

a practical guide to

USING PANEL DATA

2

DIFFERENT TYPES OF PANEL SURVEYS

Aim

Panel surveys can be collected for different purposes and, like other surveys, they have different features. In this chapter we discuss the main aspects of panel surveys: who is interviewed, how many times, how the data can be collected. We then give a short overview of some frequently used panel datasets.

2.1 Introduction

In this book we discuss four major panel surveys in the world and how to use them to implement different longitudinal estimation methods. These four surveys are: the Panel Study of Income Dynamics (PSID) in the USA, the German Socio-Economic Panel (SOEP), the British Household Panel Survey (BHPS) and Understanding Society: the UK Household Longitudinal Study (UKHLS).

The SOEP, BHPS and UKHLS started by interviewing a set of private households selected from all private households in a specific country at a particular point in time, and then continued to interview these same household members (and their descendants) at regular intervals. The PSID is slightly different since it focuses on individuals rather than households (see Section 2.4). In all four surveys most of the information collected during the interviews is related to the household members' current situation. Some of these surveys

also collect information about respondents' labour market, partnership and fertility activities before the start of the survey. In Section 2.7 we discuss the specific features of each of these surveys. To understand why these particular features were chosen it is important to understand the advantages and disadvantages of various survey features. We provide an overview of these in Sections 2.2 to 2.6.

2.2 Censuses Versus Sample Surveys

All surveys, including panel surveys, can be categorised into two main types: censuses and sample surveys. The panel surveys discussed in this book belong to the latter category; sample surveys are more popular than censuses.

The purpose of empirical analyses is often to estimate population parameters such as means, proportions and correlation between certain characteristics. Depending on the purpose of the research, a population could be all individuals, or households or firms in a country. It could also be all residents of a city, or all pupils in a school, and so on. For example, we may want to know whether in the UK the average pay of men differs from that of women, what proportion of single mothers participate in the labour market, or what is the effect of living in an economically deprived neighbourhood on the chances of getting a job.

To answer such kinds of questions, ideally we would want to gather information from every member of the population of interest. This type of survey is referred to as a census or a complete enumeration survey. Unfortunately, a census of a relatively large population geographically spread over the whole country is often too costly in terms of money and time. However, we can use information gathered from sample surveys to infer parameters and correlations of interest in the population. Sample surveys collect information from a subset of the population of interest, referred to as a sample. The four surveys discussed in this book are all sample surveys, not censuses.

A sample can be drawn from a population using different methods; in this book we discuss different types of sampling techniques. The quality of an estimator based on a sample depends on the sampling technique; we discuss these issues in Chapter 7.

The estimation techniques we discuss in the second and third parts of this book are some of the most popular methods used to estimate population parameters (quantities in the population) from sample statistics (the same quantity in the sample). For example, the average pay of adults living in the UK in 2010 is a population parameter, while the average pay of a sample drawn from this population is a sample statistic. A sample statistic used to estimate the population parameter is referred to as an estimator. Statistical techniques are then used to determine the quality of the estimate.

2.3 Surveys Versus Administrative Data

There are a variety of sources of longitudinal data; an important distinction is between surveys and administrative data. Administrative data are collected by government agencies or regulating bodies for their own administrative purposes. For example, government departments may collect certain information on people's characteristics to verify whether they are entitled to certain types of benefits, and record the actual amount of each type of benefit paid to claimants, for the whole duration of the claim. Administrative databases therefore usually include a large number of individuals; in some cases, they include information on the whole population. For example, if our population of interest is that of people on unemployment benefit, then administrative data on unemployment claimants are essentially censuses. If our population of interest is that of people who are searching for a job, including those who do not qualify or apply for unemployment benefit, then these data are a (possibly selected) sample of the population of interest. To the extent that they refer to the whole population of interest, administrative data can be considered better than survey data. Since they refer to a sample of the population, survey data are subject to sampling error.

Because administrative data are collected for administrative purposes only, they often include only basic information on the relevant individuals and quite often lack detailed socio-demographic information that may be useful for research purposes. Furthermore, because of confidentiality reasons, these types of data may be difficult to access for research purposes.

Also note that, depending on the purpose of data collection, some administrative data may be cross-section rather than longitudinal. The longitudinal data are often collected only over the period of time that the person satisfies certain criteria for inclusion. For example, data on unemployment benefits include information about individuals currently receiving unemployment benefits. Anyone who stops claiming such benefit falls out of scope and is from that moment on excluded from the data.

