SOEP The Socio-Economic Panel

898

SOEP Survey Papers

Series C – Data Documentation

SOEP – The German Socio-Economic Panel at DIW Berlin 2020

The Socio-Economic Module of the Berlin Aging Study II (SOEP-BASE): Description, Structure, and Questionnaire

Anke Boeckenhoff, Denise Sassenroth, Martin Kroh, Thomas Siedler, Peter Eibich, Gert G. Wagner

DIW SOEP

Running since 1984, 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.

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**

Please cite this paper as follows:

Anke Boeckenhoff, Denise Sassenroth, Martin Kroh, Thomas Siedler, Peter Eibich, Gert G. Wagner. 2020. The Socio-Economic Module of the Berlin Aging Study II (SOEP-BASE): Description, Structure, and Questionnaire. SOEP Survey Papers 898: Series C. Berlin: DIW/SOEP.

ISSN: 2193-5580 (online)

DIW Berlin
German Socio-Economic Panel (SOEP)
Mohrenstr. 58
10117 Berlin, Germany

Contact: soeppapers@diw.de

The Socio-Economic Module of the Berlin Aging Study II (SOEP-BASE): Description, Structure, and Questionnaire

Anke Boeckenhoff, Denise Sassenroth, Martin Kroh, Thomas Siedler, Peter Eibich, Gert G. Wagner

Berlin, 2020

Preprint 2013

The Socio-Economic Module of the Berlin Aging Study II (SOEP-BASE): Description, Structure, and Questionnaire

Anke Boeckenhoff[1], Denise Sassenroth[1], Martin Kroh[1,2], Thomas Siedler[1,3], Peter Eibich[1] and Gert G. Wagner[1,4]

[1] German Institute for Economic Research (DIW Berlin)

[2] Humboldt-Universität zu Berlin

[3] Universität Hamburg

[4] Max Planck Institute for Human Development, Berlin, and Berlin University of Technology (TUB)

Berlin, July 2013

Abstract

The Berlin Aging Study II (BASE-II) is a multidisciplinary study that allows for the investigation of how a multitude of health status factors as well as many other social and economic outcomes interplay. The sample consists of 1,600 participants aged 60 to 80, and 600 participants aged 20 to 35. The socio-economic part of BASE-II, the so called SOEP-BASE, is conducted by the SOEP Group at the DIW Berlin. The surveyed socio-economic variables are fully comparable with the variables of the long running German Socio-Economic Panel Study (SOEP), which increases the analytical power of BASE-II. The socio-economic data collected on the individual and on the household level are enriched with geo-referenced context data ("neighbourhood data") in order to disentangle the interplay between individual, societal and regional determinants on individuals' health status and other outcome variables. Furthermore, as the BASE-II study is based on a convenience sample, the SOEP Group at the DIW provides weights for the BASE-II dataset that correct for selectivity bias.

Zusammenfassung

Die Berliner Altersstudie II (BASE-II) ist eine multidisziplinäre Studie, die es erlaubt, das Zusammenspiel verschiedenster Faktoren auf den Gesundheitsstatus und viele andere soziale und ökonomische Outcomes zu untersuchen. Die Stichprobe besteht aus 1.600 Teilnehmern zwischen 60 und 80 Jahren und 600 Teilnehmern im Alter von 20 bis 35 Jahren. Der sozio-ökonomische Teil von BASE-II wird von der SOEP-Gruppe am DIW Berlin durchgeführt und die entsprechenden SOEP-BASE-Daten mit den SOEP-Daten voll vergleichbar gemacht. Daten liegen sowohl auf der Individualebene als auch auf Haushaltsebene vor und werden mit geo-referenzierten Kontext-Daten (auf der Ebene von „Nachbarschaften") angereichert, sodass das Zusammenspiel von individuellen, gesellschaftlichen und regionalen Determinanten auf den individuellen Gesundheitszustand und andere Outcome-Variablen herausgearbeitet werden kann. Da die BASE-II Studie auf einem Convenience-Sample beruht, werden extensive Selektions-Analysen durchgeführt und darauf aufbauend wird der BASE-II Datensatz mit Gewichten versehen, die für Selektionsverzerrungen korrigieren.

