneliness can be said to constitute the opposite of a state of connectedness. That is, people who are lonely lack the social capital provided by social networks. Loneliness is considered a discrete and subjective construct, in contrast to social isolation. Whereas social isolation reflects an objective social situation characterized by lack of relationships with others (Dykstra, 2009), loneliness is a marker for the quality of a person's social interactions. As such, loneliness develops when one's social relationships are not accompanied by the desired degree of intimacy (de Jong Gierveld, 1998).

"Loneliness appears to be present everywhere researchers look for it" (Perlman, 2004, p.185). Nevertheless, it has been argued that cultural factors, values, and norms substantially influence the experience of loneliness (Johnson & Mullins, 1987; Jylhä & Jokela, 1990). Hence, there is a need for cross-cultural data to verify the culture-bound assertion. Because the SHARE project includes countries from different European regions which are characterized by different cultural norms (e.g. collectivist vs. individualist societies), it provides comprehensive data for cross-cultural comparisons of loneliness. Notably, a North-South contrast was found by Sundström, Fransson, Malmberg and Davey (2009) based on SHARE wave 1 data, with Mediterranean countries reporting higher rates of loneliness. Given the importance of the phenomenon, the measurement of loneliness has been expanded in SHARE Wave four. Across the various waves of SHARE data collection, several adjustments were made in the measure of loneliness. In the first and second wave, it was measured solely by a direct single item. In the fourth wave a 3-item indirect scale was added to the single direct probe.

The direct item used in waves 1 and 2 was taken from the CES-depression scale. In the 6th CES-D self-labeling item, participants were asked how often they had experienced loneliness during the last week, on a 4-point ordinal scale ranging from 1 (almost all the time) to 4 (almost never). In the second wave the wording of the direct loneliness item was slightly changed and the response option was reduced to 2 (yes/no). The use of a direct single item is known as a direct measure because it includes the specific words "lonely" or "loneliness" in the question. The direct approach is easy to apply in social surveys; it has strong face validity, and has been adopted extensively in past as well as in contemporary research (Routasalo, Savikko, Tilvis, Strandberg, & Pitkälä, 2006; Savikko, Routasalo, Tilvis, Strandberg, & Pitkälä, 2005; Shiovitz-Ezra & Ayalon, 2010; Sundström et al., 2009; Thurston & Kubzansky, 2009).

One of the shortcomings of the direct measure relates to respondents' unwillingness to admit that he or she feels lonely. Previous experimental studies have demonstrated that lonely people are socially stigmatized and that they are perceived in a much more negative light than their counterparts who are not lonely (Lau & Gruen, 1992; Rotenberg, 1998; Rotenberg & Kmill, 1992). Therefore, the social stigma of loneliness (Crocker & Major, 1989) may discourage people from characterizing themselves as lonely (Victor, Scambler, Bond, & Bowling, 2000; Victor, Scambler, Bowling & Bond, 2005), and can result in an underestimation of the phenomenon. Moreover, there is some question about the reliability of single-item measures (Marangoni & Ickes, 1989).

To counteract these shortcomings of direct measurement of loneliness, SHARE wave four was expanded to also include indirect measurements of loneliness. The indirect approach uses multiple-item scales that do not explicitly refer to the term loneliness. The Revised UCLA Loneliness Scale (R-UCLA; Russell, 1996; Russell, Peplau, & Cutrona, 1980), which represents this approach, is one of the most widely used scales (see for example, Cacioppo, Hawkley & Thisted, 2010; Hawkley, Masi, Berry & Cacioppo, 2006; Hawkley, Thisted, & Cacioppo, 2009; Steptoe, Owen, Kunz-Ebrecht, & Brydon, 2004; and Vander Weele, Hawkley, Thisted & Cacioppo, 2011). The original form of the scale comprises 20 self-report items on which participants are asked to rate how often they experience certain feelings that implicitly capture loneliness (e.g., "How often do you feel lack of companionship?"). Responses are based on a 4-point, Likert-type scale ranging from 1 (never) to 4 (often). The R-UCLA has been tested in several samples and is proven to have good psychometric properties.

Because comprehensive social surveys such as SHARE aim to reduce interview time and minimize response burden, shortened versions of relatively long scales are needed. The shortest R-UCLA scale developed to date is made up of three items answered on a 3-point scale ranging from "often" to "hardly ever or never". The psychometric properties of the shortened version are adequate in terms of reliability and validity (Hughes, Waite, Hawkley & Cacioppo, 2004). In SHARE wave four, the 3-item R-UCLA indirect measurement scale is used in addition to one direct measure of loneliness. The full scale used in SHARE wave four is presented in Figure 3.4.1. Other leading social surveys, such as the HRS, NSHAP (in the US), and ELSA (in England), have used this shortened indirect scale to measure loneliness, thus offering the additional advantage of harmonization.

Please say how much you agree or disagree with each of the following statements. How much of the time do you...

	Often ▼	Some of the time ▼	Hardly ever or never ▼
a) ... feel you lack companionship?	☐ 1	☐ 2	☐ 3
b) ... feel left out?	☐ 1	☐ 2	☐ 3
c) ... feel isolated from others?	☐ 1	☐ 2	☐ 3
d) ... feel lonely?	☐ 1	☐ 2	☐ 3

Figure 3.4.1 Loneliness survey items in SHARE wave four

A final point of note is the adjustment of the method used to administer the loneliness items in SHARE wave four. Whereas the single direct loneliness item was incorporated in a self-administered questionnaire in wave one, it was included in the CAPI questionnaire in the second wave. In wave four, all the loneliness measures were asked in a leave-behind questionnaire. Insofar as loneliness is a private, sensitive and stigmatizing concept, lonely people might avoid admitting having experienced it to the interviewer. Self-administered data collection is considered to be more confidential, and can be expected to elicit responses that are more reliable (de Leeuw, 1992).

References

Burt, R. S. (1986). *A note on sociometric order in the general social survey network data*. Social Networks, 8, 149-174.

Burt, R. S. ,Guilarte, M. G. (1986). *A note on scaling the general social survey network item response categories*. Social Networks, 8, 387-396.

Cacioppo, J. T., Hawkley, L. C., Thisted, R. A. (2010). *Perceived social isolation makes me sad: 5-year cross-lagged analyses of loneliness and depressive symptomatology in the Chicago Health, Aging, and Social Relations Study*. Psychology and Aging, 25(2), 453-63.

Cornwell, B., Schumm, L. P., Laumann, E. O., Graber, J. (2009). *Social networks in the NSHAP Study: Rationale, measurement, and preliminary findings*. Journals of Gerontology Series B-Psychological Sciences and Social Sciences, 64, I47-I55.

Crocker, J., Major, B. (1989). *Social stigma and self-esteem: The self –protective properties of stigma*. Psychological Review, 96, 608-30.

de Jong Gierveld, J. (1998). *A review of loneliness: Concept and definitions, determinants and consequences*. Review in Clinical Gerontology 8, 73–80.

de Leeuw, E. (1992). *Data quality in mail, telephone, and face to face surveys*. Amsterdam, The Netherlands: T.T Publikaties.

Dykstra, P.A. (2009). *Older adults loneliness: Myths and realities*. European Journal of Ageing, 6(2, :91-100.

Hawkley, L. C., Masi, C. M., Berry, J. D., Cacioppo, J. T. (2006). *Loneliness is a unique predictor of age-related differences in systolic blood pressure*. Psychology and Aging, 21(1), 152-164.

Hawkley, L. C., Thisted, R. A., Cacioppo, J. T. (2009). *Loneliness predicts reduced physical activity: Cross-sectional & longitudinal analyses*. Health Psychology, 28(3), 354-363.

Hughes, M. E., Waite, L. J., Hawkley, L. C. , Cacioppo, J. T. (2004). *A short scale for measuring loneliness in large surveys: Results from two population-based studies*. Research on Aging, 26(6), 655-672.

Johnson, D. P., Mullins, L. C. (1987). *Growing old and lonely in different societies: Toward a comparative perspective*. Journal of Cross Cultural Gerontology, 2, 257-275.

Jylhä, M., Jokela, J. (1990). *Individual experiences as cultural: A cross-cultural study on loneliness among elderly*. Ageing and Society, 10, 295-315.

Lau, S., Gruen, G. E. (1992). *The social stigma of loneliness: Effects of target person's and perceiver's sex*. Personality and Social Psychology Bulletin, 18(2), 182-189.

Litwin, H. (2011). *The association between social network relationships and depressive symptoms among older Americans: What matters most?* International Psychogeriatrics, 23, 930-940.

Litwin, H. (2010). *Social networks and well-being: A comparison of older people in Mediterranean and non-Mediterranean countries.* Journal of Gerontology: Social Sciences, 65B, S599-S608.

Litwin, H. (1996). Introduction. In H. Litwin (Ed.), *The social networks of older people: A cross-national analysis*. (pp. 1-13). Westport, CT: Praeger.

Litwin, H., Landau, R. (2000). *Social network type and social support among the old-old*. Journal of Aging Studies, 14, 213-228.

Marangoni, C., Ickes, W. (1989). *Loneliness: A theoretical review with implications for measurement*. Journal of Social and Personal Relationships, 6, 93–128.

Melkas, T., Jylhä, M. (1996). Social network characteristics and social network types among elderly people in Finland. In H. Litwin (Ed.), *The social networks of older people: A cross-national anlaysis* (pp. 99-116). Westport, CT: Praeger.

Milligan, G. W., Cooper, M. C. (1987). *Methodology review: Clustering methods*. Applied Psychological Measurement, 11, 329-354.

Perlman, D. (2004). *European and Canadian studies of loneliness among seniors*. Canadian Journal of Aging, 23, 181-188.

Pescosolido, B. (2011). *Social connectedness in health, morbidity and mortality, and health care – the contributions, limits and further potential of Health and Retirement Study*. Forum for Health Economics & Policy, 14(3), Article 11.

Rapkin, B. , Luke, D. (1993). *Cluster analysis in community research: Epistemology and practice*. American Journal of Community Psychology, 21, 247-277.

Rotenberg, K. J. (1998). *Stigmatization of transitions in loneliness*. Journal of Social and Personal Relationships, 15(4), 565-576.

Rotenberg, K. J., Kmill, J. (1992). *Perception of lonely and non-lonely persons as a function of individual differences in loneliness*. Journal of Social and Personal Relationships, 9, 325-330.

Routasalo, P. E., Savikko, N., Tilvis, R. S., Strandberg, T. E. , Pitkälä, K. H. (2006). *Social contacts and their relationship to loneliness among aged people – A population-based study*. Gerontology, 52, 181-187.

Russell, D. W. (1996). *UCLA Loneliness Scale (version 3): Reliability, validity, and factor structure.* Journal of Personality Assessment, 66(1), 20-40.

Russell, D. W., Peplau, L. A., Cutrona, C. E. (1980). *The revised UCLA Loneliness Scale: Concurrent and discriminant validity evidence*. Journal of Personality and Social Psychology, 39(3), 472-480.

Savikko, N., Routasalo, P., Tilvis, R. S., Strandberg, T. E., Pitkälä, K. H. (2005). *Predictors and subjective causes of loneliness in an aged population*. Archives of Gerontology and Geriatric, 41(3), 223-233.

Shiovitz-Ezra, S., Ayalon, L. (2010*). Situational versus chronic loneliness as risk factors for all-cause mortality*. International Psychogeriatrics, 22(3), 455-462.

Steptoe, A., Owen, N., Kunz-Ebrecht, S. R., Brydon, L. (2004). *Loneliness and neuroendocrine, cardiovascular, and inflammatory stress responses in middle-aged men and women*. Psychoneuroendocrinology, 29(5). 593-611.

Stone, L., Rosenthal, C. (1996). Profiles of the social networks of Canada's elderly: An analysis of 1990 General Social Survey data. In H. Litwin (Ed.), *The social networks of older people: A cross-national analysis*. (pp. 77-97). Westport, CT: Praeger.

Sundström, G., Fransson, E., Malmberg, B., Davey, A. (2009). *Loneliness among older Europeans*. European Journal of Ageing, 6(4), 267-275.

Thurston, R. C., Kubzansky, L. D. (2009). *Women, loneliness, and incident coronary heart condition*. Psychosomatic Medicine, 71(8), 836-842.

Vander Weele, T. J., Hawkley, L. C., Thisted, R. A., Cacioppo, J. T. (2011). *Marginal structural model analysis for loneliness: Implications for intervention trials and clinical practice*. Journal of Consulting and Clinical Psychology, 79(2), 225-235.

van Tilburg, T. G. (1995). Delineation of the social network and differences in network size. In C. P. M. Knipscheer, J. de Jong Gierveld, T. G. van Tilburg, & P. A. Dykstra (Eds.), *Living Arrangements and Social Networks of Older Adults* (pp. 83-96). Amsterdam: VU University Press.

Victor, C. R., Scambler, S. J., Bond J., Bowling, A. (2000). *Being alone in later life: Loneliness, social isolation and living alone*. Review in Clinical Gerontology, 10, 407-417.

Victor, C. R., Scambler, S. J., Bowling, A., Bond, J. (2005*). The prevalence of, and risk factors for, loneliness in later life: A survey of older people in Great Britain*. Ageing & Society 25, 357-375.

Wenger, G. C. (1991). *A network typology: From theory to practice*. Journal of Aging Studies, 5, 147-162.

4 Collection of Biomarkers in the Survey of Health, Ageing and Retirement in Europe (SHARE)

Barbara Schaan, GESIS

4.1 Closing the gap – a new perspective of empirical research on aging in Germany

While research on aging has made several remarkable achievements in many areas there is still need for a better integration of previous results and future research projects across disciplines, which takes into account the complex interdependences between biological-medical and socio-economic factors in the aging process, both on an individual and the societal level.

This research gap was the starting point for a German pilot project within SHARE, which was funded by the Volkswagen Foundation and will be expanded to more participating countries in upcoming waves.

This pilot project exploited the existing data infrastructure of the Survey of Health, Ageing and Retirement in Europe in order to create a comprehensive database, consisting of:

a) Information on the current living conditions in waves one, two and four of SHARE
b) Retrospective life-histories in wave three (SHARELIFE)
c) Administrative process data from the German Pension Fund (for a detailed description see chapter 5)
d) Objectively measured biomarkers (dried blood spots, blood pressure, height, and waist circumference)

This pilot project generated a basis for the integration of interdisciplinary research on aging, the separation between medical-biological and socio-gerontological research on aging is an impediment for the development of measures that might help improving the quality of life of older individuals. This pilot project focused on intervention points (e.g. retirement) in the lifecycle. As intervention points we understand both medical and socio-economic interventions, which can – sometimes with a substantial time lag – affect morbidity, mortality, and quality of life of persons at age 50 years and older. A ´social management` of aging processes must address such intervention points; and in most cases, the effects of such interventions depend on interrelated health and socio-economic factors.

While many correlations are known, it is not well understood which causal mechanisms drive the interactions between interventions and environment; as little understood is which interventions are the most effective in improving the quality of life. A better understanding of causal mechanisms – and this implies better recommendations with respect to interventions – can only be reached by combining biomedical insights with knowledge about the socioeconomic environment of individuals. So far, examples for fruitful combinations of biomedical and socio-economic research came mostly from the United States. Two institutional examples include the Aging Center of the RAND

Corporation in Santa Monica, California, and the very successful Schools of Public Health, for example at Harvard or at Johns Hopkins University.

4.2 Biomarkers in SHARE

One recent development is the inclusion of physical measurements and biomarkers in social surveys. So far these measurements were often taken in smaller, non-representative clinical studies. In the last couple of years more and more large-scale surveys added physical measurements and biomarkers to their program since there is promising scientific value to it:

a) Measurement of respondents‘ health can be improved: Standard health questions in surveys are often subject to the respondents' own interpretation (of the question), own evaluation or perception (of health status), and own knowledge (of health status). The value of subjective health measurements is undeniable, but some research questions require objective measurements. Biomarkers enable researchers to validate respondents' self-reports and therefore to study the amount and determinants of under-, over-, and misreporting in large-scale population surveys.

b) Identification of causal relationships: biomarkers can help to understand the complex relationships between social status and health and their physiological pathways.

c) Pre-clinical information: Biomarkers allow to identify pre-disease pathways, since the physiological processes are often below the individual‘s threshold of perception.

From the first wave on, SHARE combined self-reports on health with two physical performance measurements: grip strength and walking speed. Additionally, respondents reported their height and weight. In wave two, SHARE added peak-flow (measuring lung strength) and chair stand (measuring lower body performance) to the questionnaire. In wave three (SHARELIFE) grip strength was the only physical measurement included in the questionnaire. Wave four included grip strength, peak-flow and self-reported height and weight.

Wave four also added new biomarkers to the German part of the study. Germany served as a pilot country for the inclusion of innovative biomarkers for a full-scale implementation in SHARE. The new measures included were 1) height (in addition to self-reported height of the respondents), 2) waist circumference, 3) blood pressure, and 4) dried blood spots (DBS). An overview of the objective health measures in SHARE can be found in table 4.1.

Table 4.1 Overview of physical measurements and biomarkers in SHARE

	Wave 1 (2004/05)	Wave 2 (2006/07)	Wave 3 (2008/09)	Wave 4 (2010/11)
		Performance measures		
Grip strength	yes	yes	yes	yes
Lung strength (peak flow)	-	yes	-	yes
Walking speed	yes	yes	-	-
Chair stand	-	yes	-	-
		Biomarkers (Germany only)		
Height:				
- self-reported	-	-	-	yes*
- measured	-	-	-	yes
Waist circumference	-	-	-	yes
Blood pressure (seated)	-	-	-	yes
Dried Blood Spots - HbA1c	-	-	-	yes
Dried Blood Spots - Total cholesterol	-	-	-	yes
Dried Blood Spots - C-reactive protein	-	-	-	yes

*in all countries

The measurement of height allows validating the self-reported height and thus enhances the accuracy of the computed body mass index (BMI). Waist circumference – together with body height – allows the computation of the height-to-waist-ratio (HTWR), which is an indicator for the distribution of body fat. The measurement of blood pressure allows identifying respondents with high blood pressure. All three measurements - BMI, HTWR, and blood pressure - serve as indicators for the risk of developing cardiovascular diseases.

The major goal of the DBS was to collect information on total cholesterol, C-reactive protein (CRP, which is a marker for inflammations in the body), and HbA1c (which is a measure for blood sugar levels over the last 120 days). Both, CRP and total cholesterol, are associated with the development of cardio-vascular diseases, whereas HbA1c allows the identification at respondents who have diabetes – another prevalent disease in older age. The advantages of DBS – in contrast to whole blood - are that they are minimally invasive and can be carried out by trained field interviewers1 in the home environment of the respondents. They can be obtained by pricking a respondent´s finger with a lancet. The blood drops forming at the pricking site of the finger are collected on a filter card. The procedure is very similar to the everyday procedure with which diabetic persons control their blood sugar levels. Furthermore, the DBS require no special shipment treatment (although they are sensitive to high temperatures) and are easy to store once they arrive in the laboratory.

Whereas biomarker data have been collected in smaller studies within a clinical setting, SHARE was the first study in Germany which collected biomarkers nation-wide and without medical staff. Instead, the measurements in SHARE were carried out by specially trained field interviewers in the home environment of the respondents. In order to design a training program for interviewers, SHARE representatives attended the

[1] A special interviewer training is needed, but legal restricitions on which qualifications are necessary to conduct the DBS procedure vary between countries

interviewer trainings of the “Health and Retirement Study” (HRS) in the United States. HRS implemented the collection of biomarkers in its previous waves with great success.

4.3 Interviewer selection and training

In general, interviewers play a very important role – not only in gaining cooperation but also in the accuracy of the measured data (height, waist circumference, blood pressure) and the quality of the collected blood spots (number and size). Thus, careful interviewer selection prior to interviewer training was essential. Before they were invited to the training interviewers, were already informed that their tasks would include the collection of dried blood spots and the measurement of height, waist circumference, and blood pressure. As a consequence, only those interviewers who expressed no reluctance towards the collection of these biomarkers were subsequently invited and received the training.

The SHARE interviewer training closely followed the HRS model in preparing the interviewers for their tasks. Interviewers were trained in depth for any eventualities, such as respondents who take blood thinners or respondents who do not bleed enough. The training on collecting biomarkers took about seven hours. The majority of time was dedicated to supervised hands-on training sessions.

4.4 Results

Despite the detailed and standardized training and the careful selection prior to training, the monitoring process during fieldwork revealed very large differences between interviewers in the quality of performance. A small number of interviewers had to be re-trained or was taken off the job due to constant poor performance.

Collecting biomarkers was not only new and challenging for the survey agency and the interviewers, but also for the respondents. The participation rates among respondents were within the expected range, but with a large variation between interviewers. Whereas some interviewers had participation rates of more than 80 percent for the blood spots, there were also many interviewers who had very low participation rates. All in all, more than 80 percent of the respondents allowed us to measure their height, waist circumference and blood pressure. Dried blood spots could be collected from about 60 percent of the respondents. However, in some cases the blood spots collected were too small and could not be considered for analyses. At the time of writing (October 2012), only preliminary data were available. More detailed analyses will be conducted in future releases.

4.5 Process

The biomarkers were collected in the middle of the interview. First, the interviewers explained the purpose of the biomarkers and handed an information leaflet and a consent form to the respondents. Interviewers were trained to allow sufficient time for reading the form and to answer question. Upon completion of the consent form, interviewers were instructed to carry out the measurement. Without written consent the measures could not be conducted for legal restrictions. Next, interviewers explained the procedures and asserted that the respondent understood the instructions and that the measurement was safe for the respondent. Then the measurements started. The first measurement was height. A metal tape measure and a rafter square were used for the measurement of height. Respondents were asked to take off their shoes before the measurement. The measurement itself took place next to a wall. The rafter square was placed on the head of the respondents to enhance the accuracy of the measurement.

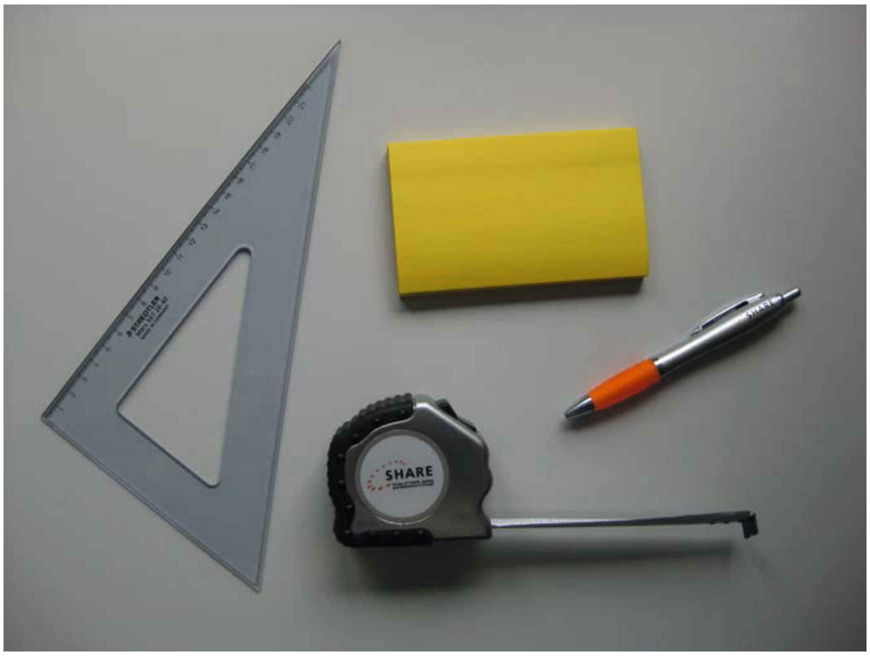


Figure 4.1 Materials for height measurement

After measuring height waist circumference was assessed. A soft tape measure was required for measuring waist circumference. First, respondents were asked to take off any bulky clothes, but it was not necessary to undress completely. Then they were asked to point to their navel. Interviewers were instructed to place the tape measure around the body at the height of the navel. Respondents were asked to breathe in, to breathe out and to hold their breath for a second. The measurement was taken while holding the breath. Respondents were allowed to measure their waist circumference themselves if they wished to do so. In these cases, the interviewer only explained the procedure to them and assisted if necessary.

After assessing waist circumference interviewers continued with measuring blood pressure. Blood pressure was measured three times in a row with about 1 minute pause between each measurement. The measurement was conducted by using an electronic blood pressure monitor. All interviewers were equipped with exactly the same type of monitor which they had to bring with them to the respondents homes. Respondents were asked to sit still and relaxed without talking. Furthermore, we instructed interviewers to assert that respondents did not sit with their legs crossed during the measurement or with a wrong arm position. Respondents who were interested in the results of their blood pressure measurements received the results only after the third measurement was finished, together with a table from the World Health Organization which classified the results into normal, prehypertension, and several stages of hypertension.

Finally, the blood spots were collected. The DBS required several different materials. We ordered pre-packed test kits, which contained two filter cards, two lancets, an alcohol wipe, a gauze pad and a bandage. Since no supplier for these test kits could be identified in Germany, we finally decided to purchase the test kit in the United

States. We further equipped interviewers with disposable rubber gloves, disinfection liquid, additional bandages and several other items like additional envelopes. Interviewers were instructed to explain the procedure of taking DBS in detail. If respondents wished to do so, they were allowed to prick their fingers themselves with the interviewer only assisting. After the blood spots were collected on the filter card, interviewers put them aside in order to let them dry and continued with the standard SHARE interview. After the end of the interview the blood spots were prepared for sending them to the laboratory. For that purpose, interviewers were equipped with pre-addressed and pre-stamped envelopes. Stickers with unique barcodes were put onto the filter cards. The barcode numbers were also entered into the CAPI questionnaire. This allowed a) to send the filter cards to the laboratory without revealing the SHARE identification number to the laboratory staff (for privacy law reasons) and b) to link the results coming back from the laboratory to the corresponding interview. The DBS were sent to the laboratory using standard mail service. Although DBS do not require any special shipping method (e.g. priority mailing), we instructed interviewers to send the DBS to the laboratory immediately after the interview. In case immediate shipping was not possible we advised interviewers to store the filter cards with the DBS at a dry and cool place until shipping was possible.

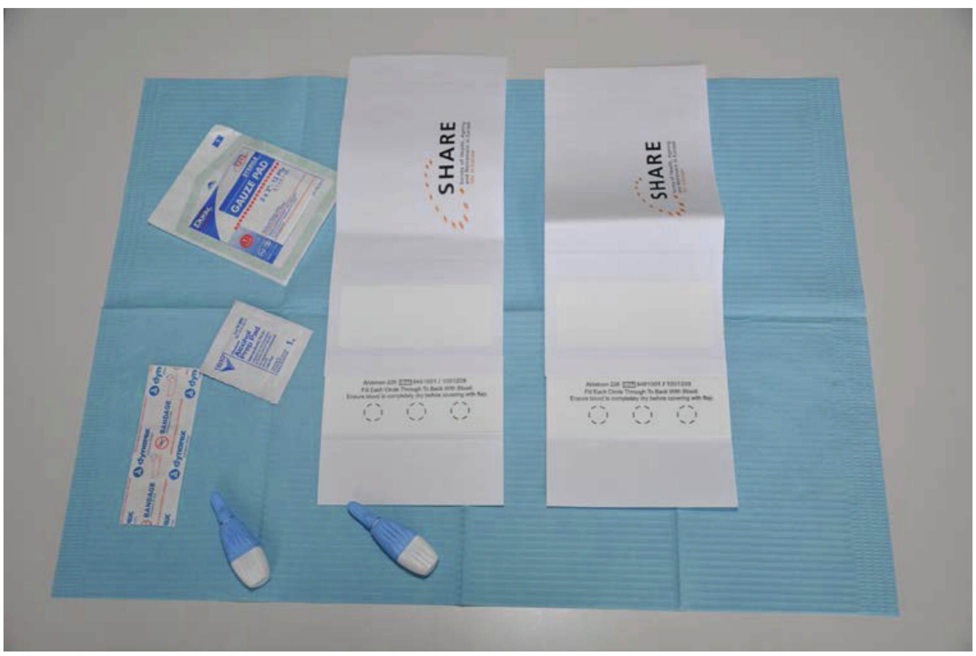


Figure 4.2 DBS test kit (filter cards, lancets, disinfection wipe, gauze pad, and plaster)

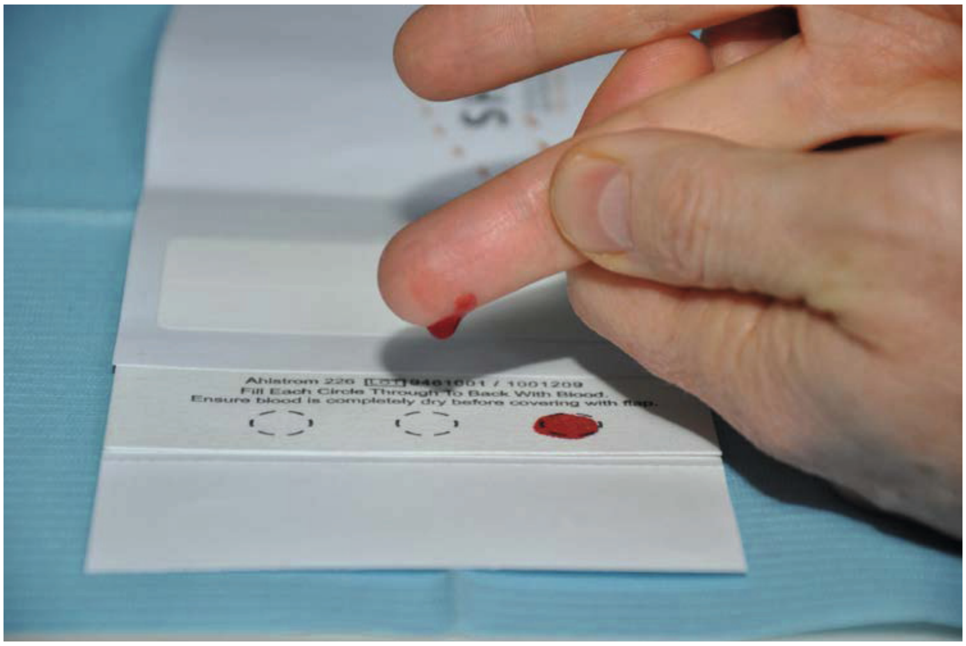


Figure 4.3 Collecting the blood on the filter card

4.6 Data sources

The biomarker part of the interview was not programmed as CAPI. All questions of the biomarker section were included in a paper-and-pencil booklet, which was filled by interviewers. All results of the measurements of height, waist circumference, and blood pressure were noted down in this booklet. After the interview, the interviewers sent the booklet to the survey agency. An external laboratory processed the filter cards with the dried blood spots and the sent us the results for HbA1c, cholesterol and CRP.

The three different data sources (booklet, lab data, and CAPI) were linked using a barcode system. We produced a set of unique barcode identifiers. Interviewers were equipped with barcode sticker sheets. Each sheet contained several stickers with the same unique barcode. Interviewers randomly chose a barcode sticker set for each respondent. The barcode stickers were put on a) the filter cards and b) the booklet. Furthermore, interviewers entered the barcode number into the SHARE CAPI. This allowed us to link the lab data to the booklet, finally link this merged dataset to the SHARE CAPI.

Special ethical considerations

The ethics review board in Germany advised to provide respondents with the results of the blood spot analyses via their general practitioner if requested. About 60 percent of the respondents who allowed us to collect their blood wished to have their general practitioner informed in case any of the parameters measured from their blood were out of range. This added further logistics to the project since respondents had to provide their general practitioners' addresses, and special letters for the general practitioners had to be designed. Furthermore, respondents had the possibility to narrow the range of possible analyses with their blood (e.g. exclusion of DNA analyses).

4.7 Challenges and opportunities

The pilot project described above took place in Germany only. As of November 2012, expanding the collection of biomarkers to more SHARE countries was planned. Before biomarkers could be implemented in other SHARE countries, we have to clarify the following questions:

1) What are the legal restrictions under which blood samples can be obtained from respondents (e.g. written consent) in each country willing to participate?
2) Are there any special requirements regarding the ethical review for the collection of biomarkers?
3) Are trained interviewers allowed to take blood samples if minimally invasive methods (e.g. collecting dried blood spots by pricking a finger) are used?
4) Can blood samples be sent across borders to be analyzed in the same laboratory?
5) Is access of researchers to biomarker data subject to specific conditions and restrictions (e.g. special user statements)?

As of October 2012, SHARE was still investigating the specific requirements and limitations of collecting biomarkers across Europe. As pointed out above, the scientific value of collecting biomarkers and linking them to representative survey data is undeniable, especially with regard to aging research. But is also offers great opportunities for methodological experiments and research. A potential cost of collecting biomarkers is the increase in respondent burden and it may reduce the willingness of respondents to cooperate in future waves. All German SHARE respondents who participated in previous waves were asked for their participation in collecting biomarkers. Since these were panel respondents we already had a great deal of information about them, which allowed us to study determinants of non-response. Collecting biomarkers is far beyond usual survey business – especially for interviewers. Not all interviewers feel comfortable doing it. Some interviewers might not be able to convince respondents to participate since they are not convinced themselves. Other interviewers might want to finish the interview quickly and thus might be too impatient to wait until a large blood drops forms (which results in blood spots that are too small to be analyzed). Other interviewers might be diabetic themselves or have a diabetic person in the family and are therefore very familiar with the procedure of pricking the finger with a lancet. Although interviewers are the key to success of a survey, often not much is known about interviewers' attitudes towards a given research project and their motives for working as an interviewer. Therefore, we asked all German SHARE interviewers to participate in a paper and pencil study. The aim of this study was to learn more about the interviewers themselves. In particular, we were interested in the work experience that they have, which strategies they apply in order to cope with initial refusals, and how they feel about the collection of biomarkers, to find out how this might affect their success in collecting the respective information and conducting the measurements. Data from that project were not available at the time of writing (November 2012).

Experiences from other countries with respect to biomarkers have been extremely positive (e.g., Crimmins & Seeman 2001; Weinstein & Willis 2001). Following the positive experience in Germany, SHARE planned using the pretest of wave 5 to test the logistics of a collection of dried blood spots in further SHARE countries. All these efforts will help with a full-scale implementation of biomarkers in a future wave.

References

Crimmins, E.M., Seeman, T. (2001). Integrating Biology into Demographic Research on Health and Aging (With a Focus on the MacArthur Study of Successful Aging), in: C.E. Finch et al. (eds.). *Cells and Surveys – Should Biological Measures Be Included in Social Research?* National Academy Press: Washington, D.C., 9-41.

Weinstein, M., Willis, R.J. (2001). Stretching Social Surveys to Include Bioindicators: Possibilities for the Health and Retirement Study, Experience from the Taiwan Study of the Elderly, in: C.E. Finch et al. (eds.). *Cells and Surveys – Should Biological Measures Be Included in Social Research?* National Academy Press: Washington, D.C., 250-275.

5 Linking SHARE Survey Data with Administrative Records: First Experiences from SHARE-Germany

Julie Korbmacher, Christin Czaplicki, Max Planck Institute

5.1 Introduction

SHARE-RV stands for the direct linkage of survey data of the Survey of Health, Aging and Retirement in Europe (SHARE: www.share-project.org) with administrative records of the German Pension Insurance[1] (DRV: Deutsche Rentenversicherung). SHARE-RV is a study within the German subsample of SHARE and started in the third wave of SHARE as a pilot study. The goal is to provide researchers from different fields with a rich database.

Despite the fact that administrative data are not primarily generated for research purposes, they have many advantages compared to survey data. They often cover nearly 100 percent of the population of interest and they are more accurate than survey data because issues such as recall errors do not play a role (Calderwood and Lessof, 2009). However, administrative data are process data collected for a specific purpose (for example, to calculate retirement claims) and therefore are limited to information needed for that specific purpose (Rehfeld, 2008). Surveys have the advantage that researchers can design the questions in such a way that data needed to answer a specific research question are collected. Here issues like unit or item nonresponse as well as recall error may reduce data quality and external validity.

Combining survey and administrative data could be a fruitful way to combine the best of both worlds and provides a wide range of research possibilities for content related research as well as methodological research. Content related research benefits from the accurateness of administrative data as collecting information in a survey with the same level of detail as can be found in administrative records is very difficult. For example, analyzing and predicting old age poverty requires information on expected pensions. Not all people know how many earning points they accumulated or how much pension they will get (the German pension system administers so-called “earning points” to compute future claims). Administrative data include that information on the individual level practically free of measurement error. However, administrative records cannot provide data on the household level and income sources other than public pensions. Combining survey and administrative data allows taking advantage of both. Another strand of research which benefits from linking these two data sets is methodological research: Users of survey data often assume that the data they use are free of measurement error and bias. But one has to take into account that some

[1] The German Pension Insurance consists of 16 independent organizations. The procedure used here has been discussed with data privacy protection officers in all organizations. Furthermore, the project has been discussed in two councils of the self-government board of the public pension insurance. The positive decision of each data privacy protection officer and the two councils is a necessary condition for data linkage projects like SHARE-RV.

variables collected in a survey are prone to errors. Different sources of error have been identified, but a prominent one is recall error, where some people misremember events which may result in “wrong” answers. To validate these errors other data sources including objective measurements of the same aspect are needed. The fact that some information is included in both data sets offers the possibility to validate survey data and identify characteristics which increase the risk of errors.