In contrast to administrative data, surveys are specifically collected for research purposes. We may distinguish between multipurpose and specific-purpose surveys. While in multipurpose surveys a large number of questions are asked on different aspects of people's lives, specific-purpose surveys may be commissioned to collect detailed information on specific aspects of people's lives (for example, environmental behaviour, consumption patterns, and so on).

The information reported by individuals in sample surveys may not always be accurate, for example when respondents do not remember certain events or facts, when they have difficulties understanding the question, when they do not feel comfortable answering sensitive questions such as questions about their income (see Groves et al. 2004). On the other hand it could be argued that administrative data may also be subject to error.

With the explicit consent of survey respondents, survey data can be linked to administrative data such as hospital episodes, benefits or educational records. This is becoming more common since such data linkage allows a set of substantial and precise information to be imported into the survey data. The linkage is more valuable the higher the proportion of survey respondents who give consent to the linkage, and the higher the success of the linkage itself (the linkage between survey and administrative records often relies on individual information such as name, year of birth, and so on, being recorded unambiguously across surveys).

2.4 Who is Included in the Survey and for How Long?

Surveys differ in terms of population of interest, which in turn determines who is included in the survey and for how long. One important distinction is between individual and household surveys.

Individual surveys ask questions to individual respondents about different aspects of their lives. Information on other members of the household may also be collected, but the interview is with only one person in the household. Household surveys instead ask questions to all people living in the household, thus producing multiple individual interviews for households with more than one adult. Household surveys add another layer of complexity for data management. Among the datasets discussed in this book, the PSID started as an individual panel survey, while the SOEP, BHPS and UKHLS started as household panel surveys.

Another important distinction is between individual panels and cohort surveys. While panel data include people of all ages living in a certain country at a certain point in time, cohort surveys focus on individuals with a common experience at a specific time period, such as being born in a particular year, or entering a country in a particular period, or graduating from high school in a particular year or decade. The most common types of cohort surveys are birth cohort surveys. These are individual surveys which collect information on individuals born in a particular short time span and follow them over their life cycle with interviews at specific points in time, often determined by age. As the sample members are of similar age, this increases comparability across respondents, but may reduce the range of research questions that can be answered with these types of data; for example, even if we combine data from different cohort surveys we may not be able to analyse the impact of the most recent recession on different age groups.

Various birth cohort surveys have been commonly used for research. The National Child Development Study follows a sample of people born in the UK in a single week in 1958, with interviews at ages 7, 11, 16, 23, 33, 42, 46, 50 and currently 55. The 1970 British Cohort Study follows a sample of people born in the UK in a single week in 1970, with interviews at ages 5, 10, 16, 26, 30, 34

and 38 and most recently 42. The Millennium Cohort Study follows a sample of people born in the UK in 2000–01, with data collected at the ages of 9 months, 3, 5, 7 and currently 11 years. The National Longitudinal Survey of Youth 1979 follows a sample of people born in the USA between January 1957 and December 1964 with annual interviews from 1979 (when the respondents were aged between 16 and 22) to 1992 and biennial interviews after that. The National Longitudinal Survey of Youth 1997 follows a sample of people born in the US between January 1980 and December 1984 with annual interviews from 1997, when the respondents were aged between 13 and 17.

In contrast to cohort surveys, panel surveys include a sample of people who are resident in a certain area at a certain time period and follows them, and sometimes their descendants as well. For example, the PSID started with a sample of people resident in the USA in 1967–8; the SOEP started with a sample of households living in West Germany in 1983–4; while the BHPS started with a sample of households living in Great Britain in 1990–91; and the UKHLS started as a sample of households living in the UK in 2008–09. The inclusion in the survey of people of all ages increases the heterogeneity of the respondents in terms of their life experiences; for example, the sample may include respondents both born and grown up during periods of boom and periods of recession.

Finally, surveys differ in terms of their length. The length of a survey is often determined a priori depending on the needs of the commissioning body and of the survey itself. In fixed life surveys respondents are followed up to a certain number of periods, after which they are typically substituted by a new set of respondents. Fixed life surveys are often used in rotating panel surveys, such as the UK Labour Force Survey (LFS). Respondents in the LFS are interviewed quarterly for up to five successive quarters, after which they are substituted by a new set of respondents sampled from the population currently living in the UK. Hence, the sample is divided into five parts: each quarter, one-fifth of the LFS respondents have their first interview, one-fifth each have their second, their third and their fourth interviews, while one-fifth have their fifth and final interview.