Keywords: *Cohort Study, Berlin Aging Study, BASE-II, SOEP*
JEL Classification: *C81, C83, I12, I14, I24, I31, I32, J14, Y80, Z13*

1 Introduction

The Berlin Aging Study II (BASE-II) is a joint project of five research institutions, namely the SOEP Group at the DIW Berlin, the Research Group Geriatrics at the Charité Berlin, the Ageing and Tumour Immunology Group (TATI) of the University of Tuebingen, the Max Planck Institute for Molecular Genetics, and the Max Planck Institute for Human Development. BASE-II is funded by the German Federal Ministry of Education and Research (BMBF).[1]

The focus of this paper solely is on the socio-economic part of the BASE-II study (in the following called SOEP-BASE[2]) conducted by the German Socio-Economic Panel Study (SOEP) group at the DIW Berlin (German Institute for Economic Research). Section 2 delineates the SOEP-BASE in detail and briefly describes the research areas of each research unit involved in the BASE-II study[3] in order to provide an overview of the whole study. The third section documents the target population and the data structure of SOEP-BASE. Since the target population of the BASE-II study is a convenience sample, analyses of self-selection of participants and weighting of the data are of particular importance for data of BASE-II. Thus section 3.3 addresses this issue and provides some basic information on weighting prospects and selectivity of the BASE-II data. Section 4 gives an overview of the concepts, instruments and inventories used in the SOEP-BASE study. Finally, the possibilities of data access for external users are presented in the fifth section.

2 Description and Background of the Berlin Aging Study II

Aging as a systemic and dynamic process is interrelated with genetic factors, and likely to be influenced by various demographic, social, economic or psychological factors. BASE-II

[1] Support by the BMBF via VDI/VDE, grant numbers #16SV5536K, #16SV5537, #16SV5538, and #16SV5837 (previous BMBF support via DLR-grant #01UW0808) until December 2014. Another source of funding is the Max Planck Institute for Human Development (MPIB), Berlin, Germany. Additional contributions (e.g. equipment, logistics, and personnel) are made from each of the other participating sites.

[2] From the SOEP's point of view BASE-II is a "related study" (cf. Siedler et al. 2009). All such related studies have acronyms that begin with "SOEP" (instead of "BASE II").

[3] Further information on the BASE-II study can be found on the program website, located at http://www.base2.mpg.de

achieves a better understanding about the conditions for “successful aging” (Rowe & Kahn, 1987).

BASE-II presents a unique longitudinal study aggregating somatic, behavioral and social dimensions of the developments in adulthood. For this, BASE-II covers not only hypothesis-driven analyses but also broad possibilities for explorative studies. BASE-II, as a truly multi-disciplinary (and multi-institutional) study, contains much potential for the robust analyzation and understanding of aging processes. However, the study concentrates on identifying the causes for healthy and unhealthy aging as well as their interdependencies and changes during the process of aging.

With the different participating project partners originating from diverse disciplines, BASE-II focuses on integrating evidence from ‘macro level’, e.g. physical health, behavior, and aspects of the environment, with data at the ‘micro level’, e.g. represented by cellular and various molecular and biochemical parameters (Bertram et al., 2013).

The study’s central topics can be summarized as follows:

- cardiovascular status, physical fitness and cognition;
- molecular genetics, social and environmental correlations with life satisfaction;
- education, metabolic disorder and cognition;
- immune biomarkers and health ; and
- mobility and prevention of drops.

The comprehensive collection of bio-measures, accomplished by the Charité's Geriatrics Research Group at the Evangelisches Geriatriezentrum Berlin, constitutes a major part of the study by providing objective measures of an individual’s health. Moreover, Charité collects data on nutrition, subjective well-being and sex life using self-administered questionnaires. One of the aims of the study is to eventually predict future health and diseases from this data.

In addition, each BASE-II participant undergoes several medical examinations which shed light on the participant’s diseases, namely heart diseases, metabolic disorders, neuropsychiatric disorders, musculoskeletal disorders, cancer and disorders of the sensory organs. In addition, the Charité collects new types of bio-measures for aging processes as well as information on drug therapies and on negative effects of medication. The idea is to disclose determinants of predispositions for late onset disorders. This data collection of bio-measures is rounded out through an analysis of the immune parameters in the BASE-II participants,

which is carried out by the Ageing and Tumour Immunology Group (TATI) at the University of Tuebingen.