5.2 Data overview of administrative records

To improve the data infrastructure between research and administrative statistics, several research data centres have been founded in Germany. One of them is the Research Data Centre of the German Pension Insurance (FDZ-RV) which was setup in 2004 as an integral part of the German Pension Insurance. The FDZ-RV supplies researchers with cross-sectional and longitudinal micro-datasets on topics like retirement, disability, and rehabilitation.

The German Pension Insurance collects information on all people who have ever been subject to social insurance contributions. This implies that not all Germans are included in the database; excluded are, for example, civil servants or freelancers. This affects only a small share of the population, however, since the majority of Germans (nearly 90 percent) do have a record (Mika, Rehfeld, and Stegmann, 2009).

The data provided by the FDZ-RV for research purposes are anonymized[2] subsamples drawn from the pool of individuals who are insured in the German Pension Insurance. The administrative data which are linked with SHARE do have the same format and content but refer to those SHARE respondents who agreed to the linkage. The FDZ-RV provides SHARE with two different data sets which will be described in the following:

The dataset of the socially insured population (VSKT: Versichertenkontenstichprobe) is a longitudinal data set which includes detailed data on peoples’ working history and the state of their pension entitlements. This dataset consists of three parts. First, socio-demographic characteristics such as information about children, the current insurance type and the place of residence (state) are included. Second, information on the calculation of the (prospective) pension, most importantly the sum of earning points collected during the working history is available. These two parts are cross-sectional and refer to the last calendar year. The third part is the main component of the VSKT: It contains respondents’ biographies beginning with the age of 15 until the maximum age of 67 (Himmelreicher and Stegmann, 2008). Information is provided on a monthly basis and comprises, periods of employment, caregiving, illness or unemployment, the earning points for these activities and much more (Stegmann, 2007).

Cross-sectional pension data (RTBN: Versichertenrentenbestand) are available for retirees only. This dataset includes the value of pensions paid by the German Pension Insurance and information about all entitlements used for the pension calculation. As the longitudinal dataset (VSKT) ends with the transition into retirement at the latest, the

[2] For further information on anonymization procedures see Stegmann et al. 2005.

additional pension data allow analyzing respondents' lives beyond their working period.

5.3 Procedure

The method of linking different data sources depends on legal and technical constraints of each dataset. In a first step one has to define the type of data linkage:

- match data sources of the same person vs.
- match data sources of people who are similar (in a statistical sense).

Both procedures have advantages and disadvantages. When linking data from the same person, the respondents' consent is necessary (Calderwood and Lessof, 2009), which often decreases the number of linkable cases. Consent is not necessary when linking persons who are statistical similar, but one has to deal with the fact that the linked data refer to a "statistical twin" only. The SHARE-RV project is based on a direct linkage procedure meaning that the records of exactly the same person (here: SHARE respondents) are linked.

In a second step one has to decide how to link the different data sets to make sure that the correct data are used. There are two types of individual-level linkage: deterministic and probabilistic linkage (see Gill, 2001, for a detailed explanation, and Calderwood and Lessof, 2009, for an overview). The main difference is whether disagreement between the matching variables is allowed. Deterministic matching (not allowing for disagreements) depends on a unique identifier available in both datasets whereas probabilistic matching (allowing for disagreements) is based on different linking variables which are all allocated with different weights.

SHARE-RV is based on deterministic matching using the Social Security Number (SSN) as a unique identifier for all Germans who have a record at the German Pension Insurance. The FDZ-RV is allowed to provide data for those respondents who gave their written consent within the SHARE interview. Given that the survey is computer assisted, an additional paper form is necessary to collect the signature. The implementation of the consent question in the CAPI instrument differs between waves. Changes in that procedure are based on problems and experiences of each wave. A pilot study was implemented in the third wave of SHARE with two steps of consent:

Step 1: The first step was a verbal consent at the end of the SHARE interview. The respondents were asked for their consent to link the information collected in the interview with the administrative data held by the German Pension Insurance.

Step 2: If the respondents gave their consent, the interviewer handed out a consent form that had to be filled out by the respondents themselves. This form recorded the respondents' SHARE-ID[3], the SSN, all information needed to generate and/or check the SSN as well as the signature. For data protection reasons, the respondents had to send this letter directly to the German Pension Insurance. The SSN was only used to identify the respondents' DRV records and is not included in the resulting dataset.

5.4 Results of the pilot study

Consent rates differed substantially between the two steps of consent. Figure 5.1 shows the results of the two consent questions.

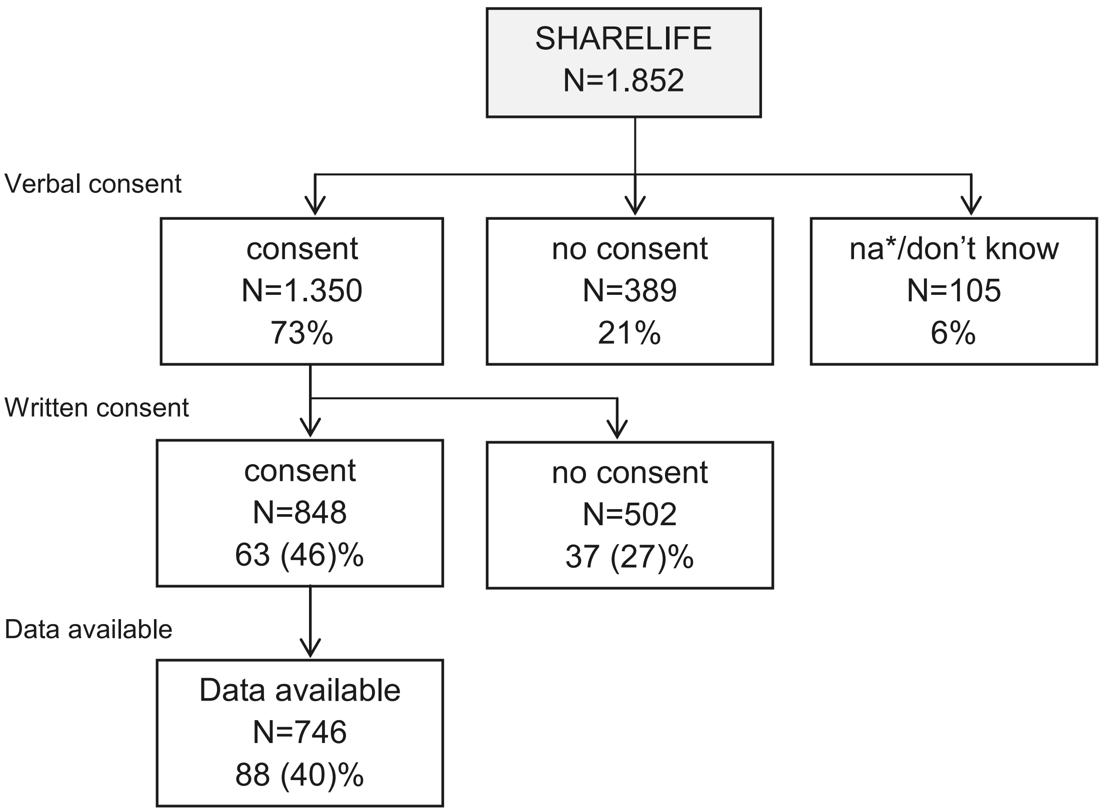


Figure 5.1 SHARELIFE Release 1.0.0 (Numbers in parentheses refer to initial sample of 1.852 cases), * na: not applicable

Seventy three percent of the respondents consented verbally to the linkage of their survey data with administrative records, whereas 21 percent refused and the remaining respondents argued that they do not have entitlements to the German Pension Insurance or answered "don't know". In the second step of consent, 63 percent of those respondents who consented in the first step sent back their signed consent form. The fact that this step was completely self-administered made it difficult to identify reasons for the gap between verbal and written consent. To decrease that gap, we implemented different changes for following waves: The two steps of consent were combined into one so that all respondents will receive the consent form and have to make a decision

[3] The SHARE-ID is necessary to assign the consent form to a SHARE interview.

only once. Further, interviewers will be more involved in the process. They will be responsible for administering the consent letters and bringing it directly to the post office to reduce respondents' burden. The consent form was modified so that providing the SSN will be optional. If all other personal information was provided, the FDZ-RV will be able to generate the SSN. If the correct records can be identified in the administrative data, they will be linked. Additional dropouts were due to data availability. Not for all respondents who consented data were available (that affected 12 percent of the consenters). Some records were (temporarily) not available for different reasons4 and some respondents provided a wrong SSN. To minimize errors in the SSN an additional step of checking the SSN was implemented. Overall, the first delivery of administrative data included 746 linkable cases which correspond to 40 percent of the German wave three respondents. Some data availability problems were solved at the time of writing (Nov. 2012) so that the linkage rate will be higher in the next release.

5.5 Data availability & outlook

The two datasets (SHARE and administrative records) will be available separately. Registered SHARE users will be able to download the survey data via the SHARE research data center. The administrative data will be available via the FDZ-RV after successful registration as a user. That dataset will include the SHARE identification number which is unique for all respondents over all waves. Independent of the time the consent was collected, the administrative data could be linked with all waves a respondent participated. For those respondents whose working history was not completed (people who were still working) the longitudinal data set can be updated and enlarged each year by information of 12 additional months. In contrast to the SHARE data, the administrative dataset will be updated every year. We will continue the project, asking all new German respondents of a refresher sample for consent. Additionally we will expand that project to other SHARE countries in further waves. As of October 2012, Austria, Sweden, Denmark and Estonia are planning to implement a link with their country specific administrative data in SHAREs fifth wave of data collection. The content of their administrative datasets might differ from what is collected in Germany, but will nevertheless be interesting and valuable.

References

Calderwood, L., Lessof, C. (2009). *Enhancing Longitudinal Surveys by Linking to Administrative Data*. In P. Lynn (Ed.), Methodology of Longitudinal Surveys). Chichester, UK: John Wiley & Sons, Ltd.

Gill, L. (2001). *Methods for Automatic Record Matching and Linkage and their Use in National Statistics*. National Statistics Methodology Series, No. 25. London: Her Majesty's Stationery Office. Retrieved October 2011 from http://www.ons.gov.uk/ons/guide-method/method-quality/specific/gss-methodology-series/index.html.

[4] http://www.share-project.org/fileadmin/pdf_documentation/SHARE-RV/1_Methodenbericht.pdf

Himmelreicher, R. K., Stegmann, M. (2008). New Possibilities for Socio-Economic Research through Longitudinal Data from the Research Data Centre of the German Federal Pension Insurance (FDZ-RV). In: *European Data Watch. Schmollers Jahrbuch 128* (2008). (pp. 647-660). Berlin: Duncker und Humblot.

Stegmann, M., Luckert, H., Mika, T. (2005). *Die Bereitstellung prozessproduzierter Daten der GRV im Forschungsdatenzentrum der Rentenversicherung (FDZ-RV)*. Grundsätze zur faktischen Anonymisierung von Mikrodaten und zu Gastwissenschaftler-Arbeitsplätzen. In: DRV 2–3/2005, 203–215.

Mika, T., Rehfeld, U. G., Stegmann, M. (2009). *Provisions for old age. Income provisions and retirement*. RatSWD Working Paper Series 112, Berlin: Council for Social and Economic Data (RatSWD).

Rehfeld, U. G. (2008). Daten und Fakten für die Rentenpolitik - von Geschäftsstatistiken zum Forschungsdatenzentrum der Rentenversicherung. In: G. Rolf, M. Zwick, G.G. Wagner (Eds.), *Fortschritte der informationellen Infrastruktur in Deutschland. Festschrift für Johann Hahlen zum 65. Geburtstag und Hans-Jürgen Krupp zum 75. Geburtstag*. (pp. 194-208). Baden-Baden: Nomos Verlagsgesellschaft.

Stegmann, M. (2007). Aufbereitung der Sondererhebung "Versicherungskontenstichprobe (VSKT)" als Scientific Use File für das FDZ-RV. In: *Die Versicherungskontenstichprobe als Scientific Use File*. Workshop des Forschungsdatenzentrums der Rentenversicherung (FDZ-RV) am 30. und 31. Oktober 2007 in Würzburg. DRV-Schriften Band 79, 17-33.

6 Investigating Response Behavior

Axel Börsch-Supan, Max Planck Institute

Ulrich Krieger, University of Mannheim

6.1 Introduction

Response to the major longitudinal surveys in Europe and the US has steadily declined over the recent 10-15 years. This is worrisome since non-response may be selective, introducing biases such as an overrepresentation of middle-income classes and healthy individuals. In order to examine the effects of respondent incentives on survey participation, to measure the nonresponse bias by conducting an nonresponse follow up study, and to investigate how additional training affects interviewer performance, a controlled experiment was introduced in the refreshment sample of the fourth wave of SHARE-Germany1. With the SHARE survey already funded, implemented and electronically monitored, no extra fieldwork cost had to be covered by the funders of this project. In addition, this setting had the advantage of testing the projects hypotheses in a realistic survey environment.

6.2 Experimental design

When the project was started in September 2010, the two-part design of the experiments was finalized in cooperation with the team running the SHARE survey at the MEA and the agency contracted for fielding the survey, Institut für angewandte Sozialforschung, Bonn (infas). All experiments, as all other SHARE fieldwork procedures in general, have been submitted to, and approved by, the ethics committee at the University of Mannheim which was the legally responsible entity for SHARE during wave four.

6.2.1 Incentive experiment

This first part of the study evaluated if prepaid cash incentives did increase cooperation rates and how different amounts of cash incentives influenced success rates. The implementation of this experiment was no trivial matter as unconditional prepaid cash incentives are an uncommon procedure to increase response rates in Germany. In contrast to the original plan of running the incentive experiment with a double blind approach, we decided to run the experiment half blind, informing interviewers of the treatment condition of respondents but leaving the respondents uninformed. The reasons for this decision were twofold. First, the agency had doubts about the practical implications of not informing the interviewers about such a considerable design feature of such a large study. It is important for the interviewers to be fully informed about all aspects of the survey to present the study accurately and positively at the doorstep interaction when establishing contact. It may have been disadvantageous for the interviewer (and agency and the study sponsor) if a respondent had fielded questions to an interviewer about the prepaid cash incentive and that interviewer being completely unaware of this design feature.

Second, the project team came to the conclusion that informing interviewers about the incentives would create a more realistic scenario. When using incentives in

[1] The refresher sample was not finished due to capacity limits and was consequently not released.

subsequent waves (without experimental control), interviewers will be well aware of the presence of incentives. They will be able to refer to the incentives and use those as a reference when making contact. So in a non-laboratory setting, incentives will always jointly affect interviewers and respondents. It was decided that running the experiment half-blind would increase the transferability of our findings to future survey research projects. The cash incentives were send out to sample members together with the advance letter. The selection of addresses into treatment and control group was done by the project staff. The entire refreshment sample was divided in four different replicates (or "batches") to be fielded sequentially. To maximize the duration of fieldwork for sample members in the treatment group, the experiment was run in the first batch of addresses that were send out right at the start of fieldwork.

Sample members from sample points in communities with less than 9000 inhabitants were excluded from the experiment because there was a considerable risk of respondents finding out about other respondents' incentives, and thus confounding the treatment conditions in these smaller towns. The refreshment sample was drawn from a total of 210 sample points, 54 sample points were excluded for this reason.

After the addresses had been drawn from the register, the survey agency delivered the sample to the project team. Treatment and control group status were randomly allocated in all of the 156 larger sample points. Therefore, all sample points in the experimental treatment contained sample members of the treatment and the control group. As a consequence, the experiment was run in all German regions and all interviewers had sample members in 6.1 gives a summary of the experimental setup, all experimental conditions, as interviewers usually work in a certain sample point. Table especially the three treatment groups which were given 10, 20 and 40 Euros in cash, respectively.

Table 6.1 Summary of sample size

	Sample points	Addresses
Full sample	210	9235
First batch	210	5247
Excluded (small communities)	54	1347
Part of the experiment	156	3900
Experimental conditions		
40 Euros	156	750
20 Euros	156	750
10 Euros	156	1375
No prepaid Incentive	156	1.025

The size of the treatment groups was computed to have enough statistical power for significance testing.

6.2.2 Nonresponse follow-up

As second part of the project, the interviewers administered an ultra-short questionnaire to those respondents who refused to participate in the main SHARE survey. Drawing on work by Lynn (2003) on the Pre-Emptive Doorstep Administration of Key Survey Items (PEDAKSI) method and following the example set by a study on a doorstep questionnaire survey (DQS) in the European Social Survey (Matsuo et al. 2010), a very short questionnaire including basic demographics and five items from the main survey were used in the study. The goal of this short questionnaire was to compare responses of respondents in the SHARE survey to respondents that only answered the questions of the short questionnaire. Without the doorstep questionnaire these respondents would have fallen in category of nonrespondents. Thus, using the DQS allowed drawing conclusions on potential nonresponse bias. In the planning stage, four questions (one with a follow up) were selected that closely resemble key survey questions of the full questionnaire:

- How many persons live in this household?
- If more than 1 person: Are children or a partner among them?
- How many children do you have?
- How would you rate your health status?
- What would best describe your current occupational status?

The questionnaire was designed to fit on less than one page to signal respondents the ease and minimal effort of completing it. Interviewers were advised to approach all sample members who finally refused participation in the main survey with a request to participate in the DQS. Completed DQS were then send back to the survey agency, coded electronically and made available to the project team.

6.3 Results

Survey operations conducted by infas did not proceed as planned. We encountered severe capacity limitations, huge delays, and a massively reduced number of interviewers supplied by infas. Hence, the final sample size was lower than planned (1900 rather than 4000) as not all sample members had been contacted. Moreover, not all SHARE quality standards have been adhered to by infas (e.g. a minimum of 8 personal visits to sample members). Nevertheless, the integrity of the two experiments was retained, although with a severe loss of statistical power. In order to maintain a clean sample design, SHARE decided to repeat the refreshment sample in wave five.

6.3.1 Incentive experiment

The main effect of unconditional prepaid cash incentives on response rates was overall positive. Moreover, the magnitude of the effect was higher the higher the incentive was. Controlling for age, gender and population size of the community of the sampling point, sample members receiving an incentive cooperated more with the survey request than those respondents from the control group. This analysis was restricted to addresses that have a final disposition code (out of sample, refusal or interview) or could not be contacted during the field period while exhausting the minimum number of 8 personal visits to the housing address. Out of the 3900 addresses in the experiment this condition was met for 2241 sample members. The key dependent variable was cooperation of households. While SHARE is a survey of individuals and incentives have been targeted to the respondent drawn from the German population register, the survey also targets the cohabitating household members. As the incentive may have also had an influence on partners we looked at cooperation of at least one household member. Table 6.2 shows the results from the three treatment conditions and the control group.

Table 6.2 Effect of incentives on cooperation

Incentives	Not Cooperating	Cooperating	Total
Control	398 (73%)	150 (27%)	548 (100%)
€ 10	494 (62%)	307 (38%)	801 (100%)
€ 20	258 (59%)	178 41%	436 (100%)
€ 40	209 (46%)	247 (54%)	456 (100%)

The proportion of sample members cooperating with the survey request increased monotonously with the amount of incentives paid. The control group had the lowest likelihood of cooperation of the four groups. The difference of the 40€ group to the control group was about 25 percentage points. Paying 10€ cash unconditionally to sample members still increased the likelihood of taking part in the survey by over 10 percentage points compared to the control group that received no incentive. The difference between 10€ and 20€ unconditional cash incentive was not statistically significant (the difference being about 2 percentage points in likelihood to participate). The results show clearly that paying larger incentive amounts increased the participation to a larger extent than smaller cash amounts. Moreover, and most importantly, all differences to the control group are statistically significant. We observe a percentage point difference of over 13 percent between the 20€ and the 40€ incentive treatments. This contradicts the work by Scherpenzeel (2008) who found a diminishing return of large incentive offerings in the LISS panel. Our findings are in line with a meta-study conducted by Singer et al. (1999) that found a significant linear effect for incentive sizes.

To check if these findings hold in multivariate analysis a fixed-effects logistic regression on household cooperation with the interview request was conducted. As households are nested within interviewers, we used interviewer fixed-effects. Explanatory variables in the regression are gender, age group, having received the second version of the advance letter (see section 6.2.2), and the incentive treatment. Population size of the municipalities of the sampling points was controlled by adding dummy variables for the lower third (9.000-35.0000 inhabitants) and the upper third of the population distribution (more than 200.000 inhabitants), respectively. Table 6.3 shows the results.

Apart from the effects of incentive treatment, two control variables yielded significant results: gender and living in smaller municipalities. Households of male target persons are 25 percent more likely to cooperate than if the target person is female. The odds of households in communities with less than 35.000 inhabitants cooperating were reduced by about 50 percent compared to households from towns that had between 35.000 and 200.000 inhabitants. Regarding the main effect, we show the contrasts of the control group to treatment groups. All those effects are positive and significant, meaning that the odds of responding were larger for all incentive groups than the control group. For the 10€ and 20 €, the odds were about twice as large as in the control group.

In the 40 € group the odds of responding were four times as large as in the control group.

Table 6.3 Fixed-effects logistic regression on household cooperation

	Odds ratio coefficients (standard errors in brackets)
10 € incentive	1.91*** (0.27)
20 € incentive	2.18*** (0.38)
40 € incentive	4.07*** (0.72)
Male	1.25* (0.14)
50-54 years old	1.10 (0.14)
2nd Version of advance letter	0.89 (0.20)
Smaller sample point	0.49* (0.15)
Large sample point	0.93 (0.25)

N= 2751, *McFadden's R*= 0.05, 107 interviewers

Thus, the positive effect of incentives on response propensity could hold up in a multivariate analysis since fieldwork of the refreshment sample was discontinued. We also check for possible biases due to selective interviewer effort.

Interviewers might have tried to contact the sample members from the incentive conditions first. Given that interviewers were aware of the incentive treatments and getting primarily paid for completed interviews, interviewers may have concentrated their efforts on those cases that have been offered an incentive due to the anticipated higher chances of a successful interview.

We therefore analyzed the effect of incentive treatment on the likelihood of households being contacted. If interviewers had prioritized contacting incentivized households we would expect to find effects of incentives in this analysis. As above, we used gender, age, the advance letter version and sampling point population as control variables. Individual differences in interviewers' success of contact were again controlled by using interviewer fixed effects. Table 6.4 below shows the results of the fixed-effects logit regression on contact. For this analysis all 3900 cases from the experiment were included in the analysis, however, 778 addresses form 29 interviewers were dropped from the analysis because there was no variation on the dependent variable among these cases: they either contacted all addresses or no addresses at all.

The only significant effect in this analysis was the negative effect of being in the age cohort of 50 to 54 years. The odds of those respondents being contacted were 38 percent smaller than those of the older respondents. There was no significant effect of the incentive treatment groups, indicating that the likelihood of being contacted did not depend on receiving a prepaid incentive or not. Obviously, the interviewers did not use the information on incentive treatment for their contact strategies. This is plausible since we know from talking to the agency and the interviewers that driving distances between households is their foremost consideration when planning their work. While interviewers get some compensation for travel cost, their main income results from finalized interviews. Thus, they try to optimize driving distances between addresses. Given that incentivized addresses were distributed randomly, optimizing driving distances trumped potential considerations on contacting incentivized households.

Table 6.4 Fixed-effects logistic regression on household contact

	Odds ratio coefficients (standard errors in brackets)
10 € incentive	1.05 (0.18)
20 € incentive	0.96 (0.18)
40 € incentive	1.01 (0.20)
Male	0.87 (0.10)
50-54 years old	0.58***(0.07)
2nd version of advance letter	1.16 (0.27)
Smaller sample point	0.62 (0.25)
Large sample point	1.70 (0.72)

N= 2824, *McFadden's* R^2= 0.02, 88 interviewers

As a negative side effect of the incentive experiment, hotline calls and subsequent refusals occurred more often in incentive groups. The proportion of sample members refusing was about equal across treatment groups, about 6.5 percent. The differences between treatment groups were not significant. We interpret this result as an indication of resentment towards the survey resulting from peoples' general opposition to receiving cash in the mail, not necessarily the cash amount. Refusal in the control group was considerably lower at 1.5 percent of sample members. The loss of these sample members has to be put into perspective to the large gains in cooperation in the treatment conditions. Moreover, while one has to take the hotline refusals seriously, the great majority (almost 95 percent of sample members) did not voice concerns about the incentive treatment and cooperated more than the members of the control group.

6.3.2 Nonresponse follow-up

Data from the nonresponse DQS were delivered to the project team in late June 2012. As in the incentive experiment, we only analyzed finalized cases. Out of the 2990 finalized addresses, 889 addresses were eligible for the "Doorstep Questionnaire Survey" (DSQ) as a consequence of "final" refusal to the SHARE survey request. Of all eligible cases, 197 respondents completed the DQS, yielding a response rate of 22 percent. This is in line with findings of Lynn (2003) who reported a cooperation rate of 25 percent. It is lower than the response rate of 45 percent of the DQS survey run in the ESS. We conjecture that the early end of fieldwork has played an important role here as well, as hard refusals and requests for DQS participation typically occur in later stages of the fieldwork when all other attempts to gain respondents cooperation in the main survey have failed.

For a comparison of SHARE and DQS respondents we focused on the sample of those addresses that have been part of the experiment. Figure 6.1 gives an overview of the results split by experiment conditions. Incentivized cases were compared to the control group.

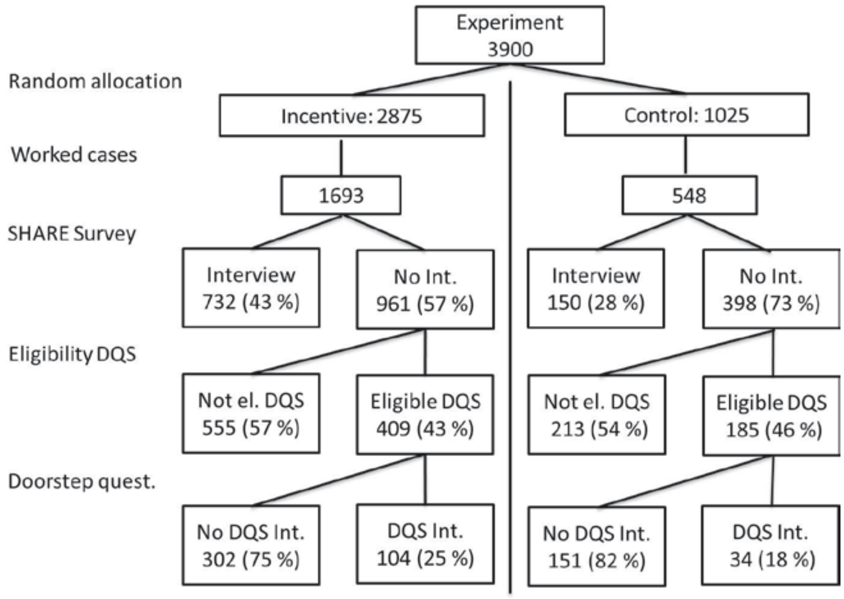


Figure 6.1 Results of SHARE and Doorstep Questionnaire Surveys

Comparing those sample members eligible for the DQS who took part in the DQS to those that refused we found no significant effects for age and gender (analysis not shown). The cooperation rate was larger among those respondents that did receive an incentive.

Having information from the respondents of the DQS allowed for a comparison of their answers to SHARE respondents more specifically an assessment of nonresponse bias of the main SHARE survey. We could investigate the difference of answers to key survey questions between respondents and nonrespondents. Further, we could differentiate between nonrespondents and SHARE respondents with and without a prepaid incentive treatment.

Table 6.5 gives the result of mean comparisons between SHARE respondents with and without incentive treatment, the DQS respondents and those that refused both requests. These three groups are split into treatment and control group. The information about gender and age were available for all targeted households from the population register. Health, occupational status and number of children were taken from the SHARE interview or the Doorstep Questionnaire. Household income was available for those respondents that completed the SHARE interview only.

Table 6.5 comparison of means between groups, standard deviations in brackets

	Incentive			No Incentive		
	SHARE	DQS	Frame	SHARE	DQS	Frame
Male	.49 (.50)	.49 (.50)	.44 (.50)	.49 (.50)	.45 (.51)	.43 (.50)
Age 50-55	.17 (.38)	.20 (.40)	.18 (.38)	.24 (.43)	.21 (.42)	.18 (.38)
Health: fair	.30 (.46)	.21 (.41)	-	.24 (.43)	.15 (.36)	-
Health: bad	.13 (.33)	.17 (.38)	-	.12 (.33)	.15 (.36)	-
Working	.31 (.46)	.22 (.42)	-	.36 (.48)	.29 (.46)	-
No. of children	1.86 (1.35)	1.65 (1.28)	-	1.89 (1.39)	1.4 (.88)	-
HH income	4737 (7635)	-	-	5028 (7088)	-	-
N	732	104	754	150	34	364

When comparing SHARE respondents with and without incentives, a significant difference in age composition was found. The proportion of younger sample members (50 to 55 years) was larger in the control group (t=1.9). The mean household income was lower in the incentive group than in the control group but not significant. There were no significant mean differences between the two DQS groups.

Were respondents different from nonrespondents? When comparing gender and age composition of SHARE and DQS respondents to those that did not complete an interview, the only significant mean difference is the larger portion of males among SHARE respondents than the nonrespondents (t=1.9). The mean differences between SHARE respondents and DQS respondents were not significant when controlling for experimental condition. When collapsing incentive and control group, the difference of 9 percentage point in the proportion of respondents rating their health condition as fair became significant (t=-2.3).

We conclude that there is little evidence for nonresponse bias amongst the variables we examined here. The comparison between SHARE respondents across the experimental conditions is very limited as we only looked at household income. Incentivized respondents reported a mean household income that is about 150€ lower than the mean household income reported by control group households. Further analyses with this data will be necessary to determine if this difference is an indication of low income populations being brought into the survey as mentioned by others (e.g. Singer and Kulka, 2000). Most importantly, however, is the evidence that the incentive treatment did not introduce bias into the study.

6.4 Acknowledgements

This project was conceived by Mathis Schröder and Axel Börsch-Supan and supervised by Dr. Annelies Blom (until December 2010) and Ulrich Krieger. We acknowledge the help provided by our student research assistants Ute Hoffstätter, Anna Krüger, Elisa Leonhard and Maximilian Weiß. At the Institut für angewandte Sozialforschung (infas) we thank Anne Kersting and Birgit Jesske for their time and efforts. We thank Frauke Kreuter, Josef Brüderl, Arie Kapteyn, Mick Couper, Peter Lynn, Peter Lugtig, Julie Korbmacher and Stephanie Eckman for helpful comments on the manuscript.

References

Lynn, P. (2003). *PEDASKI: Methodology for Collecting Data about Survey Non-Respondents*. Quantity and Quality, Vol. 37, pp. 239-261.

Matsuo, H., Billiet, J., Loosveldt, G., Berglund, F, Kleven, O. (2010). *Measurement and adjustment of non-response bias based on non-response surveys: the case of Belgium and Norway in the European Social Survey Round 3*. Survey Research Methods, Vol. 4, pp. 166-178.

Scherpenzeel, A. (2008). *Recruiting a probability sample for an online panel: effects of contact mode, incentives and information*. Working Paper, University of Tilburg.

Singer, E., Van Hoewyk, J., Gebler, N., Raghunathan, T., McGonagle, K. (1999). *The Effect of Incentives on Response Rates in Interviewer-Mediated Surveys*. Journal of Official Statistics, Vol. 15, pp. 217-230.

Singer, E., Kulka R.A. (2000). *Paying respondents for survey participation*. Survey Methodology Program Working Paper No.092, Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor.

7 Software Innovation in SHARE Wave Four

Arnaud Wijnant, Maurice Martens, Eric Balster, Marcel Das, CentERdata, Tilburg University

7.1 Introduction

In SHARE wave four, survey instruments were updated and developed further to improve their functioning and make all aspects of conducting the survey a smoother experience. These can be categorized into four aspects: the revision of the Computer-Assisted Personal Interview software (CAPI), updates to the translation tool to obtain national-language interview software, the sample management software and data delivery procedures that ultimately yield the scientific data. These four aspects will be discussed next.

7.2 CAPI instrument

The development of the wave four CAPI instrument started with a review of the wave two CAPI instrument. The reason was that the wave three interview was different as it assessed the life histories of the SHARE panel members. The survey questions and routing of items was to a large part identical to wave two. Routing for baseline versus longitudinal respondents was brought back and only some sections and questions were removed or added (such as the brand new Social Network Module, see chapter 3).

The most fundamental adjustments in the CAPI instrument were: 1) the way in which feed-forward information from previous waves was handled, 2) the introduction of a Social Network module (for details see chapter 3), and 3) the architecture of the children module.

When the questionnaire was started for the first time to conduct an interview, the feed-forward information was loaded by the Sample Management System (SMS) from a Blaise database. Blaise is a programming tool commonly used to implement CAPI software. In this way, some basic background information about the respondent was loaded into the questionnaire software, such as the date of the previous interview and the list of children that were reported in previous waves.

The structure of the new Social Network module is basically a loop of questions that asks if there is a person or persons that belong to the social network of the respondent (for details see chapter 3). The maximum number of persons that could be mentioned was seven. After listing these persons, and asking follow-up questions, an overview was given to identify and remove duplicate entries. This turned out to be an important step because later in the interview the social network list was included in the response options of a question. Based on the relations found in the Social Network module it was possible to shorten the questionnaire. It was no longer necessary to ask for parents or siblings if they were already identified in this module. The identified children, however, posed a problem. There could already be a list of children available for longitudinal respondents contained in the feed-forward information from the panel preload file in the Blaise database. Obviously, the two groups of children could overlap. It was decided to match the Social Network children with the preloaded children. Initially, automatic identification with a computer algorithm was tested, but this did not yield an acceptable solution. Therefore, it was decided that the respondent should point out which children listed in the Social Network module matched with the preloaded children in the children section. To guide this process both sets of children were shown on the screen in one grid (see Figure 7.1).

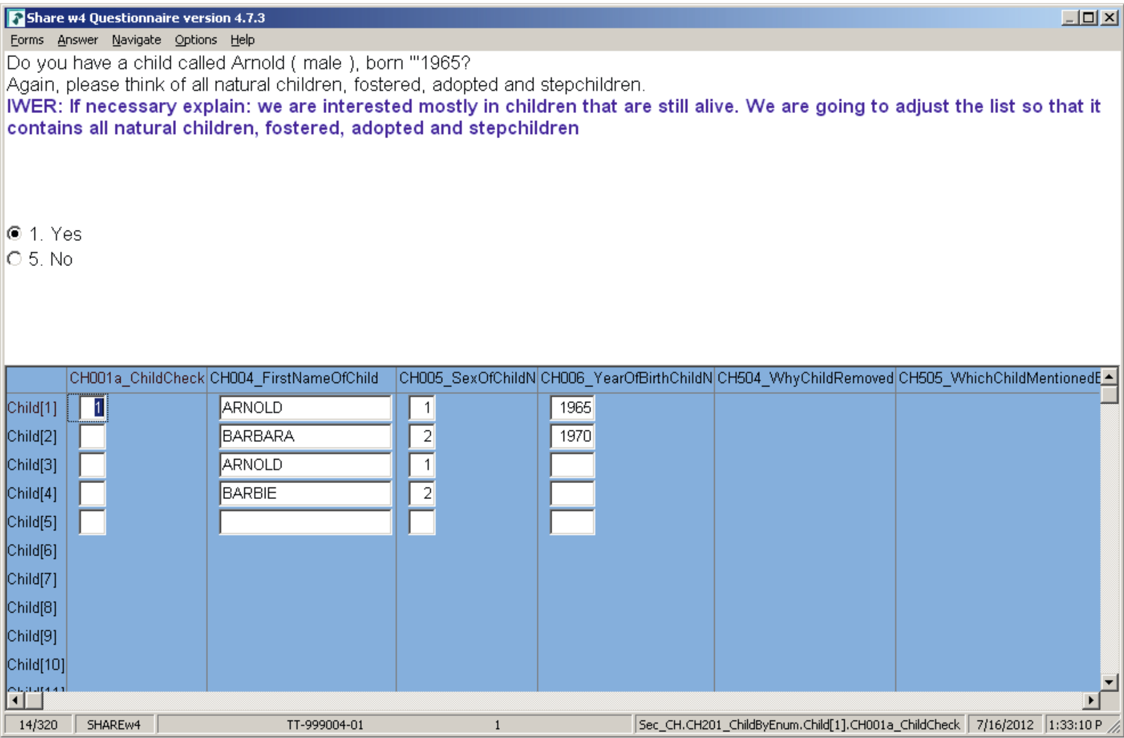


Figure 7.1 Child grid

Notes: The first two child names were preloaded; the third and fourth were entered in the social network module. Other children could also be added to this grid starting with the fifth position up to 20 children.

The first question: “Do you have a child called NAME (SEX), born YEAR” was actually meant to find out if a certain child existed. However, if the child was already identified earlier in this list the interviewer still had to tick ‘no’ as an answer to this question. This meant the wording shown above could not be applied in a straight-forward fashion. After answering ‘no’, the next question was why the child had to be taken out of the list (see Figure 7.2). The respondent could indicate the child had to be taken out of the children list for various reasons: 1. Child of partner from whom the respondent separated, 2. Child died, 3. Child unknown or 4. Already mentioned. If reason 4 was chosen, the respondent was asked to identify which children were identical. After this identification, the answers to the location and contact frequency questions were copied from the Social Network into the child data.