In contrast to fixed life surveys, perpetual surveys have indefinite life, with no scheduled end: individuals are interviewed as long as they remain within the scope of the survey (for example, as long as they live in the country). The household panel surveys discussed in this book are all perpetual life surveys (the BHPS officially ended after 18 waves but the surviving sample became part of Understanding Society after that).

2.5 Methods of Data Collection

The main methods of data collection are face-to-face personal interviews, telephone interviews and self-completion interviews (where the respondents

complete a questionnaire by themselves). Self-completion questionnaires are often mailed to respondents, although are now increasingly implemented via the Web. Each of these methods has different implications on cost, non-response and quality in general; see Roberts (2007) for a review of advantages and disadvantages of the different modes of data collection. Nowadays surveys are increasingly being administered in multiple modes (De Leeuw 2005). Some parts of the interviews may be conducted in a different mode, or some of the sample members may be interviewed in a different mode. For example, most BHPS sample households are interviewed face-to-face but around 500 are interviewed by telephone, and those interviewed face-to-face also receive self-completion paper questionnaires to fill in.

As discussed in Chapter 1, retrospective and prospective surveys are the common types of longitudinal surveys, deriving their names from the different data collection methods employed.

In retrospective surveys, sample members are interviewed (often only once) and asked about their past history, so that the longitudinal element of the survey (the repeated information) is built very quickly and relatively cheaply, as generally only one interview is necessary.

In prospective surveys, sample members are interviewed every year (or every few years) and are asked about their current situation each time they are interviewed. In this case the longitudinal element of the survey is constructed (slowly) year by year via successive yearly interviews, although note that not all prospective surveys are annual, or even take place at regular intervals. Hence, prospective surveys are more expensive than retrospective ones. However, compared with retrospective surveys, prospective surveys are more likely to produce better quality data (see Chapter 1 of Taris 2000). The first reason is that people may forget details or the exact dates of events, especially when these happened a long time in the past; retrospective surveys are more likely to suffer from recollection bias than prospective surveys. The second reason is that retrospective surveys also suffer from survivor bias. By design, retrospective surveys are collected from those who have survived (in the literal sense or more generally those who are still present in the population of interest) until the time of the interview. If those who survived are systematically different from those who did not in terms of the variable of interest then estimates based on these data will be biased, hence survivor bias. Take for example a retrospective survey on a sample of migrants to study outcomes such as labour market success. If, by the time the survey is conducted, those migrants who had poor labour market success have returned to their home country (hence they left the population of interest or did not survive) then estimates of labour market success of migrants based on these data will be biased upwards.

It is worth noting that most prospective surveys also collect some data retrospectively (see the online appendix for details on the data collection of some commonly used panel datasets).

2.6 Attrition and Refreshment Samples

As already mentioned in Chapter 1, not all sample members are interviewed in every interview wave (unit non-response). For example, some people may change their residence without communicating the change to the survey organisers, who may therefore be unable to locate them for further interviews. Even when people can be located, the survey organiser may not be able to get in touch with them to fix an interview date (for example, some people may be away or too ill to answer the door or the phone). Some of them may have decided they no longer want to be part of the survey and refuse to give further interviews. In Chapter 1 we also discussed the difference between wave non-response and attrition: the former refers to intermittent non-response and the latter to dropping out of the survey permanently.

Non-response or attrition is a serious problem for sample-based estimations. First, it reduces sample size, thus reducing the precision of estimates based on these data. Secondly, attriters and non-respondents may be systematically different from respondents so that the estimates based on respondent reports may also be biased. In Chapter 7 we discuss the relevance of non-response and attrition in more details. If the sample size reduces to such a level that it greatly reduces the precision of estimates based on these data, then sometimes refresher samples are drawn from the original population to replenish the reduced sample.

Since panel and cohort surveys collect information from a sample of residents which is drawn at one particular point in time, these surveys may miss out on population changes. New births are usually incorporated in the sample design by determining who, from the original sample, should be followed: for example, respondents and their current and future offspring. Nevertheless, new waves of immigration are likely to be missed by many sample designs (exceptions are rotating panels, which incorporate regular refresher samples into the sample design). Hence, sometimes refresher samples drawn from the new population (for example, new immigrants) may be added to the original sample at a later stage in the life of the panel to account for the population changes.