The Max Planck Institute for Molecular Genetics (MPIMG) aims at identifying the genetic factors that influence the aging process. For this purpose, a genome-wide screening is conducted by microarray genotyping and imputation of unobserved genotypes using reference panels based on whole genome sequencing.

The Max Planck Institute for Human Development (MPIB) conducts a comprehensive cognitive battery to test different cognitive domains, such as fluid intelligence, episodic memory, working memory, attention, executive control, verbal knowledge, reading skills, as well as decision making. Additionally, BASE-II participants are asked to fill in a questionnaire comprising different aspects of subjective indicators of health and wellbeing as well as the "Subjective Health questionnaire (SHH)", which measures the individuals' expectations regarding the future development of their health status. The SHH postulates that a broad subjective health horizon is a main source of one's own motivation to engage in an active and explorative lifestyle, which, in turn, is assumed to affect cognitive performance and health in aging. Research aim of this part of the BASE-II study is to analyse age-related differences in cognitive performance and the relation of subjective health expectations on cognitive skills and health-conscious behaviour.

BASE-II connects with the first Berlin Aging Study – BASE-I (Baltes and Mayer 2001; Lindenberger, Smith et al. 2009), which was initiated in 1988 by some of the current project partners.[4] At that time the target population of the Berlin Aging Study consisted of West Berlin residents aged 70 to over 100 years. Between 1990 and 1993 "516 individuals were […] broadly examined regarding their mental and physical health, psychological functioning, and social as well as economic situation" (Bertram et al., 2013: 5).

Thus, the thematic link between the two Berlin Aging Studies is very clear – just like the differences: Between the 1990s and today there has been a large quantity of new scientific findings, especially in the field of molecular genetics. For this purpose, it is almost a necessity to use these findings for further research and to rerun the Berlin Aging study for a more detailed understanding of the aging process.

[4] Further information about the BASE-I study can be found on the following website: http://www.base-berlin.mpg.de/Introduction.html

Furthermore, the follow-up study BASE-II enables to overcome some shortcomings from the first study, for example the small sample size. Another benefit of the current BASE-II study is the expansion of the age cohort. Not only an older "target cohort" but a young "control cohort" make up the sample.

Concerning the socio-economic part, for example, it can be an interesting challenge to continue researching well-being and life satisfaction of the elderly. Research based on BASE-I showed that differences in life satisfaction are not really distinct by age, just as little as the connection between subjective well-being and objective conditions. Instead, it was assumed that adaption processes to health cutbacks might be finished before the age of 70 (Smith et al, 2009). This assumption can now be studied with the BASE-II sample, since the sample is aged 60 and older and has a younger reference group. In addition, the longitudinal design of BASE-II offers the possibility to analyze the aging process and its outcomes over a period of several years.

The SOEP Group at the DIW conducts the socio-economic part of the BASE-II study. For this purpose, BASE-II participants and, in addition their family members and other persons living together with them are asked to fill in an individual questionnaire concerning their biographical data, socio-economic and socio-demographic characteristics, their self-rated psychological characteristics and their living conditions. Furthermore, one member per household is asked to fill in a household questionnaire. This procedure is adopted from the SOEP core study (see Wagner et al. 2007). The combination of the two types of information on the individual and on the household level enables researchers to disentangle the interplay between individual and societal determinants on individuals' health status. Moreover, the transformation from an individual sample to a household sample offers a sample structure that is comparable to that of the SOEP Core study (Wagner et al. 2007).

Furthermore, the SOEP Group at the DIW equips the BASE-II dataset with geo-referenced data including information on the social structure, regional structure, and infrastructure, frequency of inhabitants' relocations and inhabitants' purchasing power of the residential areas of the BASE-II participants.[5] The consideration of geo-referenced data allows analysing the effects of regional characteristics like air pollution on the individuals' health status and provides therefore a more complete picture on health-related determinants.

[5] See Goebel et al. (2007) for more information on the data file and Wurm (2013) for a first exemplary analysis based on data of SOEP-BASE.