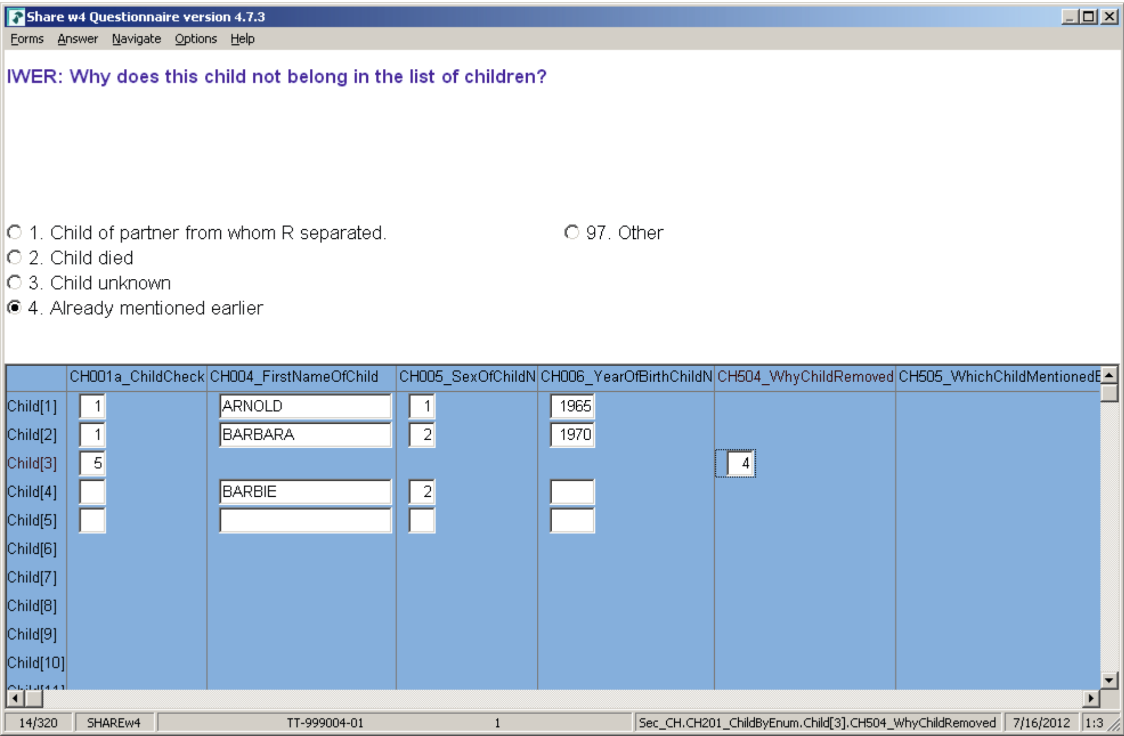


Figure 7.2 Linking the children

Notes: Child 3 already existed and is now linked to child 1.

If the respondent had other children that were not included in the feed-forward information or not mentioned in the Social Network module, they could still be added. For 'verified' preloaded children the follow-up questions were reworded to check if any changes had occurred recently (rather than asking the complete loop of follow-up questions).

7.3 Translation

The translation process in SHARE is managed by a web-accessible tool, the so-called Language Management Utility (LMU). During various SHARE waves, this tool has constantly been revised and improved. Details about the evolution of the tool can be found in Börsch-Supan et al. (2008). In wave four, the visible translation screens were kept almost identical to the third wave screens. The focus was now on improving the management of the translation process. Therefore, an administration layer was added. This had an immense impact on the underlying data structure.

In SHARE, translation is linked closely to the questionnaire development. During the generic questionnaire development there are various phases in which the translators are asked to translate, verify and re-translate. This is done in iterations, always improving on earlier versions. Translations can be used to review the question. Problems with translations could indicate that there was a problem with the conceptualization and/or wording of the generic English source questionnaire. The translation process was formalized and made configurable. The formalization was done by introducing concepts like roles, workflows, states, elements, assignments, modules, and languages. These concepts were linked together to create a robust configurable environment for translation processes. This allowed for the possibility to configure a translation process tailored to the SHARE development methods.

The LMU is centered round the concept of elements that need translation. Such an element could for instance be a question text: “What is your sex?”. This text has an answer category attached to it; “Male/Female”. To manage the translation process, initially a status “Awaiting translation” for all languages was attached to the translatable element by a programmer. During translation a translator typed in the translation but also changed the state to another state “Translated, ready for check”. Now a checker came in who checked whether the translation of the translator made sense. He either moved the translation to “Check failed” or “Check OK”. If “Check failed” a translator needed to come back in, reviewed and proposed a new wording. If the status “Check OK” was set the translation was done. All users could attach messages to a translatable element during the translation process. To further optimize the translation process other roles and statuses were defined. An admin role was defined that could move elements to and from all states. An imported status was given to translated elements that were already translated in another wave of the study. In the future waves other roles, statuses and workflows will be defined dynamically in the LMU, thereby opening up the possibility to further professionalize the translation process.

Id	Name	Read Locked	Edit Locked	Read Unlocked	Edit Unlocked	Email All Status	Email My Status	View Problem Reports	Icon	Color	Actions
1	Admin	1	1	1	1	0	0	1			View Edit Delete
3	Translator	0	0	1	1	0	0	0			View Edit Delete
4	Checker	0	0	1	0	0	0	0			View Edit Delete
5	Programmer	1	1	1	1	0	0	0			View Edit Delete

Figure 7.3 Roles used in the LMU

These roles were created to allow spreading the responsibilities of the translation process over multiple persons. Each element was assigned a status (see Figure 7.4). A role could be assigned a workflow. The workflow was a matrix that specified the rights a certain role had to move translatable elements between statuses.

name	initial	final	editable	visible	icon	color	Actions
Under development	1	0	0	0			View Edit Delete
Awaiting translation	1	0	1	1	1.		View Edit Delete
Translated, ready for check	0	0	0	1	2.		View Edit Delete
Check failed	0	0	1	1	3.		View Edit Delete
No translation needed	1	1	0	0	4.		View Edit Delete
Imported	1	0	1	1	5.		View Edit Delete
Check ok	0	1	1	1			View Edit Delete
Imported from other translation	1	0	1	1			View Edit Delete

Figure 7.4 Statuses in the LMU

A programmer was typically given the rights to manipulate the English (Generic) source questionnaire. A translator could make changes to a translated version. A checker couldn’t change a question but could give comments and reject a translation. An admin role could be configured to move between statuses without limitations.

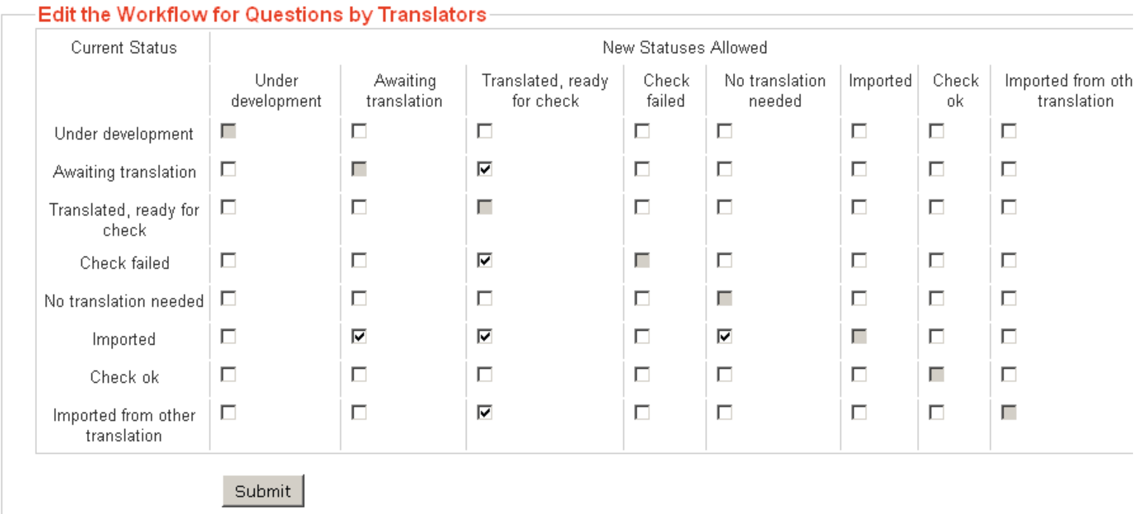


Figure 7.5 Defining a workflow

An example of a translator's workflow definition is shown in Figure 7.5. In this definition the translator could:

1. Move elements with status "Imported" to "Awaiting translation",
2. Move elements awaiting translation to "Translated, ready for check",
3. Move "check failed" to "Translated, ready for check",
4. Move imported elements to "Translated, ready for check",
5. Move imported from other translation elements to "Translated, ready for check",
6. Move imported to "No translation needed".

A translatable element typically started with the status "Under development" and should have ended with that status "Check ok". A translator could not move the translatable elements through the complete workflow alone. The elements would have needed a user with "Programmer"-role to change the translation status from "Under development" to "Awaiting translation" before they could be manipulated. And they would have needed a "Checker" to move the status to "Check ok". Using the workflow and roles to guide a translatable element through various statuses ensured that several persons had seen a given translation before it was signed off. To have an "emergency escape" from this division of workflow an "Admin" role was introduced that had the right to move to and from any status. In the system, every change in status was logged. Which user moved which translatable element to which status at what time was recorded. The system was very flexible and highly interactive. Statuses could be added or removed, workflows could be rephrased, and roles could be added.

A user would see a set of statuses that could be accessed (see Figure 7.6), and an interface to read the history of messages and write new messages.

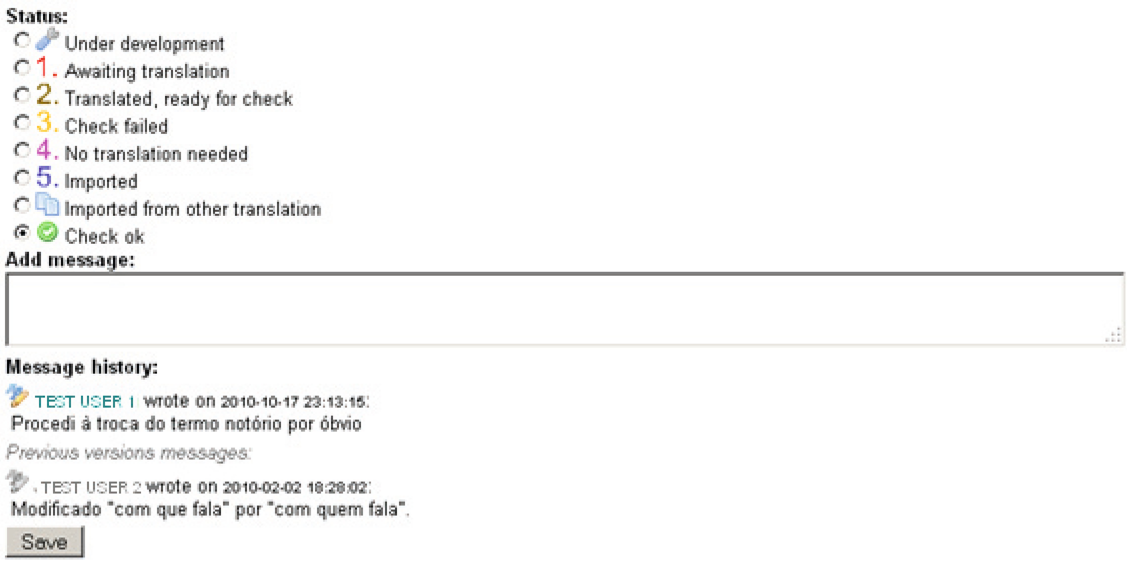


Figure 7.6 Assigning a status

A good configuration of the workflows and the assignment of roles to a translation team could guide the translation process while also providing a useful interface for managing this process.

Various graphical overviews could be derived based on the status of the translatable elements allowing easy overview of the translations (see Figure 7.7). This provided comparability between languages and made it easier to identify problems in the overall process.

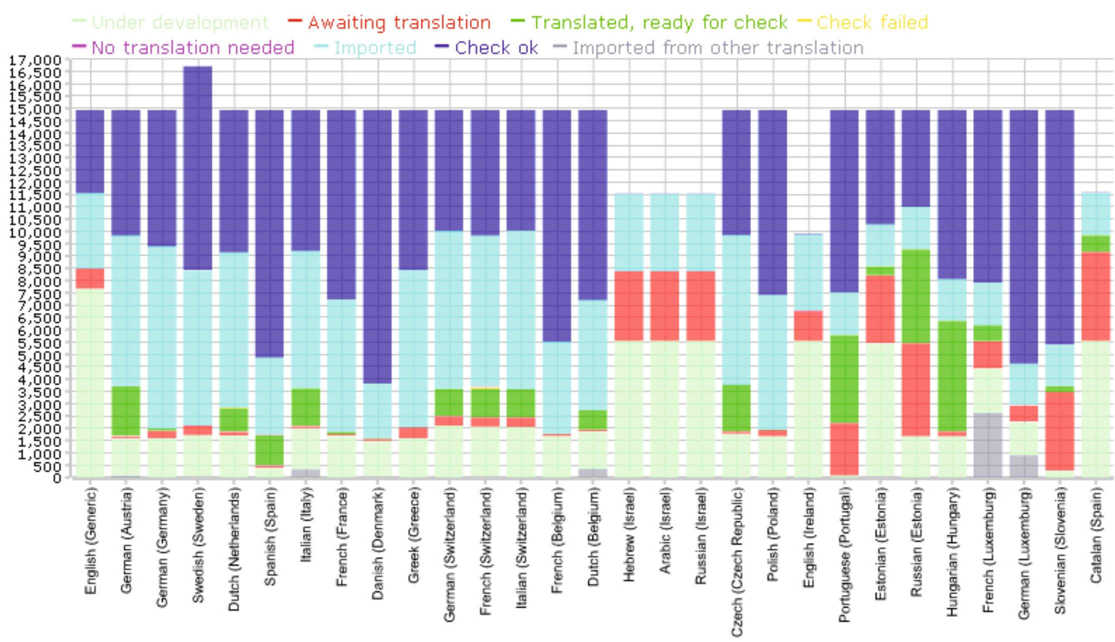


Figure 7.7 Management overview

7.4 Fieldwork sample management

The Sample Distributor (SD) was the software package used by the different survey agencies in the SHARE project to manage their fieldwork of a SHARE wave. The SD is a program that contains the complete household sample of a country. SD administrators can assign and unassign households of the sample to laptops. The SHARE Sample Management System (SMS) application was installed on these laptops. The SMS application enabled interviewers to contact and interview the assigned households. Because interviewers were quite satisfied with the SMS application in the third wave (SHARE Life), only minor changes were made to the fourth wave SMS version. Once an interview was finished, the SD collected the data of the SMS applications in the field and sent them to CentERdata where the data were processed for further analysis.

The SD was designed in such a way that every modification by an interviewer resulted in a backup (separately saved copy) of the affected households. Interviews were not only saved in a database but were also written to log files. Data security was an important part of the features of the SD, as well. All data that were being synchronized by different sites were encrypted in order to prevent misuse of these confidential data. In addition, the use of sample data was split across different roles in the SHARE project. Only survey agencies had the addresses of respondents; addresses were never included in exports to CentERdata. Country teams and the MEA database management team are the only teams that have characteristics of the sampled households (these are used to be preloaded in the questionnaires), but they also never had access to any address data. The SD was designed a closed system. Changing the internal data was allowed only through the SD. This could be done directly by using the SD user interface (see Figure 7.8 for an example of this user interface) or indirectly by importing a correction file that had been issued by CentERdata (to solve advanced issues).

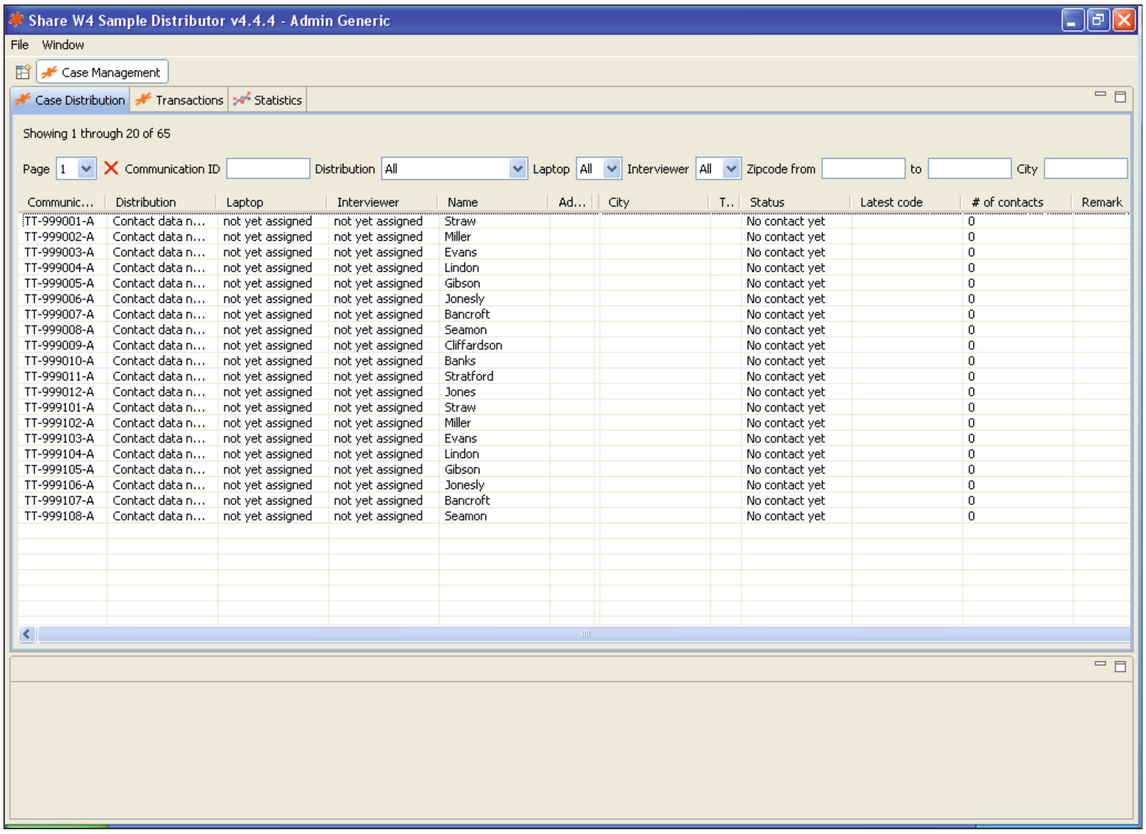


Figure 7.8. Sample Distributor user interface.

Notes: Via this screen administrators can manage the sample of a country.

Interviewer laptops played an important role in the SHARE fieldwork. Since the interviewers in SHARE were the persons who contacted the households, the SD design needed to accommodate for that situation. Once a household was assigned to an interviewer in the SD, the laptop of the interviewer was the only equipment where changes (like adding contacts, updating contact information but also conducting interviews) to the household could be registered. There were some exceptions to this situation, for example if the agency wanted to update the address information of a household. If this situation occurred, the agency could add an update to the household. This update, however, was not applied to the household directly. Instead, the SD sent a signal to the laptop to add this contact. Once the laptop received this signal, the update was executed and synchronized with the SD. This mechanism was ideal to maintain the household information and interview states of a household in their correct status.

The wave four SD software was based on the SD that has been used in the third wave (SHARE Life). Although the process of the SHARELife fieldwork looked very similar to that of the fourth wave, there were some major changes that affect the SD. One of these changes was that the sample size had increased significantly. Many countries added very large refreshment samples. The total number of interviews doubled from 30,000 interviews in wave three to more than 60,000 interviews in wave four. Also the number of variables and existing information about the respondents preloaded to the SD had increased significantly since the last wave.

The selection process of eligible respondents changed because of the refreshment samples. In wave four, it was possible for agencies to prescreen households to determine whether or not they contain a respondent that needs to be interviewed to fit into the sample design. This screening could both be done on the SD server and the interviewer laptop. In addition, the logging the status of drop-off questionnaires was added as a functionality to the SD. In Germany, an additional functionality was added to support the administration of an experiment with dried blood spots (for details see chapter 4). Besides programming these changes, the programming team of CentERdata invested a lot of effort in improving the performance of the SD without losing any functionality or user friendliness of the program.

Creating an optimal Sample Distributor was a trade-off between the performance, sample sizes, scope of functionalities and user friendliness. For wave four, a lot of effort was put into to improving the performance, but other challenges (like sample size and functionality) also increased. Figure 7.9 shows which areas were improved during the wave three and wave four software development. It is worth pointing out that improving one of the four would have been simple if it had been done at the expense of the other areas. The real challenge was to improve these areas while the others were held constant or improved, as well.

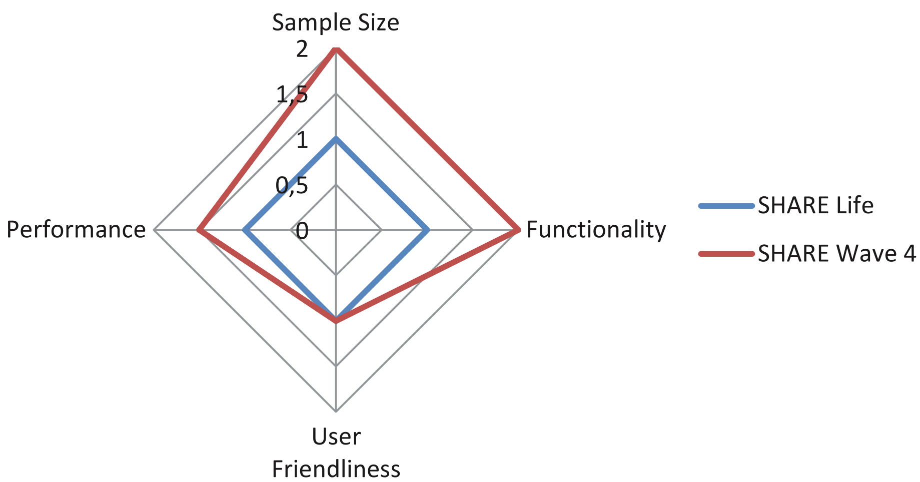


Figure 7.9 Improvement areas during SHARE Life and SHARE wave four

Unfortunately, improvements to the performance were not sufficient for some countries to prevent the SD from running slow. The database system used in the fourth wave was easy to install (simply running the installer, no special IT skills were required). However, some actions could take up to a few minutes (or even longer), especially towards the end of the fieldwork. This was caused by the size of the database. The larger the size of a database file, the more time it took to find an item in this file. In addition, more items needed to be found in a large database. For this reason, the software would have to be adjusted for future waves to cope with larger database sizes.

A solution would have been to replace the database system in the background with a better one. An option in the future could be using MySQL, a database management system with evidently better performance characteristics (especially for large datasets) than the current included database system. However, the installation of such a database management system will require more knowledge and effort to install and configure the system. A comparison of the characteristics of both future versions is shown in Figure 7.10.

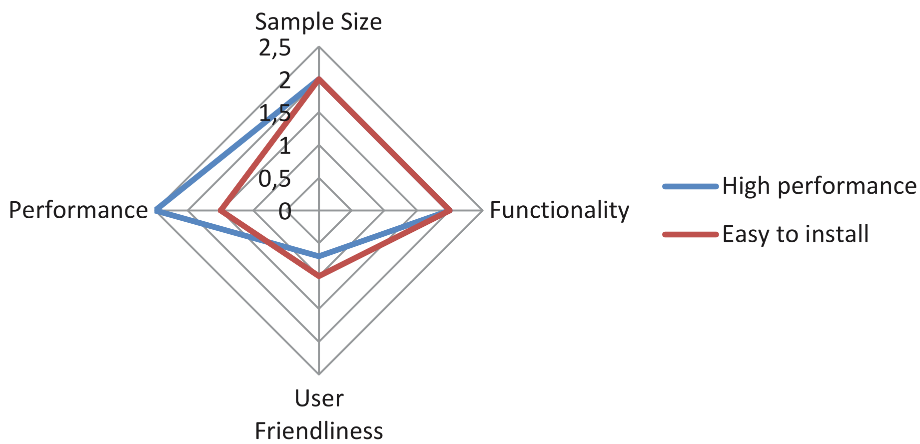


Figure 7.90 Spider diagram of the trade-offs for the two different versions of the SHARE SD for the fieldwork in the future.

Notes: One high performance version which is not easy to install (indicated in blue) and one easy to install version for small sample sizes and small fieldwork agencies (the red line).

During the fieldwork of the fourth wave there was an increasing demand for agencies to do tasks with the sample on a different system than the laptop of the interviewer (for example adding a contact with a respondent made by a call-center). For this reason one future plan is to split the database structure containing the households into two different independent parts. One part will be meant for the agency to add/remove or update the households. The other part will be for the interviewers' laptops. The system should be created in such a way that both the interviewers and SD administrators will be able to add data to the households individually and do not need to wait for each other. This will be especially helpful with the administration of drop-offs or the biomarker experiment which may run in some of the SHARE countries in future waves.

7.5 Data delivery

The data delivery tools were improved in wave four. Once the data had been collected at the agencies via the SD, agencies could send the data to CentERdata. CentERdata used the SHARE Data Delivery Service to create two different types of data from the Sample Distributor exports. The first type concerned the data files. These contained all the answers given by respondents during fieldwork. The second type of data concerned the monitoring files, which contained statistics about the fieldwork itself. For example, the interviewer performance could be found in these monitoring files. The process from SD export to data and monitoring files consisted of many steps. These steps are shown in Figure 7.11 in a stylized way. When SD exports were available at the CentERdata server, the first step of the data delivery was to unpack and decrypt the data files. CentERdata is the only organization within the SHARE network that has the key to decrypt the export files that were created by the Sample Distributors in each country.

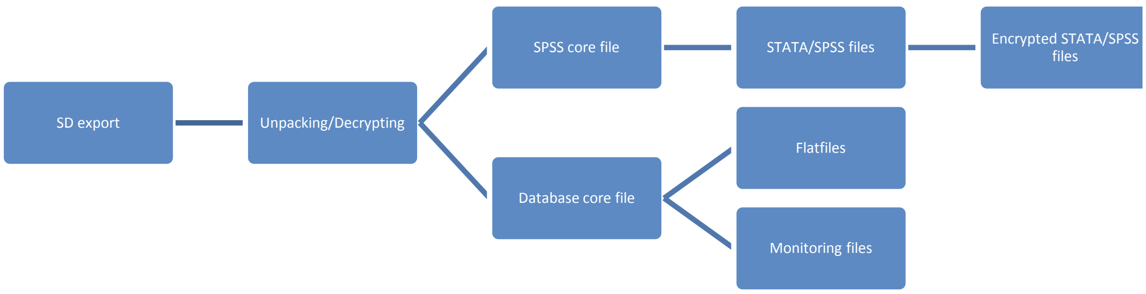


Figure 7.10 Data delivery process workflow

Once the files were unpacked and decrypted the first conversion took place. The Blaise databases were converted to SPSS data files. Parallel to this, another conversion took place. The anonymized SD database that is part of the SD export was converted to a database core file that could be read by SPSS scripts. The core SPSS files were then used to generate final products like STATA/SPSS datasets (that contain files per country and per module) for the country teams or encrypted STATA/SPSS datasets to be processed further. The second branch of Figure 7.11 symbolize the database core files used to generate flat files that contain the most up-to-date version of the sample state and monitoring files that contained the most up-to-date information of the fieldwork progress. All these files were distributed automatically to a password protected website to share with researchers who have the credentials to log on.

7.6 Final remarks

CentERdata encountered several issues during the fieldwork of wave four that could –despite concerted efforts- not be fully resolved. These will need to be fixed during the development process of the fifth wave. We plan improving the CAPI instrument’s routing and use preloading to further shorten the questionnaire and lower respondent burden.

To improve and further professionalize the translation process it will be necessary to make the LMU more user-friendly and to invest in enabling different and more enhanced translation methods. The performance of the SD and SMS software was a concern. Not only the algorithms and the data communication layer will have to be further optimized, but also an advanced version should be implemented. This advanced version should guarantee better performance than the regular version, but will require more IT knowledge to install. The suggested improvements should provide the means to have an even smoother fieldwork period in wave five.

References

Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J.,Weber, G. (2008). *Health, ageing and retirement in Europe (2004-2007). Starting the longitudinal dimension.* Mannheim: Mannheim Research Institute for the Economics of Aging (MEA).

8 Sample Design in SHARE Wave Four

Peter Lynn, University of Essex

Giuseppe De Luca, University of Palermo

Matthias Ganninger, Sabine Häder, GESIS

8.1 Introduction

This chapter outlines the design of the samples that are included in SHARE wave four. We begin by defining the population that SHARE aims to represent and explaining why this definition was adopted. We then set out the objectives of the sample design and summarise the approach that was taken to meet these objectives, thus placing the samples selected at wave four in the context of the samples selected at previous waves. We include a description of the process by which sample designs were developed and agreed and we describe the nature of the sample designs implemented, including discussion of the role played by sampling frames, stratification, sample clustering and variation in selection probabilities. The chapter ends with a description of the process of developing the weights that have been provided for use by data analysts. The weights adjust both for variation in selection probabilities by design and for variation in participation probabilities caused by non-response and analysts are strongly encouraged to use them.

8.2 What population does SHARE represent?

The target population for inference from SHARE is the European population aged 50 and older. However, the study design must also take into account practical considerations relating to the ability to sample and collect data from respondents. Two restrictions are introduced as a consequence. The first is that the study population is restricted to those people who are resident in a private household at the time of sampling and at the time of fieldwork. Residents of institutions are excluded, with the exception of countries using as a sampling frame a population register in which residents of residential and nursing homes are included. In such cases, those residents were included. The second restriction is imposed by the practicalities of interviewing in different languages. The study is restricted to people who speak (one of) the national language(s). Also, as the household context is important the spouses/partners of sample members are included, regardless of their own age. Thus, the definition of the study population for SHARE wave four is:

Persons born in 1960 or earlier, and persons who are a spouse/partner of a person born in 1960 or earlier, who speak the official language(s) of the country and who are residents within private households, regardless of nationality and citizenship.

To achieve representation of this population, SHARE employs a sample design which involves baseline samples of the household population aged 50 and older at a particular point in time in each country, supplemented by regular refreshment samples of the sub-population of people who have turned 50 since the original baseline sample was selected. The design and implementation of these baseline and refreshment samples are described in the next section.

8.3 The basic sample design

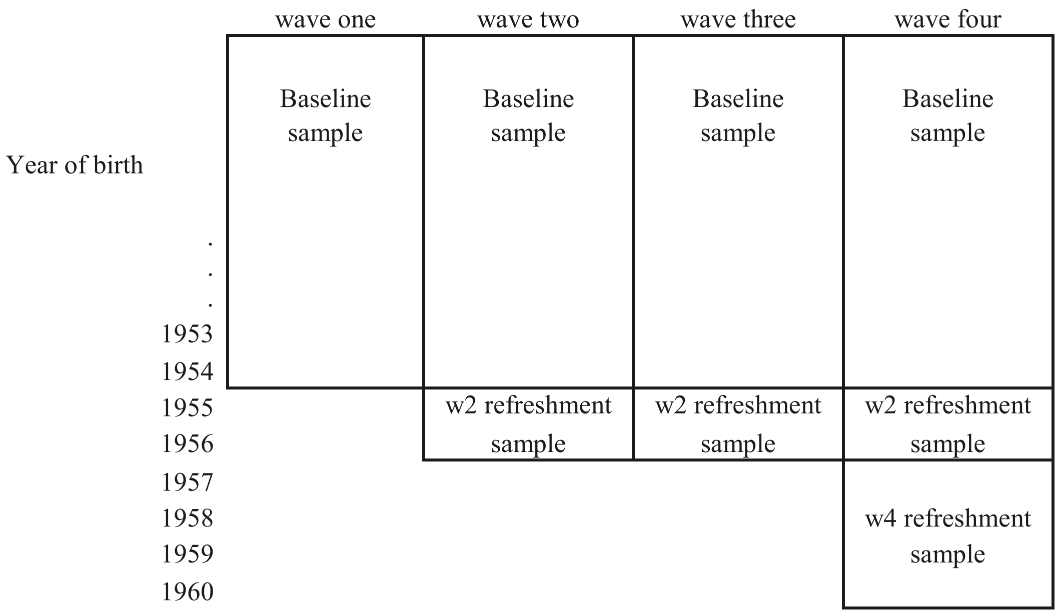
The sampling rationale for baseline and refreshment samples was the same that all sophisticated cross-national survey programs apply at present. Kish (1994, p.173) provides the underlying idea: “Sample designs may be chosen flexibly and there is no need for similarity of sample designs. Flexibility of 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.” This encapsulates the idea that to facilitate inference to the population of interest, it is necessary that the survey is based upon probability samples with full population coverage. SHARE therefore insists on the use of probability sampling, with known selection probabilities for each individual. The extent to which full population coverage is strictly possible depends on the quality of sampling frames available in each country (see next section), but in all cases close to full coverage was achieved. The details of the sample design varies between countries, as discussed later in this chapter, but the basic principles of probability-based selection and maximal population coverage underpin all the designs used.

Probability sampling and the absence of under-coverage ensure that a sample can provide unbiased estimates. But in addition to the avoidance of bias, it is necessary that samples provide sufficient precision to enable meaningful estimation. This requires adequate sample sizes, minimal clustering and minimal variation in selection probabilities. Precision can also be aided by the use of sample stratification, so this is encouraged where possible. For this reason the procedures adopted by SHARE address each of these components of the design: sample size, clustering, variation in selection probabilities and stratification. Regarding sample size the target for each country is to conduct 6,000 interviews overall at each wave, baseline and refreshment samples combined. Regarding the other elements of sample design, advice is provided to participating countries by means of the “SHARE Sampling Guide” and through bilateral discussion with a member of the SHARE Sampling Panel. The results and the implications are outlined in a later section of this chapter.

Even with a well-designed sample selection process, the sample of respondents can become unrepresentative of the target population due to non-response. A final important ingredient in order to achieve the inferential aims of the study is therefore to achieve high response rates. The extent to which this was achieved can be found in chapter 10.

Four new countries entered SHARE for the first time in the fourth wave. These countries - Estonia, Hungary, Portugal and Slovenia – therefore had to construct baseline samples that will ultimately form their “first wave” panel cases. Other countries had to select refreshment samples of people born between 1957 and 1960 to add to their existing sample of people born in 1956 or earlier. (Wave 1 baseline samples consisted of people born in 1954 or earlier; Wave two refreshment samples contained people born in 1955 and 1956; No refreshment samples were added at wave 3.) For some countries – where no refreshment sample had been added at wave two – the wave four refreshment sample included people born between 1955 and 1960. Additionally, many of the countries which had participated in any of the previous three waves conducted so far were faced with a sample size problem. Due to panel mortality, the number of cases in the initial sample has decreased from wave to wave. Consequently, many countries deemed it necessary to implement a refreshment sample across the full age range of people born in 1960 or earlier, in order to have a large enough sample size for subgroup analyses such as by age groups. Where possible, these full-range refreshment samples included an over-sampling of persons born in 1957 to 1960 (or 1955 to 1960 if the country had no wave two refreshment sample), to maintain the statistical efficiency of the overall sample. Figure 8.1 illustrates examples of the different kinds of sample combinations that can be found in the SHARE data.