2.7 Household Panel Surveys Around the World

There are various panel surveys collected around the world. Below we discuss the main characteristics of some of the most commonly used ones: namely, the PSID, SOEP, BHPS and UKHLS.

Although most of the examples in this book are based on the BHPS, in Chapters 3 to 6 and in Chapter 12 we discuss all four datasets. Below we focus on the most important characteristics of these surveys and refer to the online appendix for more details.

Further details about each of the four surveys can be accessed through their user guide as well as through technical papers focusing on specific aspects of the survey. There is also excellent online interactive documentation which is a good place to start. The questionnaires and other materials used during surveys, user guides and technical papers are also available online. Some people find it helpful to join the respective user groups and attend training workshops offered by the data managers.

These surveys are multipurpose surveys and collect information on socio-demographic characteristics (such as gender, age, ethnicity or race, region and country of birth), family background (such as living arrangement and parents' employment when respondent was a child), education and training, labour market activities, partnership and fertility, income, wealth and assets, self-reported physical and mental health, and values and opinions. In addition, the UKHLS includes information on bio-markers for a subsample.

2.7.1 The Panel Study of Income Dynamics (PSID)

The PSID is one of the oldest ongoing household panel surveys. The survey started in 1968 with a sample of about 18,000 individuals in 4,802 families living in the USA, with data collected annually from the head of the household. This is the male for married couples, and either the male or the female for other families (for a recent overview see McGonagle et al. 2012).

The respondent provides information about him- or herself and other members of the family, although in specific supplements (like the ones about retrospective histories) both members of the couple are interviewed. From the 1970s, however, the PSID collects the same details also from the spouse of the head of the household. The survey intends to follow members of the original family unit and their offspring (note that the PSID focuses on families while the other three datasets discussed below focus on households).

The 1968 PSID sample comprised two independent subsamples: the 'SRC' (Survey Research Center) and the 'SEO' (Survey of Economic Opportunity), referred to as the 'core samples'. The SRC sample was collected by the Survey Research Center, at the Institute for Social Research, University of Michigan, and the SEO sample was collected by the Bureau of the Census for the Office of Economic Opportunity.

The SRC sample was a sample representative of the US population, with equal selection probabilities in the 48 coterminous states. This sample produced 2,930 interviews. In the SEO sample, low-income families had comparatively higher selection probabilities. The SEO sample included only people living in Standard Metropolitan Statistical Areas (SMSAs), and people living in non-SMSAs in the southern region. This sample produced 1,872 interviews. The 18,000 individuals living in these interviewed households became permanent sample members or were said to have the 'PSID gene'. These individuals and all

their biological or adopted descendants would be followed over time. In 1990, an additional sample of 2,043 Latino individuals (Mexican, Cuban and Puerto Rican) was added but dropped after 1995 due to lack of funding and under-representation of immigrants of Asian descent (this short-lived subsample is rarely used by researchers for analysis). Two additional immigrant samples were added in 1997 and 1999, respectively, and are referred to as the 'immigrant samples'. To reduce costs, 57% of the SEO sample was also dropped in 1997; this was also the year when the interview frequency changed from annual to biennial. Interviews in the PSID were conducted face-to-face from 1968 to 1972 and by phone from 1973. This change led to much shorter interviews.

2.7.2 The German Socio-Economic Panel (SOEP)

The SOEP is a household panel survey of the non-institutionalised resident population in Germany. It started in 1984 with a sample of private households in West Germany which included a sample of households with non-ethnic minority household heads (called subsample A) and an oversample of households with household heads of Turkish, Greek, Yugoslavian, Spanish or Italian origin (called subsample B).

In 1990 an oversample of private households with a German Democratic Republic citizen as head of household was added (subsample C), followed by an immigrant boost sample in 1994 and 1995 (subsample D), a refreshment sample with similar design as subsample A (subsample E) in 1998 and a boost sample of high-income earners in 2002 (subsample G). In 2000 an innovation sample with a similar design as subsamples A and E was added.

All household members aged 16 and over are eligible for personal interviews, which are conducted primarily face-to-face with a small self-completion component. Information is collected prospectively, with some retrospective elements such as employment, partnership and fertility histories. In addition to the information collected in personal interviews, the head of the household is also asked questions about the household, parents about their children, and interviewers about the interview process.

