

SOEP Survey Papers

Series G – General Issues and Teaching Materials

SOEP Glossary

Running since 1984, the German Socio-Economic Panel study (SOEP) is a wide-ranging representative longitudinal study of private households, located at the German Institute for Economic Research, DIW Berlin.

The aim of the SOEP Survey Papers Series is to thoroughly document the survey's data collection and data processing.

The SOEP Survey Papers is comprised of the following series:

Series A – Survey Instruments (Erhebungsinstrumente)

Series B – Survey Reports (Methodenberichte)

Series C – Data Documentation (Datendokumentationen)

Series D – Variable Descriptions and Coding

Series E – SOEPmonitors

Series F – SOEP Newsletters

Series G – General Issues and Teaching Materials

The SOEP Survey Papers are available at
<http://www.diw.de/soepsurveypapers>

Editors:

Dr. Jan Goebel, DIW Berlin

Prof. Dr. Martin Kroh, DIW Berlin and Humboldt Universität Berlin

Prof. Dr. Carsten Schröder, DIW Berlin and Freie Universität Berlin

Prof. Dr. Jürgen Schupp, DIW Berlin and Freie Universität Berlin

Please cite this paper as follows:

SOEP Group. 2015. SOEP Glossary. SOEP Survey Papers 285: Series G. Berlin: DIW/SOEP



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/). © 2015 by SOEP

ISSN: 2193-5580 (online)

German Socio-Economic Panel (SOEP) | DIW Berlin
Mohrenstr. 58
10117 Berlin, Germany
Contact: soepapers@diw.de

SOEP GROUP

SOEP GLOSSARY

Berlin, 2015

SOEP GLOSSARY

SOEP-Core	2
Contents of SOEP-Core	2
SOEPlong.	3
SOEPregio	3
SOEP Pretest	3
SOEP-LEE	3
CNEF–Cross-National Equivalent File of the SOEP	4
LIS	4
SOEP-IS (Innovation Sample)	5
SOEP-IS	5
SOEP-RS (Related Studies)	6
FiD data (Families in Germany)	6
BASE II (Berlin Aging Study II)	6
PIAAC-L.	6
SOEP-ECEC Quality (K ² ID-SOEP)	6
IAB-SOEP Migration Sample	6
TwinLife (cooperation study)	6
SOEP Service	7
SOEPnewsletter	7
SOEPlit	7
SOEPcampus	7
SOEPmonitor	7
SOEP-in-Residence	7
SOEP Re-Analysis	7
SOEPdataFAQ	7
Digital Object Identifiers (DOI)	8

SOEP-Core

The German Socio-Economic Panel (SOEP) is a wide-ranging representative longitudinal study of private households, located at the German Institute for Economic Research, DIW Berlin. Every year, there were nearly 11,000 households, and more than 20,000 persons sampled by the fieldwork organization TNS Infratest Sozialforschung. The SOEP study is available in the two formats „SOEPCore“ and „SOEPlong“.

Contents of SOEP-Core

The SOEP started in 1984 as a longitudinal survey of private households in the Federal Republic of Germany. The central aim then and now is to collect representative micro-data to measure stability and change in living conditions by following a micro-economic approach enriched with variables from sociology and political science. Therefore the central survey instruments are a household questionnaire, which is responded by the head of a