Example A: Countries which had a refreshment sample at wave two



Example B: Countries which had no refreshment sample at wave two

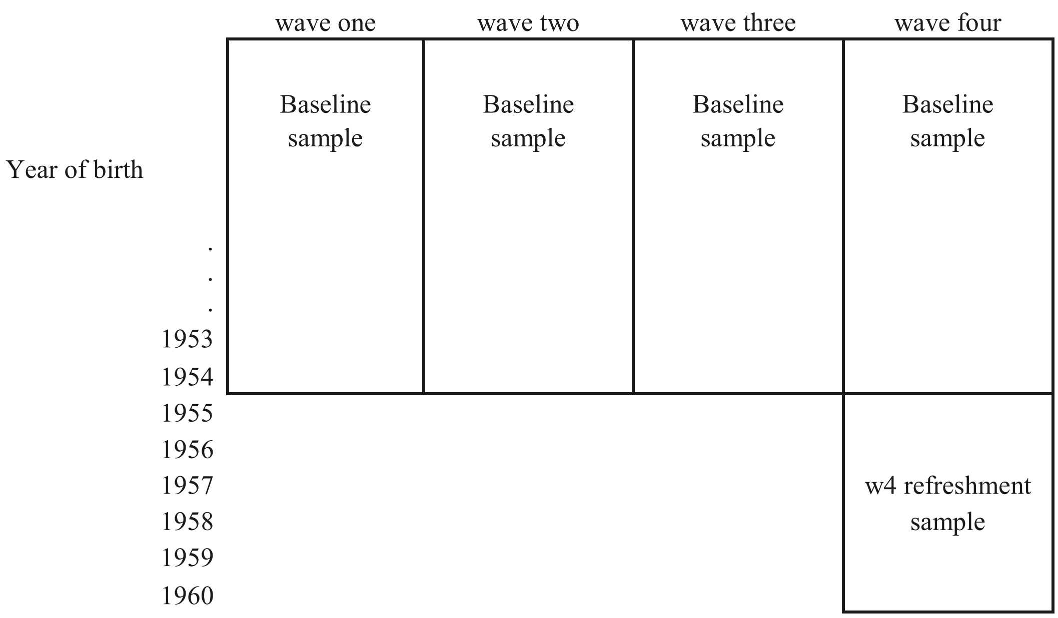


Figure 8.1A Relationship between samples and waves

Example C: Countries which had a refreshment sample at wave two and full age range refreshment at wave four

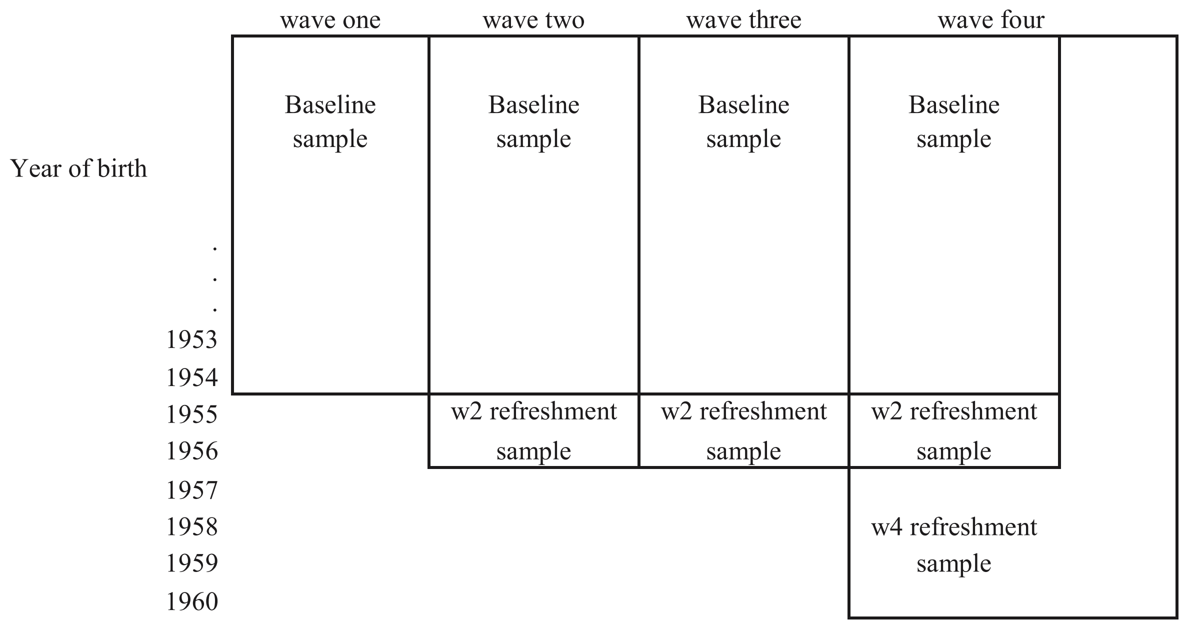


Figure 8.1B Relationship between samples and waves

8.4 How was the sample design controlled?

The sample design requirements for SHARE are set out in the “SHARE Sampling Guide” and were widely disseminated and discussed with the country teams. For wave 4, SHARE created for the first time a “Sampling Panel”, consisting of four international experts on survey sampling (the authors of this chapter), all with experience of cross-national comparative surveys. The role of the panel was to discuss the proposed sample designs for both baseline and refreshment samples with each country team, to suggest improvements, and ultimately to assess the acceptability of the design. One panel member was assigned to each country to provide technical assistance during the entire sampling process. This approach gave country teams access to expert assistance in developing efficient and appropriate designs and also increased the likelihood of consistent decisions being made across countries. The process was generally deemed to have been a success, though of course a limitation is that it was not possible to influence the design of samples that had already been selected at earlier waves.

8.5 National variations in design

In developing national sampling designs, the first task was to find the most suitable sampling frame in each country. The sampling experts and the national country teams were looking for frames with minimum under-coverage and minimum over-coverage, i.e. the most often updated frames from the most trustworthy sources. An important characteristic any candidate frame had to fulfil was the availability of reliable information on age for the frame population since the target population comprised only those persons born in the year 1960 or earlier. If this information was not available from a given frame, a screening procedure had to be applied. The table below shows a summary of sampling frames. More details on sample frames and screening procedures can be found in Appendix 2.

Table 8.1 Description of sampling frames in countries with baseline/refreshment samples in wave four

Country	Description of frame	Units
Austria	List of all dwellings with corresponding p.o. boxes	A
Belgium	National population register	I
Czech Republic	Electoral register	A
Denmark	National population register	I
Estonia	National population register	I
France	The rolling population census	I
Germany	Local population registries	I
Hungary	Population registry of Hungary	I
Netherlands	Refreshment Sample from 26 municipalities	I
Portugal	National Health System register	H
Slovenia	Central register of population	I
Spain	Population register based on census and municipal registers	I
Sweden	Population register NAVET of the Swedish tax authority	I
Switzerland	Population register	I

A-Addresses, H-Households, I-Individuals

Due to privacy or legal restrictions it was not always possible to use the best existing frame in a given country. For example, Austria has a modern, computer-based population register. But this register was and still is (as of late 2012) unfortunately not accessible for survey sampling. On the other hand, SHARE was the first survey that was allowed to use the Swiss population register which is known to be of excellent quality. As a rule the sampling experts did not insist on taking the same frame as in the previous SHARE wave but instead countries were allowed to find the best one. In general, finding suitable sampling frames for sample selection is a very difficult, challenging and time consuming step in cross-national survey sampling. SHARE is no exception to this rule.

The next step was the design of the samples given the frames in each country. Usually the sampling experts recommended a regional stratification scheme to ensure a good representation of different geographical areas of the country. If further relevant characteristics were available on the sampling frame – such as age in the case of population registers – countries were advised to also use them for stratification. As in other survey programs, such as the European Social Survey (ESS) or the Programme for the International Assessment of Adult Competencies (PIAAC), a guiding principle is to design sampling plans which yield minimum variation in inclusion probabilities and a minimum amount of clustering. This is because these two design characteristics directly influence the precision of estimates based on the underlying samples. Finding a sampling frame which allows for such a design is, however, not always possible.

Such a scenario applies, for example, if a country team only has access to a list of households and an eligible person has to be selected from all eligible target persons of a sampled household. In this case, variation in inclusion probabilities cannot be avoided. This procedure introduces a so called “design effect due to unequal inclusion probabilities” ($Deff_p$). Other studies (e.g. ESS) have shown that $Deff_p$ usually ranges between 1.20 and 1.25 for designs that involve the random selection of one adult per household, depending on the variation of household sizes in a country. This variation in inclusion probabilities has to be taken into account by a design weight which is just the inverse of the inclusion probability. For SHARE, $Deff_p$ should tend to be smaller than this, as it depends on the distribution of 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 or over or a couple containing at least one person aged 50 or over. In most countries, few households contain more than one age-eligible unit and very few have more than two.

Fortunately, many countries had access to population registers, e.g. Denmark, Slovenia, Switzerland and Germany. In these countries sample designs could be implemented which yielded equal inclusion probabilities for all elements. In Germany, however, SHARE had to use a two-stage clustered sample design as the population registers are locally administered by the municipalities. Therefore, a number of municipalities had to be selected at the first stage and age eligible persons at the second stage. In such a case, an additional component of the design effect emerges. It is the design effect due to clustering ($Deff_c$). Usually, $Deff_c$ is larger than 1 since both the mean cluster size of the primary sampling units (municipalities, in the case of Germany) and the intraclass correlation determine its magnitude. Therefore, by design, the mean cluster size had to be chosen as small as possible and as many primary sampling units as possible had to be selected. This is at odds with the interests of the survey agencies for which an increase in the number of primary sampling units is associated with increased costs.

The refreshment sample of France shall serve as an example of how design weights in part determined the design effect. The following figure shows the distribution of design weights in the French refreshment sample by region (region was a one of two stratification variables in the French sample design). Although the sample design was chosen such that the overall variation in inclusion probabilities would be as small as possible, the household selection still caused variation in inclusion probabilities as Figure 8.2 clearly shows. This lead to a design effect due to unequal inclusion probabilities of 1.33 in France.

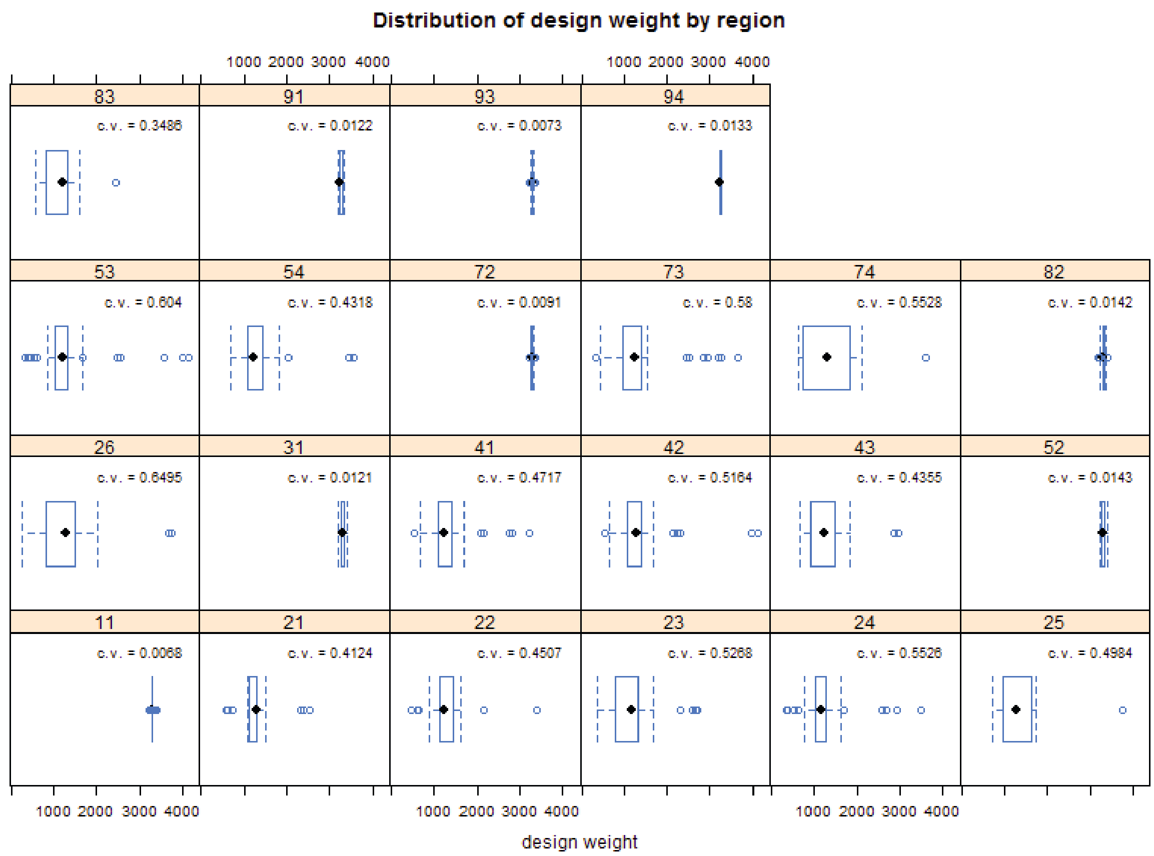


Figure 8.2 Distribution of design weights in the French refreshment sample.

8.6 Sample size and response rates

In contrast to many cross-sectional survey programs SHARE did not define a minimum net sample size (like for example PIAAC does) or a minimum effective sample size (like in the ESS) because the size of the refreshment sample should be determined by the size of the surviving initial panel sample, i.e. the smaller the sample size of the surviving initial panel sample, the larger the size of the refreshment sample should be. The ultimate guideline is to conduct 6000 individual interviews overall, if panel respondents and refreshment respondents are combined at the end of fieldwork. For baseline samples, SHARE has the rule that the net sample size should be as large as possible, given the cost restrictions in the country.

The resulting net sample size was difficult to estimate in advance, mainly due to these reasons:

- In countries where no age information from the frame was available, a screening procedure had to be conducted, i.e. a contact person in the household had to be interviewed how many people belonging to the target population lived in the household. The response rate of these contact persons was difficult to anticipate. An example would be Austria.
- The response rate of the selected persons within the household was difficult to estimate in advance.

- Within a selected household, one age-eligible member plus his/her partner/spouse had to be interviewed in addition. Whether there was a partner/spouse to be interviewed was not known from the frames. Thus, this percentage had to be estimated and evidence from previous waves indicated that it differs between countries. Furthermore, the response rate of the partners/spouses was also difficult to estimate in advance.
- The ineligibility rate, i.e. deficiencies of the frame had to be assessed in advance. If the estimated eligibility rate was too small this reduced – given a fixed gross sample - the resulting net sample size.

Some of these problems are clearly illustrated in the example of the baseline sample size calculation for Slovenia. The following is an extract from the Slovenian sampling design form (see Appendix 2 for details):

"The gross sample size will be n_{gross} = 4.200 (21 primary respondents in 200 PSUs). With a response rate of about 60% and an eligibility rate of 90%, assuming 60% of primary respondents have partners, and assuming 50% response rate of partners, this leads to a net sample size n_{net}=2948 (2268 primary respondents + 680 partners). This net sample means about 15 interviews completed per PSU. To summarize, that means that we start with a gross sample of 4200 individuals from the register, to reach estimated 2948 completed interviews, which include the partners. "

Thus, estimation in advance of the study of the net sample size that would result from any given gross sample size was subject to substantial uncertainty (especially in countries without a frame of individuals) as it relied on several more or less weak assumptions. Details about response rates and retention rates can be found in chapter 10.

8.7 Analysis weights

Sampling design weights, defined as the inverse of the probability of being included in the sample of any specific wave, compensate for unequal selection probabilities of the various sample units. Without such weights it is not possible to obtain unbiased estimators of population parameters of interest. However, even with such weights, estimators are unbiased only under the ideal situation of complete response. Unfortunately, survey data are always affected by unit nonresponse (i.e., eligible sample units fail to participate in the survey because of either noncontact or explicit refusal to cooperate). Such nonresponse occurs at each wave, resulting in panel attrition (i.e., responding units in a given wave of the panel drop out in a subsequent wave). Therefore, estimators constructed using sample design weights alone, and ignoring unit nonresponse and attrition, may be biased (Lessler and Kalsbeek 1992). Although sample design weights are included in the public release of the SHARE data, we strongly discourage users to rely on these weights unless they are used for the implementation of specific statistical methods which account for nonresponse errors in other ways, or for other specific purposes.

The strategy used by SHARE to cope with the potential selection bias generated by unit nonresponse and panel attrition relies on the ex-post calibration procedure of Deville and Särndal (1992). As discussed in Appendix 1, this statistical re-weighting procedure gives calibrated weights which are as close as possible, according to a given distance measure, to the original design weights while also respecting a set of known population totals (the calibration margins). Under certain assumptions about the missing data process, calibrated weights may help reduce the potential selectivity bias generated by unit nonresponse and panel attrition. The key assumption is that, after conditioning on a set of variables (the calibration variables), there is no relation between the response probability and the other key survey variables excluded from the conditioning set. Using the terminology introduced by Rubin (1987) this corresponds to assuming that the process generating missing observations is missing-at-random (MAR). This assumption could be relaxed by considering more sophisticated approaches where the process for the outcome of interest and the response process are estimated jointly (see, for example, De Luca and Peracchi 2012). However, these approaches are generally specific to the research questions under investigation and they require auxiliary information on all eligible sample units. Thus, depending on the purpose of the analysis to be performed, users should decide whether calibrated weights provided in the public release of the SHARE data are enough to compensate for the potential selectivity bias associated with unit nonresponse and panel attrition.

As in the previous waves, the public release of the wave four SHARE data includes calibrated cross-sectional weights to be used in the context of cross-sectional analyses and calibrated longitudinal weights to be used for longitudinal analyses. Since the basic units of analysis can be either individuals or households, both types of 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.

Calibrated cross-sectional weights are defined for the sample of 50+ respondents (either individuals or households) in wave four by ignoring the distinction between longitudinal and refreshment samples. At the individual level, each 50+ respondent receives a calibrated weight that depends on the household design weight and the respondent's set of calibration variables. At the household level, each interviewed household member receives a common calibrated weight that depends on the household design weight and the calibration variables of all 50+ respondents in the same household. Calibrated weights are always computed separately by country in order to match the size of national populations of individuals born in 1960 or earlier. Within each country, we used a set of calibration margins for the size of the target population across 8 gender-age groups (i.e. males and females with year of birth in the classes (-1930], [1931-40], [1941-50], [1951-60]) and across NUTS1 regional areas. For each type of calibrated weight, we also provide a flag variable which is equal to 1 whenever the corresponding calibrated weight is missing. This occurs for respondents younger than 50 years (i.e. age-ineligible partners of an age-eligible respondent), those with missing information on the set of calibration variables (i.e. year of birth, gender and NUTS1 code), and those with missing sampling design weights (i.e., respondents with missing sampling frame information).

Calibrated longitudinal weights differ from calibrated cross-sectional weights in three important respects. First, these weights are only defined for the balanced sample of eligible units who participated in two or more waves of the panel. Second, calibrated longitudinal weights take into account mortality of the original target population across waves. Mortality affects both the sample and the population. Thus, the target population for longitudinal analyses is the original population at the beginning of the time reference period that survives up to the end of period. Third, since the SHARE panel consists of four waves, one can compute thirty different types of calibrated longitudinal weights depending on the selected combination of the waves (i.e., 1-2, 1-3,…,3-4, 1-2-3,…,2-3-4, 1-2-3-4) and the basic unit of analysis (either individuals or households). To simplify the structure of the public release of the data, SHARE provides calibrated longitudinal weights only for the fully balanced panel sample (i.e. the sample of 50+ respondents participating to all waves). These calibrated weights are computed separately by country in order to match the size of the national populations of individuals born in 1954 or earlier that survive up to 2011. We used a set of calibration margins for the size of the target population across eight gender-age groups (i.e. males and females with year of birth in the classes (-1924], [1925-34], [1935-44], [1945-54]) and across NUTS1 regional areas. Mortality is accounted for by subtracting from each population margin the estimated number of deaths between 2004 and 2011. Calibrated longitudinal weights are available at the individual and the household level. Notice that, for the weights at the household level, we only require that there is at least one eligible respondent in each wave. Thus, households with one partner participating in the first wave and the other partner participating in the other waves belong to the balanced sample of households, even if neither partner belongs to the balanced panel of individuals.

For longitudinal analyses based on other possible combinations of waves, users can compute their own calibrated longitudinal weights. To support users in this methodological task, SHARE provides a Stata command called cweight.ado which implements the calibration procedure by Deville and Särndal (1992), a Stata do-file weighting.do which illustrates step-by-step how to compute calibrated longitudinal weights at the individual and the household level, and tables of country specific information needed to compute the population calibration margins.

References

De Luca, G., Peracchi, F. (2012). *Estimating Engel Curves under Unit and Item Non-response*. Journal of Applied Economics

Deville, J-C, Särndal, C-E. (1992). *Calibration Estimators in Survey Sampling*. Journal of
the American Statistical Association, 87(418): 376-382

Kish, L. (1994). *Multipopulation survey designs*. International Statistical Review, 62: 167-186

Lessler, J., Kalsbeek, W. (1992). *Nonsampling Error in Survey*. New York: John Wiley & Sons

Rubin, D.B. (1987). *Multiple Imputation for Nonresponse in Surveys.* New York: Wiley

Appendix 1 The calibration procedure

This appendix provides a formal description of the new theoretical framework used since the third wave of SHARE to compute calibrated weights. Calibrated weights of the first two waves have being also updated through the public release 2.4.0. Additional methodological details on the calibration procedure can be found in Devile and Särndal (1992).

Consider a finite population P={1, …, k, …, N} from which a probability sample $S \subset P$ is drawn according to a given sampling design. Let w_k be the original sampling design weight of the kth sample unit, and assume that only a sub-sample of respondents $R \subseteq S$ agree to participate to the survey. Following Devile and Särndal (1992), calibrated weights w_k^* can be obtained minimizing the sum of the distances

$$\sum_{k \in R} G(w_k^*, w_k)$$

subject to a set of J calibration equations

$$t_x = \sum_{k \in R} w_k^* x_k ,$$

where $x_k = (x_{k1}, \ldots, x_{kJ})$ and $t_x = (t_1, \ldots, x_{kJ})$ are J-dimensional vectors of calibration variables and known population totals, respectively. Before release 2.4.0, the distance function between the original sampling design weight w_k and the calibrated weight w_k^* was arbitrarily taken as a chi-square distance function of the form,

$$G(w_k^*, w_k) = (w_k^* - w_k)^2 / w_k ,$$

On the one hand, this was a convenient choice since this distance function guarantees that calibrated weights exist with probability 1 and they have the following closed form expression

$$w_k^* = w_k \left[1 + \left(t_x - \sum_{k \in R} w_k x_k \right)^T \left(\sum_{k \in R} w_k x_k x_k^T \right) x_k \right].$$

On the other hand, however, it was recognized that this distance function is unbounded and hence it is likely to give problems with the range of feasible values that calibrated weights can take. Depending on the chosen calibration margins, calibrated weights can be indeed negative or extremely large. Negative weights are inadmissible, while extremely large weights may lead to unrealistic estimates of various population domains.

To overcome these theoretical problems of the chi-square distance function, the new version of the calibration procedure is based on a bounded distance function of the following form (case 6 in Devile and Särndal 1992)

$$G(q_k) = (q_k - L) \log\left[(1-L)^{-1} (q_k - L) \right] + (U - q_k) \log\left[(U-1)^{-1} (U - q_k) \right],$$

where $q_k = w_k^* / w_k$, and L and U are constant coefficients such that $L < 1 < U$. If $G'(\cdot)$ is the first partial derivative of $G(\cdot)$ with respect to w_k^* and $F(\cdot)$ is the inverse of $G'(\cdot)$, then one can show that

$$F(v) = \frac{L(U-1) + U(1-L)\exp(Av)}{(U-1) + (1-L)\exp(Av)},$$

with $A = (U-L)/[(1-L)(U-1)]$. As shown by Devile and Särndal (1992), calibrated weights can be then computed in two steps. In the first step, one determines the vector of the Lagrange multipliers $\lambda = (\lambda_1, \ldots, \lambda_J)$ solving the system of first order conditions,

$$\sum_{k \in R} w_k \left[F(x_k^T \lambda) - 1\right] x_k = t_x - \sum_{k \in R} w_k x_k.$$

In the second step, one computes the calibrated weights using the following expression

$$w_k^* = w_k F(x_k^T \lambda).$$

Unlike the chi-square distance function, this distance function guarantees by construction that calibrated weights are bounded between Lw_k and Uw_k. The main drawback is that a solution to the optimization problem may not exist and in any case it depends on the choice of the distance function through the coefficients L and U. To handle this problem we use more than 400 distance function by choosing a grid alternative combinations of L and U. Among the distance function which lead to a solution to the above optimization problem, we then selected the pair (L, U) that gives calibrated weights w_k^* with minimum standard deviation.

Appendix 2 – National Sampling Design Forms[1]

Austria
Refreshment or baseline sample: Refreshment
Survey Institute: IFES GmbH
Country sampling contact: Andreas Bugnar
SHARE sampling expert: Sabine Häder
Reference survey: Finanzielle Lage österreichischer privater Haushalte (OeNB)
Date: 11 January 2011

Target population, Population coverage	All German speaking residents born 1960 or earlier and their spouses / partners at the time of interview independent of the spouse's/partner's age. The target population does not include those living in institutions.
Screening frame (if applicable)	Not applicable
Screening frame problems (if applicable)	Not applicable
Screening design (if applicable)	Not applicable
Remarks	Not applicable
Sampling frame	Stage 1: list of all Austrian Zählbezirke (Enumeration Areas) Stage 2: list of all dwellings with corresponding p.o boxes
Sampling frame problems	In Austria there are no addresses with data regarding persons aged 50+ available, therefore a screening inside the randomly drawn households and oversampling in general is necessary (because persons 50+ do not exist in every household). Due to legal reasons there is no access to the central household register (Zentrales Melderegister).
Sampling design	Stratified two stage probability sample: **First stage:** Random draws without replacement (inside Strata) from 8.745 Zählbezirke (Enumeration Areas). *Zählbezirk: smallest territorial unit of a collection of dwellings, a Zählbezirk contains on average around 450 dwellings* Zählbezirke are stratified according to NUTS 3 regions X sizes of settlement in Bundesland (chart below). Vienna is one NUTS 3 Region and is therefore divided in 23 districts. Sum is 193 Strata. Allocation of sample points is done proportional to population. **Second Stage:** Random draws without replacement (inside Zählbezirke) from dwellings (p.o. box code), source for p.o. box code is the Austrian Address Information System (Address register, including the numbers of all p.o. boxes), Identification of household corresponding to p.o. box code via companies databases or name bought from adress providers. Inside households with more persons aged 50+ we will choose the youngest within this group to compensate the panel-effect. According to data from Statistik Austria in about one of two households there is a person aged 50+. The regional distribution of the Panel-Households will be taken into consideration within the refreshment-sampling.

[1] Some forms displayed here may have been appended by national teams after they were gathered for publication.

<table>
<tr><td>Remarks</td><td>Overview NUTS 3 Regions in Austria:
<table>
<tr><th></th><th></th><th colspan="9">Population Size of Municipality</th></tr>
<tr><th>Province</th><th>NUTS3 Regions (districts for Vienna)</th><th>-2,000</th><th>-3,000</th><th>-5,000</th><th>-10,000</th><th>-20,000</th><th>-50,000</th><th>- 1 Mio</th><th>+ 1 Mio</th><th>Total</th></tr>
<tr><td>Vienna</td><td>23</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>23</td><td>23</td></tr>
<tr><td>Lower Austria</td><td>7</td><td>7</td><td>7</td><td>7</td><td>7</td><td>6</td><td>6</td><td></td><td></td><td>40</td></tr>
<tr><td>Burgenland</td><td>3</td><td>3</td><td>3</td><td>3</td><td>2</td><td>1</td><td></td><td></td><td></td><td>12</td></tr>
<tr><td>Styria</td><td>6</td><td>6</td><td>6</td><td>5</td><td>6</td><td>3</td><td>1</td><td>1</td><td></td><td>28</td></tr>
<tr><td>Carinthia</td><td>3</td><td>3</td><td>3</td><td>3</td><td>3</td><td>2</td><td>1</td><td>1</td><td></td><td>16</td></tr>
<tr><td>Upper Austria</td><td>5</td><td>5</td><td>5</td><td>5</td><td>5</td><td>3</td><td>2</td><td>1</td><td></td><td>26</td></tr>
<tr><td>Salzburg</td><td>3</td><td>3</td><td>3</td><td>3</td><td>3</td><td>2</td><td></td><td>1</td><td></td><td>15</td></tr>
<tr><td>Tyrol</td><td>5</td><td>5</td><td>5</td><td>4</td><td>4</td><td>3</td><td></td><td>1</td><td></td><td>22</td></tr>
<tr><td>Vorarlberg</td><td>2</td><td>2</td><td>2</td><td>2</td><td>2</td><td>2</td><td>1</td><td></td><td></td><td>11</td></tr>
<tr><td>Total</td><td>57</td><td>34</td><td>34</td><td>32</td><td>32</td><td>22</td><td>10</td><td>6</td><td>23</td><td>193</td></tr>
</table>
</td></tr>
<tr><td>Auxiliary frame data that can be used by SHARE</td><td>Not applicable</td></tr>
<tr><td>Selection probabilities</td><td>to be determined later</td></tr>
<tr><td>Design weights</td><td>to be determined later</td></tr>
<tr><td>Expected individual response rate (for sampling purposes)</td><td>Vienna 50%, rest of Austria 65%

Based on the assumption of delivery of net 4.000 interviews are net ca. 2.500 households (factor for calculation of number of households is based on experience from last SHARE survey) and expected individual response rate (Vienna 50%, rest of Austria 65%).</td></tr>
<tr><td>Target sample sizes</td><td>Gross sample without oversampling: 4.113 households in 395 sample points (8 adresses in Vienna per sample point, 12 adresses rest of Austria per sample point).

Gross sample including oversampling: 8.226 households in 790 sample points (8 adresses in Vienna per sample point, 12 adresses rest of Austria per sample point).</td></tr>
</table>

Belgium
Refreshment or baseline: refreshment
Survey Institute: CELLO, University of Antwerp
Country sampling contact: Karel Van den Bosch
SHARE sampling expert: Peter Lynn
Reference survey:
Date: 27 October 2010

Target population, Population coverage	All residents speaking French or Dutch born 1962 or earlier, and their spouses/partners at the time of interview, living in the Belgian regions Brussels, Wallonia or Flanders. The target population does not include individuals living in the German-speaking communities in the east of Belgium (0.6% of the population). The target population does include individuals living in 'collective households', i.e. homes for the elderly.
Screening frame (if applicable)	No screening is necessary in Belgium
Screening frame problems (if applicable)	
Screening design (if applicable)	
Remarks	
Sampling frame	Stage 1: List of all municipalities in Wallonia, Flanders and Brussels (excluding the German-speaking municipalities); Stage 2: National Register of all persons resident in Wallonia, Flanders and Brussels.
Sampling frame problems	Persons do not always actually live at the registered address. Register information might be outdated since there is a time-lag between moving house and registering the new address.
Sampling design	**Two stage sampling of Refreshment sample born in 1960 or earlier.** ***Stage 1: Selection of municipalities.*** Data on the number of persons born in 1960 or earlier by municipality are used. These data are provided by Statistics Belgium. Municipalities are distributed across 11 strata, according to region and size, as follows: Brussels (capital region): one stratum. Flanders: five strata. The big cities Antwerp and Gent each form one stratum, the other municipalities are distributed across three strata, such that these strata have equal size in terms of the target population (born 1960 or earlier). The criterion for assigning municipalities to one of the three strata was its size in terms of the target population. Wallonia: five strata. The big cities Liège and Charleroi each form one stratum, the other municipalities are distributed across three strata, such that these strata have equal size in terms of the target population (born 1960 or earlier). The criterion for assigning municipalities to one of the three strata was its size in terms of the target population. The target sample sizes are distributed across the strata in proportion to the size of the strata in terms of the target population (born 1960 or earlier). This gives the target sample size n_s within each stratum s. The number of municipalities to be selected within each stratum m_s (except the four one-city strata) is determined by the formula: $m_s = \text{ROUND}(n_s / 25)$, where 25 is the target cluster size. Within each of the strata (except in the four one-city strata) municipalities were selected proportional to size in terms of the target population, and

	without replacement. ***Stage 2: Selection of individuals / couples / households within each municipality.*** The sampling frame (national register) has information on the age, sex and 'relation to the reference person' of all individuals within each household. The program to select persons must be written and executed by the programmers of the National Register, which charges the costs of this programming to us. Given, therefore: - the need for a simple sampling procedure (to reduce costs and errors) - the information available - and the goal of an EPSEM sample (Equal Probability of Selection Probability) we devised the following method: - within each municipality, persons within the target population are sampled by simple random sampling (without replacement) - the spouse / partner of each selected person is identified, and his/her age is determined - if the spouse/partner belongs to the target population, the (original) person is marked as 'target-couple', otherwise she/he is marked as 'target-single'. - from the group marked as 'target-couple', half are deleted from the sample by simple random sampling - the selected persons, as well as their spouse/partners, if they belong to the target population, are retained as the final sample. It would have been more efficient (from a statistical point of view) to divide in advance the whole population within each municipality into two strata, 'target-couples' and 'target-singles', but this appeared not to be feasible. Two stage sampling of Refreshment sample: - born in 1957 - 1960 living in Wallonia or Brussels. n = 200. - born in 1955 - 1960 living in Flanders. n = 450. These individuals are selected within the municipalities selected for the original (2004) sample. The target sample numbers will be distributed across municipalities proportional to the size of the municipalities in terms of the target population (born in indicated age bracket). Selection of individuals will proceed in the same way as for the general refreshment sample.
Remarks	Given the limited budget for the survey, we minimize travel costs by using clusters of at least 25 individuals.
Auxiliary frame data that can be used by SHARE	Information on the structure of the population can be found at the website of the national statistical institute. These numbers can be useful when calculating calibrated weights since one can then take into account not only the population size, but also the age, sex and marital status distribution in the total population.
Selection probabilities (sampling plus screening, if applicable)	
Design weights	
Target response rate (for sampling purposes)	The minimum response rate is estimated to be 40% (based on refreshment sample response rates of previous waves).
Target sample size	The gross sample size equals 4422 individuals. Out of these 4422 individuals

	1654 persons are to be interviewed in the Dutch-speaking part of Belgium (Flanders), 2768 in the French-speaking part of Belgium (Wallonia and Brussels). The target net sample size is 2800 refreshment interviews in Belgium as a whole.

Switzerland
Refreshment or baseline sample: Refreshment sample
Survey Institute: Link Institute
Country sampling contact: Bryce Weaver
SHARE sampling expert: Sabine Haeder
Reference survey:
Date: 11 February 2011

Target population, Population coverage	Persons living in private homes will be considered. The main refreshment sample will be randomly selected from individuals born in 1960 or earlier. The supplementary or corrective sample will be selected amongst those born in the years 1957 - 1960, to compensate for their ineligibility in the previous refreshment sample.
Screening frame (if applicable)	Not applicable
Screening frame problems (if applicable)	Not applicable
Screening design (if applicable)	Not applicable
Remarks	
Sampling frame	The sampling frame is the Stichprobenrahmen für Personen- und Haushaltserhebungen (SRPH) managed by the 'Bundesamt für Statistik' (BFS). Due to concerns for the protection of data, the BFS will do the actual selection of the samples. The SRPH is a registry of all residents in Switzerland compiled from communal registries. As birth year is available in this registry, the samples will be randomly selected (in general without replacement) directly from the eligible population. This alleviates the need for screening, the saved resources will be applied toward increasing the effective numbers. For all individuals, the address is known in the registry. The selected individuals will be cross-referenced with telephone registries by the BFS. The addresses of all individuals will be delivered and, when legal by Swiss law to do so, the cross-referenced telephone numbers will be as well. The field work will be done by the Link Institute (LINK) in Switzerland.
Sampling frame problems	The biggest weakness with the SRPH is that it is new and currently untested. This is making it hard to predict response rates for SHARE. Other problems with the SRPH's relative youth is that certain, otherwise useful variables, are not yet reliable within some communities. This eliminates some alternative sampling options that would reduce the variance in selection probabilities. The main intrinsic weakness is that family relations are not known, and (in our context) that the variable for marital status is the legal one and not the co-residence indicator we use to define the partner. Given the uncertainties of this sampling frame there are several unknowns that are difficult to estimate reliably. This uncertainty is taken into account when developing the sampling plan. The main strategy is to include packets that can be released if certain response rates are not achieved.
Sampling design	Main refreshment sample: stratified one-stage random sample
Remarks	Because of the uncertainties about the frame, the two refreshment samples are divided into a certain number of packets. The first packet will be relatively larger than the subsequent reserve packets. A reserve packet is released when the response rate of the selected individual in the first packet falls below the threshold given in the tables (column “primary RR” in tables 8.2 and 8.1 in the document attached). The first threshold is set above what is believed to be attainable (as the first packet is automatically and is not considered a reserve packet).
Auxiliary frame data that can be used by SHARE	Not applicable

Selection probabilities (sampling plus screening, if applicable)	Not applicable
Design weights	to be determined later

Target response rate (for sampling purposes)	66.2% - 34.3%
Target sample size	Corrective sample: We wish to select a number of respondents, n_{corr} , that will give us the same order of magnitude for the weights between the original sample and this one. To do this, we estimate the number of individuals that are 54+ versus the number that are 50 - 53 by using compiled 2009 data from the official 'Statistique de l'état annuel de la population' (ESPOP). The respective numbers (needed only for a ratio) are 2220493 individuals 54+ and 439666 individuals 50-53. Using the number of projected respondents from the original sample, we seek that ncorr satisfies $n_{corr}/1100 = 439666/2220493$. $n_{corr}= 218$, which we round to the nearest 10 giving $n_{corr}=220$ Main sample: $n_{main}= 2300 - n_{corr} = 2080$
Remarks	The total number of interviews that can be conducted, given budget constraints, in both of the refreshment samples, is n=2300. The rate at which a responding individual will have a responding partner, is 0.33, the eligibility rate is 0.95.