3 Data Description of the SOEP-BASE

3.1 Target Population(s)

The BASE-II target population consists of elderly persons aged between 60 and 80. As a reference group, young persons aged 20 to 35 are also included in the study. The aim was to sample 1600 older and 600 younger residents of the greater Berlin metropolitan area for the study. For an analysis of the power of the sample see Bertram et al (2013, p. 7). Information on age and location are related to the point in time of recruitment. Consequently, BASE-II participants might be older than 80. Additionally, they might have moved into another region – an aspect that is particularly true for the younger group, which consists of individuals more willing to relocate.

As a distinctive characteristic of the BASE-II study, the target population has not been randomly sampled. Instead, participants of BASE-II were recruited by the MPIB on the basis of stored address files from three previously conducted studies at the MPIB that have had a focus on neuro-cognition. Therefore, the BASE-II population consists of three subgroups (in the following called laboratory samples L1, L2 and L3) which differ with regard to their study provenience. L1 participants came from the formerly conducted COGITO study (Schmiedek et al. 2010), while L2 participants are mainly from the Dopamine study (Li et al. 2009) but also from the SPACE study (Loevdén et al. 2012; Wenger et al. 2012). Although recruited by the MPIB, most L3 participants had not yet participated in any MPIB study. In addition, L3 is still open to recruitment of additional participants, who should replace participants lost due to sample attrition.

It is crucial to know that up to 2013, not all recruited participants had participated in the SOEP-part of BASE-II. Accordingly the sample size of L1 to L3 is smaller than the number of recruited respondents. Moreover not all participants who were asked to participate in the SOEP-BASE agreed to do so. Therefore, the number of cases who are already surveyed SOEP-alike is smaller than the target of 2,200 participants in BASE-II.

A differentiation between the three subgroups is important, as they also differ with regard to their entrance date in the BASE-II study and, as a further consequence, also with regard to their number of waves of participation in SOEP-BASE. While sample L1 was

surveyed SOEP-alike as early as 2008, sample L2 was first surveyed in 2009 and sample L3 was surveyed by means of SOEP-instruments in 2012 for the first time.

Participants of these three studies had been recruited by means of advertisements in local newspapers and the Berlin public transport system. Furthermore, individuals participating received a remarkable expense allowance from the MPIB that far outruns the typical incentives received by social survey respondents. This led to approximately 10,000 responders of whom 2,875 were invited for an additional screening (either in-house or by telephone), leading to 2,262 individuals eligible for inclusion in BASE-II, i.e. 79% of those who were initially invited.

From those, the BASE-II management team selected 2,200 individuals to represent the BASE-II baseline cohort based on their age and sex as follows (Table 1). A total of 1,600 participants belong to the older cohort aged between 60 and 80 years, whereas the remaining 600 individuals belong to a younger subgroup (serving as a reference population) aged between 20 and 35 years. By design, each age subgroup contains equal numbers of males and females. See table 1 for other socio-demographic details of the BASE-II baseline cohort.

Thus, not only have the BASE-II participants selected themselves into the study, but they also are expected to represent a very specific subpopulation of the Berlin population. Moreover, potential participants were drawn from a pool of voluntary individuals originally recruited by the Max-Planck-Institute for Human Development as part of earlier projects focusing on neuro-cognition. The potential participants were included in the studies if they met the specific inclusion/exclusion criteria. Although the L3 participants had not taken part in a MPI study at the time of inclusion here, the inclusion and exclusion criteria of the Dopamine study were also applied for to scientific reasons (e.g. sample coherence, comparability).

Exclusion criteria for the participants were mainly based on the magnetic resonance tomography (MRT) eligibility, which were conducted in all of the three studies (Dopamine, COGITO, SPACE). Some L2 participants originating from the SPACE study differed in some inclusion/exclusion criteria. As the SPACE participants represent a minor subgroup of the BASE-II participants (N<100), only the criteria from the COGITO and the Dopamine study are reported here. For further information of the SPACE criteria see Loevdén et al. (2012) and Wenger et al. (2012).

The Dopamine and COGITO study MRT-exclusion criteria were: claustrophobia, overweight (> 120 kg), tinnitus, or a walking impairment that makes sitting or lying for more

than 2 hours difficult. In addition, participants with tattoos, or other metallic implants or materials, like a dental prosthesis, a dental brace, a joint prosthesis, or metallic fragments were not included in the study.