SOEP-Core Topics

Year	Wave number	Wave letter	Topic
1986	3	C	Residential environment and neighborhood
1987	4	D	Social security, transition to retirement
1988	5	E	Household finances and wealth
1989	6	F	Further occupational training and professional qualifications
1990	7	G	Time use and time preferences; Labor market and subjective indicators
1991	8	H	Family and social networks
1992	9	I	Social security (2nd measurement)
1993	10	J	Further occupational training (2nd)
1994	11	K	Residential environment and neighborhood (2nd); Working conditions; Expectations for the future
1995	12	L	Time use (2nd)
1996	13	M	Family and social networks (2nd)
1997	14	N	Social security (3rd)
1998	15	O	Transportation and energy use; Time use (3rd)
1999	16	P	Residential environment and neighborhood (3rd); Expectations for the future (2nd)
2000	17	Q	Further occupational training (3rd)
2001	18	R	Family and social networks (3rd)
2002	19	S	Wealth and assets (2nd); Social security (4th); Health (SF12, BMI)
2003	20	T	Transportation and energy use (2nd); Trust; Time use (4th)
2004	21	U	Residential environment and neighborhood (4th); Further occupational training (4th); Risk aversion; Health (2nd)
2005	22	V	Expectations for the future (3rd); Big Five; Reciprocity
2006	23	W	Family and social networks (4th); Working conditions (ERI); Health (3rd); Grip strength
2007	24	X	Wealth and assets (3rd); Social security (5th)
2008	25	Y	Further occupational training (5th); Health (4th); Grip strength (2nd); Trust (2nd); Time use (5th)
2009	26	Z	Residential environment and neighborhood (5th); Risk aversion (2nd); Big Five (2nd); Globalization and transnationalization; Diseases
2010	27	BA	Consumption and saving; Reciprocity (2nd); Health (5th); Grip strength (3rd)
2011	28	BB	Family and social networks (5th); Working conditions (ERI) (2nd); Diseases (2nd)
2012	29	BC	Wealth and assets (4th); Social security (6th); Health (6th); Grip strength (4th)
2013	30	BD	Big Five (3rd); Trust (3rd); Loneliness; Working conditions (ERI) (3rd); Diseases (3rd)
2014	31	BE	Health (7th); Risk aversion (3rd); Globalization and transnationalization (2nd); Residential environment and neighborhood (6th);

household and an individual questionnaire, which each household member is intended to answer. Furthermore beginning with 1997, there are wave-specific \$LELA files (Lebenslauf - engl. life course) containing the biography information as collected in the respective year.

A rather stable set of core questions is asked every year covering the most essential areas of interest of the SOEP:

- population and demography
- education, training, and qualification
- labor market and occupational dynamics
- earnings, income and social security
- housing
- health
- household production
- preferences and values
- satisfaction with life in general and certain aspects of life.

Additionally, yearly topical modules enhance the basic information in (at least) one of these areas by asking detailed questions as documented in the following table. These modules for the main part appear in the personal questionnaires; only some of them are additions to the household questionnaire. Starting in the year 2001, the data have become even richer by including several different health measures and well-known psychological concepts as well as age specific questionnaires.

SOEPlong

SOEPlong is a highly compressed, easily analyzable version of the SOEP data that is much simpler to handle than the usual SOEP-Core version. It contains a significantly reduced number of datasets and number of variables.

The data are no longer provided as wave-specific individual files but rather pooled across all available years (in “long” format). In some cases, variables are harmonized to ensure that they are defined consistently over time. For example, the income information up to 2001 is provided in euros, and categories are modified over time when versions of the questionnaire are changed. All these modifications are clearly documented and described for ease of understanding. In the case of recoding or integration of data (for example, datasets specific to East German or foreign populations), documentation is generated automatically and all modified variables are provided in their original form as well. SOEPlong thus provides a well-documented compilation of all variables and data that is consistent over time.

SOEPregio

SOEP offers diverse possibilities for regional and spatial analysis. With the anonymized regional information on the residences of SOEP respondents (households and individuals), it is possible to link numerous regional indicators on the levels of the states (Bundesländer), spatial planning regions, districts, and postal codes with the SOEP data on these households. However, specific security provisions must be observed due to the sensitivity of the data under data protection law ([see overview](#)). Accordingly, users are not allowed to make statements on, e.g., place of residence or administrative district in their analyses, but the data does provide valuable background information.

SOEP Pretest

Within the framework of SOEP, the questionnaires are pretested before being fielded each year. The aim of the pretests is to test new sets of questions or modifications to certain questions. Furthermore, behavioral experiments are prepared and tested and sometimes even included in the main SOEP survey. A pretest in the SOEP usually includes about 1,000 respondents. The samples are representative by approximation for the population aged 16 years and older in Germany. Data are collected by Infratest and passed on to the SOEP, which makes the data available to external users. <https://paneldata.org/studies/5>.