Czech Republic
Refreshment or baseline sample: Baseline and Refreshment
Survey Institute: SC&C Ltd.
Country sampling contact: Pavlina Varutti, Michal Svoboda
SHARE sampling expert: Giuseppe De Luca
Reference survey:
Date: 07 December 2010

Target population, Population coverage	All households with at least one Czech speaking member born 1960 or earlier. All Czech speaking residents born 1960 or earlier and their spouses/partners at the time of the interview independent of the spouse's/ partner's age.
Screening frame (if applicable)	The sampling frame provides only information on the address of residence. A preliminary screening phase in the field is then needed to assess age-eligibility of the sampled units.
Screening frame problems (if applicable)	Not applicable
Screening design (if applicable)	Not applicable
Remarks	Individuals living in institutions for the elderly are excluded from the target population.
Sampling frame	The sampling frame is a list of all electoral districts in Czech Republic (Czech Statistical Office, 2009) plus a list of households/addresses in the selected electoral districts.
Sampling frame problems	The electoral register does not cover people living in institutions (homes for elderly, prisons or similar institutions), nationals who have lost their voting rights and non-citizens.
Sampling design	Czech republic is one of the countries who jointed SHARE in the 2006 wave of the study. The sample from the 2006 wave is a representative sample of the population born 1956 or earlier. It includes a main sub-sample of 4171 households and a vignette sub-sample of 2004 households. Both sub-samples were drawn using a three-stage sampling with selection of electoral districts in the first stage, selection of households/addresses in the second stage and screening for age-eligibility in the third stage. In the first stage, the 12466 electoral districts of Czech republic were classified in 21 strata by using the non-empty combinations of NUTS2 regional code (8 regions) and size of the municipality (3 groups: regional, middle and small municipalities).[2] After performing a preliminary factor analysis using the available and relevant socio-political information, the electoral districts of each stratum were ordered on the basis of their factor scores and then selected by systematic sampling with a fix step. The number of districts selected in each stratum was proportional to the total number of electoral districts, which was in turn strongly correlated with the size of the population in each stratum. In the second stage, a sample of about 40 households was drawn by simple random sampling within electoral each district selected in the first stage. Of these, about 27 households were randomly assigned to main sample and the remaining to the vignette sample. In few electoral districts where the size of the population was lower than 40 households, the overall district was included into the sample. In the third stage, a preliminary screening phase in the field was conducted by the interviewers to assess households with at least one individual born 1956 or earlier. All age-eligible household members, plus their spouses/partners independent of age, were considered to be eligible for the SHARE interview. The sample of the 2008 wave is just a follow-up of the sample from the 2006 wave and it does not include any new refreshment sample. The sample of the 2010 wave is a representative sample of the population born 1960 or earlier. In addition to the main and the vignette sub-samples from the

[2] Prague and Central Bohemia were classified in one and two strata respectively.

	2006 wave, it includes a new refreshment sub-sample of 8376 households drawn using a three-stage sampling design similar to that adopted in the 2006 wave. In the first stage, the 13194 electoral districts of Czech republic were classified in 23 strata by using the non-empty combinations of NUTS2 regional code (8 regions) and size of municipality (3 groups: regional, middle and small municipalities).[3] Within each stratum, electoral districts were ordered on the basis of their factor scores and selected by systematic sampling with a fix step. The number of districts selected in each stratum was again proportional to the total number of electoral districts. In the second stage, a sample of about 70 households/addresses was drawn by simple random sampling from each electoral each district selected in the first stage. In few electoral districts where the size of the population was lower than 70 households, the overall district was included into the sample. In the third stage, a preliminary screening phase in the field was conducted by the interviewers to assess households with at least one individual born 1960 or earlier. For households with more than one age-eligible person, the target person to be interviewed plus his/her partner/spouse (independent of age) were selected randomly by the Sample Management System. The other household members were not interviewed, even if age-eligible.
Remarks	Selection probabilities can only be computed for households completing the screening phase.
Auxiliary frame data that can be used by SHARE	None
Selection probabilities (sampling plus screening, if applicable)	Let $\pi_{ih}(w)$ be the probability of including person *i* of household *h* into the sample of wave *w* and denote by $\pi_h(w)$ the same probability for the whole household *h*. The probability of being included in the joint sample (i.e. main plus vignette) from the 2006 wave is $$\pi_{ih}(w=1)=\pi_h(w=1)=\frac{d_t}{D_t}\frac{a_{dt}}{A_{dt}}I\left(n_h^{56}>0\right)$$ where d_t and D_t are the target number of districts and the total number of districts in stratum t, a_{dt} and A_{dt} are the target number of addresses and the total number of addresses in district d of stratum t, n_h^{56} is the number of household members born 1956 or earlier, and $I(A)$ is the indicator function of the event A. Notice that, in the 2006 wave, all age–eligible household members were considered to be eligible for the SHARE interview. The probability of being included in the refreshment sample from the 2010 wave is: $$\pi_{ih}^*(w=2)=\pi_h^*(w=2)=\frac{d_t}{D_t}\frac{a_{dt}}{A_{dt}}\frac{n_{ih}^{60}}{n_h^{60}}I\left(n_h^{60}>0\right)$$ where n_h^{60} is the number of household members born 1960 or earlier, and $n_{ih}^{60}=1$ if the household member selected during the screening phase is single and $n_{ih}^{60}=2$ otherwise.
Design weights	Design weights in wave w are computed as the inverse of the underlying selection probability: $$W_{ih}(w)=W_h(w)=\frac{1}{\pi_h(w)}$$

[3] In this case, Prague and Central Bohemia were classified in three and two strata respectively.

Target response rate (for sampling purposes)	The expected household response rate is 58%.
Target sample size	The target sample size is 6000 interviews. The estimated number of longitudinal interviews is 1600, the expected response rate is 58%, the expected share of households with at least one individual born 1960 or earlier is 60% and 2 interviews are expected from about 50% of households. Thus, the size of the gross refreshment sample in wave 4 is 8376.

Denmark
Refreshment or baseline sample: Refreshment (Cohort)
Survey Institute: SFI-Survey
Country sampling contact: Karen Andersen-Ranberg
SHARE sampling expert: Peter Lynn
Reference survey: SHARE Wave 2
Date: 13 January 2011

Target population, Population coverage	All persons resident in Denmark in January 2011 and born in 1957-1960
Screening frame (if applicable)	Not applicable
Screening frame problems (if applicable)	Not applicable
Screening design (if applicable)	Not applicable
Remarks	
Sampling frame	Danish Population Register
Sampling frame problems	No serious problems. Some persons on the register (12%) are excluded from the frame as they have registered not to take part in research (so, some undercoverage).
Sampling design	Simple random sample of n = 563 persons (gross sample)
Remarks	
Auxiliary frame data that can be used by SHARE	
Selection probabilities (sampling plus screening, if applicable)	Equal probability, using same overall sampling fraction as wave 2 sample.
Design weights	W = 1.0 (relative to w2 sample)
Target response rate (for sampling purposes)	70%
Target sample size	390 interviewed (net sample)

Germany
Refreshment or baseline sample: Refreshment
Survey Institute: Infas (Institut für angewandte Sozialwissenschaft GmbH)
Country sampling contact: Birgit Jeske (Infas) / Annelies Blom (MEA)
SHARE sampling expert: Sabine Haeder
Reference survey: SHARE Wave 1
Date: 15 December 2010

Target population, Population coverage	All German speaking residents born 1960 or earlier and their spouses / partners at the time of interview independent of the spouse's/partner's age. The target population does not include those living in institutions.
Screening frame (if applicable)	Not applicable (no screening is necessary in Germany)
Screening frame problems (if applicable)	Not applicable
Screening design (if applicable)	Not applicable
Remarks	Not applicable
Sampling frame	Stage 1: List of all German municipalities Stage 2: Municipal population register
Sampling frame problems	Population figures used in the first sampling stage date from 31st December 2008. The proportion of persons living with a spouse or partner has been estimated fromSHARE wave 1. The municipal list of residents might include people who have moved away, but never informed the municipality about their move (especially if people moved abroad). The census test showed an over-/undercoverage of this frame of about 2%. Whether there is a partner/spouse to be interviewed is not known from the frame.
Sampling design	Stratified two-stage probability sampling Stratification: districts × regional size categories; 1,460 strata The data basis for the resident population will be provided by the Federal Statistical Institute. - Stage 1: Selection of 200 municipalities (PSUs). The municipalities are selected with probability proportional to the population size of the community (aged 50+ at 31st December 2008). This allocation ends up with 219 sample points since large cities have more than one sample point. The allocation is done by a controlled rounding procedure (Cox 1987). - Stage 2: In each of the sampling points, an equal size of individuals (44 per sample point, 9 born 1957-1960 and 35 born 1956 and earlier) will be selected (gross sample size = 9,636) from the local population register. 44 addresses per sample point should end up, with a response rate of about 31% and an ineligible rate of 10% in 4000/219= 18.3 interviews per sample point. Assumed is a factor of 1.5 interviews per individual address, i.e. additional interviews with a spouse or partner for 50% of the sampled individuals. Oversampling age cohort 1957-1960: In wave 1 individuals born 1954 and earlier were sampled. In wave 2 the refreshment sample oversampled persons born 1955 and 1956. In wave 4 we therefore oversample those born between 1957 and 1960. 14.62% of the population fall into this age bracket. In wave 1 the gross sample contained 3050 persons drawn from the register; in wave 2 the refreshment gross sample contained 1000 persons drawn from the register. In wave 4 we sample a total of 44*219=9636 individuals from the register, 35*219=7665 persons born 1956 and earlier (79.5%) and additonal 9*219=1971 persons born 1957-1960 (20.4%).

Remarks	MEA receives the full gross sample (including all names and addresses) before the start of fieldwork. This information was used to conduct checks on the contacting and interviewing procedures of the interviewers. SHARE Germany conducted respondent incentives, interviewer training and biomarker experiments. The gross sample was used to allocate the experimental groups.
Auxiliary frame data that can be used by SHARE	For the selected individuals of the gross sample sex, age and in some municipalities nationality. In addition, regional indicators.
Target sample sizes	Gross sample drawn from the register size n_gross_reg=44*219=9636 Expected gross sample size with partners n_gross=9636*1.5=14454 Expected response rate: 31% Ineligible rate: 10% n_net_reg=9636*0.31*0.9=2689 n_net=14454*0.31*0.9=4033 i.e. 4033/219=18.4 interviews per sample point (12.3 persons drawn from register and 6.1 partner/spouses) Oversampling (within above sample): From 44 individuals drawn in each sample point 9 have to be born 1957-1960 n_gross_reg_over=9*219=1971 Expected gross oversample size with partners n_gross=1971*1.5=2957 Expected response rate: 31% Ineligible rate: 10% n_net_reg_over=1971*0.31*0.9=550 n_net=2957*0.31*0.9=825 i.e. 825/219=3.7 interviews per sample point (2.5 persons drawn from register and 1.2 partner/spouses)

Estonia
Refreshment or baseline sample: baseline
Survey Institute: National Statistical Office of Estonia
Country sampling contact: Enn Laansoo Jr, Julia Aru
SHARE sampling expert: Annelies Blom
Reference survey:
Date: 28. July 2010

Target population, Population coverage	All Estonian speaking residents born 1960 or earlier and their spouses/ partners at the time of interview independent of the spouse's/ partner's age. Those living in institutions are not included.
Screening frame (if applicable)	Not applicable (no screening needed)
Screening frame problems (if applicable)	Not applicable
Screening design (if applicable)	Not applicable
Remarks	Not applicable
Sampling frame	Population Register. The frame includes all registered residents as of July 2010 born in 1960 or earlier. Persons with imprecise address are not included (ca. 1,2%).
Sampling frame problems	• The address on which an individual is registered is not always the address where the person lives. • The sampling frame does not include telephone numbers. They have to be found using various directories. • No frame information about household size.
Sampling design	Stratified sampling with simple random sampling of individuals within strata was used. Stratification was done by gender and year of birth. (Table 8.1 below) Within each gender-age stratum records are sorted by region to get better geographical allocation
Remarks	Prior to fieldwork sample will be double-checked with deaths register to exclude any possible deaths happened after sampling.
Auxiliary frame data that can be used by SHARE	Sex, age, address/region, number of persons aged 50+ living at the same address.

Table 8.1. Sample and population size by stratum

Gender	Year of birth	Sample size (n_t)	Population size (N_t)
Male	- 1930	133	14546
Male	1931-1940	358	39174
Male	1941-1950	525	57509
Male	1951-1960	746	81610
Female	- 1930	410	44865
Female	1931-1940	686	75083
Female	1941-1950	747	81773
Female	1951-1960	895	97934

Selection probabilities (sampling plus screening, if applicable)	Let π_{ih} be the probability to include person *i* in household *h* into the sample and π_h the same probability for the whole household *h*. Note that here and after by household we mean a couple of a person selected from register and his/her spouse/partner (just single selected person in case he/she doesn't have a spouse/partner). So any other age-eligible persons living together with those two are not considered as part of their household. Let selected person belong to stratum *a* and his spouse/partner to stratum *b*. Recognising that strata are large and that any individual in a household has the same inclusion probability as its household, we have $\pi_{ih} = \pi_h = \begin{cases} \frac{n_a}{N_a}, \text{if no partner or partner not age-eligible;} \\ \frac{n_a}{N_a} + \frac{n_b}{N_b} - \frac{n_a}{N_a}\frac{n_b}{N_b}, \text{if partner is age-eligible} \end{cases}$
Design weights	$w_{ih} = w_h = 1/\pi_h$
Target response rate (for sampling purposes)	60% (including frame errors)
Target sample size	Target sample size is 3500 interviews. Expected response rate is about 60-65% and 2 interviews are expected from about 30% of households. Thus gross sample of 4500 persons is ordered from the register (plus ca 10% reserve to cover lower response rate if needed).

Spain
Refreshment or baseline sample: Refreshment
Survey Institute: TNS-Demoscopia/Instituto Nacional de Estadística (INE)
Country sampling contact: Laura Crespo, Pedro Mira
SHARE sampling expert: Giuseppe De Luca
Reference survey:
Date: 07 December 2010

Target population, Population coverage	All households with at least one Spanish speaking member born 1960 or earlier. All Spanish speaking residents born 1960 or earlier and their spouses/partners at the time of the interview independent of the spouse's/ partner's age.
Screening frame (if applicable)	Not applicable (no screening needed)
Screening frame problems (if applicable)	Not applicable
Screening design (if applicable)	Not applicable
Remarks	
Sampling frame	The sampling frame is a list of all census sections by municipality (in total some 33000) plus a population register of individuals born 1960 or earlier based on census and municipal registers managed by the National Statistical Office (INE).
Sampling frame problems	Dwellings with more than 20 individuals are removed from the sampling frame, so prisons and similar institutions do not appear. Small institutions for the elderly could instead be on the list. The sampling frame does not include information on household size and telephone numbers.
Sampling design	The sample of the 2004 wave is a representative sample of the population born 1954 or earlier. It includes a main sub-sample of 2849 individuals and a vignette sub-sample of 760 individuals. Both sub-samples were drawn using a two-stage sampling with selection of census sections in the first stage and selection of age-eligible individuals in the second stage. In the first stage, municipalities were classified in 7 strata on the basis of their population size. A stratified sample of 328 census sections was drawn using, within each stratum, systematic sampling with a random start and inclusion probabilities proportional to the population size of each census section. Of these, 259 census sections were assigned to the main sub-sample and the remaining were assigned to the vignette sub-sample. In the second stage, a sample of 11 age-eligible individuals was drawn using systematic sampling with a random start from each census section selected in the first stage. The sample of the 2006 wave is a representative sample of the population born 1956 or earlier. In addition to the two sub-samples from the 2004 wave, it includes a refreshment sub-sample of 506 individuals. The sampling design is similar to that adopted in the 2004 wave. In the first stage, a stratified sample of 46 census sections was drawn using a systematic sampling with a random start and inclusion probabilities proportional to the population size of each census section. In the second stage, a sample of 11 age-eligible individuals (of which 6 born 1954 or earlier and 5 born between 1955 and 1956) was drawn using systematic sampling with a random start from each census section selected in the first stage. Overall, the refreshment sub-sample from the 2006 wave includes 276 individuals born 1954 or earlier and 230 individuals born between 1955 and 1956. The sub-sample of 276 individuals born 1954 or earlier was entirely assigned to the vignette refreshment sub-sample. Of the 230 individuals born between 1955 and 1956, 173 were randomly assigned to the main refreshment sub-sample and 57 to the vignette refreshment sub-sample.

	The sample of the 2008 wave is just a follow-up of the sample from the 2006 wave and it does not include any new refreshment sample. The sample of the 2010 wave is a representative sample of the population born 1960 or earlier. In addition to the two sub-samples from the 2004 wave and the refreshment sub-samples from the 2006 wave, it includes a new refreshment sub-sample of 2131 individuals drawn using a sampling design similar to that adopted in the previous waves. The sample of primary sampling units consists of 118 census sections. In the second stage, a sample of 18 age-eligible individuals (of which 14 born 1956 or earlier and 4 born between 1957 and 1960) was drawn by systematic sampling with a random start from each census section selected in the first stage. Overall, the refreshment sample from the 2010 wave includes 1652 individuals born 1956 or earlier and 472 individuals born between 1957 and 1960.
Remarks	Selection probabilities can only be computed for responding households.
Auxiliary frame data that can be used by SHARE	Gender, year of birth, and province.
Selection probabilities (sampling plus screening, if applicable)	Let $\pi_{ih}(s;w)$ be the probability of including person *i* of household *h* into the sub-sample *s* of wave *w* and denote by $\pi_h(s;w)$ the same probability for the whole household *h*. The probability of being included in the sample from the 2004 wave is equal to the joint probability of being included in either the main or the vignette sub-samples. Assuming that the list of individuals adopted in the second stage of the sampling design was in random order, the probability of being included in sub-sample j (with $j=1$ for the main sub-sample and $j=2$ for the vignette sub-sample) is given by $\pi_{ih}(w=1,s=j)=\pi_h(w=1,s=j)=1-\left[1-n_t^j\frac{N_{ct}}{N_t}\frac{n_{ct}^{j,54}}{N_{ct}^{54}}\right]^{n_h^{54}}$ where n_t^1 and n_t^2 are the numbers of census sections drawn in stratum t for the main and the vignette sub-samples, N_{ct} is the total population size of census section c in stratum t, N_t is the total population size of stratum t, $n_{ct}^{1,54}=n_{ct}^{2,54}=11$ is target sample size of the second stage, N_{ct}^{54} is the size of the population born 1954 or earlier in census section c of stratum t, and n_h^{54} is the number of household members born 1954 or earlier. By treating the census sections of the main and the vignette sub-samples as drawn simultaneously, the selection probability for the joint sample of the 2004 wave is given by $\pi_{ih}(w=1)=\pi_h(w=1)=1-\left[1-(n_t^1+n_t^2)\frac{N_{ct}}{N_t}\frac{n_{ct}^{1,54}}{N_{ct}^{54}}\right]^{n_h^{54}}$ The probability of being included in the refreshment sub-sample from the 2006 wave ($j=3$) is equal to $\pi_{ih}(w=2,s=3)=\pi_h(w=2,s=3)$ $=1-\prod_{i=1}^{n_h^{56}}\left[1-n_t^3\frac{N_{ct}}{N_t}\left(I(A_i^{54})\frac{n_{ct}^{3,54}}{N_{ct}^{54}}+I(A_i^{55-56})\frac{n_{ct}^{3,55-56}}{N_{ct}^{55-56}}\right)\right]$ where n_h^{56} is the number of household members born 1956 or earlier, n_t^3 is the number of census sections drawn in stratum t for this refreshment sub-sample,

	$I\left(A_i^{54}\right)$ and $I\left(A_i^{55-56}\right)$ are binary indicators for individuals born 1954 or earlier and between 1955 and 1956, $n_{ct}^{3,54}=6$ and $n_{ct}^{3,55-56}=5$ are the target sample sizes adopted in the second stage for individuals born 1954 or earlier and between 1955 and 1956, and N_{ct}^{55-56} is the size of the population born between 1955 and 1956 in census section c of stratum t. The probability of being included in the refreshment sub-sample from the 2010 wave ($j=4$) is equal to $\pi_{ih}\left(w=4,s=4\right)=\pi_h\left(w=4,s=4\right)$ $=1-\prod_{i=1}^{n_h^{60}}\left[1-n_t^4\frac{N_{ct}}{N_t}\left(I\left(A_i^{56}\right)\frac{n_{ct}^{4,56}}{N_{ct}^{56}}+I\left(A_i^{57-60}\right)\frac{n_{ct}^{4,57-60}}{N_{ct}^{57-60}}\right)\right]$ where n_h^{60} is the number of household members born 1960 or earlier[4], n_t^4 is the number of census sections selected in stratum t for this refreshment sub-sample, $I\left(A_i^{56}\right)$ and $I\left(A_i^{57-60}\right)$ are binary indicators for individuals born 1956 or earlier and between 1957 and 1960, $n_{ct}^{4,56}=14$ and $n_{ct}^{4,57-60}=4$ are the sample sizes adopted in the second stage for individuals born 1956 or earlier and between 1957 and 1960, and N_{ct}^{56} and N_{ct}^{57-60} are the underlying population sizes.
Design weights	Design weights for wave w are computed as the inverse of the underlying selection probability: $W_{ih}\left(w\right)=W_h\left(w\right)=\frac{1}{\pi_h\left(w\right)}$
Target response rate (for sampling purposes)	60% (including frame errors)
Target sample size	The target sample size is 4000 interviews. The estimated number of longitudinal interviews is 2256, the expected response rate is 60%, the expected percentage of non-sample units is 9% and 2 interviews are expected from about 50% of households. Thus, the size of the gross refreshment sample in wave 4 is 2131.

[4] Notice that, according to the fieldwork rule of the 2010 wave, the household members considered to eligible for the interview are the age-eligible sampled person and his/her partner independent of age. Other age-eligible persons living in the same household are not eligible for the 2010 SHARE interview. This fieldwork rule implies that for single and couples with only one age-eligible partner and $n_h^{60}=2$ for couples with two age-eligible partners.

France
Country: France
Refreshment or baseline: Refreshment
Survey Institute: INSEE (Institut National de la Statistique et des Etudes Economiques)
Country sampling contact: VIGLINO Lionel (INSEE) / QUENUM Sylvain (INSEE)
SHARE sampling expert: Matthias Ganninger
Reference survey:
Date: 31 January 2011

Target population, Population coverage	- Eight regions in France: Île-de-France, Rhône-Alpes, Provence-Alpes-Côte d'Azur, Pays de la Loire, Aquitaine, Nord-Pas-de-Calais, Languedoc-Roussillon, Corse. - All individuals born between 1957 and 1960, and their spouses / partners at the time of interview whatever the spouse's / partner's age. - The target population does not include those living in institutions.
Screening frame (if applicable)	From the census, the birth date is used to select the individuals born between 01/01/1957 and 31/12/1960.
Screening frame problems (if applicable)	The Census does not provide the names of the individuals, but only the addresses of the dwellings.
Screening design (if applicable)	- The sample is a draw of dwellings, in which the interviewer will choose one inhabitant born between 1957 and 1960, and his eventual spouse/partner. - A dwelling represents the primary home all through this document. - A dwelling can be sampled if there is at least one inhabitant born between 1957 and 1960.
Remarks	Since the last edition of the panel Share in 2008, INSEE has built a new master sample based on the annual census and with new Interviewer Action Areas. But the interviewers have to return in the previous areas to re-interview the SHARE panel individuals, which represent the largest part of the sample. So it was necessary to keep those previous areas, hence it was not possible to use the current new master sample.
Sampling frame	The rolling population census. The 2009 annual census is itself a sample, with various weights within a large municipality. Nevertheless, each dwelling has the same weight.
Sampling frame problems	- The current census in France is an annual rolling one. And demographic results are built by compiling 5 annual censuses. So, the rolling areas of the census 2009 did not necessarily overlap with the former PUs from the 1999 master sample 1999, from which SHARE W1 and W2 samples were drawn. But it was possible to build an expansion coefficient with the last 2007 compiled results to pass from the dwellings in the rolling areas of census 2009, to the number of all dwellings that composed the PUs 1999.
Sampling design	The sample is drawn in 3-stages from a list of dwellings recorded at the 2009 annual census survey. In the first stage, the dwellings listed are located in primary units (PUs) constituting the Interviewer Action Areas in the former master sample from the 1999 census. <u>Stage 1: the primary units</u> The master sample 1999 is drawn from the 1999 general population census. The French territory is first divided into regions, then in five strata of primary units : - SG0 : rural communities or groupings of contiguous rural communities of 1800 to 3600 dwellings - SG1 : urban communities with less than 20000 inhabitants or groupings of such urban communities with at least 1800 dwellings - SG2 : urban units of 20000 to 100000 inhabitants - SG3 : urban units of more than 100000 inhabitants (except Paris) - SG4 : the urban unit of Paris.

	The PUs are drawn in each of these regional strata proportionally to the number of dwellings they have. (: the drawing procedure ensures an equal repartition between regions). The number of primary units drawn in SG0, SG1 and SG2 are respectively 128, 75 and 93 (i.e. a sampling rate of nearly 9%). For SG3 and SG4, all PUs are kept. Stage 2: the secondary units are the dwellings First, the resampling of the dwellings from census 2009 in order to give them the same weight within each PU. This resampling represents the new sampling frame. Then the dwellings are drawn in a stratified two-step sample, trying to obtain a self-weighted one. A first-step sample (sized 5000), which is drawn in a general population, is used for the coordination with the others surveys' samples. The second step sample (sized 500) is restricted to dwellings with at least one inhabitant born between 1957 and 1960. Stage 3: individuals The CAPI instrument selects a Kish individual born between 1957 and 1960, and his/her eventual spouse/partner.
Remarks	The PUs (and their selection probabilities) represent the demographic situation of France in 1999. The dwellings are drawn in the 2009 rolling annual census survey, and the expansion coefficients are built with the census compiled results 2007. So the final weights of Share should be calibrated.
Auxiliary frame data that can be used by SHARE	At the dwelling level: apartment building/ house, owner/tenant, house's surface, number of inhabitants
Selection probabilities (sampling plus screening, if applicable)	Definitions: - *i* represents one of the eight France regions covered by Share; - *j* represents one of the five strata of PUs ; - *k* represents a PU; - *m* represents a municipality. 1) Stage 1: Selection probability of PU k: π_{ijk} (they are given with the master sample 1999, and they were been calculated like that: $\boxed{\pi_{ijk} = M_{ij}\frac{X_{ijk}^{99}}{X_{ij}^{99}}}$ with: - M_{ij} : number of PUs drawn in the master sample 1999, within the region i and the stratum j ; - $X_{ijk}^{99} = \sum_{m} l_{ijkm}^{99}$: number of dwellings in the UP k (from census 1999) ; - $X_{ij}^{99} = \sum_{k,m} l_{ijkm}^{99}$: number of dwellings in the the region i and the stratum j (from census 1999).) 2) Stage 2: expansion coefficient within the UP k : r_{ijk} As the 2009 annual census 2009 is a rolling one, it does not contain all the municipalities that composed the UP k. So it is necessary to use an expansion coefficient :

$$r_{ijk} = \frac{\sum_m l_{ijkm}^{09}}{\sum_m l_{ijkm}^{07}}$$: the ratio of dwelling's number in the new sampling frame (from 2009 census) within the UP k, to the dwellings' number in all the municipalities that composed the UP k (from the 2007 census compiled results).

expansion coefficient within the region i and stratum j: t_{ij}

As the 2009 annual census 2009 is a rolling one, there are some UPs where no municipality is concerned by 2009 census . So it is necessary to use an expansion coefficient :

$$t_{ij} = \frac{\sum_{k \in 09, m} \frac{l_{ijkm}^{07}}{\pi_{ijk}}}{\sum_{k,m} \frac{l_{ijkm}^{07}}{\pi_{ijk}}}$$: the ratio of the dwellings' weighted sum in the UPs reached by the 2009 census within the region i and stratum j, to the dwellings' weighted sum in all the UPs contained in the region i and stratum j. The two numbers are calculated from the 2007 compiled census results.

2-a) Step 1:

Selection probability of a dwelling in the stratum ijk :

$$f_{ijk} = \pi_{ijk} \times r_{ijk} \times t_{ij} \times \frac{n_{ijk}}{X_{ijk}^{09}}$$

With:

- n_{ijk} : number of dwellings to draw in the stratum ijk ;
- $X_{ijk}^{09} = \sum_m l_{ijkm}^{09}$: number of dwellings in the new sampling frame (from census 2009) within the UP k.

Self-weighted sample:

To impose a self-weighted sample: $\forall i, \forall j, \forall k,\ f_{ijk} = f = \dfrac{5000}{\sum_{i,m} l_{im}^{07}}$

With :
- 5000 : all dwellings drawn in the first-stage sample;
- $\sum_{i,m} l_{im}^{07}$: all dwellings in the 8 regions (from census compiled results 2007).

So it's possible to calculate the size to draw in each stratum ijk :

$$n_{ijk} = \frac{5000 \times X_{ijk}^{09}}{\pi_{ijk} \times r_{ijk} \times t_{ij} \times \sum_{i,m} l_{im}^{07}}$$

	Finally, it is necessary to get a rounded size within each stratum ijk before making a simple random sampling. The rounding process does not modify much the self-weighted propriety. <u>2-b) Step 2:</u> Within this 5000 sized sample, there are 539 dwellings with at least one inhabitant born between 1957 et 1960. Only these 539 dwellings are kept. The final sample is drawn by a simple random sampling in each stratum ijk with the same sampling rate $\frac{500}{539}$. The final dwellings' weight is the product of the weights of the two steps : $P_{ijk} = \underbrace{\left(\frac{1}{\pi_{ijk}} \frac{X_{ijk}^{09}}{n_{ijk}} \frac{1}{r_{ijk}} \frac{1}{t_{ij}}\right)}_{\boxed{\text{Step 1}}} \underbrace{\left(\frac{539}{500}\right)}_{\boxed{\text{Step 2}}}$ The 500 sized sample is nearly a self-weighted one. <u>3) Stage 3 :</u> The CAPI instrument selects one inhabitant within the n^{ind} ones born between 1957 and 1960. The selection probability of an individual at the 3rd stage is $\frac{1}{n^{ind}}$ The probability of living with a spouse or partner for the age group 50+ in France is 0.4 (from census results).
Design weights	The final weight of the dwellings' sample is P_{ijk} For an individual born between 1957 and 1960 : $P_{ijk} \times \frac{n^{ind}}{1}$ For a spouse/partner : $\frac{P_{ijk} \times n^{ind}}{0.4}$
Target response rate (for sampling purposes)	69%
Target sample size	Net sample: 481 individuals (including partners/spouses) Gross sample: 500 persons (living in the 500 dwellings sampled from the 2009 rolling census) + approx. 200 partners/spouses

Hungary
Country: HUNGARY
Refreshment or baseline: BASELINE SAMPLE
Survey Institute: TARKI
Country sampling contact: Gabor Kezdi
SHARE sampling expert: Matthias Ganninger
Reference survey:
Date: 14 November 2012

Target population, Population coverage	The target population is the set of age-eligible individuals with Hungarian residence who speak Hungarian (non-Hungarian speakers are a negligible fraction of all residents). These are individuals who were born before December 31, 1960. The frame includes both institutionalized and non-institutionalized individuals.
Screening frame (if applicable)	Not applicable
Screening frame problems (if applicable)	Not applicable
Screening design (if applicable)	Not applicable
Remarks	Not applicable
Sampling frame	The Sampling frame is the list of individuals in the current population registry. The population registry of Hungary contains the name, address, gender and age of each resident of Hungary. The registry is based on the last census (from year 2001) and is updated by registered births, deaths and migration. It includes residents in private households as well as residents in institutional "households".
Sampling frame problems	There were no sampling frame problems. As the registry is based on the census, it gives us accurate information.
Sampling design	The sampling design is a stratified two-stage procedure, in which the inclusion probabilities are equal across strata.
Remarks	Stratification is by 2 dimensions: NUTS2 region and type of residence (city/town/village). The districts of Budapest (there are 23 of them) are treated as separate "towns" in the sense that the sample is stratified to them as well. The first stage is a sample of cities/towns/villages; the second stage is a sample of individuals (and their spouses). The first-stage sampling is with probability proportional to size (population) of the city/town/village. All large cities (i.e. settlements with an importance weight of 1 and over) are selected (including all districts of the capital, Budapest, separately), and the first-stage inclusion probability is proportional to the population size in the case of the smaller towns and the villages. The second stage inclusion probabilities compensate for the first-stage differences in order to get uniform (ex-ante) sampling probabilities at the household level. Recall that our sampling frame consists of age-eligible individuals. Therefore the sample we'll draw consists of individuals as well. Thus our case fits in group a) in the "SHARE Sampling Guide 2010" document (page 2.). Therefore the sample will consist of people in the original sample (of age-eligible individuals) and their spouses (regardless of their age).
Auxiliary frame data that can be used by SHARE	NUTS3

<table>
<tr><td>Selection probabilities (sampling plus screening, if applicable)</td><td>
<table>
<tr><td>Region</td><td colspan="4">Target sampling probabilities</td></tr>
<tr><td></td><td>Budapest</td><td>Cities</td><td>Smaller towns</td><td>Villages</td></tr>
<tr><td>1</td><td>0.0005533</td><td></td><td>0.0005544</td><td>0.0005524</td></tr>
<tr><td>2</td><td></td><td>0.0005464</td><td>0.0005515</td><td>0.0005570</td></tr>
<tr><td>3</td><td></td><td>0.0005591</td><td>0.0005548</td><td>0.0005506</td></tr>
<tr><td>4</td><td></td><td>0.0005526</td><td>0.0005531</td><td>0.0005546</td></tr>
<tr><td>5</td><td></td><td>0.0005582</td><td>0.0005531</td><td>0.0005555</td></tr>
<tr><td>6</td><td></td><td>0.0005571</td><td>0.0005556</td><td>0.0005526</td></tr>
<tr><td>7</td><td></td><td>0.0005486</td><td>0.0005555</td><td>0.0005513</td></tr>
<tr><td>Total</td><td></td><td>0.0005543</td><td>0.0005541</td><td>0.0005535</td></tr>
<tr><td>overall</td><td></td><td></td><td></td><td>0.0005538</td></tr>
</table>
</td></tr>
<tr><td>Design weights</td><td></td></tr>
<tr><td>Target response rate (for sampling purposes)</td><td>We targeted a response rate higher than 60%. Empirical evidence also suggested that the response rate for Budapest and cities is lower than for smaller towns and villages.</td></tr>
<tr><td>Target sample size</td><td>2000 households, 3000 individuals</td></tr>
<tr><td>REMARKS</td><td>During the fieldwork we realized that the quality of some of our interviews were low. Households with low quality interviews have been substituted in the same NUTS2 region and the same type of residence (in Budapest the same district). For having these new sample members we extended our whole sample. Both stages of our (stratified two-stage) sampling design were repeated again with fewer cases: the first stage was a sample of settlements (the selection of cities, towns and villages) and the second stage was the sample of individuals. Thus our gross sample size was boosted, but the selection probabilities remained the same as in the table above.</td></tr>
</table>

Portugal
Refreshment or baseline sample: Baseline
Survey Institute: GfK
Country sampling contact: Alice Delerue A. Matos
SHARE sampling expert: Giuseppe De Luca
Reference survey:
Date: 17 November 2010

Target population, Population coverage	All households with at least one Portuguese speaking member born 1960 or earlier. All Portuguese speaking residents born 1960 or earlier and their spouses/partners at the time of the interview independent of the spouse's/ partner's age.
Screening frame (if applicable)	Not applicable (no screening needed)
Screening frame problems (if applicable)	Not applicable
Screening design (if applicable)	Not applicable
Remarks	
Sampling frame	The sampling frame is a population register of individuals born 1960 or earlier from the National Health System.
Sampling frame problems	For each unit of the sampling frame there is an address attached. However, the sampling frame does not include the names of individuals using the National Health System. This means that it is not possible to identify the household member originally selected for the interview and the auxiliary sampling frame information does not necessarily refer the person effectively selected for the interview. Because of this problem, the Portuguese sampling frame is treated as a sampling frame of households. For each sampled households, we only know that there should be at least one person age 50+. The address on the health register may not coincide with the address where people live. For example, some people may change address without updating the health register because they want to avoid medical appointments in a different medical centre or with a different doctor. The address (i.e. the 7 digits zip code) is missing for about 5.8 percent of the units originally included in the sampling frame. These units were excluded from the sampling frame because it is not possible to know the region to which they belong. According to some consistency tests, there are no statistically significant differences in the age and sex distributions of the units included and excluded from the sampling frame. The sampling frame may not cover eligible individuals who are not registered on the Nation Health System. The extent of this coverage error is unknown, but it is expected to be very small. The sampling frame includes people living in institutions. The sampling framing does not always contain information on telephone numbers. The sampling framing does not contain information on household size.
Sampling design	Portugal is one of the countries who jointed SHARE in the 2010 wave of the study. The sampling design is a five-stage sampling with selection of 4-digit zip codes in the first stage, selection of parishes in the second stage, selection of 7-digit zip codes in the third stage, selection of addresses in the fourth stage, and screening for age-eligibility in the fifth stage. Details on these five stages are given below: **Stage 1:** Portugal was stratified in 22 sub-regions by using the 20 non-empty combinations of region (7 regions) and size of the 50+ population within the region (3 groups: less than 10000 inhabitants, between 10000 and 20000 inhabitants, and more than 20000 inhabitants), plus Madeira and Azores which

	were treated as separate strata.[5] One 4-digit zip code was then selected by simple random sampling from each sub-region. **Stage 2:** From each 4-digit zip code selected in stage 1, a sample of parishes was drawn with probability proportional to the number of their 7-digit zip codes. Notice that a parish may in general belong to more than one 4-digit zip code. In these cases, we considered only the portion of the parish which belongs to the 4-digit zip code selected in stage 1. **Stage 3:** From each parish selected in stage 2, a sample of 7-digit zip codes was drawn using simple random sampling.[6] **Stage 4:** From each 7-digit zip code selected in stage 3, a sample of (no more than 20) addresses was drawn using systematic sampling with a random start. Overall, the size of the samples drawn at stages 2, 3 and 4 was determined such that the number of addresses in each sub-region was proportional to the size of the corresponding 50+ population. The only exceptions are: (i) the two sub-regions of Madeira and Azores where number of selected addresses is proportional to five times the size of the 50+ population, and (ii) the two sub-regions of "Sul Interior" where number of selected addresses is proportional to two times the size of the 50+ population. **Stage 5:** A screening phase in the field was carried out by the interviewers through the SHARE Sample Management System in order to select randomly the age-eligible household member to be interviewed. The partner/spouse of the selected household member was interviewed independent of age, while the other household members were not interviewed even if age-eligible.
Remarks	
Auxiliary frame data that can be used by SHARE	Date of birth, gender and region.
Selection probabilities (sampling plus screening, if applicable)	Let π_{ih} be the unconditional inclusion probability of individual *i* in household *h* and denote by π_h the same probability for the whole household *h.* We also further denote by *s* an indicator for strata, *z* an indicator for 4-digit zip codes, *p* an indicator for parishes, *t* an indicator for 7-digit zip codes and *a* an indicator for addresses. In stage 1, the inclusion probability for the 4-digit zip code *z* in stratum *s* is given by $\pi_{z\|s} = \frac{1}{Z_s}$, where Z_s denotes the total number of 4-digit zip codes in stratum *s*. In stage 2, the inclusion probability for the parish *p* in (*z*, *s*) is given by $\pi_{p\|zs} = p_{zs}\frac{T_{pzs}}{T_{zs}}$, where p_{zs} is the number of parishes selected in (*z*, *s*), T_{pzs} is the total number of 7-digit zip codes in (*p*, *z*, *s*), and T_{zs} is the total number of 7-digit zip codes in (*z*, *s*). Notice that, for parishes belonging to more than one 4-digit zip code, T_{pzs} refers to the number of 7-digit zip codes of parish *p* which also belong to the 4-digit zip code *z* of stratum *s*.

