

1 Retrospective Data Collection in the Survey of Health, Ageing and Retirement in Europe

Axel Börsch-Supan and Mathis Schröder

Collecting individual data for social research is one of the essential tasks of the social sciences. Suitable data allow for meaningful policy research, e.g., deriving and evaluating welfare programmes, and therefore help to increase the general wellbeing of individuals. Longitudinal data are especially advantageous because they document changes over time – for example how people move from work into retirement when they become older.

One of the principal difficulties faced by longitudinal surveys is how to deal with the “initial conditions” – the lives of respondents before the baseline year of a survey. This is especially problematic for surveys that start monitoring events at later ages. While sometimes, this problem can be overcome when administrative records are available and can be linked to the survey, many domains of people’s lives cannot be linked, as no administrative records exist. In those cases the only realistic option for retrieving such information on the initial conditions is to obtain it directly, albeit retrospectively, from the respondents.

This method has many benefits over regular longitudinal data collection: it is faster, as the respondents give their answers to things that might have happened decades ago. A conventional longitudinal survey would have to wait very long – actually a generation’s time – to link the present to those events in the past. This also means that the retrospective data collection is less costly in obtaining the same information. In addition, the risk of respondents dropping out of the study is much smaller than in longitudinal study.

Some drawbacks vis-à-vis prospective data collection are clear: the retrospective data collection may suffer from recall bias, as respondents may err on when an event actually happened, or on how an event exactly took place. Also, less information can be extracted by means of a retrospective data collection because memory fades.

With its third wave of data collection, the Survey of Health, Ageing, and Retirement in Europe, SHARE, combines the two worlds of retrospective and prospective data collection. SHARE is a longitudinal study that started in 11 European countries in 2004, asking about 30,000 respondents aged 50 or older about their contemporaneous living situation. A second wave was fielded in 2006, which included two new countries as well as refresher samples to keep the initial sample size. SHARE covers multiple aspects, reaching from economic variables and demographics to health variables. For more details on SHARE, see the “First Results Books” by Börsch-Supan et al. (2005) and Börsch-Supan et al. (2008), as well as the “Methodology Book” by Börsch-Supan and Jürges (2005).

A central challenge in SHARE has been from the beginning that all sampled individuals are at least 50 years of age, and thus, many of the “initial conditions” – i.e., all experiences during the first 50 years of each respondents’ life – are unknown to the researcher. This is especially worrying, since a lot of research has proven the importance of early life events for later life outcomes – be it childhood

health for adult health, parental socio economic status on own financial wellbeing or employment history on pension income.

The third wave of SHARE, called SHARELIFE, has been implemented to collect the retrospective histories of the SHARE respondents. This volume provides the methodological background to the data collection and complements the more substantive analyses in Börsch-Supan et al. (2011). It is meant as a support manual to research activities in which the SHARELIFE data are used. It also documents that the latest advances of methodological research were successfully implemented in the process of this large data collection effort. We cover the technical implementation as well as important issues that surround any data collection effort – e.g. fieldwork monitoring, quality assurance, and data management.

The second chapter, written by *Mathis Schröder*, provides an overview of the literature both with respect to recall errors and ways to overcome them. The topics of the SHARELIFE study are then briefly discussed in this context.

The third chapter shows how the (electronic) implementation of previous research was achieved in a large field study such as SHARELIFE. *Marcel Das, Maurice Martens, and Arnaud Wijnant* give a programmers' look at problems in the development of a survey instrument and show the different requirements a CAPI programme needs to fulfil.

The fourth chapter shows how the fieldwork was done during the field phase. This chapter, written by *Barbara Schaan*, presents how the technical achievements – especially the sample management system – connects to the coordination of SHARE's thirteen countries and fourteen survey agencies.

The provision and implementation of certain quality standards is another essential part of the operation. In chapter five, *Kirsten Alcer, Heidi Guyer, and Grant Benson* provide some background and an overview of the results of implementing the first quality profile in SHARE, which was used to achieve a common and comparable quality standard in data collection across all countries.

Data collection in SHARELIFE not only implied more than 1,000 interviewers in contact with more than 25,000 people, but also brings about a tremendous amount of data management work. In chapter six, *Christian Hunkler, Thorsten Kneip, Julie Korbmacher, Stephanie Stuck and Sabrina Zuber* document the processes from SHARE's raw data, coming directly from the field, until the moment the final data used by the research community are released.

Chapter seven then provides an overview of the first important results – the survey participation. *Annelies Blom and Mathis Schröder* explore how the countries differ in their participation rate both on an individual and on a household level. They look at determinants of participation such as gender and age, but also show how previous participation of individuals affects current participation.