SOEP-LEE

There is increasing consensus in the economic and social sciences that the workplace plays a crucial role in individual life outcomes. This is true in the economic and sociological labor market research, network and social capital research, health research, the research on educational and competency acquisition processes, wage information, and the work-life interface, as well as in the inequality research as a whole. For this reason, there has been increasing interest in what are known as “linked employer-employee” (LEE) datasets, in which employees' individual data are linked with information on their employers.

The workplace data collected in the framework of the project SOEP-LEE will substantially expand the information on the work contexts and working conditions of respondents to the SOEP survey. The project has been implemented by asking all dependent employees in all of the SOEP samples to provide local contact information to their employer in 2011. The employer contact data then formed the basis for a standardized employer survey conducted separately from the rest of the SOEP survey. This employer information can be linked with the individual and household data from the SOEP study.

The new linked employer-employee dataset opens up new opportunities for wide-ranging forms of secondary analysis with innovative questions from wide range of disciplines in the social and economic sciences. An additional unique feature of SOEP-LEE is the analysis of employer survey data quality, carried out through the measurement of meta- and paradata over the course of data collection. As a result, this project also contributes to the ongoing development and refinement of survey methodology in the field of organizational studies. http://www.diw.de/en/diw_01.c.433198.en/soep_lee.html

CNEF—Cross-National Equivalent File of the SOEP

The International Science Use Version of the SOEP (95% version) can be used worldwide. The Research Data Center SOEP is providing it upon request for free via secure download.

CNEF data will no longer be distributed by Cornell University, but by Ohio State University. At the moment, an order form is not available, but the conditions are unchanged: \$125 one-time charge at first order. More information is given here: Cross-National Equivalent File Project <http://cnef.ehe.osu.edu/>

LIS

LIS, the cross-national data center in Luxembourg—formerly known as the Luxembourg Income Study—will soon turn 32 years old. While LIS' mission and core work have not changed since its inception—that is, to acquire and harmonize high-quality microdatasets and to make them available to researchers around the world—LIS is constantly evolving and growing, as is its user community which currently numbers in the thousands.

LIS' data holdings are organized into two databases. The longstanding **Luxembourg Income Study (LIS) Database**, which is focused on income data, will soon contain over 300 datasets from more than 50 high- and middle-income countries. The smaller and newer **Luxembourg Wealth Study (LWS) Database** contains microdata on assets and debt; *LWS* now includes 20 datasets from 12 countries. (Germany was one of the earliest participating countries; the *LIS* and *LWS Databases* contain 11 and 2 datasets from Germany, respectively.) www.lisdatacenter.org

SOEP-IS (Innovation Sample)

The research infrastructure SOEP at DIW Berlin established a longitudinal Innovation Sample (SOEP-IS) in 2011 for particularly innovative research projects. The SOEP-IS is primarily available for methodical and thematic research that involves too high a risk of non-response for the long-term SOEP study.

SOEP-IS

- is based on an evaluation conducted by the German Council of Science and Humanities.
- is a longitudinal sample for particularly innovative survey methods and behavioral experiments.
- will be further developed in the period from 2012 to 2017 and should be fully developed by 2017.

The annual fieldwork runs from September to December of each year. The first wave of the first subsample of the SOEP-IS started in September 2011, with a newly developed core questionnaire “SOEP Innovations” and new methods to measure gender stereotypes.

The overall volume and costs of the surveys conducted in the SOEP-IS are lower than if “fresh” samples were used: central household and individual characteristics, invariant over time, are already available and do not have to be collected again.

A two-step module of Governance is established to regulate topics and question modules: first, the SOEP survey management runs a basic methodological test to establish whether the size, format, and survey mode outlined in a proposal seem appropriate for implementation in the SOEP-IS. The SOEP Survey Committee then checks the content of proposals received and prioritizes these for selection purposes.

Information about SOEP-IS in general and about the application process is published in: SOEP Innovation Sample (SOEP-IS) - Description, Structure and Documentation by David Richter and Jürgen Schupp (SOEPpaper 463).

www.diw.de/soep-is

paneldata.org/studies/4

SOEP-RS (Related Studies)

FiD data (Families in Germany)

The project *Familien in Deutschland* (Families in Germany) – is a longitudinal panel study financed by the German Federal Ministry for Family Affairs, Senior Citizens, Women and Youth (BMFSFJ) and the German Federal Ministry of Finance (BMF). Its main purpose is to provide researchers with new and better data on specific groups in the German population: low-income families, families with more than two children, single parent families, as well as families with young children. The data are the backbone of the first large-scale evaluation of family policy measures in Germany on behalf of the two involved ministries. In 2014 FiD has been fully integrated into SOEP-Core.