Since the main focus of all MPI studies was on investigating different aspects of cognitive functioning, participants had to speak German as their mother tongue, could not have participated in any other study in the last two years prior to recruitment, and could not have had any walking impairments, rheumatism, or severe dorsal pain that made it impossible to take part at the cognitive tests in the study rooms. To that end, all potential participants who met characteristics that may influence cognitive functioning were excluded from the studies. Thus, participants had to be right-hand dominant, non-smokers and not take psychotropic drugs. Additionally, participants could not suffer from neurodegenerative and neurological diseases such as Parkinson's disease, seizure, giddy spells or imbalance. Also, participants who had suffered from heart attacks, apoplectic strokes, severe cardiac arrhythmia, diabetes, or operations of heart, head, or vessels were excluded from the studies as these health states interfere normal cognitive processes.

Table 1 gives an overview of the target sample. The sample members are already recruited and for all of them some basic characteristics are known, but up to now the baseline assessments have not been completed for all the target respondents. In other words: BASE-II is still in its initial phase. On the other hand, for the subgroups L1 und L2 (coming from the studies Dopamine, COGITO and SPACE) some longitudinal data is already available.

Table 1. Basic socio-demographic characteristics of the BASE-II cohort and comparison with representative samples from Berlin and Germany.

	Young (Age 20-35)			**Old (Age 60+)**		
	BASE II	Berlin	Germany	BASE II	Berlin	Germany
Number of observations	600	173	3,802	1,600	234	6,487
Age	27.32	27.83	27.64	66.76	69.62**	70.51**
Female	0.50	0.55	0.53	0.50	0.50*	0.52
German nationality	0.99	0.96*	0.93**	0.99	0.96**	0.96*
Family status[1]						
Married or living together	0.16	0.19	0.27**	0.57	0.65*	0.69**
Single	0.84	0.80	0.67**	0.09	0.06	0.03**
Divorced/Separated	0	0.01	0.03**	0.29	0.14**	0.08**
Widowed	0	0	0	0.05	0.16**	0.19**
Highest school degree[2]						
Elementary school	0.01	0.08**	0.17**	0.16	0.31**	0.54**
Intermediate school	0.12	0.21**	0.33**	0.27	0.36**	0.20**
High school	0.86	0.62**	0.42**	0.51	0.29**	0.18**
No school or other school	0.01	0.09**	0.08**	0.06	0.05	0.08**
Employment status						
Employed	0.44	0.71**	0.72**	0.14	0.20**	0.16**
Self-rated health[3]						
Very good	0.17	0.12	0.19	0.08	0.03**	0.03**
Good	0.29	0.49**	0.53**	0.32	0.26	0.25**
Fair	0.18	0.28**	0.21*	0.33	0.47**	0.43**
Poor or very poor	0.36	0.12**	0.07	0.28	0.24	0.29
Satisfaction with[4]						
Life in general†	7.01	7.06	7.12	7.49	6.83**	7.00**

Legend to Table 1: Numbers are means for continuous variables and proportions for dichotomous variables. "Age" is calculated using 2009 as reference. Data sources: BASE-II, SOEP (v28), and unpublished data. All variables for BASE-II are derived from the full baseline cohort (n = 2,200), except where labeled with 1(n=2,155), 2(n=2,172), 3(n=2,188), 4(n=2,079). †As measured on a Likert scale 19 ranging from 0 ("completely dissatisfied") to 10 ("completely satisfied"). P-values are based on two-sample t-tests of proportions (for binary outcome variables) and two-sample mean comparison t-tests (for continuous outcome variables) comparing BASE-II with SOEP data for Berlin and Germany (not overlapping with BASE-II): * P-value ≤ 0.05, ** P-value ≤ 0.01.

Source: BASE-II, SOEP, Bertram et al. (2013: 7).

3.2 Structure

The socioeconomic part of BASE-II (SOEP-BASE) started in 2008. At this time, the target population consisted only of the first sample L1 (COGITO), which was then followed for

three subsequent waves – 2009, 2010 and 2012. In 2009 the second sample (mainly the Dopamine-Study) was integrated in the sample and re-interviewed 2012. In 2012 the samples were enriched by a third sample L3 (including SPACE respondents) for targeting the final population.