The eighth chapter gives a first very brief look at the quality of the data, where *Christelle Garrouste and Omar Paccagnella* are concerned with the congruency of the earlier two SHARE waves with the retrospective information collected in SHARELIFE. They exemplify their analysis by reporting on three variables which are collected both in SHARELIFE and in SHARE's wave 1 and

2. The differences they find are contributed to various circumstances – not the last being the fact, that questions are asked differently across waves, and hence the measurement of difference may not be perfect. However, they conclude that the data are of good quality.

Finally, this book contains an important detail for the researcher – the appendix with the actual questionnaire. While the questions are shown in their English version, of course each country had a specific language version, sometimes even more than one. In addition, the flow of questions, with their respective filtering, is also provided in graphical form, and may help researchers in understanding how the interview situation itself took place. It should be noted, that the website (www.sharelife-project.org) provides all information provided in this book and more – for example country specific questionnaires, latest results and, even more important, the actual data for download.

We hope that with this book, those familiar with the Survey of Health, Ageing and Retirement in Europe find a valuable extension to their already existing research agenda. To those new to SHARE, we hope that SHARELIFE provides an exciting input and inspiration to join the growing community of SHARE users.

Acknowledgements

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.

Collecting these data has been possible through a sequence of contracts by the European Commission and the U.S. National Institute on Aging, as well as support by many of the member states.

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). Further support by the European Commission through the 6th framework programme (projects SHARE-I3, RII-CT-2006-062193, as an Integrated Infrastructure Initiative, COMPARE, CIT5-CT-2005-028857, as a project in Priority 7, Citizens and Governance in a Knowledge Based Society, and SHARE-LIFE (No 028812 CIT4)) and through the 7th framework programme (SHARE-PREP (No 211909) and SHARE-LEAP (No 227822)) is gratefully acknowledged. We thank, in alphabetical order, Giulia Amaducci, Kevin McCarthy, Hervé Pero, Ian Perry, Robert-Jan Smits, Dominik Sobczak and Maria Theofilatou 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, Ruth Paserman, Fritz von Nordheim, and Jérôme Vignon, and by DG Economic and Financial Affairs through Declan Costello, Bartosz Pzrywara and Klaus Regling.

Substantial co-funding for add-ons such as the intensive training programme for SHARE interviewers came from the US National Institute on Ageing (U01

AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, R21 AG025169, 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.

Some SHARE countries had national co-funding which was important to carry out the study. Sweden was supported by the Swedish Social Insurance Agency and Spain acknowledges gratefully the support from Instituto Nacional de Estadística and IMSERSO. Austria (through the Austrian Science Foundation, FWF) and Belgium (through the Belgian Science Policy Administration and the Flemish agency for Innovation by Science and Technology) were mainly nationally funded. Switzerland received additional funding from the University of Lausanne, the Département Universitaire de Médecine et Santé Communautaires (DUMSC) and HEC Lausanne (Faculté des Hautes Etudes Commerciales). Data collection for wave 1 was nationally funded in France through the Caisse Nationale d'Assurance Maladie, Caisse Nationale d'Assurance Vieillesse, Conseil d'Orientation des Retraites, Direction de la Recherche, des Etudes, de l'Evaluation et des Statistiques du ministère de la santé, Direction de l'Animation de la Recherche, des Etudes et des Statistiques du ministère du Travail, Caisse des Dépôts et Consignations, and Commissariat Général du Plan. INSEE (Institut National de la Statistique et des Etudes Economiques) co-founded all 3 waves.

SHARELIFE was a different type of survey than the previous two rounds of interviews, requiring new technologies to be developed and used. Programming and software development for the SHARELIFE survey was done by CentERdata at Tilburg. We want to thank Alerk Amin, Maarten Brouwer, Marcel Das, Maurice Martens, Corrie Vis, Bas Weerman, Erwin Werkers, and Arnaud Wijnant for their support, patience and time. Kirsten Alcser, Grant Benson, and Heidi Guyer at the Survey Research Center (SRC) of the University of Michigan Ann Arbor again provided the Train-the-Trainer programme for SHARELIFE, and invested tremendous amounts of time and work to develop the prototype of a quality profile for the data collection, which included visiting the sites of the national interviewer trainings in participating countries. Kate Cox, Elisabeth Hacker, and Carli Lessof from the National Centre for Social Research (NatCen) gave helpful input in designing the questionnaire and pointed out the retrospective specifics in the interview process. We always kept in close contact with the professional survey agencies – IFES (AT), PSBH, Univ. de Liège (BE), Link (CH), SC&C (CZ), Infas (DE), SFI Survey (DK), Demoscopia (ES), INSEE (FR), KAPA Research (GR), DOXA (IT), TNS NIPO (NL), TNS OBOP (PL), and Intervjubolaget (SE) – and thank their representatives for a fruitful cooperation. Especially the work of the more than 1000 interviewers across Europe was essential to this project.