BASE II (Berlin Aging Study II)

The Berlin Aging Study II (BASE-II) is an extension and expansion of the Berlin Aging Study (BASE). This new study with more than 2,200 participants of different ages aims to complement the analysis of cognitive development across the lifespan by including socio-economic and biological factors such as living conditions, health, and genetic preconditions. The study is funded by the Federal Ministry of Education and Research. Participants are involved in the annual survey of the German Socio-Economic Panel (SOEP) and provide information about their life situation and living conditions.

paneldata.org/studies/3

PIAAC-L

The Programme for the International Assessment of Adult Competencies (PIAAC), carried out on behalf of the OECD, examines the basic skills that are necessary for adults to participate successfully in society and working life. Findings from the 2011/2012 wave of the PIAAC study were released in October 2013.

Around 98% of the approximately 5,400 PIAAC survey respondents in Germany agreed to participate in further surveys. PIAAC-L is a cooperative project of GESIS, the National Educational Panel Survey (NEPS) at the Leibniz Institute for Educational Trajectories (LifBi), and the Socio-Economic Panel (SOEP) at DIW Berlin, whose aim is to convert the PIAAC study into a longitudinal study with three waves. This will create one of the world's first internationally comparable longitudinal studies on competencies and their significance across the life course.

SOEP-ECEC Quality (K²ID-SOEP)

Are some groups of parents in Germany more likely to choose high-quality education and care institutions for their children than others, e.g. whether due to a lack of information or varying preferences? Are mothers whose children attend high-quality settings more satisfied and more likely to be employed?

These are some of the questions studied as part of the project “Early childhood education and care quality in the Socio-Economic Panel (K²ID-SOEP)—direct and indirect effects on child development, socio-economic selection and information asymmetries.” The three-year project launched in September 2013 is funded by the Jacobs Foundation (see <http://kid2id.de>).

IAB-SOEP Migration Sample

The IAB-SOEP Migration Sample is a joint project of the Institute for Employment Research (IAB) and the Socio-Economic Panel (SOEP) at the German Institute for Economic Research (DIW Berlin). The project attempts to overcome limitations of previous datasets through a sample that takes into account changes in the structure of migration to Germany since 1995. The dataset is an additional sample for the SOEP-Core study and therefore completely harmonized with the SOEP and integrated into SOEP v30 (identical questionnaire with additional questions on the respondent's migration situation). The study opens up new perspectives for migration research and gives insights on the living situations of new immigrants to Germany. Data collector: TNS Infratest Sozialforschung GmbH.

paneldata.org/studies/6

TwinLife (cooperation study)

TwinLife is a 12-year representative behavioral genetic study investigating the development of social inequality. The long-term project has begun in 2014 and will survey more than 4,000 pairs of twins and their families regarding their different stages of life on a yearly basis. All of the subjects reside in Germany. Not only social, but also genetic mechanisms as well as covariations and interactions between these two parameters can be examined with the help of identical and fraternal twins. In order to document the individual development of different parameters it is important to examine a family extensively over the course of several years. The focus is on five important contextual points: Education and academic performance, career and labor market attainment, integration and participation in social, cultural and political life.

<http://www.twin-life.de/en>

SOEP Service

SOEPnewsletter

Above and beyond the comprehensive documentation and the various user support programs, the SOEP Research Data Center publishes the quarterly SOEPnewsletter, containing the latest updates on data, conferences, and related information, and distributes it by email to the constantly growing international SOEP user community.

www.diw.de/SOEPnewsletter

SOEPlit

Many of the research findings and publications based on SOEP data are archived at DIW Berlin. You will find the bibliographic descriptions in our SOEPlit database. In addition, we collect publications based on the European Community Household Panel (ECHP) and the Luxembourg Income Study (LIS), as the data on Germany contained within these international comparable data sets are partly generated from SOEP data.