[5] This stratification resulted in only one empty sub-region, namely "Sul interior" with population size greater than 20000 inhabitants aged 50+.
[6] Of course, 7-digit zip codes are finer geographical partitions of the 4-digit zip codes selected in stage 1.

	In stage 3, the selection probability for the 7-digit zip code t in (p, z, s) is given by $\pi_{t\|pzs} = \frac{t_{pzs}}{T_{pzs}}$, where t_{pzs} is the number of 7-digit zip codes selected in (p, z, s). In stage 4, the inclusion probability for address a in (t, p, z, s) is given by $\pi_{a\|tpzs} = \frac{a_{tpzs}}{A_{tpzs}}$, where $a_{tpzs}=\min(20, A_{tpzs})$ is the number of addresses selected in (t, p, z, s), and A_{tpzs} is the total number of addresses in (t, p, z, s). In stage 5, the inclusion probability of individual i in (a, t, p, z, s) is given by $\pi_{i\|atpzs} = \frac{n_{atpzs}}{N_{atpzs}}$, where N_{atpzs} is the number of age-eligible individuals living in (a, t, p, z, s) and n_{atpzs} is equal to 1 if the age-eligible household member selected during the screening phase is single, and is equal to 2 otherwise. The unconditional inclusion probabilities of individual i and household h can be obtained by multiplying the conditional probabilities of these five stages: $\pi_{ih} = \pi_h = \pi_{i\|atpzs}\pi_{a\|tpzs}\pi_{t\|pzs}\pi_{p\|zs}\pi_{z\|s}$.
Design weights	$w_{ih} = w_h = 1/\pi_h$
Target response rate (for sampling purposes)	The target response rate is 60% at the individual level
Target sample size	The target sample size is 2000 interviews. The expected response rate is 60%, the expected percentage of non-sample units is 10% and 2 interviews are expected from about 50% of households. The size of the gross sample is 2507.

Italy
Refreshment or baseline: Refreshment
Survey Institute: DOXA S.p.A
Country sampling contact: Danilo Cavapozzi
SHARE sampling expert: Peter Lynn
Reference survey: SHARE wave 1 and wave 2
Date: 10 June, 2011

Target population, Population coverage	The target population of individuals consists of all Italian-speaking residents born in 1960 or earlier and their spouses/partners. The target population does not cover individuals who are incarcerated, hospitalized, institutionalized or out of Italy during the whole fieldwork period.
Screening frame (if applicable)	Not applicable (no screening is necessary in Italy)
Screening frame problems (if applicable)	Not applicable
Screening design (if applicable)	Not applicable
Remarks	The sampling design for the wave 4 refreshment sample derives from the one used in the previous SHARE data-collections run in Italy. See Paccagnellla and Bowater (2004), *SHARE: The Italian Sampling Design – Wave 1*, mimeo.
Sampling frame	Stage 1: List of all Italian municipalities Stage 2: List of electoral divisions from the Italian Ministry of Interior Stage 3: Gender specific municipal electoral registers
Sampling frame problems	The electoral registers do not cover people in institutions such as hospitals and nursing homes (unless they officially reside at their old address), nationals who have lost their voting rights (e.g. convicted criminals), non-citizens and does not capture temporary changes of address. Overall, the excluded individuals amount to about 5% of the total Italian population, but a large share is below 50 years of age.

Sampling design	Three-stage sampling Stage 1: Selection of municipalities Municipalities are stratified by population size 50+ as of 2009 (large, medium and small municipalities) and by geographical location (North-West, North-East, Centre, South and Islands). Overall, we define 15 strata.. We select 93 municipalities. The large municipalities included in the gross-sample are the 11 largest municipalities in terms of 50+ population. The remaining 82 medium and small municipalities to select are drawn by simple random sampling without replacement from each stratum. Stage 2: Selection of electoral divisions within municipalities For electoral purposes, municipalities are divided up into smaller regions known as electoral divisions, containing roughly the same number of people eligible to vote. The general aim is to select 4 electoral divisions by simple random sampling without replacement from the divisions in each sampled municipality. For the large municipalities of Rome, Milan, Turin and Naples we select 16, 12, 8 and 8 electoral divisions respectively. These sample size ratios have been used since the wave1 baseline data-collection. They are based on the 50+ population resident in these municipalities. If a municipality is made up of 1, 2 or 4 electoral divisions, all these divisions will be selected. If a municipality is made up of 3 electoral divisions, 2 divisions will be selected. Selection is always made using simple random sampling without replacement. Stage 3: 2-phase sampling of individuals within electoral divisions In each electoral division, electoral registers are gender specific and include all individuals eligible to vote, regardless of their age. It should be noticed that for each individual in the electoral registers we know name, age, gender and the address. The information in the electoral registers is updated on a regular basis (about every 6 months). *First phase* In the first phase we use simple random sampling without replacement to select a sample of 30 males and a sample of 30 females of any age from each electoral division. Finally, individuals aged less than 50 are deleted from the list of individuals sampled. *Second phase* Simple random sampling without replacement is used to select the individuals to include in the gross-sample from the list obtained at the end of the first phase. In general, the gross- sample at the end of this second phase will include 25 individuals (11 males and 14 females) from each selected municipality. In the cases of Rome, Milan, Turin and Naples we will include, respectively, 100 individuals (44 males and 56 females), 75 individuals (33 males and 42 females), 50 invididuals (22 males and 28 females) and 50 individuals (22 males and 28 females) in the gross-sample. Our sampling design is name based. We use the name of the initially-selected elector aged 50 or over as the sample person and then include in the survey the household containing that person. In each household, we will consider eligible for the interview the initially selected elector aged 50 or over and her

spouse/partner regardless of his age. No other household members will be interviewed, even if aged 50 or over.

Throughout this document, the sampling units are then defined as
the initially selected elector if she is single;
the initially selected elector and her spouse if the selected elector has a cohabiting partner.

Oversampling of the cohorts 1957-1960

The refreshment sample has two components: a random sample representative of the population of individuals 50+ in 2010 and the oversampling of the cohorts 1957-1960 needed to keep the overall gross-sample (refreshment + longitudinal) representative of the 50+ population in 2010. Indeed, individuals in these cohort were excluded from the sampling design of previous waves since they were no age-eligible at that time.

Let *R* be the the number of individuals included in the gross- sample at the end of the three-stage sampling, *Y* the number of those born between 1957 and 1960 (it includes the oversampling for these cohorts) and *A* the number of those born between 1956 or earlier.

$$R \quad Y \quad A$$

where

$$Y \quad \frac{P}{1 \quad P} n^{06}_{56} \quad Pn^{10}_{60}$$

$$A \quad (1 \quad P)n^{10}_{60}$$

$$P \quad \frac{N^{06}_{57,58,59,60}}{N^{06}_{57,58,59,60} \quad N^{06}_{56}}$$

n^{06}_{-56} is the number of households in the current longitudinal gross-sample of wave4 (2369);
n^{10}_{-60} is the number of households in the gross-sample representative of the 50+ population in 2010 (i.e. the refreshment gross-sample minus the oversampling);

$N^{06}_{57,58,59,60}$ is the number of individuals born between 1957 and in 1960 living in Italy in 2006;
N^{06}_{-56} is the number of individuals born in 1956 or earlier living in Italy in 2006.

In Italy, *P*=0.13 (calculation based on the numbers of the National Statistical Institute).

To achieve our targets (see below), we estimate *R* to be equal to 2500 (i.e. we draw 2500 individuals according to the three- stage sampling scheme described above):
601 individuals born between 1957 and 1960;
1899 individuals born in 1956 or earlier.

Remarks	Our interviewers are instructed to contact sampling units living in medium and small municipalities who have moved to a new address in the same municipality (whenever the new address is retrieved); sampling units living in large municipalities who have moved to a new address in the same district (whenever the new address is retrieved).
Auxiliary frame data that can be used by SHARE	Gender, year of birth.
Selection probabilities (sampling plus screening, if applicable)	<u>Stage 1: Selection of municipalities</u> Let us define M_s as the number of municipalities in a given stratum s =1,…15. From each stratum s we select m_s municipalities using simple random sampling without replacement. The selection probability for a given municipality m in a stratum s is $P_{m\|s}$ and depends on m_s and M_s. For each stratum s we know both m_s and M_s, then $P_{m\|s}$ can be always calculated. Note that $P_{m\|s}$ is equal to 1 for all municipalities in the strata including large municipalities. The number of municipalities m_s to be drawn in each stratum s is proportional to the share of Italian 50+ living in that stratum. <u>Stage 2: Selection of electoral divisions within municipalities</u> Let us define D_m as the number of electoral divisions in a given municipality m selected at the first stage. From each municipality m we select d_m electoral divisions using simple random sampling without replacement. The selection probability for a given electoral division d in a municipality m is $P_{d\|m}$ and depends on d_m and D_m. This probability is conditional on the selection of the municipality m at the first stage. For each selected municipality m, we know both d_m and

	D_m, then $P_{d\|m}$ can be always calculated. $P_{d\|m}$ =1 if D_m=1, 2 or 4 =2/3 if D_m=3 =4/D_m if D_m>4 <u>Stage 3: 2-phase sampling of individuals within electoral divisions</u> The third-stage selection probability of a sampling unit h included in the electoral division d is named $P_{h\|d}$ and depends on I_d= total number of persons aged 50-53 on the electoral registers in the selected electoral division d in municipality m; J_d= total number of persons aged 54+ on the electoral registers in the selected electoral division d in municipality m; i_d= number of persons aged 50-53 selected in to the gross-sample in municipality m; j_d= number of persons aged 54+ selected in to the gross-sample in municipality m. $P_{h\|d}$ =2i_d/I_d if the sampling unit h consists of two persons aged 50-53; =2j_d/J_d if the sampling unit h consists of two persons aged 54+; =(i_d/I_d)+(j_d/J_d) if the sampling unit h consists of two persons, one aged 50-53 and one 54+; =i_d/I_d if the sampling unit h consists of a single person aged 50-53; =j_d/J_d if the sampling unit h consists of a single person aged 54+. We record 1. the number of people included in each electoral register considered; 2. the number of people 49-, 50-53 and 54+ selected from each electoral register considered; 3. the number of people 50-53 and 54+ selected from each electoral register considered and included in the gross-sample (this allows calculating i_d and j_d). Notably, for 21 out of 93 sampled municipalities, the electoral registers are available in electronic format. For all these municipalities we can record I_d and J_d directly. For the remaining municipalities, I_d and J_d are not available but they can still be estimated on the basis of the recorded information described at points 1 and 2 of the list reported above. By the design of SHARE, the probability of selecting the sampling unit members is equal to the probability of selecting the sampling unit*. If we define $P_{h\|dms}=P_{m\|s}\cdot P_{d\|m}\cdot P_{h\|d}$ as the probability of selecting a given sampling unit h from the electoral division d of the municipality m, the probability of selecting each sampling unit member j is $P_{jh\|dms}=P_{h\|dms}$.

	* see De Luca and Rossetti (2008), *Sampling Design and Weighting Strategies in the Second wave of SHARE*, in Health, Ageing and Retirement in Europe (2004-2007). Starting the Longitudinal Dimension, editors A. Börsch-Supan, A. Brugiavini, H. Jürges, A. Kapteyn, J. Mackenbach, J. Siegrist and G. Weber, pp. 331-336, Mannheim: MEA.
Design weights	The design weight for a given sampling unit h included in the electoral division d of the municipality m is the inverse of its selection probability. If $w_{h\|dms}$ is the design weight of the sampling unit h, we have $w_{h\|dms}=1/P_{h\|dms}$. The design weight for each sampling unit member in a given sampling unit h is equal to $w_{h\|dms}$.
Target household response rate (for sampling purposes)	Our target is the achievement of a household response rate of 60%.
Target sample size	Our target is to conduct 1,929 individual interviews.

The Netherlands
Refreshment or baseline sample: Refreshment
Survey Institute: TNS NIPO
Country sampling contact: Marjolein Zonjee (TNS NIPO)
SHARE sampling expert: Matthias Ganninger
Reference survey: not applicable
Date: 15 September 2011

Target population, Population coverage	Sample from 26 municipalities in the Netherlands. Alkmaar Alphen ad Rijn Bergen op Zoom Enschede Goes Hilversum Lelystad Steenwijkerland Tytsjerksteradiel Amersfoort Beemster Bernheze Den Haag Den Helder Emmen Kampen Kerkrade Leeuwarden Leidschendam-Voorburg Moordrecht Nijkerk Rotterdam Slochteren Tilburg Utrecht Zwolle Dutch speaking residents of the 26 municipalities Born 1960 or earlier at the time of interview and their partners, independent of age. The target population includes those living in institutions.
Screening frame (if applicable)	Not applicable (no screening is necessary in the Netherlands, because names, addresses and ages of household are known in municipal administration).
Screening frame problems (if applicable)	Not applicable.
Screening design (if applicable)	Not applicable
Remarks	Not applicable
Sampling frame	29 municipalities are part of the longitudinal sample. We've asked these 29 municipalities to deliver a refreshment sample for Wave 4. 26 municipalities were willing to deliver a refreshment sample. The samples from the municipalities contained information about the sex and age of the target persons. So simple random sampling of individuals was possible.
Sampling frame problems	3 municipalities were not willing to deliver a sample

Sampling design	Stage 1: Contacting 29 municipalities. Stage 2: 26 municipalities were willing to deliver a refreshment sample with: - 25 households with at least 1 person born 1957 - 1960 - 34 households with at least 1 person born in or before 1956 Stage 3: Preparing refreshment gross sample: random selection of 22 or 23 households per municipality for cohort 1957 – 1960 random selection of 31 or 32 households per municipality for cohort <1957 Stage 4: Checking and preparing addresses for fieldwork.
Remarks	We've sent CentERdata the sample data to check the sample.
Auxiliary frame data that can be used by SHARE	
Selection probabilities (sampling plus screening, if applicable)	
Design weights	
Target response rate (for sampling purposes)	50% response on individual level for refreshment sample 1,5 individuals per household
Target sample size	Net sample refreshment: 1045 individuals Gross sample refreshment: 1395 households (592 cohort 1957-1960, 803 cohort <1957).

Sweden
Refreshment or baseline sample: Refreshment
Survey Institute: Intervjubolaget IMRI
Country sampling contact: Per Johansson, Daniel Hallberg
SHARE sampling expert: Giuseppe De Luca
Reference survey:
Date: 01 October 2010

Target population, Population coverage	All households with at least one Swedish speaking member born 1960 or earlier. All Swedish speaking residents born 1960 or earlier and their spouses/partners at the time of the interview independent of the spouse's/ partner's age.
Screening frame (if applicable)	Not applicable (no screening needed)
Screening frame problems (if applicable)	Not applicable
Screening design (if applicable)	Not applicable
Remarks	The target population includes individuals living in institutions for elderly, but not those who live in prisons or similar institutions.
Sampling frame	The sampling frame is the population register NAVET of the Swedish tax authority (Skatteverket). It includes all registered residents as of 2011-02-16 born in 1956 or earlier.
Sampling frame problems	The sampling frame does not include individuals who have a protected and secret identity and address (less than 0.1 percent of the total population). The address on which a person is registered is not always the address where the person lives. For instance, immigrants may *de facto* have returned to their home countries but are still registered as residents in Sweden. Another example is persons in bad health who live somewhere else than their old home at the registered address. The sampling frame does not include information on household size and telephone numbers. The latter have to be found using various directories. In case of household split, the population register NAVET cannot be used to find contact information on the new household of a spouse/partner who was age-eligible at the time of sampling. In such circumstances, contact information must be obtained during field work by approaching the original sampled person.
Sampling design	The sample of the 2004 wave is a representative sample of the population born 1954 or earlier. It includes a main sub-sample of 3150 individuals, a supplementary sub-sample of 950 individuals and a vignette sub-sample of 600 individuals. Main and supplementary sub-samples were drawn in two different periods using stratified sampling with simple random sampling of individuals within strata. Stratification was done by gender and year of birth. Sample and population size by strata are provided in Tables 8.1 and 8.2 respectively. The vignette sample was drawn using a stratified two-stage sampling design with regions as primary sampling units and individuals born 1954 or earlier as secondary sampling. In the first stage, primary sampling units were stratified in 5 strata (Stockholm, Gothenburg, Malmö, plus the southern and the northern parts of Sweden). The three largest regions (Stockholm, Gothenburg and Malmö) formed three separate strata and were included with certainty. From the fourth stratum (i.e. the southern part of Sweden) 9 of 48 regions were randomly selected, while 4 of 19 regions were randomly selected from the fifth stratum (i.e. the northern part of Sweden). In the second stage, individuals were randomly drawn from each region selected in the first stage. The sample size used in the second stage was constant for all regions within the same stratum $(n_1 = 107; n_2 = 53; n_3 = 34; n_4 = 34; n_5 = 25)$.

	The sample of the 2006 wave is a representative sample of the population born 1956 or earlier. In addition to the three sub-samples from the 2004 wave, it includes a new refreshment sub-sample of 624 individuals which was drawn using a stratified sampling with simple random sampling of individuals within strata.[7] As for the first wave, stratification was done by gender and year of birth. However, it was based on a finer partition of the year of birth to account for oversampling of individuals born between 1955 and 1956. The relevant sampling design information is provided in Table 8.3. The sample of the 2008 wave is just a follow-up of the sub-samples from the first two waves and it does not include any new refreshment sample. Due to lack of funds, the sample of the 2010 wave is also a follow-up of the sub-samples from the first two waves and it does not include any new refreshment sample.
Remarks	The sample of the 2010 wave does not include a refreshment sample because of lack of funds.
Auxiliary frame data that can be used by SHARE	Gender, year of birth, marital status, number of children, if immigrant and country of origin.
Selection probabilities (sampling plus screening, if applicable)	Let $\pi_{ih}(s;w)$ be the probability of including person *i* of household *h* into the sub-sample *s* of wave *w* and denote by $\pi_h(s;w)$ the same probability for the whole household *h*. The probability of being included in the sample from the 2004 wave is equal to the joint probability of being included in one of its three sub-samples: main, supplementary and vignette. As for the main and the supplementary sub-samples, it is worth noticing that strata are large, household members can belong to different strata, and any age-eligible household member has the same inclusion probability as the whole household. Thus, the probability of being included in one of these two sub-samples ($s=1$ for the main sub-sample and $s=2$ for the supplementary sub-sample) is $$\pi_{ih}(w=1;s)=\pi_h(w=1;s)=\left[1-\prod_{i=1}^{n^*_{h,54}}\left(1-\frac{n_{t(i)\|s}}{N_{t(i)\|s}}\right)\right]I\left(n^*_{h,54}>0\right),$$ where $n^*_{h,54}$ is the number of household members born in 1954 or earlier, $n_{t(i)\|s}$ and $N_{t(i)\|s}$ are the target number of sample units and the total number of population units in stratum $t(i)$ for sub-sample *s* (see Tables 8.1 and 8.2), and $I(A)$ is the indicator function of the event A . Here, strata are functions of *i* because they depend on gender and year of birth of the age-eligible household members. As for the vignette sub-sample, it is worth noticing that all age-eligible household members belong to the same region. Thus, the probability of being included into the vignette sub-sample is $$\pi_{ih}(w=1;s=3)=\pi_h(w=1;s=3)=\left[1-\left(1-\frac{r_{t(h)}}{R_{t(h)}}\frac{n_{r(h)}}{N_{r(h)}}\right)^{n^*_{h,54}}\right]$$ where $r_{t(h)}$ and $R_{t(h)}$ are the target number of regions and the total number of regions in stratum $t(h)$, $n_{r(h)}$ and $N_{r(h)}$ are the target number of 50+ individuals and the total number of 50+ individuals in region $r(h)$, and $n^*_{h,54}$ is the number of household members born in 1954 or earlier. The inclusion

7 The sample of the second wave does not include any refreshment for the supplementary sub-sample because it was considered as part of the main sub-sample.

9 Fieldwork Management and Monitoring in SHARE Wave Four

Frederic Malter, Max Planck Institute

9.1 Introduction

Fieldwork monitoring is a set of activities aimed at minimizing survey errors from various sources while data collection is still underway and corrective action is still possible. In SHARE, an added layer of complexity results from the fact that most countries contract for-profit survey agencies to conduct fieldwork and manage interviewers. The role of the central coordination team of SHARE at the Max Planck-Institute in Munich was thus to inform the contracted businesses and scientific country teams on a number of relevant indicators of fieldwork progress and data quality. Most indicators were set out as quality indicators in the so-called model contract that specifies what good performance is according to SHARE.

The most salient problem to be minimized, using the total survey error concept (Lepkowski and Couper, 2002), was unit nonresponse which has quite unfavorable consequences for panel studies (Watson and Wooden, 2009, p. 159). Unit non-response, be it from lack of locating the respondent, lack of establishing contact or lacking willingness to cooperate (Lepkowski and Couper, 2002) is the main cause of attrition in panel samples. The same factors also influence low response rates in refreshment samples, albeit to a different degree (Lepkowski and Couper, 2002).

During SHARE wave four, we focused fieldwork monitoring on activities to minimize the following three causes of unit non-response a) lack of contacting households, b) gaining respondent cooperation, and c) dealing with cases of initial refusal. A second set of monitoring activities was geared at reducing measurement error (or even bias) due to non-standardized interviewing. The source of error we focused on was undesired interviewer behavior, more specifically not reading question texts properly. Finally, interview length, or rather the deviation from average interview length, was used as a proxy for implausibly short interviews and thus improperly conducted interviewing. Feedback to survey agencies on these indicators was intended to stimulate corrective action, i.e. making agency managers relay these findings to interviewers. The hope was that making interviewers aware of their being monitored would guide their behavior towards more successful and proper interviewing.

9.2 SHARE interview tools, data types and data delivery

All eighteen countries collecting data during SHARE wave four used the same standardized electronic contact protocol, the Sample Management System (SMS). The SMS was installed on every interviewer laptop. It was designed in an iterative process since wave one and enabled interviewers to manage their assigned households (details about the evolution and functionality of this software can be found elsewhere, e.g. (Das et al., 2011). Briefly, the SMS application on a laptop contains all households to be interviewed by a specific interviewer. It allowed the interviewer to record the entire contact sequence from first attempts to completed interview. All contact attempts and contacts were supposed to be entered into the SMS by interviewers, using predefined result codes.

After the composition of the household was assessed per SMS, the actual interview software started. The Computer-Assisted Personal Interviewing (CAPI) software that stores the interview responses was implemented using Blaise code and contained a functionality of logging keystrokes. These keystroke records allowed the assessment of critical indicators of proper interviewing, such as overall length of interviews, length of modules or individual items (broken down by countries, sample types or interviewer, depending on the purpose). Longer introduction texts were implemented as separate items. Thus, average length of reading these longer texts could be compared with normative length to read the same texts.

Data recorded by the SMS-CAPI application was then synchronized with servers of the survey agency. The software that collects synchronized data from interviewer laptops and contains all households of a country is called the Sample Distributor (SD). It was used – among other things – by survey agency administrators to assign households to interviewer laptops. All laptop data that was synchronized with the agency SD at specified dates was then sent to Centerdata servers. After the first step of processing at Centerdata, SHARE central coordination received interview, SMS- and keystroke data on a fortnightly basis. All dates of data transmission from agency servers to Centerdata servers were fixed before fieldwork started to ensure a synchronized availability of fieldwork data. The central fieldwork monitoring team then combined data of all countries and generated reports that were sent to all country teams and contracted survey agencies. In these reports, the current state of fieldwork and relevant statistics on fieldwork progress and integrity of data collection were laid out. Specific problems were pointed out with suggested solutions. This represented a unique feature of the SHARE data collection effort: data on the state of fieldwork is available in real time.

9.3 Survey agencies

In wave four, most countries contracted fieldwork out to the same for-profit survey agency as in wave 3. Table 9.1 below shows contracted agencies over time. Belgium represents an exception as it had two survey agencies: one for the Dutch-speaking part and one for the French-speaking part of Belgium.

Table 9.1 Survey agencies of SHARE waves one to four of countries participating in wave four

Country	Wave 1	Wave 2	Wave 3	Wave 4
Austria	IMAS	IMAS	IFES	IFES
Belgium (French-speaking)	PSBH, University of Liège	PSBH, University of Liège	PSBH, University of Liège	PSBH, University of Liège
Belgium (Dutch-speaking)	PSBH Antwerp Univ.	PSBH Antwerp Univ.	CELLO - Antwerp Univ.	CELLO - Antwerp Univ.
Czech Republic	-	SC & C	SC & C	SC & C
Denmark	SFI Survey	SFI Survey	SFI Survey	SFI Survey
Estonia	-	-	-	Statistics Estonia
France	INSEE	INSEE	INSEE	INSEE/ GFK-ISL
Germany	infas GmbH	infas GmbH	infas GmbH	infas GmbH
Hungary	-	-	-	TARKI
Italy	DOXA S.p.A.	DOXA S.p.A.	DOXA S.p.A.	DOXA S.p.A.
Netherlands	TNS NIPO	TNS NIPO	TNS NIPO	TNS NIPO
Poland	-	TNS-OBOP	TNS-OBOP	TNS-OBOP
Portugal	-	-	-	GfK Metris
Spain	TNS Demoscopia	TNS Demoscopia	TNS Demoscopia	TNS Demoscopia
Slovenia	-	-	-	CJMMK
Sweden	Intervjubolaget	Intervjubolaget	Intervjubolaget	Intervjubolaget
Switzerland	MIS Trend	LINK	LINK	LINK

9.4 The fieldwork period

Most countries of wave four had a refreshment sample in addition to their panel sample: Austria, Belgium, Switzerland, Czech Republic, Denmark, Spain, France, Italy and the Netherlands. Belgium was counted as two countries, with a French-speaking and Dutch-speaking part. Countries with a panel sample only were Germany, Poland and Sweden. Further, four countries joined SHARE in the fourth wave. Accordingly, they did baseline interviews only: Estonia, Hungary, Portugal, and Slovenia. Israel and Greece, two previous SHARE countries, were not part of wave four.

The largest challenge for standardized, harmonized monitoring and management of fieldwork in wave four was the highly asynchronous fieldwork periods between countries. This was largely a result of decentralized funding and associated delays in the start of fieldwork. As can be seen in Figure 9.1, the most extreme case was Poland which only secured funding for wave four in December 2011. By the time first interviews were conducted in Poland, Estonia had already completed fieldwork six months earlier. Figure 9.1 contains the most important “milestones to be passed” to get SHARE fieldwork underway. From the respondents' point of view, however, the first encounter with each SHARE wave happens through an advance letter, sent prior to any contact attempts by actual interviewers. The technological pre-requisite for survey agencies to start fieldwork is the availability of the Sample Distributor software that contains all households of the longitudinal and refreshment gross samples (highlighted blue). This tool was installed on servers of the survey agency and was used to assign a set of households to an interviewer laptop. The next milestone was the first complete interview. The time lapse between receiving the SD and conducting the first interview is spent with setting up the SD software, assigning households to laptops and interviewers contacting households. Obviously, the last interview date (black highlighted) signals the end of fieldwork. The time between first interview and last interview was considered the fieldwork period (highlighted grey). The number of weeks of fieldwork is given by the white number. Germany needed the most time (54 weeks) and Poland was quickest (14 weeks).

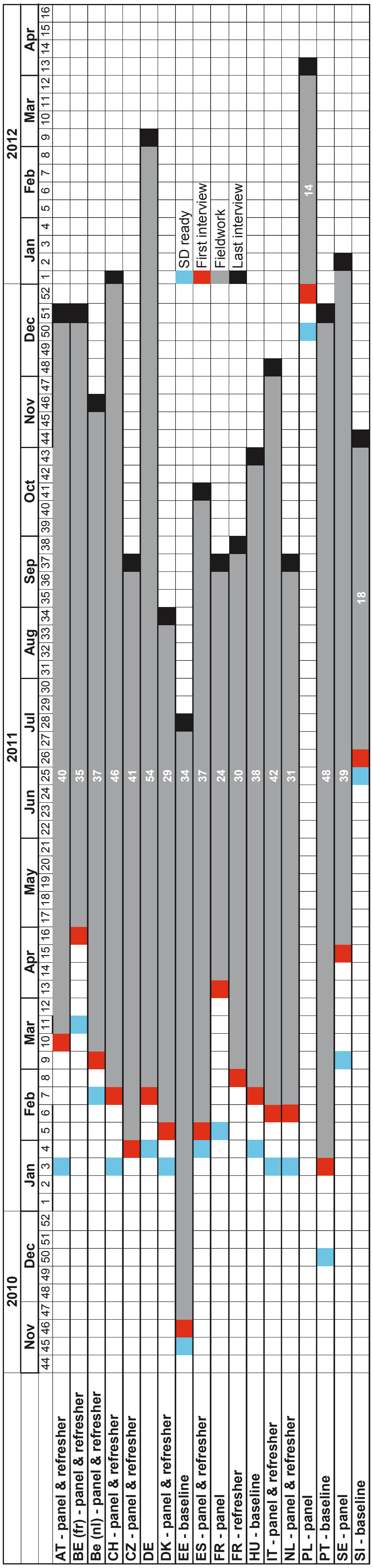


Figure 9.1 Fieldwork periods of countries in SHARE wave four

9.5 Fieldwork indicators

Most indicators reported in the following sections were part of each fortnightly report. Depending on the stage of fieldwork (just started, advanced, almost done), indicators relevant at that stage were included. Here again, asynchronous timing of fieldwork periods made some information more relevant in some countries than others because by the time country X was in an advanced stage of fieldwork, country Y might have just started. For some indicators, their choice was driven by their continuous relevance to assess progress of fieldwork. Those indicators were carried forward from one reporting period to the next. Some other indicators were only assessed once and reported back to survey agencies in order to highlight areas of improvement. Details on final indicators can be found in chapter 10.

9.5.1 Total number of interviews

Figure 9.2 below shows the progress in absolute number of interviews over the entire fieldwork period. Countries differed markedly in their slopes. Some gathered high numbers of interviews over a fairly short period of time (such as EE, PL and SI), whereas others took a long time to accumulate a substantial number of interviews (such as DE, PT, SE).

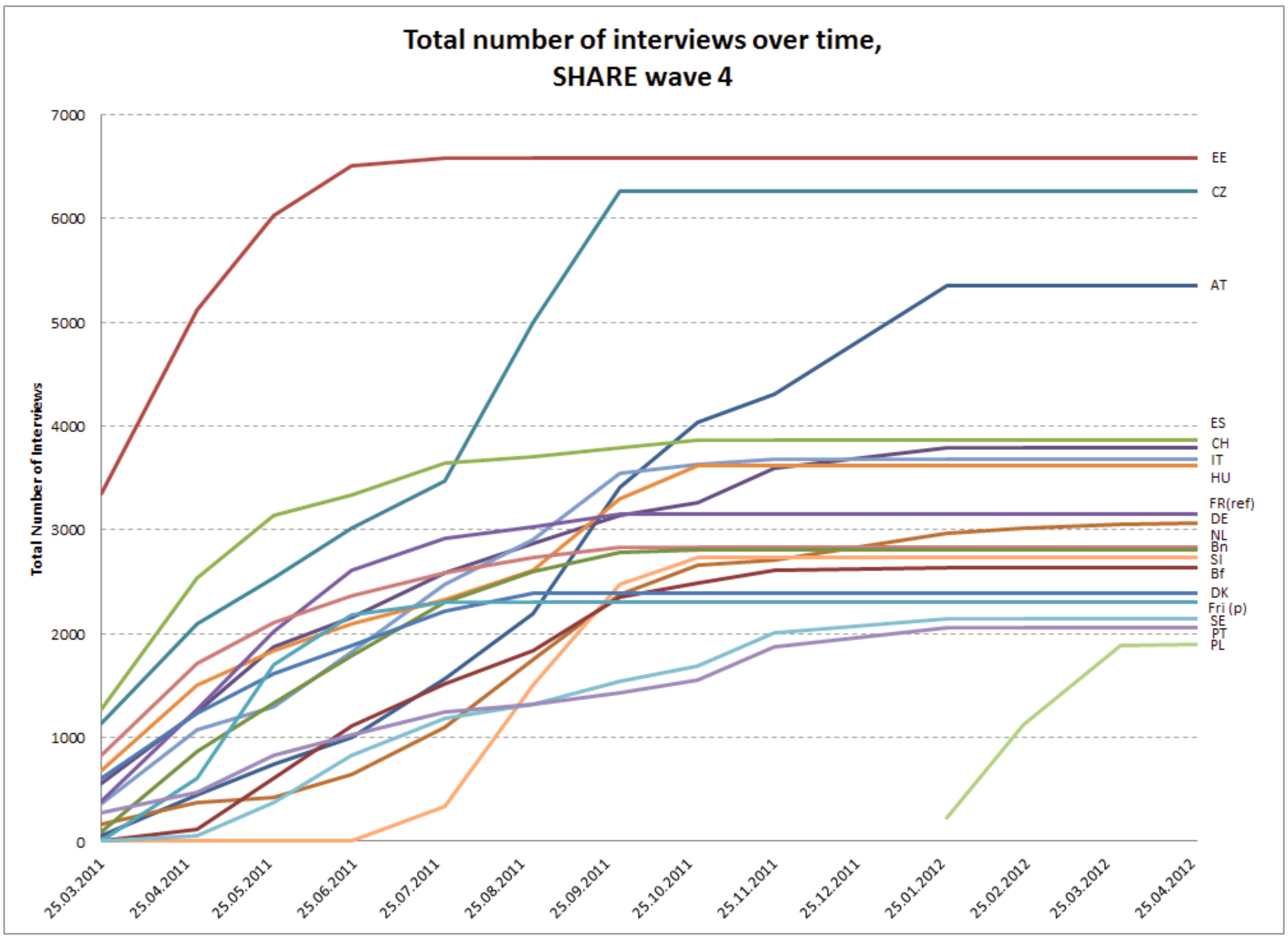


Figure 9.2 Progress in absolute number of interviews over the entire fieldwork period

9.5.2 Contacting households

Attempting to contact households was the first action any interviewer had to do after advance letters had been sent to all prospective respondents in the gross sample. The rate of gross sample households that were actually contacted is one of the two logical ceilings of final response/retention rates. Cooperation rates (see chapter 10) represent the second set of logical ceilings to final response/retention rates. As can be seen in Figure 9.3, countries differed in their strategies of contacting households. Some countries had very steep increases from the get-go, whereas others only very gradually increased their contact attempts.