All members of households who participated in the first wave and their descendants are followed as long as they live in Germany; people who join these households are also interviewed and since 1998 are also followed as long as they live in Germany.

2.7.3 The British Household Panel Survey (BHPS)

Similar to the SOEP, the BHPS is a household panel survey, of people living in the UK. The data collection began in 1991 on a sample of approximately 5,500

households and 10,000 adult respondents in England, Wales and (most of) Scotland.

To allow analyses separately for the three British countries, since 1999 the survey includes additional samples of about 1,500 households living in Scotland, and a further 1,500 households living in Wales. In 2001 a sample of about 2,000 households living in Northern Ireland was added to the BHPS sample.

Interviews are carried out annually with each adult member (aged 16 and over) of the selected households, with youth self-completion questionnaires for children aged 11–15 added from 1994. All members of households who participated in the first wave and their descendants, referred to as Original Sample Members (OSMs), are followed as long as they live in the UK. Respondents who join these households after the first wave, referred to as Temporary Sample Members (TSMs), are also interviewed, but only as long as they are living with the original household members. When TSMs have children with OSMs, they become Permanent Sample Members (PSMs) and are followed even when they stop living with an OSM.

Among the panel surveys discussed in this book, the BHPS is the only one which is not currently ongoing. As a stand-alone survey, the BHPS finished in 2008. From 2010 the sample of surviving households from the BHPS has been incorporated into a new and larger UK household longitudinal survey Understanding Society.

2.7.4 Understanding Society: The UK Household Longitudinal Study (UKHLS)

The UKHLS is a household panel survey drawn from the non-institutionalised resident population of the UK. The UKHLS includes a sample of approximately 26,000 households living in the UK, and a sample boost of about 4,000 households belonging to selected ethnic minorities. The BHPS sample of 6,500 households that survived after the eighteenth wave became part of the UKHLS sample from the second wave onwards.

The large sample size is a peculiar characteristic of the UKHLS compared with other panel surveys. This allows analysis of specific subpopulations such as pre-retirement workers, single mothers, and so on; the ethnic minority boost sample makes it possible to analyse the selected ethnic minority groups separately.

All household members who are aged 16 and over are eligible for personal face-to-face and self-completion interviews; young people aged 10 to 15 are asked to complete a self-completion questionnaire. Most information is collected prospectively, with some retrospective elements such as partnership and fertility histories. Similar to the BHPS and SOEP, the household reference persons answer questions regarding the household, and parents answer questions about their children. In addition, interviewers provide information about the interview process.

The following rules for the UKHLS are similar to those of the BHPS with some exceptions: only children of OSMs who are women become OSMs; only TSMs who are men and who have children with an OSM become PSMs and are followed as long as they are in the UK. While the BHPS sample members maintain their sample status when they join the UKHLS, once they become part of the UKHLS, the following rules of UKHLS apply.

An additional sample, called the Innovation Panel, with a sample of approximately 1,500 households is interviewed one year prior to the UKHLS main survey. The Innovation Panel is used for research on survey methodology issues and survey 'experiments'. It is essentially a different survey from the main UKHLS dataset, and should not be used in conjunction with the main UKHLS sample.

Compared with the other three surveys discussed above, the UKHLS puts a much greater emphasis on health measures, particularly direct health measures such as grip strength, blood pressure, lung function test, and so on. Cognitive ability tests have also been administered. These measures have been collected at one point in time from subsamples of respondents.

2.7.5 Other Panel Surveys

Quite a few countries have now started their own household panel surveys, which often have a structure very similar to the SOEP and BHPS. However, these surveys are also rather heterogeneous across countries since questions are often asked differently in each national survey. In an attempt to produce datasets that allow cross-country comparisons, the Development of Policy Analysis and Management at Cornell University in collaboration with a number of institutes around the world has developed the so-called Cross-National Equivalent File (CNEF). The 1970–2009 CNEF contains harmonised variables for the PSID, the SOEP, the BHPS, the Household Income and Labour Dynamics in Australia, the Canadian Survey of Labour and Income Dynamics, and the Swiss Household Panel. Recent additions to the CNEF are the Korean Labor and Income Panel Study and the Russian Longitudinal Monitoring Survey.