To keep this service up to date, we ask all authors to send us copies of all of their publications based on SOEP data by e-mail to: soeplit@diw.de

www.diw.de/SOEPlit

SOEPCampus

The SOEP is working to strengthen methodological training in the use of SOEP data—especially for young scholars in the disciplines of sociology, economics, and psychology. In addition to holding workshops at universities, we list workshops and lectures providing introductions to the use of the SOEP data or dealing with particular issues of data use on our website at:

<http://www.diw.de/soepcampus>.

SOEPmonitor

The SOEPmonitor compiles time series since the mid 1990's for chosen indicators, calculated on basis of the SOEP data. The most important function of the SOEPmonitor—aside from reporting detailed information on the situations of individuals and households—is to give SOEP users a benchmark for their own studies. With the figures contained in the SOEPmonitor, we offer an important reference point to evaluate the results of your own research.

Simultaneously the numerical series of the SOEPmonitor represent social indicators. With every issue of the

SOEPmonitor, we provide data series for the years 1984 to the current wave disaggregated for East and West Germany since 1990 households and persons. Since the 2007 SOEPmonitor tables are in English as well.

www.diw.de/SOEPmonitor

SOEP-in-Residence

In addition to offering SOEP users the standard Scientific Use File (via secured download), a special mode of online access (via SOEPremote), and advice over the SOEPHotline, we also provide the opportunity to conduct research during a stay in the SOEP Department at DIW Berlin. Direct discussion with SOEP team members and our user-friendly environment provide fruitful input and support, enabling visiting scholars to work effectively on research projects and bring them to successful completion. For the use of small-scale coded geodata, a research stay at the SOEP Data Research Center located at DIW Berlin is mandatory. SOEP also provides research stays to address special research questions and topics. Furthermore, research visit to SOEP's field organization, TNS Infratest Sozialforschung, are also possible.

https://www.diw.de/en/diw_02.c.222617.en/soep_in_residence/html

SOEP Re-Analysis

Data protection issues are of utmost importance to SOEP and CNEF users as well. First, data protection comprises part of the (implicit) contract between the survey and the respondent. Second, in order to access the data, users are required to address data protection issues thoroughly. Ultimately, all these precautions are crucial to ensure future participation by panel respondents. As such, making SOEP and CNEF data available for re-analyses while maintaining the highest levels of data protection can present a major-challenge. Whenever such a microdata set is not considered completely anonymous from a legal point of view, we—as data producers—are not permitted to allow archiving without setting and guaranteeing adherence to clear-cut access regulations. http://www.diw.de/en/diw_01.c.340858.en/soep_re_analyses.html

SOEPdataFAQ

The FAQ is updated with the answers to the most frequent questions that were submitted to the SOEP-Hotline and to the researchers.

http://www.diw.de/en/diw_02.c.222860.en/faq_questions_about_data_analyses.html

Digital Object Identifiers (DOI)

The need for replicability of findings makes it necessary to be able to identify and cite the particular SOEP data used in research. One way of doing this is through the system of Digital Object Identifiers (DOI), which is already being used for numerous publications. It is also well-suited for research data, and is therefore now being used for the SOEP data as well.

Digital identifiers provide a form of permanent identification for digital objects and thus guarantee that they can be found again on the Internet. They are a basic requirement for citing and finding research data on the Internet, even when the location (URL) has changed. A series of meta-data are linked with each DOI (defined in the "metadata schema") in order to guarantee improved description and recognition of the data.

The SOEP RDC, as a publication agent, will be assigned the prefix 5684 in each DOI registered via da|ra. It is important for SOEP users to know that this does not change anything about our proposed mode of citation for the SOEP data. Rather, this provides you with the additional possibility to add a unique DOI to your citations.

Because precise references to data sources are becoming increasingly important in the scientific research community, the SOEP group recommends citing the SOEP data as follows.

English:

Socio-Economic Panel (SOEP), data for years 1984-2013, version 30, SOEP, 2015, doi:10.5684/soep.v30.

German:

Sozio-oekonomisches Panel (SOEP), Daten für die Jahre 1984-2013, Version 30, SOEP, 2015, doi:10.5684/soep.v30.

Short Version:

SOEP v30