Comparisons with representative survey data from Berlin and Germany, ascertained via the SOEP questionnaire (see table 1), found that BASE-II participants are characterized by higher education and better self-reported health status than the general population of Berlin and Germany (table 1). In addition, BASE-II participants in the older subgroup report a significantly higher divorce/separation rate than participants in the age-matched reference populations. For convenience samples, such as BASE-II, this is a commonly observed phenomenon.

Table 2 displays the structure of the basic SOEP-BASE sample from a more technical point of view. The table shows the gross number of eligible households per subpopulation as well as the corresponding net number of participating households. The response rates are calculated on the basis of the differences between targeted ("gross" amount) and surveyed households ("net" amount). The last row of the table provides the overall response rates per wave (see also Bohlender, Siegel 2013). The above average response rates reflect the peculiarity of the study, namely that the BASE-II participants have selected themselves into the study and are thus highly motivated to take part in SOEP-BASE as well.

Table 2: Participating households, per wave and sample

	2008	**2009**	**2010**	**2011**	**2012**
L1				*not surveyed*	
Targeted	193	192	176		137
Surveyed	170	161	136		102
Response rate	88.1%	83.9%	77.3%		74.5%
L2	*not surveyed*		*not surveyed*	*not surveyed*	
Targeted		1,476			1,391
Surveyed		1,158			1,008
Response rate		78.5%			72.47%
L3	*not surveyed*	*not surveyed*	*not surveyed*	*not surveyed*	
Targeted					448
Surveyed					346
Response rate					77.2%
Total N				*not surveyed*	
Targeted	193	1,668	176		1,976
Surveyed	170	1,319	136		1,456
Response Rate	88.1%	79.8%	77.3%		73.7%

Source: SOEP-BASE; own calculations.

The data collection mode varied by wave. In 2008, all interviews were conducted via computer assisted personal interviewing (CAPI). For the same subpopulation (L1), the main survey mode in 2009 was again CAPI, while the remaining six households chose another option for being surveyed, namely a face-to-face interview or a self-administered interview. In contrast, subpopulation L2 was mainly surveyed by computer-assisted web-interviews (CAWI) and mailed questionnaires in 2009. In 2010, participants from L1 were surveyed for the third time. The exclusive survey mode in this year was a self-administered interview via mailed questionnaire. In 2012, all participants could choose between a CAWI and a mail survey mode.[6] Table 3 provides an overview of the modes per survey year.

Table 3: Main Survey Modes by Wave and Subpopulation

	2008	2009	2010	2012
L1	CAPI	CAPI	Mail	CAWI/Mail
L2	*Not surveyed*	CAWI/Mail	*Not surveyed*	CAWI/Mail
L3	*Not surveyed*	*Not surveyed*	*Not surveyed*	CAWI/Mail

3.3 Weighting

Convenience samples like the BASE-II population are characterised by a high risk of biased data resulting from the self-selection of participants into the sample. Moreover, given that a random selection of participants is required for common statistical inferences, convenience samples allow generalizing findings only after an assessment of a potential self-selection bias is carried out.

Siedler and Sonnenberg (2010: 7-8) highlight the benefits of combining convenience samples with representative surveys arguing that not only such potential biases can be assessed, but also knowledge on the generalizability of findings based on convenience

[6] For a survey-methodological "side study" the choice of survey mode in 2012 was manipulated by offering monetary incentives that varied in the amount conditional on the choice of data collection mode. Participants were assigned either to a CAWI group or to a mail group. An incentive of 15 Euros was offered to individuals of the CAWI group who chose the CAWI mode and to individuals of the mail group who chose the mail mode. An incentive of 10 Euros was offered to those individuals who chose the mode contrary to their group assignment. The incentives were offered conditional on having filled in the individual questionnaire, while no incentive was provided for filling in the household questionnaire. This type of incentive design differs from that in the other survey years. In 2008, all participants who completed an individual questionnaire were provided with a personalised lottery ticket of 5 €. The same conditional incentive was provided in 2009 and 2010 to individuals who participated for the first time. For those, who had already participated in the previous wave, the lottery ticket was sent unconditionally before the interview was conducted.

samples can be gained. Given that the SOEP serves as such a representative survey (Siedler et al. 2009), it is possible to compare the BASE-II participants with the German population of the same age groups surveyed in the SOEP. Indeed, table 1 shows, in such a comparison, that the BASE-II participants are characterized by a higher educational background and a more positive outlook on life compared to other inhabitants of Berlin and Germany. The comparison illustrates that the target population of the BASE-II study does not resemble a representative sample (Bertram et al. 2013: 7). As a consequence, weighting of the data is of particular importance in this study.