The CNEF provides a simplified version of the national panels and guidelines for formulating equivalent variables across countries. The CNEF files also include a set of constructed variables that are not directly available in the original surveys and that can be merged into the national surveys and included in the analyses.

In Europe, some panel datasets have been collected in harmonised ways across countries. One of these is the European Community Household Panel (ECHP). Data are available annually from 1994 to 2001, when the survey officially ended. ECHP data have been collected using a standardised questionnaire. The first wave of data was collected from about 60,500 households (130,000 adults aged 16 years and over) in the 12 European Community Member States (Belgium, France, Italy, Luxembourg, the Netherlands, Germany, Denmark,

Ireland, the United Kingdom, Greece, Portugal and Spain). Data collection for Austria started in 1995; data for Finland have been added since 1996, and for Sweden since 1997.

Cross-country comparability in the ECHP is achieved by means of common survey structure and procedures, common questionnaires and common standards for data processing.

A more recent source of household panel data harmonised across EU countries is the European Union Statistics on Income and Living Conditions (EU-SILC). EU-SILC data cover the 25 EU Member States, Norway and Iceland. The EU-SILC is intended to be a finite-life sample, in which households are interviewed for up to a maximum of four (consecutive) years. In contrast to the ECHP, EU-SILC data are not based on a common questionnaire, but on a list of harmonised target variables, collected following common guidelines, procedures and concepts.

2.8 Summary and Suggestions for Further Reading

In this chapter we have discussed the different types of longitudinal data that can be used for analysis. In the rest of the book we focus on panel surveys rather than census or administrative data, although most of the data management and econometric techniques discussed are relevant for most types of longitudinal datasets. The book mostly uses the BHPS in its examples, but data management techniques (and the online appendix) also include detailed explanations of PSID, SOEP and UKHLS data.

Key points

- Surveys differ in many aspects: who is interviewed (households or individuals), how the data are collected (prospectively or retrospectively), and so on.
- There are numerous panel surveys around the world, with different degrees of cross-country comparability.
- Attempts have been made – and are ongoing – to produce data that are comparable across countries.

Suggestions for further reading

- For a good primer on survey methodology including inference, sampling, non-response, interview modes and questionnaire design, and see Groves, R.M., Fowler, J.F.J., Couper, M.P., Lepkowski, J.M., Singer, E. and Tourangeau, R. (2004) *Survey Methodology*. Hoboken, NJ, Wiley-Interscience.

- There is extensive research on different methods of data collection and some that recommend using mixed modes and designing surveys so as to minimise any differences in response due to differences in interview mode (unimode design). See:
 - De Leeuw, E. (2005) To Mix or Not to Mix Data Collection Modes in Surveys. *Journal of Official Statistics* 21(2): 235–55.
 - Dillman, D.A. (2007) *Mail and Internet Surveys: The Tailored Design Method*. Hoboken, NJ, Wiley.
 - Roberts, C. (2007) Mixing Modes of Data Collection in Surveys: A Methodological Review. NCRM Methods Review Papers (NCRM/008).
- A concise overview of different data collection methods, specifically administrative and sample surveys, their advantages and disadvantages, can be found on the Statistics Canada website: www.statcan.gc.ca/edu/power-pouvoir/ch2/5214911-eng.htm.
- The Administrative Data Liaison Service provides support and help to researchers using UK administrative data for their research. See www.adls.ac.uk/.
- Nowadays the most detailed user guides for the panel datasets we have briefly discussed here are all online. See the US PSID: <http://psidonline.isr.umich.edu/>. On this website, you will find that the PSID also offers short video podcasts which clearly explain some key aspects of the study: <http://psidonline.isr.umich.edu/videos.aspx>. Although created to train interviewers, it is quite useful for new users as well. A short overview outlining the main features is available from www.uncg.edu/bae/people/ribar/teaching/ECO725/notes/intro_PSID.pdf. See also the SOEP: www.diw.de/en/soep; the UK BHPS: <https://www.iser.essex.ac.uk/bhps> and UKHLS: www.understandingsociety.org.uk/; the CNEF: www.human.cornell.edu/pam/research/centers-programs/german-panel/cnef.cfm; the ECHP: epp.eurostat.ec.europa.eu/portal/page/portal/microdata/echp; and the EU-SILC: epp.eurostat.ec.europa.eu/portal/page/portal/microdata/eu_silc.