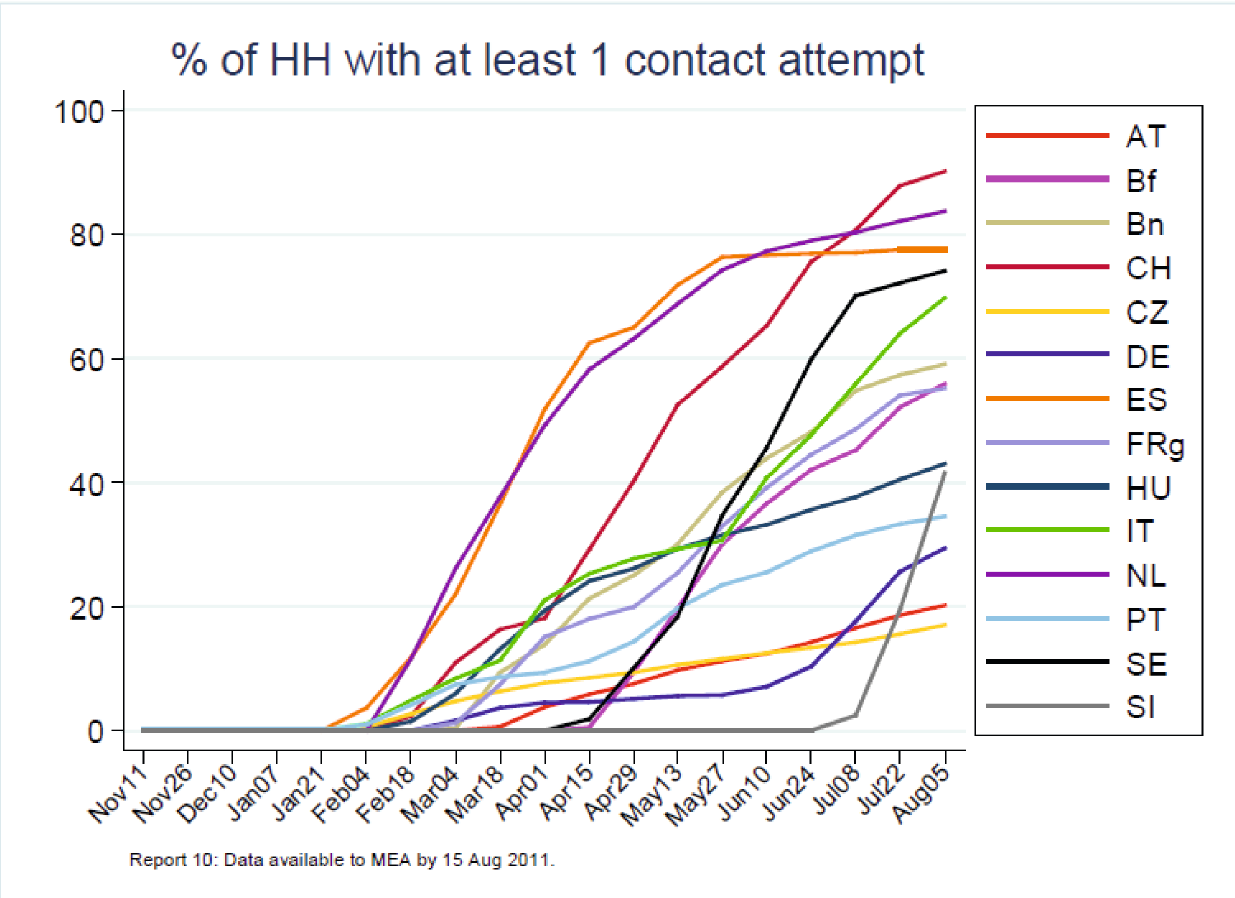


Figure 9.3 Countries' strategies of contacting households

9.5.3 Gaining cooperation

Once contact has been established, obtaining respondent cooperation is the next challenge to be overcome. Figure 9.4 below shows conditional cooperation rates, i.e. based on those individuals that have already been contacted. The number served as a first indication of the general strategy of the survey agencies. Low rates suggested a high number of contacts over a relatively low number of complete interviews. This indicator could not be carried forward in a straightforward manner because a change of this rate could have resulted from an increase in contacts (denominator), complete interviews (numerator) or both.

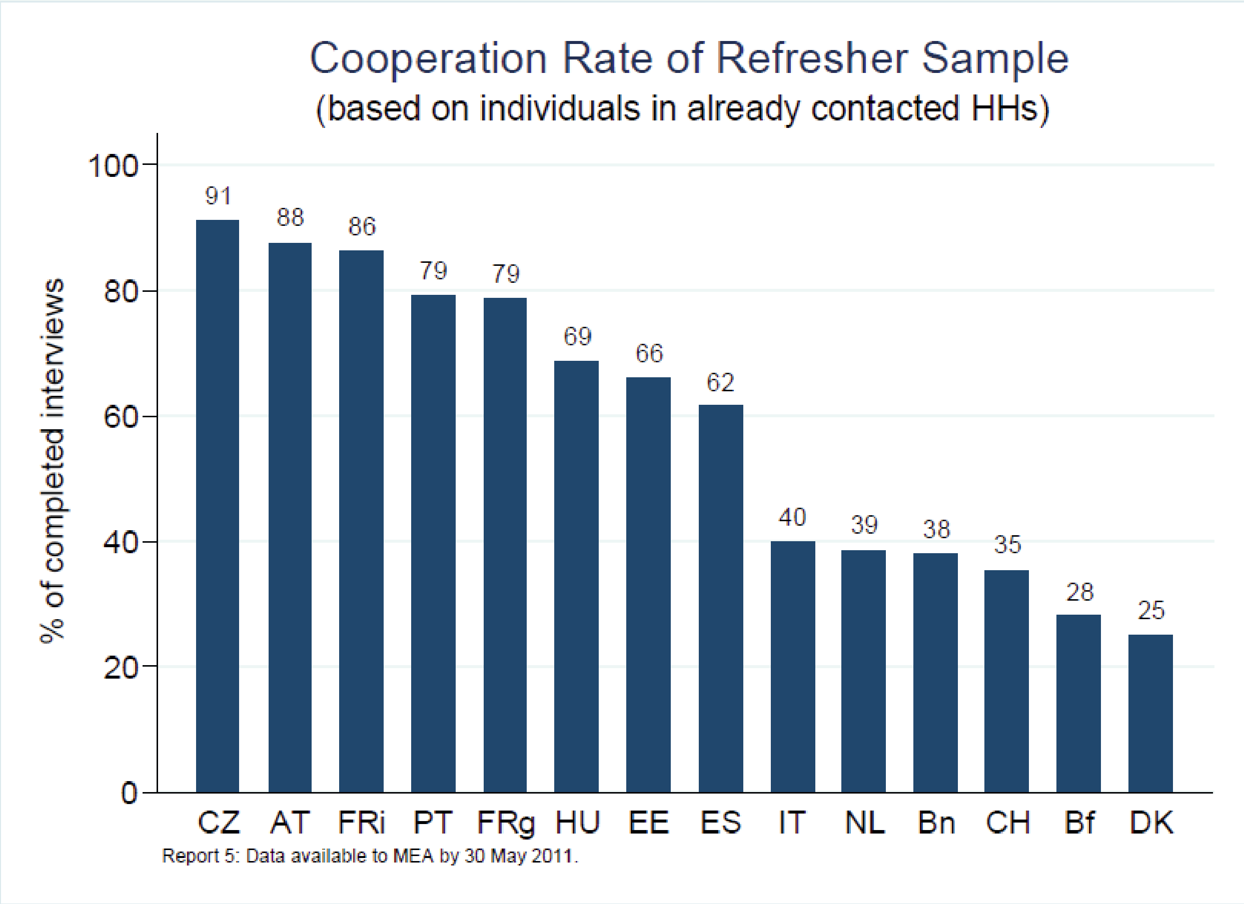


Figure 9.4 Conditional cooperation rates

9.5.4 Refusals

Refusing survey participation is the main source of non-cooperation. In the graph below it can be seen that histograms of refusal types differed between countries, indicating different use of these codes by interviewers and/or different refusal patterns of approached respondents. Overall, a lack of interest or disapproval of surveys (code 207) was the most prevalent in all countries.

When we provided this information to survey agencies, we pointed out different strategies of re-approaching households dependent on their refusal reason.

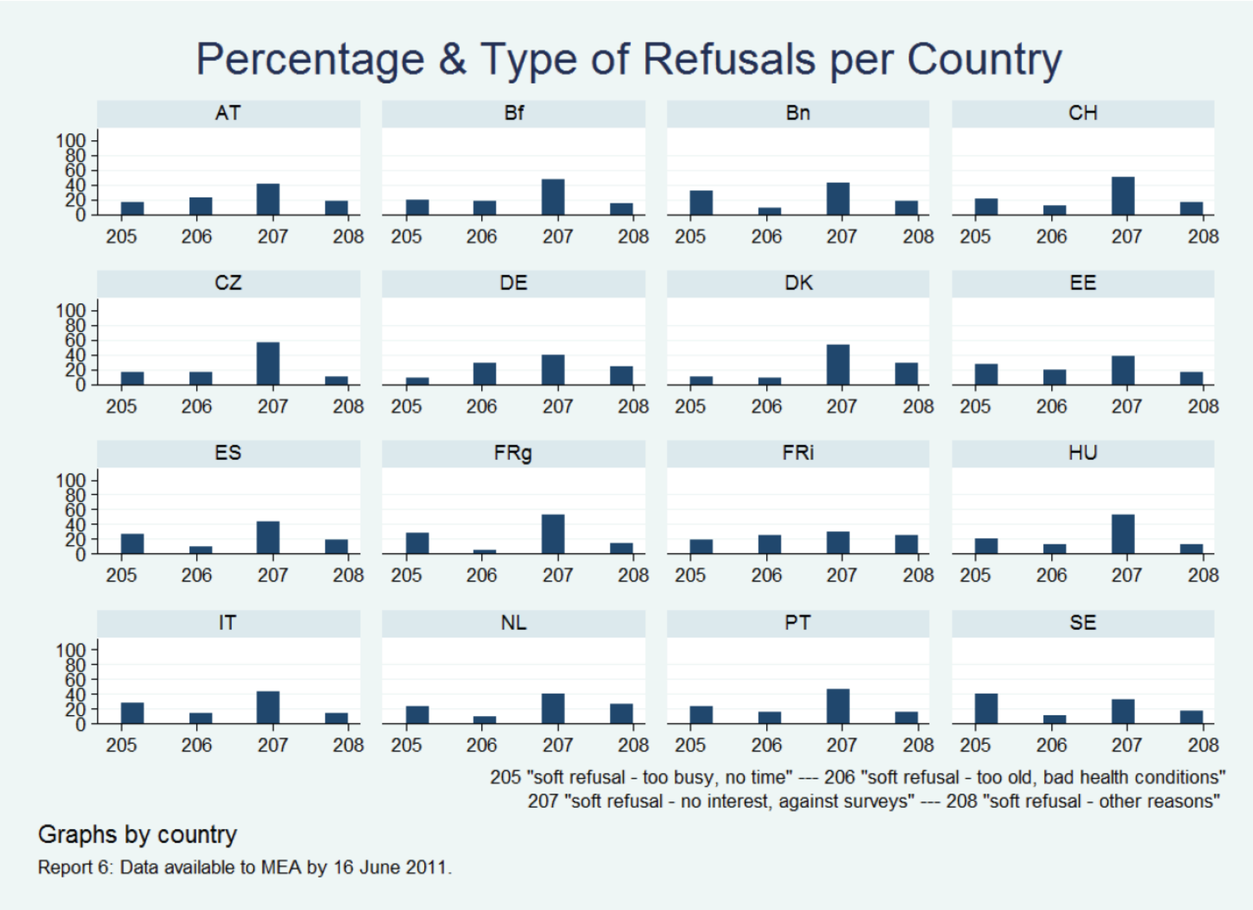


Figure 9.5 Percentage of refusals per country

9.5.5 Panel retention rates

In SHARE, special emphasis was put on respondents who participated in the previous wave ("sub-sample A") and respondents who have not participated themselves in the previous wave but lived in a household where their spouse or partner had participated in the previous wave ("sub-sample B"). The snapshot in Figure 9.6 shows different trajectories of conducting interviews. Some survey agencies collected interviews very quickly (CH, FR) while others made barely any progress at all over the time displayed in Figure 9.6 (DE).

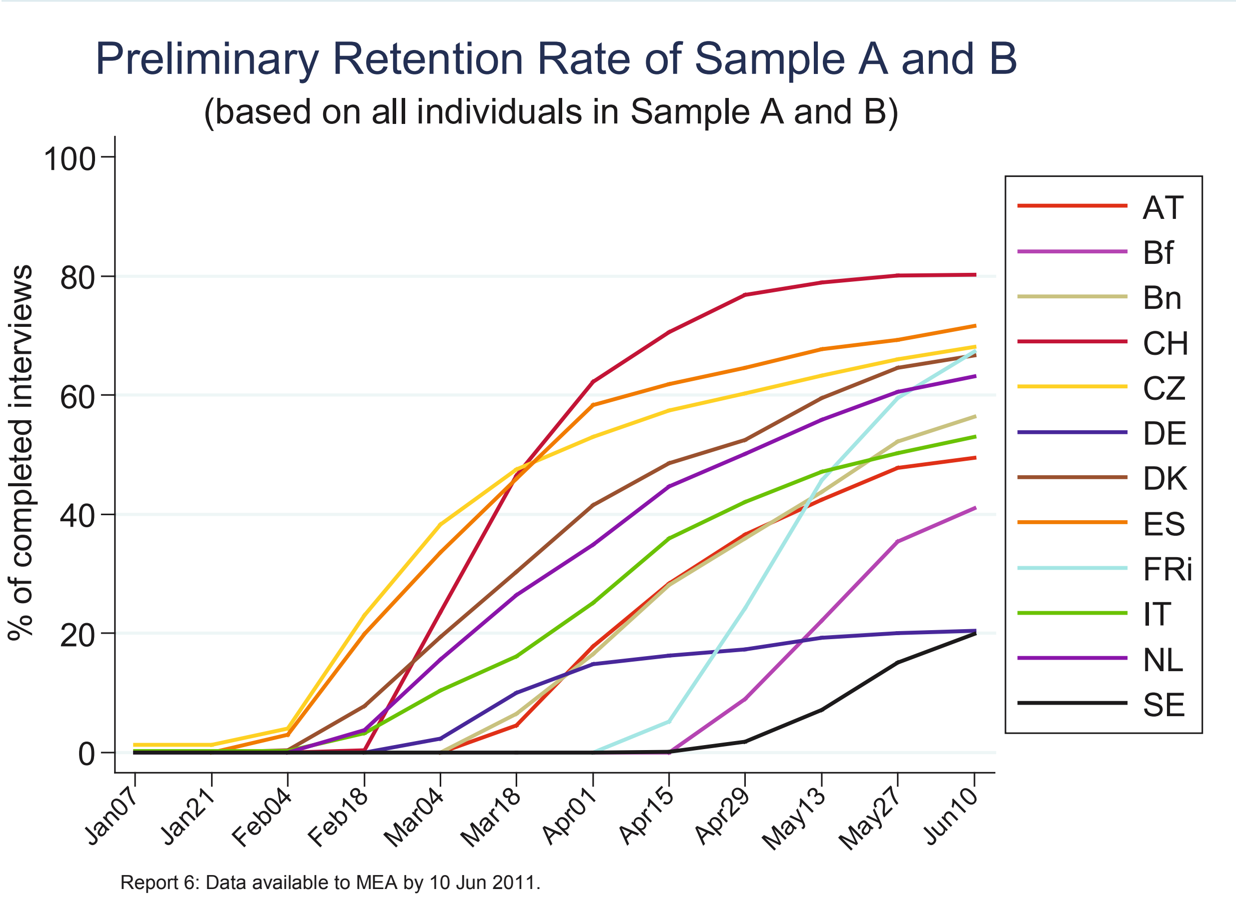


Figure 9.6 Different trajectories of conducting interviews

Survey agencies were constantly reminded and encouraged to reach at least a target retention rate of 80 percent. In addition, personal communication was intensified with agencies during the end of fieldwork to discuss solutions to improve retention rates. While not part of regular monitoring feedback, interviews with panel respondents that had not participated in one or more previous waves ("sub-sample C") were paid an additional 25% of the contractual unit costs. Details on final outcomes can be found in chapter 10.

9.5.6 Contacting and cooperation revisited: "Unused potential"

Towards the end of fieldwork we analyzed all SMS codes to identify "unused or under-used potential" (see Figure 9.7). Un- or underused potential was identified as follows: every household where either the entire household or at least one eligible person within the household had a code out of the three code categories listed below:

a) HH/individuals with unsuccessful contact attempts, i.e. where no one answered (SMS code 201)

b) HH/individuals with successful contacts but without interviews (SMS codes 202-204)

c) HH/individuals with soft refusal codes (SMS codes 205-208).

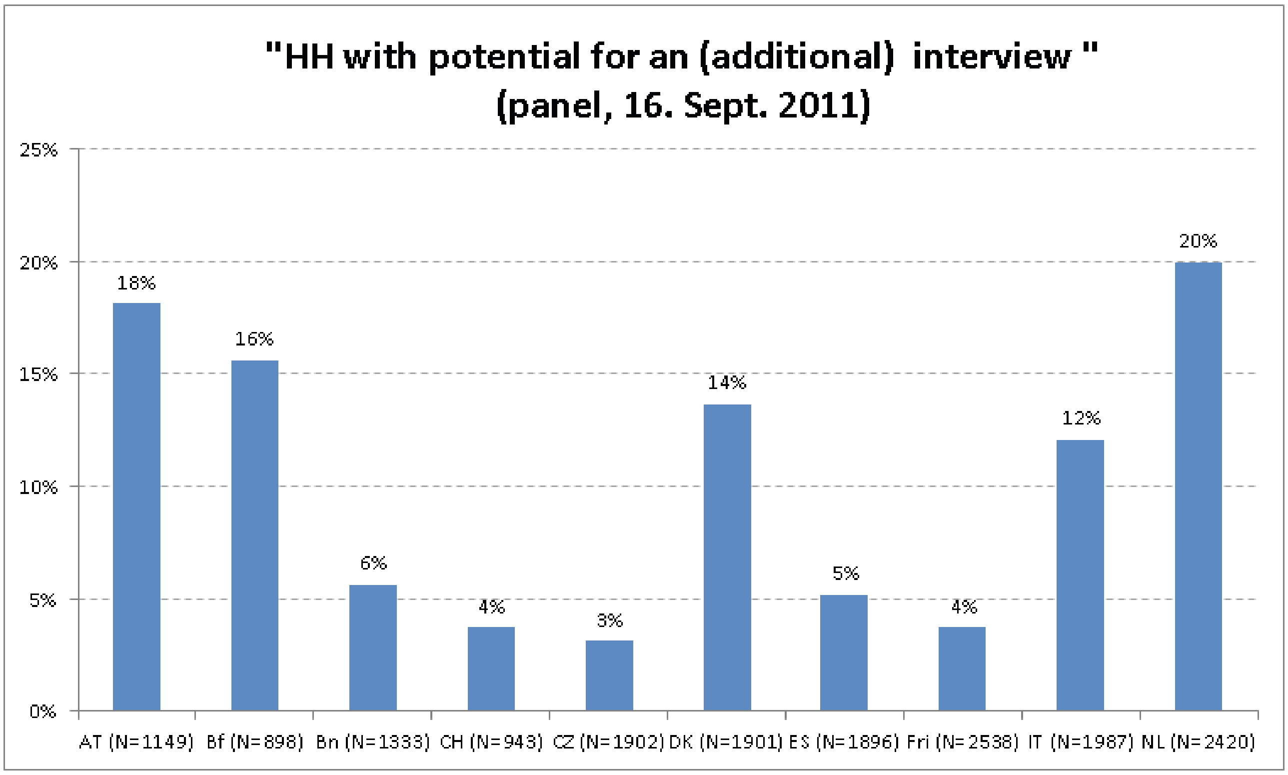


Figure 9.7 Panel households with SMS codes indicating potential for an (additional) interview, as of Sept. 16, 2011

Stark between-country variation was found. The rates in the figure above constitute the upper limit of households to be re-contacted as it must be assumed that at least some households under the category "soft refusal" could as well have been classified as final (hard) refusal by interviewers.

9.5.7 Percent of active interviewers

Duration of fieldwork (expressed in weeks) is dependent on roughly six parameters:

a) Gross sample size (number of households/individuals to be contacted),

b) Expected response and/or retention rates

c) Total number of interviews to be conducted which equals the net sample size of respondents, calculated by multiplying the gross sample size by the expected response/retention rates

d) Average number of active interviewers per week

e) Average number of households with a final contact status per interviewer per week

f) Average number of interviews per interviewer per week

The higher the average number of active interviewers, the sooner a fieldwork is complete. Any survey benefits from a consistently high number of active interviewers, as dragging fieldwork on for too long reduces the chance of obtaining interviews[1]. This has many reasons. For example, interviewers get a routine in conducting a specific interview and lose this routine if they take longer breaks from a specific study to work on other studies. Another reason is that all advance letters may be sent out at once. A long time between receiving the advance letter and being contacted by an interviewer may reduce the target person's willingness to cooperate (ESS, 2012), e.g. due to forgetting the receipt of the letter (Link and Mokdad, 2005).

As can be seen in Figure 9.8, different strategies were also observed for bringing on the full interview staff. Of all "early starter countries" Portugal took the longest to bring on at least 50 interviewers and even this number is only the peak. The Czech Republic, on the other hand, had their full interviewer staff active from the start of fieldwork and consistently throughout it. Naturally, the rate of active interviewers goes down at the end of fieldwork, exemplified by the trajectory of Spain in the graph below (for clarity, only a selected number of countries are shown here).

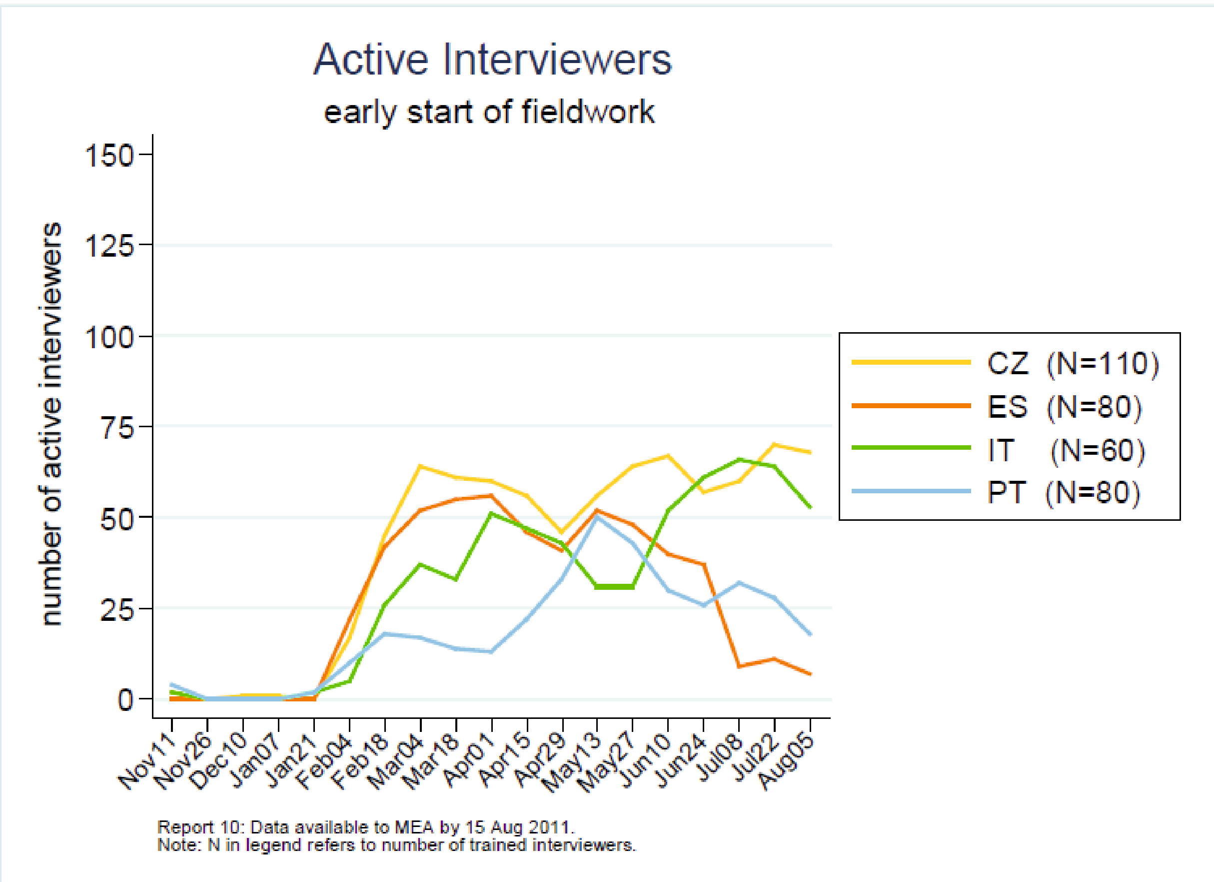


Figure 9.8 Active interviewers of all „early starter countries"

[1] An exception may be the time-consuming re-uptake of „difficult" households, e.g. initial refusals, at the end of the fieldwork period by more experienced interviewers.

9.5.8 Interview length

During fieldwork, we also checked the length of the interviews to detect implausibly short ones. Very short interviews might indicate illegitimate skipping of items by an interviewer. Figure 9.9 displays box plots of interview length by country, sorted by median length of country.

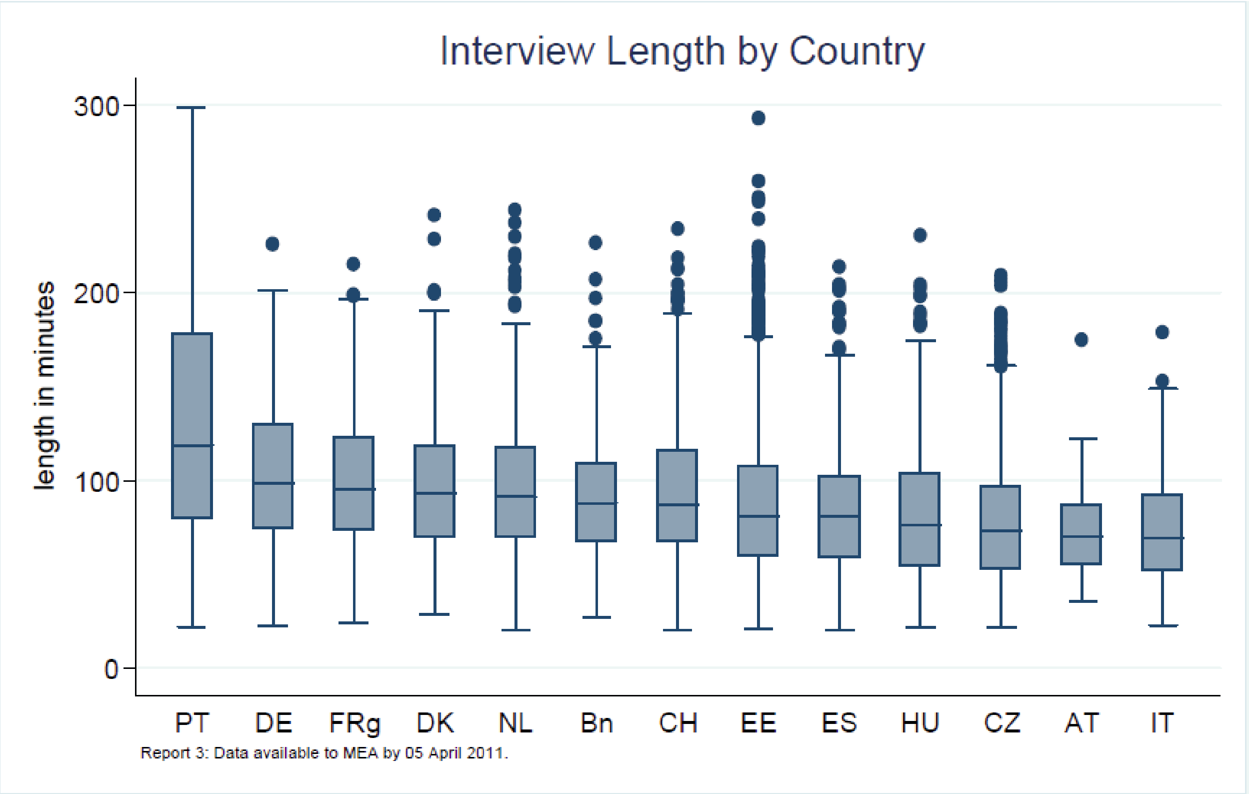


Figure 9.9 Interview length by country

While median length shows considerable variation between countries, some of that variability must be credited to difference in "language speed". It is as noteworthy, however, how much variability can be found within countries. Some of this may be credited to "interviewer effects", indicating a heterogeneous experience of interviewer staff with conducting complex face-to-face interviews such as SHARE. In addition to the country-level results shown here, many countries used interview length to monitor their interviewers' performance.

9.5.9 Reading times of introduction texts

Proper reading of the entire question texts (including introductions) is a key requirement of standardized interviewing (Fowler, 1991; Fowler and Mangione, 1990; Houtkoop-Steenstra, 2000). We computed the time to read out several fairly long introduction texts to identify possible deviations from a normative standard: the time to read out the English generic introduction text at proper reading speed. As can be seen in Figure 9.10, we found strong country differences in reading out long introduction texts between countries. But even within countries, there was large variability between interviews. In our report to survey agencies, we highlighted that more than half of all countries have very high positive skew, meaning that the bulk of interviewers took very little time to read the intro to the social networks module (variable SN001). It took most interviewers just a few seconds. The red line indicates how long it took to properly read out the English text of SN001. Language differences alone cannot explain these stark differences and right-skewed distributions. Clearly, interviewers cut the intro text short or skipped it entirely. Very similar findings emerged for three other "long" question texts that we checked.

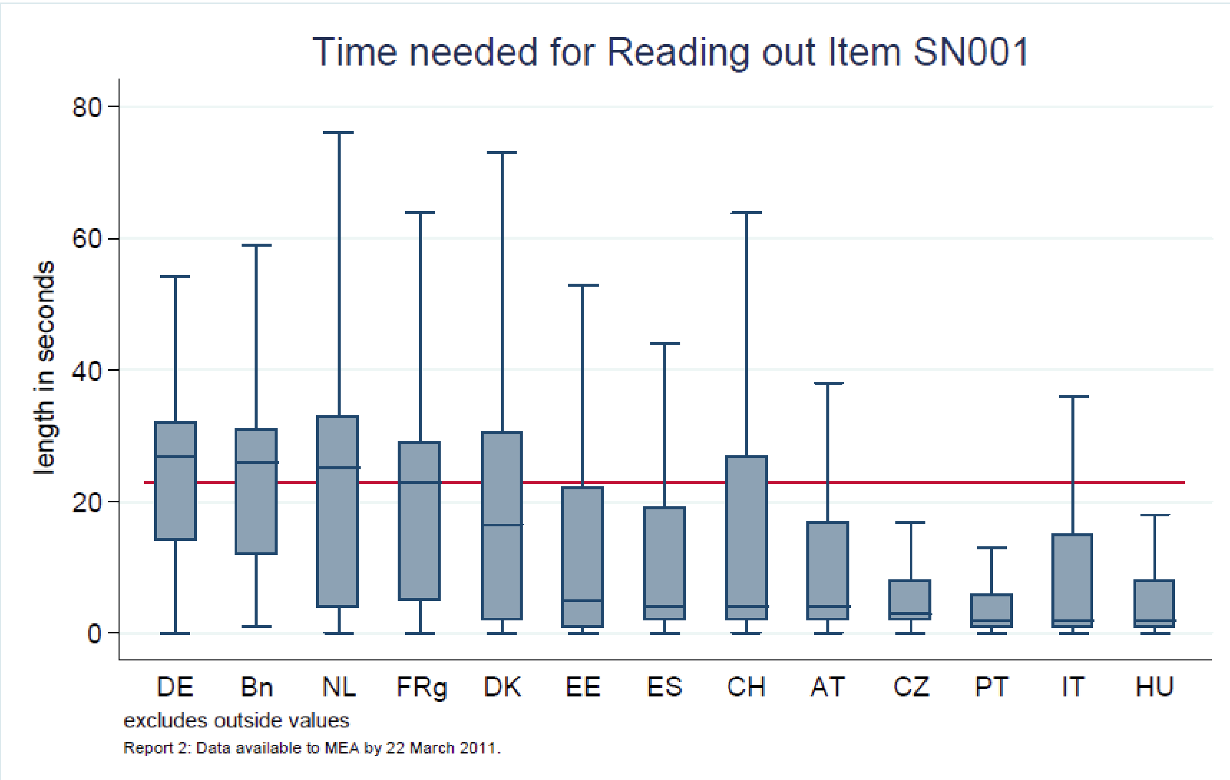


Figure 9.10 Time needed to read the intro to the social networks module (variable SN001)

9.6 Summary and discussion

The purpose of fieldwork monitoring is to minimize total survey error as much as possible. In SHARE wave four, we focused mostly on giving feedback to survey agencies to minimize unit nonresponse. One of the key challenges in giving feedback about the status of fieldwork and suggesting corrective actions is the principal-agent problem inherent in the governance structure of SHARE. The central coordination team had no direct interaction with the actual interviewers. In other words, all feedback to the interviewers was mediated by the management of the survey agencies or, in some cases, by the staff of scientific university teams. The complex IT system that allowed for real-time fieldwork status still had a lead time of about two weeks, meaning that by the time the reports where sent out, fieldwork had already progressed about another week, leading to some inconsistencies between fieldwork information of agencies and current reports written by the central coordination team.

Accurateness of the data was dependent on a well-functioning IT system at the survey agencies, and cooperative interviewers with internet-capable laptops. At least one survey agency did not allow internet on their interviewers' laptop. This created serious problems in data transmission and led to an ongoing uncertainty of the actual state of fieldwork in that country. The reason was that in this case interviewers had to manually copy various files to USB sticks and email those from personal computers with internet access to the survey agency or even send the USB sticks per mail to the agency.

At the end of fieldwork, when large quantities of interview and contact information was accumulated at the Sample Distributor of an agency, performance issues of the SD tool put a strain on finishing fieldwork. Un-assigning and re-assigning households to different interviewer laptops could take a very long time, as could synchronizing the SMS client with the Sample Distributor. A technical solution was identified for the next wave (see chapter seven).

An IT infrastructure capable of delivering real-time data from the field had its costs. Substantial resources in terms of staff and time went into developing, testing and training the IT tools. Every survey agency had to train their interviewer staff in proper handling of the SMS client.

The great advantage of having real-time data from the field is apparent: problems with response and retention rates could be identified without delay and were addressed swiftly. Especially bad outcomes, such as insufficient contact rates of households were thereby avoided and total survey error likely reduced. An empirical test of this claim will be forthcoming, as a second advantage of a fully electronical contact protocol and survey software is the availability of "paradata", i.e. contact data, interviewer characteristics such as preferred working hours or days and many more. These rich paradata are currently (November 2012) processed and will be made available in future data releases to the scientific community.

References

Das, M., Martens, M. Wijnant, A. (2011). Survey Instruments in SHARELIFE. In M. Schröder (Ed.), *Retrospective Data Collection in the Survey of Health, Ageing and Retirement in Europe*. SHARELIFE Methodology (pp. 20-27). Mannheim: MEA.

ESS. (2012). *Field Procedures in the European Social Survey: Enhancing Response Rates* (pp. 1-10): European Social Survey.

Fowler, F. J., Jr. (1991). Reducing Interviewer-Related Error through Interviewer Training, Supervision, and Other Means. In P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, S. Sudman (Eds.), *Measurement Errors in Surveys* (pp. 259-278). New York: John Wiley & Sons.

Fowler, F. J., Jr., Mangione, T. W. (1990). *Standardized Survey Interviewing: Minimizing Interviewer-Related Error*. Newbury Park, CA: Sage Publications, Inc.

Houtkoop-Steenstra, H. (2000). *Interaction and the Standardized Survey Interview: The Living Questionnaire*. Cambridge: Cambridge University Press.

Lepkowski, J. M., Couper, M. P. (2002). *Nonresponse in the Second Wave of Longitudinal Household Surveys*. In R. M. Groves, D. A. Dillman, J. L. Eltinge, R. J. A. Little (Eds.), Survey Nonresponse (pp. 259-272). New York: John Wiley & Sons, Inc.

Link, M. W., Mokdad, A. (2005). *Advance Letters as a Means of Improving Respondent Cooperation in Random Digit Dial Studies: A Multistate Experiment*. Public Opinion Quarterly, 69(4), 572-587. doi: 10.1093/poq/nfi055

Watson, N., Wooden, M. (2009). Identifying Factors Affecting Longitudinal Survey Response. In P. Lynn (Ed.), *Methodology of Longitudinal Surveys* (pp. 157-182). Chichester: John Wiley & Sons Ltd.

10 Survey Participation in the Fourth Wave of SHARE

Thorsten Kneip, Max Planck Institute

10.1 Introduction

This chapter describes the patterns of survey participation in the fourth wave of SHARE. The term survey participation is used to describe the proportion of households and individuals of the initial gross sample which delivered completed interviews, which were found to be ineligible, and which did not respond. In the following, survey participation patterns are presented separately for the longitudinal samples in countries that have been participating in SHARE since wave one or two,[1] and wave four baseline samples, which comprise refreshment samples as well as initial samples in newly participating countries (for details see chapter 8). While in the context of baseline and refreshment samples the focus is on response behavior to the initial survey request, for the longitudinal samples the focus is on response behavior at subsequent waves, i.e. on panel retention.

Studying survey participation is important for at least two reasons: First, ineligibility or non-response involve a larger gross sample size and, thus, higher survey costs in order to obtain a target number of interviews. Second, non-response is a main source of non-sampling error (Lessler & Kalsbeck 1992). Although response rates alone are not sufficient to evaluate the impact of non-response error, they contain crucial information for understanding the sources of non-response bias and for assessing overall data quality (c.f. De Luca and Peracchi 2005). A third aspect is vital in the context of a panel study: Analyses of panel data look at change over time. Therefore, respondents need to be observed at various points in time. With high attrition rates, however, the number of such observations will decrease rapidly.