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. Input into the concepts of retrospective questionnaires came from Robert Belli and David Blane. The life history questionnaire has been implemented first in the ELSA study, and without the help of people involved there (James Banks, Carli Lessof, Michael Marmot and James

Nazroo), SHARELIFE could not have been created in such a short time. SHARELIFE has also greatly profited from detailed advice given by Michael Hurd, Jim Smith, David Weir and Bob Willis from the HRS as well as by the members of the SHARE scientific monitoring board: Orazio Attanasio, Lisa Berkman, Nicholas Christakis, Mick Couper, Michael Hurd, Daniel McFadden, Norbert Schwarz and Andrew Steptoe, chaired by Arie Kapteyn. Without their intellectual and practical advice, and their continuing encouragement and support, SHARE would not be where it is now.

Since SHARELIFE was an entirely newly designed questionnaire, the work of developing and constructing the questions was immense. We are very grateful to the contributions of the eight working groups involved in this process. Specifically, Agar Brugiavini, Lisa Calligaro, Enrica Croda, Giacomo Pasini, and Elisabetta Trevisan developed the module for financial incentives of pension systems. Johannes Siegrist and Morten Wahrendorf provided input for the module on quality of work and retirement. The development of questions for the part of disability insurance and labour force participation of older workers was responsibility of Hendrik Jürges, whereas the health and retirement section was constructed by Johan Mackenbach and Mauricio Avendano. Preventive care, health services utilisation, and retirement fell into the realm of Brigitte Santos-Eggimann and Sarah Cornaz, and Karsten Hank provided his input for the gender, family, and retirement section. Wealth and retirement questions were designed by Guglielmo Weber and Omar Paccagnella, and finally, questions on health risk, health insurance, and saving for retirement were integrated by Tullio Japelli and Dimitri Christelis.

A large enterprise with 150 researchers in thirteen countries entails also a large amount of day-to-day work, which is easily understated. We would like to thank Kathrin Axt, Maria Dauer, Marie-Louise Kemperman, Tatjana Schäffner, and Eva Schneider at the MEA in Mannheim for their administrative support throughout various phases of the project. Annelies Blom, Martina Brandt, Karsten Hank, Hendrik Jürges, Dörte Naumann, and Mathis Schröder provided the backbone work in coordinating, developing, and organizing the SHARELIFE project. 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 Christian Hunkler, Thorsten Kneip, Julie Korbmacher, Barbara Schaan, Stephanie Stuck, and Sabrina Zuber for data cleaning and monitoring services at the MEA in Mannheim, and Guiseppe de Luca and Dimitri Christelis for weight calculations and imputations in Padua, Salerno and Venice. Thorsten Kneip and Marco Hambuch provided the essentials for the flowcharts and the questionnaire representation in this book, while Elisa Sept was responsible for the design work around the book.

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); Karel van den Bosch, Sergio Perelman, Claire Maréchal, Laurant Nisen, Jerome Schoenemaekers, Greet Sleurs and Aaron van den Heede (Belgium); Radim Bohacek, Michal Kejak and Jan Kroupa (Czech Republic); Karen Andersen

Ranberg, Henriette Engberg, and Mikael Thinggaard (Denmark); Anne Laferrère, Nicolas Briant, Pascal Godefroy, Marie-Camille Lenormand and Nicolas Sirven (France); Axel Börsch-Supan and Karsten Hank (Germany); Antigone Lyberaki, Platon Tinios, Thomas Georgiadis and George Papadoudis (Greece); Guglielmo Weber, Danilo Cavapozzi, Loretta Dobrescu, Christelle Garrouste and Omar Paccagnella (Italy); Frank van der Duyn Shouten, Arthur van Soest, Manon de Groot, Adriaan Kalwij and Irina Suanet (Netherlands); Michał Myck, Malgorzata Kalbardczyk and Anna Nicinska (Poland); Pedro Mira and Laura Crespo (Spain); Kristian Bolin and Thomas Eriksson (Sweden); Alberto Holly, Karine Moschetti, Pascal Paschoud and Boris Wernli (Switzerland).

References

- Börsch-Supan, A., Brugiavini, A., Jürges, H., Mackenbach, J., Siegrist, J., Weber, G. (eds.) (2005). *Health, Ageing and Retirement in Europe – First Results from the Survey of Health, Ageing and Retirement in Europe*. MEA: Mannheim.
- Börsch-Supan, A., Jürges, H. (eds.) (2005). *The Survey of Health, Ageing and Retirement in Europe – Methodology*. MEA: Mannheim.
- Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G. (eds.) (2008). *Health, Ageing and Retirement in Europe – Starting the Longitudinal Dimension (2004-2007)*. MEA: Mannheim.
- Börsch-Supan, A., Brandt, M., Hank, K., Schröder, M. (eds.) (2011). *The Individual and the Welfare State: Life Histories in Europe*. Springer: Heidelberg.