All numbers and figures reported in this chapter are based on information from the SHARE sample management system (SMS) and additional national gross sample information. As the post-processing of data was still going on when the present volume went to press (November 2012) this report is, to some extent, preliminary. Furthermore, this analysis does not include the Austrian and the German refreshment samples. In Austria, the gross sample is still under construction. In Germany, the refresher sample was not finished due to capacity limits and was consequently not released. Not included are furthermore Greece, Ireland, and Luxembourg as they did not participate in the main survey of wave four, and Israel which has a different schedule of waves. Finally, Belgium is accounted for as two entities as there were two separate samples for Flanders (Bn) and Wallonia (Bf), which were also independently administered by different survey agencies.

The remainder of this chapter is organized as follows: Section 10.2 defines the target population of SHARE and discusses issues related to the assessment of the sample units' eligibility. Section 10.3 presents the composition of the longitudinal and baseline/refreshment samples with regard to units' eligibility status. Section 10.4 reports patterns of survey participation in the baseline and refreshment samples, both on the household and individual level. Similarly, section 10.5 describes retention patterns in the longitudinal samples, again for households and individual respondents. Section 10.6 summarizes the main results and offers some conclusions. Additional information on

[1] In wave three (life histories), no additional respondents were sampled.

the assignment of final household states as well as on the computation of the presented rates is provided in the appendix.

10.2 Target population and eligibility criteria

In each country, the target population consists of people living in residential households who are 50 years or older plus their partners, irrespective of age.2 The target population is further restricted by additional eligibility criteria: People who are currently in prison, moved abroad, or are unable to speak the specific language of the national questionnaire are considered ineligible. Additionally, persons deceased, hospitalized, moved to an unknown address, or not residing at the sampled address are considered ineligible for the initial sample. All originally selected sample units which could not be located due to errors in the sampling frame (e.g. non-existent address or vacant house) were considered ineligible. For panel respondents, however, these criteria do not lead to ineligibility. The target population is then redefined implicitly according to these definitions of eligibility.

The way in which eligibility of an initially sampled household is determined depends on two conditions: "age-eligibility" (that is, whether or not the household contains at least one person who is 50+) and other eligibility criteria. In principle, age-eligibility may be determined through the very first part of the interview, the so-called cover screen (CV), has been completed. The CV is a brief interview on household composition before the actual SHARE interview starts. In practice, the CV is incomplete for non-responding households (i.e. households that were not contacted or refused to complete the CV) and the gathered information does not allow assessing the age-eligibility of all sampled households. This problem, which is common to all countries, has different solutions depending on the nature of the sampling frame adopted. In one group of countries (Belgium, Switzerland, Denmark, Estonia, Spain, Hungary, Italy, Netherlands, and Slovenia), the sampling frame contains information on the age of the sampled household member. For this first group of countries (type 1), age-eligibility is determined directly from the information provided by the sampling frame. Likewise, for longitudinal households age-eligibility is determined from information of previous waves. They can thus, for all countries, be regarded as type 1. In another group of countries (Austria, Czech Republic, France, and Portugal) the sampling frame does not contain any information on age. For this second group of countries (type 2), a screening phase is required in order to assess the age-eligibility status of the sampled households. For both groups of countries, the other eligibility criteria are instead determined through information provided by interviewers on the non-sample categories described above.

10.3 Classification of sample units

The American Association for Public Opinion Research provides guidelines for a final classification of sample units. On this basis (AAPOR 2011) a variety of indicators on participation behavior, like, for instance, response rates, can be calculated. Following these guidelines, data from the SHARE Sample Management System (SMS) was used to classify the longitudinal as well as the baseline/refreshment gross sample of each country into three main categories: (I) eligible households, (ii) ineligible households, and (iii) households of unknown eligibility. The SMS data contain event history information that allows classifying the outcome of each contact attempt into exhaustive

[2] While in the first wave *all* age-eligible persons per sampled household plus their partners where selected for an interview, only *one* age-eligible person per household (plus partner) has been selected from wave two onward.

and mutually exclusive categories. Table 10.A.1 in the appendix lists the detailed categories and the final SMS codes used for both longitudinal and baseline/refreshment samples.

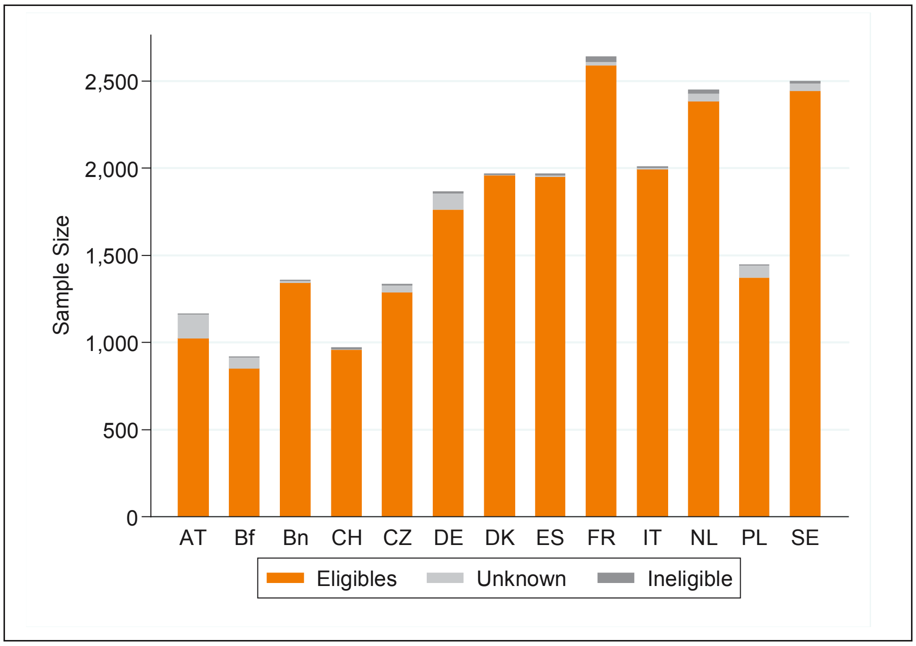


Figure 10.1 Longitudinal samples by classification of sample units

Figure 10.1 shows the size of the longitudinal part of the sample in each country and how it is composed regarding household eligibility status. 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 ever interviewed in any previous SHARE wave. The longitudinal gross samples almost only comprise eligible cases (97 percent). Households in the longitudinal sample could only turn ineligible due to incarceration during the whole fieldwork period or moving abroad of all eligible household members or due to language barriers. On average, this applies to 0.6 percent of households in the longitudinal samples. Death did not lead to ineligibility. Instead, a proxy respondent is supposed to respond to an end-of-life interview about the deceased person. Households where no contact was attempted are considered as being of unknown eligibility. Thus, a relatively large share of households with unknown eligibility can be interpreted as an indicator for poor fieldwork. However, it has to be noted that the presented figures are solely based on SMS information. Where contact attempts were not properly entered into the system, this could have led to a misclassification of household eligibility. On average, 2.4 percent of longitudinal households appear to have not been attempted for contact.

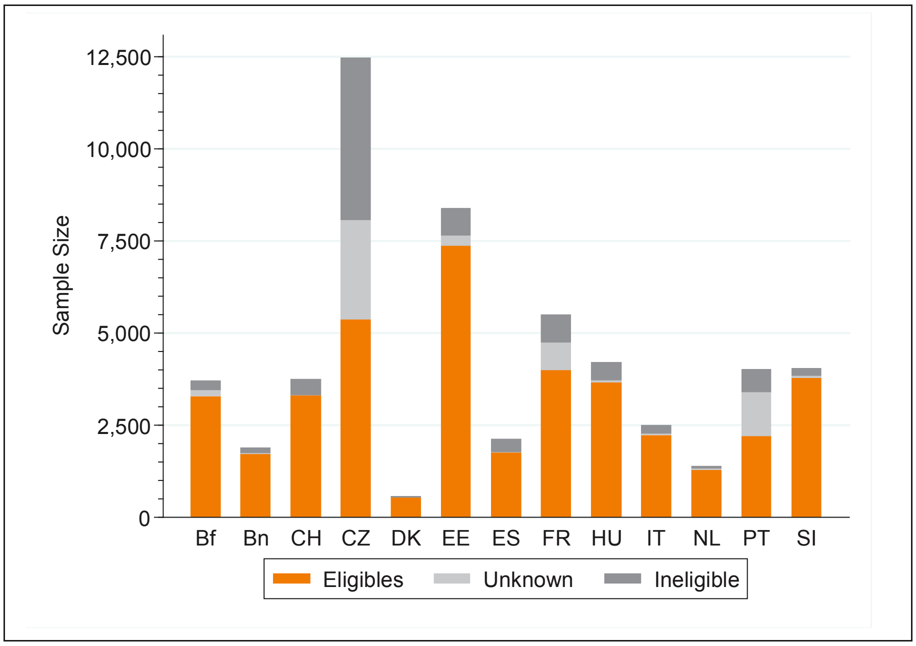


Figure 10.2 Baseline and refreshment samples by classification of sample units

Figure 10.2 shows the size of the refreshment part of the sample or the size of the baseline sample in newly participating countries (EE, HU, PT, SI), respectively. In each country, the size of the gross sample was determined as a function of the target number of interviews and the predetermined minimum eligibility and assumed response rates. Averaging across countries, 74 percent of the gross sample was eligible, 16 percent were ineligible, and 10 percent were with undetermined eligibility. Additional to the aforementioned reasons leading to ineligibility in the longitudinal sample, households are also considered ineligible in cases of death, in-patient treatment during field time, address unknown or invalid, as well as if it turned out during the cover screen interview that there were no eligible persons in the household. In countries where the sample had to be screened first due to no or limited register information (CZ, FR, and PT), (age-) ineligibility could also be an outcome of a screening contact. Thus, in the baseline/refreshment sample the fraction of ineligible households reflects the availability and quality of register information on which sampling was grounded. The fraction of ineligibles was highest in the Czech Republic, where the sample was based on the whole population and the sample frame did not contain any information on the household's age composition. In addition to failure to attempt a household, any form of screening non-response (non-contact, refusal, other non-response) led to classifying a household as unknown with respect to eligibility. Again, this fraction was most pronounced in the Czech Republic. In all countries with a substantial fraction of households with unknown eligibility, this was mainly due to screening non-response. Like in the longitudinal sample, the share of non-attempted cases is rather low.

10.4 Survey participation in the baseline and refreshment sample

10.4.1 Household participation

In order to get households to respond interviewers face the challenge of accomplishing two tasks: first, to successfully establish a contact and second, to gain cooperation of eligible household members. Response can be seen as the outcome of the two sequential events "contact" and "cooperation". The household response rate can thus be decomposed as the product of the household contact rate (i.e. the proportion of eligible households that were contacted) and the household cooperation rate (i.e. the proportion of contacted households that in which at least one interview was completed). There are several ways in which these rates may be computed, depending on how cases of unknown eligibility are handled. They could be considered as entirely eligible, entirely ineligible, or partially eligible. For the rates presented in this chapter it was assumed that only a fraction p of households with unknown eligibility was in fact eligible. For each country, p was estimated by the fraction of eligible households among the cases with known eligibility, which corresponds to assuming that the fraction of eligible households does not depend on whether the eligibility status is known or not. The exact formulas for the different outcome rates are presented in table 10.A.2 in the appendix.

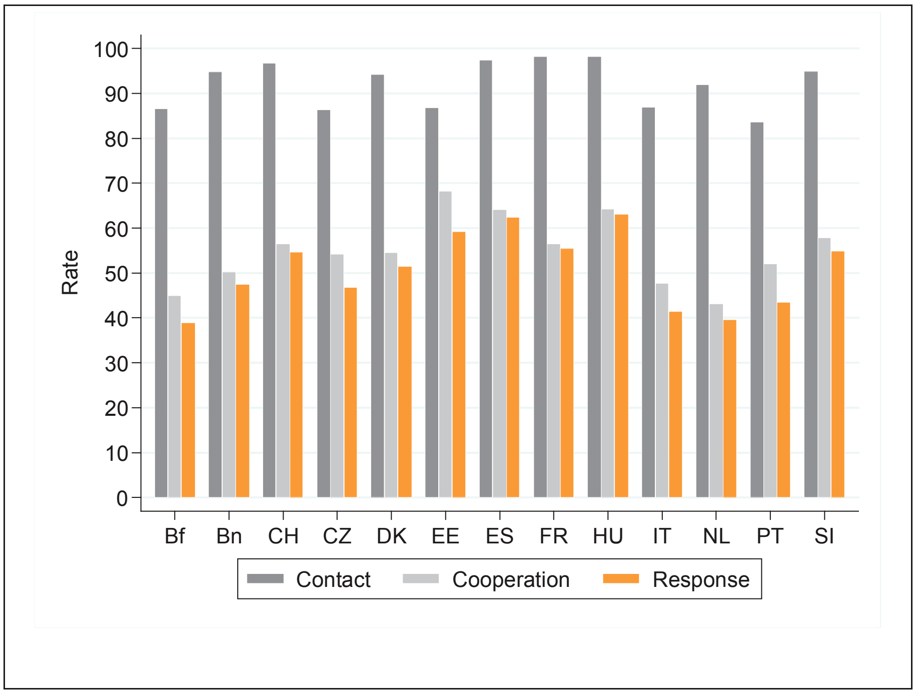


Figure 10.3 Contact, cooperation, and response rates for baseline/refreshment households

Figure 10.3 shows household contact, cooperation, and – as a result – response rates for all countries with sufficient information (i.e. all except Austria). In most countries, establishing a first contact was rather unproblematic for interviewers as reflected by contact rates well above 90 percent. There were some exceptions, though: Contact rates are rather low in the Czech Republic and Portugal – two countries which had to engage in a screening phase due to lacking register information. Contact rates were also below 90 percent in Wallonia (Bf), Estonia, and Italy. Cooperation rates ranged from about 43 percent (NL) to 68 percent (EE). Household response rates varied between about 39 percent for Wallonia (Bf) and about 63 percent for Hungary.

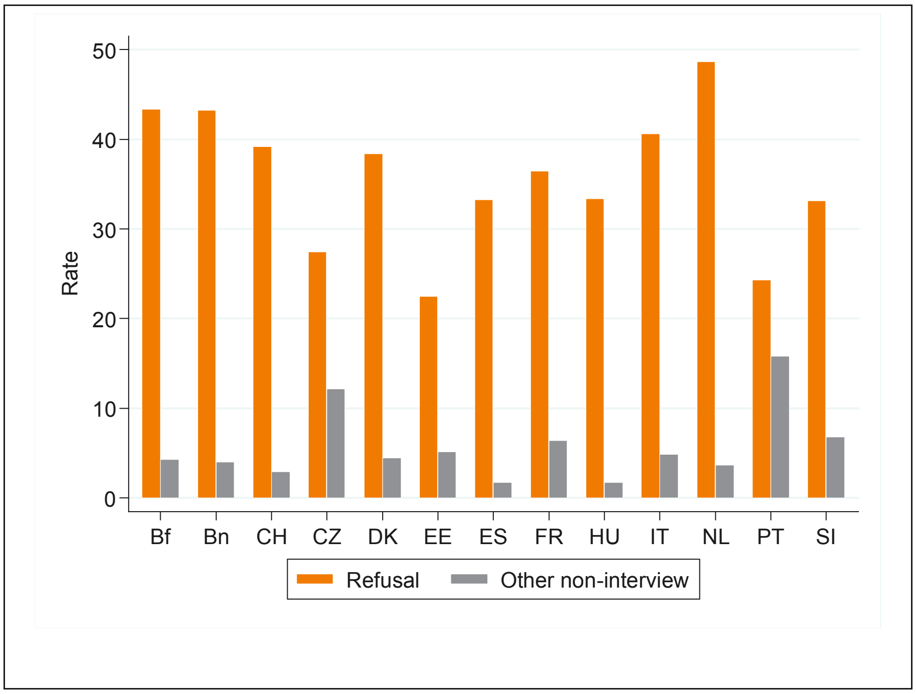


Figure 10.4 Household non-cooperation in the baseline/ refreshment samples

Household non-cooperation can be further classified into refusal versus other reasons leading to non-interview. Refusal was the predominant reason for non-cooperation, yet with a considerable variation in refusal rates among countries ranging from about 22 percent in Estonia to about 49 percent in the Netherlands (see figure 10.4). For some countries, also other non-interview reasons than refusal account for a part of non-cooperation. Non-ineligible households are categorized as other non-interview if they have been successfully contacted without ever refusing cooperation. One instance would be an appointment for an interview that then fell through for any number of reasons (i.e. not result in a completed interview during the field period). This occurred particularly often in Portugal, where the other non-interview rate was about 16 percent. It was lowest (<2 percent) for Spain and Hungary and averaged about 6 percent.

10.4.2 Individual participation

For the above reported patterns of household participation, households were considered as participating if at least one eligible household member was successfully interviewed. Another way of looking at survey participation is to study the response behavior of eligible individuals. This requires restricting the sample to the set of eligible households, and defining the response rate as the proportion of eligible individuals that actually responded. Several definitions of individual response rates are possible depending on how households with unknown eligibility are treated. In addition, the number of eligible individuals in households with an incomplete CV (i.e. where household composition could not be assessed) had to be determined. These households may or may not contain eligible individuals, and different assumptions made about their number directly affect the denominator of the response rate.

If households with known eligibility are divided into those with complete and incomplete CV (H_{CV} and $H_{\overline{CV}}$, respectively), and it is further assumed that only a fraction p of the households with unknown eligibility are in fact eligible, then the number of eligible individuals n is given by

$$n = \bar{n}_{CV} H_{CV} + \bar{n}_{\overline{CV}}(H_{\overline{CV}} + pUE),$$

where $\bar{n}_{CV}$ is the average number of eligible persons in H_{CV} and $\bar{n}_{\overline{CV}}$ is the average number of eligible persons in $(H_{\overline{CV}} + pUE)$. Because $\bar{n}_{\overline{CV}}$ is unknown, an estimate is needed. The numbers presented in this chapter are based on the assumption that, in each country, the average number of eligible persons in $(H_{\overline{CV}} + pUE)$ is the same as in H_{CV}. Accordingly, the total number of eligible persons can be estimated by

$$\hat{n} = \bar{n}_{CV} H_{CV} + \bar{n}_{CV}(H_{\overline{CV}} + pUE) = \bar{n}_{CV}(E + pUE).$$

The average number $\bar{n}_{CV}$ of eligible persons in H_{CV} ranges between a minimum of 1.51 in Wallonia (Bf) and a maximum of 1.77 in Denmark.

The household response rate gives the logical upper boundary of the individual response rate. If all eligible persons in participating households actually participated, both rates would coincide. As figure 10.5 shows, individual response rates are in fact very close to household response rates, indicating that the study managed to interview all eligible persons within a household in a large proportion of cases. On average, 89 percent of all eligible persons in eligible households could be interviewed. This ratio was highest in Hungary (97 percent) and lowest in Slovenia (78 percent). The resulting individual response rates for the baseline/refreshment samples range from about 34 percent in Wallonia (Bf) and the Netherlands to about 61 percent in Hungary.

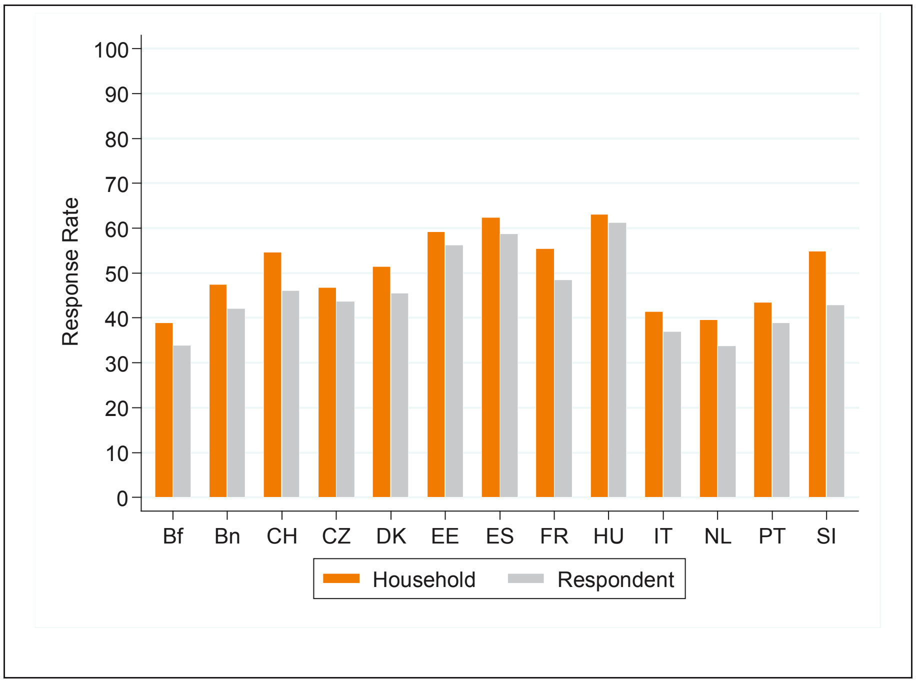


Figure 10.5 Household and respondent-level survey participation in the baseline/refreshment sample

10.5 Survey participation in the longitudinal sample

For a panel study like SHARE, it is most important to keep former respondents participating in the survey. This section investigates participation patterns of households as well as individuals that are part of the longitudinal sample.

10.5.1 Household participation

Retention can, just as response in refreshment samples, be seen as the result of contact and subsequent cooperation of households and individuals who have been successfully interviewed before, either in the wave immediately preceding the current wave, or earlier waves, or any combination thereof. The reported outcome rates are then constructed in the same way as described above for the baseline/refreshment sample (see table 10.A.2). Above, households without any contact attempts were categorized as "unknown eligibility". Yet, they could also be treated as completely eligible. In practice, this does not lead to any difference as the fraction p of non- attempted households, which are in fact eligible, was estimated to be about 99 percent or above in every country. In addition, unknown eligibility is very small in the panel samples.

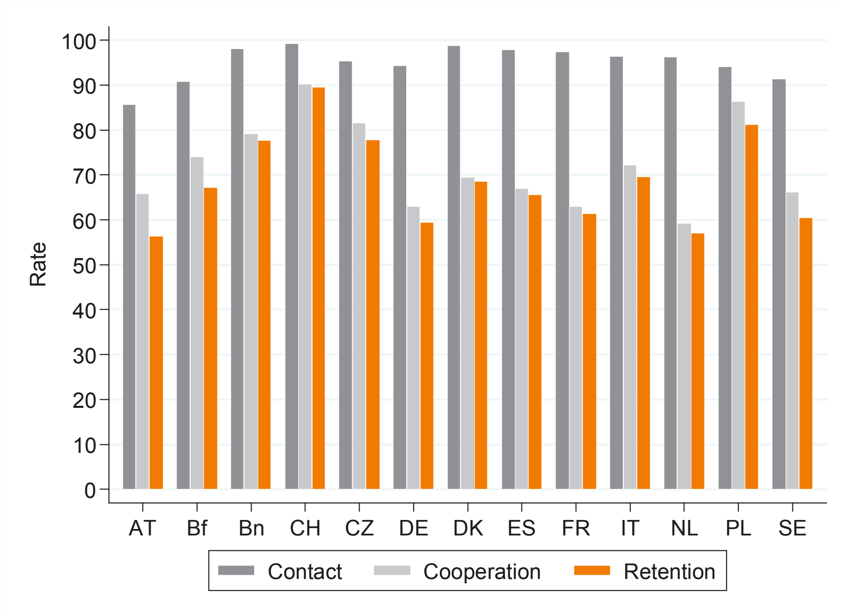


Figure 10.6 Contact, cooperation, and retention rates for longitudinal households

As can be seen in figure 10.6, interviewers' first task of establishing contact was well accomplished in most countries. Apart from Austria, contact rates were consistently above 90 percent with an average of about 95 percent. Household cooperation showed greater variation across countries than contact. Cooperation rates varied between about 60 percent in the Netherlands and about 90 percent in Switzerland. Hence, retention rates, which combine contact and cooperation, varied between below 60 percent (AT, DE, and NL) and about 90 percent (CH). Note that this retention rate refers to the participation in wave four, given that the household has been successfully interviewed in at least one previous wave, not necessarily wave three. Differences among countries are thus partially due to differences in sample composition with regard to participation behavior in previous waves. These, in turn, are a consequence of SHARE's general aim to re-attempt households not participating in a previous wave. Naturally, this only applies if legal restrictions in the participating countries allow for such re-attempts.

Figure 10.7 splits the retention rate of Figure 10.6 into two components: retention of households that participated in wave three, and recovery of households that participated in wave one and/or wave two but not in wave three. When only households that actually took part in the third wave (SHARELIFE) were considered, the resulting retention rates were more homogeneous across countries with an average rate of above 80 percent (see figure 10.7). Recovery rates of households that did not participate in wave three were about 30 percent on average. The efforts in wave four to retrieve previously lost households (see the discussion of subsamples B and C below) were thus quite successful.

Figure 10.7 Household retention by participation status in wave three

As in the baseline/refreshment samples, the main reason for non-cooperation of longitudinal households was refusal to participate. On average this applied to about 20 percent of all longitudinal households but with large differences across countries (see Figure 10.8). Refusal rates in the longitudinal sample were lowest in Switzerland (about 9 percent) and highest in Germany (about 29 percent) However, for most countries non-response was also due to other non-interview reasons. Again, other non-interview reasons were most uncommon in Switzerland (about 1 percent); they were most frequent in the Netherland (about 12 percent). On average, the other non-interview rate for longitudinal samples was about 6 percent.

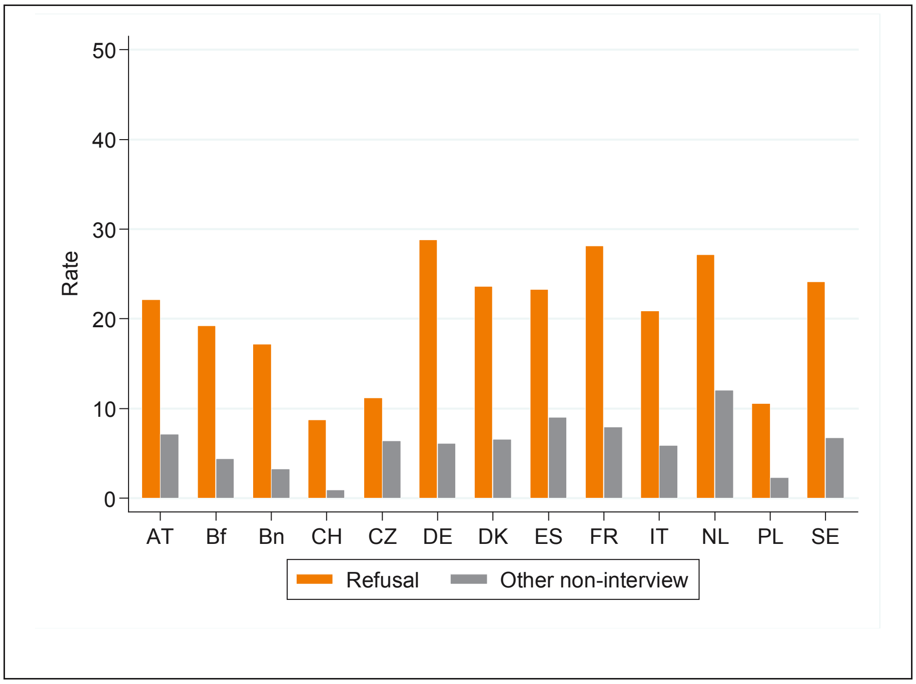


Figure 10.8 Household non-cooperation in the longitudinal samples

10.5.2 Individual participation

Retention rates at the individual level were in general close to those at the household level. Thus, like in the baseline and refreshment samples, most often all eligible members in cooperating households could be interviewed. On average, cooperation was gained from 93 percent of respondents in participating households. The ratio was highest in Poland (about 99 percent) and lowest in Sweden (about 87 percent). The resulting individual retention rates ranged from about 50 percent in the Netherlands to about 80 percent in Switzerland and Poland.

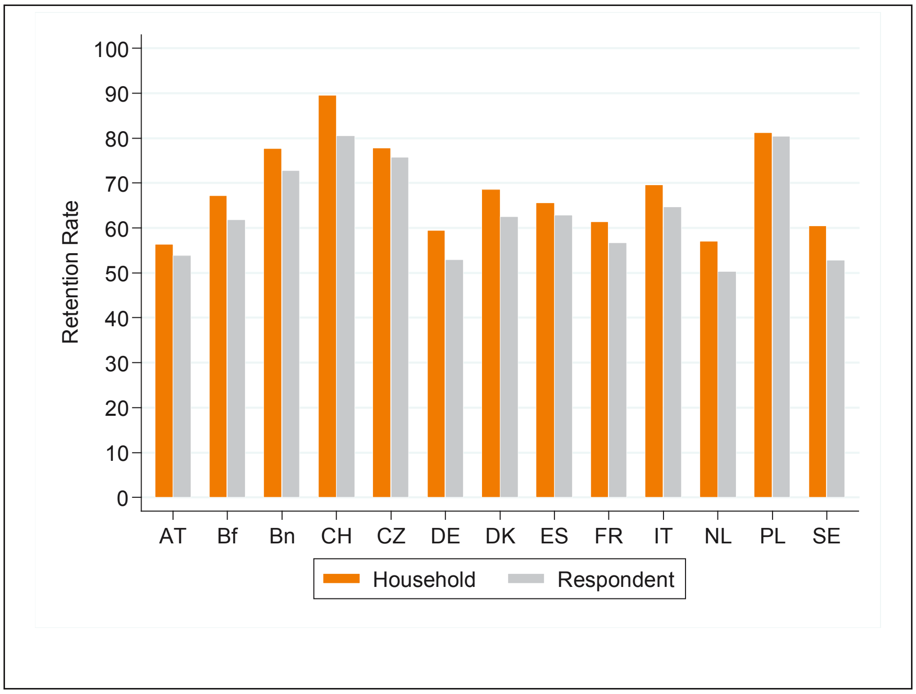


Figure 10.9 Household and Respondent Level Retention

As in the case of household retention, individual retention can be studied conditional on previous participation. At the individual level, longitudinal samples can be divided into four subsamples: Subsample A includes all respondents who participated in the previous wave of the SHARE survey. Subsample B includes those respondents who ever 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. Subsample C includes respondents who ever participated, but not in the previous wave, and do not live in a household where at least one household member participated in the previous wave. Subsample D includes missing and new partners who have not participated in SHARE so far. Thus, individual-level retention in the narrow sense is given by the proportion of respondents in subsample A participating in the fourth wave. Additionally, retention in subsamples B and C informs about how well SHARE managed to get respondents back in who had already dropped out of the study. Finally, response in subsample D is relating to eligible persons in longitudinal households never interviewed before (i.e. either new sample members or eligible sample members for which reluctance to participate was finally overcome after refusals in previous waves).

The individual retention rate in subsample A was considerably higher than individual retention with respect to all eligible household members in the longitudinal sample, irrespective of previous participation behavior. Even in Austria, where subsample A retention was lowest, it amounted to about 74 percent. It is highest in Switzerland and Poland, where about 89 percent of respondents from wave three could be re-interviewed. For subsample B, retention rates varied between about 9 (DK) and 59 percent (PL), the average was about 31 percent. It has to be noted, however, that national subsamples B contained on average only 18 individuals. For subsample C, retention rates ranged from about 15 (IT) to 71 percent (CZ) with an average of about 28 percent. The average sample size in subsample C was 160 individuals per country. For subsample D, retention – or rather response – rates varied between about 7 (DE) and

53 percent (PL) with an average of about 19 percent. The average sample size was 33 individuals per country and thus again rather small.

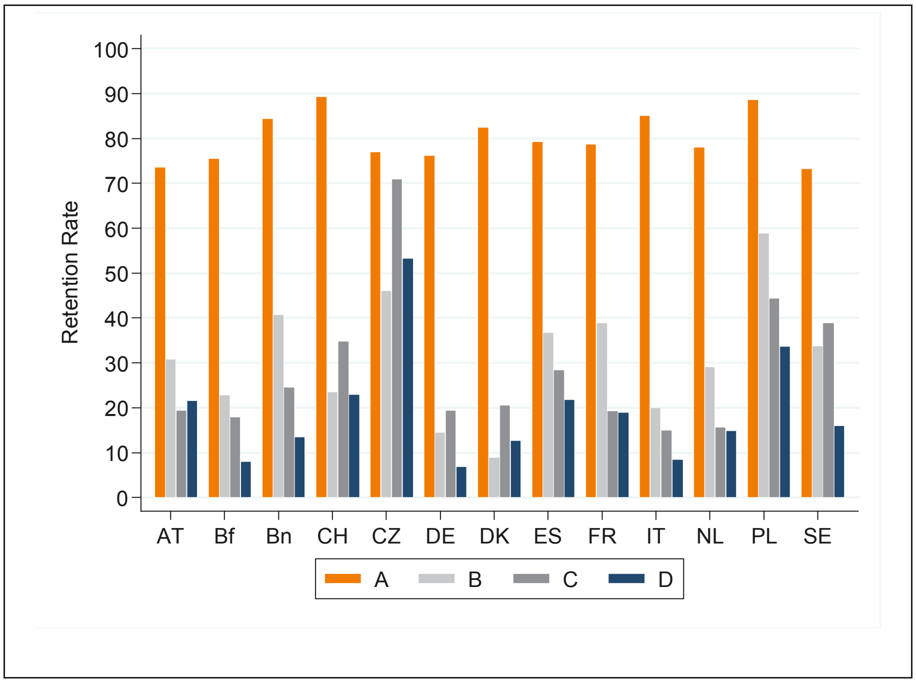


Figure 10.10 Individual retention by type of sample

10.6 Conclusions

Survey participation may be viewed as the result of a sequential process involving eligibility, contact with the eligible units, and cooperation of contacted units. For the fourth wave of SHARE, the analysis of survey participation crucially depended on whether or not the sampling frame contained preliminary information on the eligibility status of the sample units. The probability of selecting ineligible sample units was higher if the sampling frame provided only limited or no information on age eligibility as, e.g., in the case of telephone directories or registers of dwellings. However, once the effects of the different frames were taken into account, it was possible to compare response rates across all countries involved in the project.

Based on our preliminary calculations, the household response rates across those countries which had baseline or refreshment samples in the fourth wave, range from 63 percent in Hungary to 39 percent in Wallonia, averaging slightly below 50 percent. Individual response rates for the baseline/refreshment samples were on average about 6 percentage points lower than household response rates.

For the longitudinal part of the sample, household retention rates of those countries, which participated in wave three of SHARE, averaged 81 percent at the household level and 80 percent at the individual level. In addition, SHARE attempted to recover cases which had already dropped out of the panel. It is known that such cases are particularly hard to bring back. This was particularly successful in the Czech Republic, and least successful in Italy, averaging about 30 percent. All wave retention (i.e., retention of sample members who have participated in any of the earlier waves) ranged from 90 percent in Switzerland to 56 percent in Austria. Individual retention was about 5 percentage points lower.

Focusing attention on the reasons for household non-response, refusal to participate in the survey was the main reason, both in the baseline/refreshment (35 percent) and in the longitudinal sample (20 percent), although a non-negligible fraction of non-participation was also due to non-contact (9 percent in the baseline/refresher sample; 5 percent in the longitudinal sample) or other non-interview reasons (6 percent in the baseline/refreshment sample and in the longitudinal sample).

Understanding non-response behavior is important because it may represent an important source of non-sampling error. Further investigation is needed on how much respondents and non-respondents differ in order to understand whether the sample selection caused by unit non-response is a source of serious bias.

References

De Luca, G., Peracchi, F. (2005). Survey Participation in the First Wave of SAHRE. In A. Börsch-Supan, H. Jürges (Eds.), *The Survey of Health, Ageing, and Retirement in Europe – Methodology*. Mannheim: MEA.

Lessler, J., Kalsbeek, W. (1992). *Nonsampling Error in Survey*. New York: John Wiley & Sons.

Appendix

Table 10.A.1 Detailed list of potentially final SMS Codes

	Final Code
Eligible	E
Completed interview (incl. end-of-life interview)	CI
Partial interview (GS completed)	PI
Not interviewed	NI
Noncontact[1]	NC
Refusal[2]	R
Too busy, no time	
Too old, bad health conditions	
No interest, against surveys	
Other reasons	
Interrupted interview	II
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	
Ineligible	NE
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]	
Unknown Eligibility	UE
No contact attempted	UE_{NCA}
Screening non-response	UE_{NR}
Screening non-contact	UE_{NC}
Screening refusal	UE_{R}
Other screening non-cooperation	UE_{O}

Notes:

[1] Noncontact for the eligible part of the sample does not apply to the baseline/refreshment sample in the Czech Republic, France, and Portugal.

[2] 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.

[3] This led to ineligibility only in the baseline/refreshment sample, but 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, institutionalized cases were always considered eligible.
[5] Subcategories are: age ineligible household, problems with phone/address non-existent, language barriers.

Table 10.A.2 Outcome Rate Formulas

Estimated proportion of eligible households	$p = \frac{E}{E + NE}$
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]	$\frac{(CI_r + PI_r)}{\bar{n}_{CV}(E + p \cdot UE)}$

Notes:

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