The weights for the BASE-II study are inspired by the weighting procedures conducted in the SOEP. The selective nature of mortality and panel attrition is compensated for by longitudinal weights constructed in analogy to the longitudinal weights for the SOEP (see Kroh 2012).

In contrast to the design weights of the SOEP Core study, no design weights are provided for the BASE-II study as no sampling design had been used for the selection of participants. Instead, selectivity analyses serve as basis for the weights that shall compensate for self-selection processes in the study.

For BASE-II longitudinal weights as well as cross-sectional weights will be provided for households and individuals. In contrast to the SOEP Core study, SOEP-BASE weighting procedures focus on the individual level in order to accommodate the individual sample character of the BASE-II study. Therefore, individual weights are designed and then transferred to the household level (for a more detailed overview see Sassenroth et al. 2013).

Given that data users are interested in combining the SOEP-BASE data with other data collected for the BASE-II study, weights concerning solely the BASE-II participants are relevant. In the case that researchers intend to combine the SOEP-BASE data with the SOEP Core data, household level weights are more appropriate. Therefore different weights for the divergent research questions will be provided for users. Moreover, differentiations between the young and the old cohort will be considered in the weights.

Table 4: Weights for the SOEP-BASE and all BASE-II Data Files

	Cross-Sectional	Longitudinal
Young Age Cohort	individual level household level	individual level household level
Old Age Cohort	individual level household level	individual level household level

4 Survey Instruments SOEP-BASE

For saving the possibility to produce data compatible with the regular SOEP, in 2008, 2009 and 2010 the BASE-II PIs decided to use the regular SOEP questionnaires for the BASE-II participants. In 2012 a newly developed SOEP-BASE-Questionnaire was in the field. Table 5 shows in detail, which questionnaires were used for each subsample.

Table 5: Subsamples and Questionnaires per year

	Household Questionnaire			Individual Questionnaire		Individual Psychological Questionnaire	Individual Biography Questionnaire	Integrated Individual Questionnaire		
2008	L1			L1						
2009	L1	L2		L1	L2	L2	L1			
2010	L1			L1						
2012	L1	L2	L3					L1	L2	L3

In the years 2008 and 2009, the respondents of L1 were treated like regular SOEP respondents. This means for 2008 that each BASE-II respondent filled in the individual questionnaire from the SOEP.[7] Additionally, the regular household questionnaire was answered, too. Furthermore, every additional adult in the household of the BASE-II respondents was asked to respond to an individual questionnaire as well. In 2009, the Core SOEP questionnaires were used again (household and individual questionnaire). Additionally, the L1 participants filled in the SOEP biography questionnaire as well, while the L2 participants filled in a short supplementary questionnaire about personality.

In 2010, the household questionnaire as well as the individual questionnaire was used again. For 2012 the BASE group within the SOEP-group at DIW Berlin developed an integrated individual questionnaire for all three subsamples, including the L3-originating sample, in order to achieve a homogenous questionnaire system. With the purpose to save the longitudinal character of the BASE-II study, the integrated questionnaire covers questions from the regular SOEP questionnaires as well as those from the former waves.

[7] The SOEP Core questionnaires are free for download from the following webpage: http://www.diw.de/en/diw_02.c.238114.en/questionnaires_fieldwork_documents.html

The creation of an integrated questionnaire is oriented at the SOEP Innovation Sample (SOEP-IS) (see Richter und Schupp 2012) and combines questions from the regular SOEP, meaning the individual questionnaire, as well as questions about life history and questions from the SOEP-IS. Additionally, the SOEP-BASE questionnaire 2012 covers specific topics for the Berlin residents, for example questions for former residents in the German Democratic Republic (GDR).[8] Table 6 displays all topics of the household questionnaires that were asked in this wave. Topics and instruments in the integrated individual questionnaire are shown in Table 7.

All questionnaires used in the BASE-II study by TNS Infratest are provided by TNS Infratest Sozialforschung (2013). All the regular SOEP questionnaires are also available on the SOEP website.[9] The integrated questionnaire for 2012 and the short questionnaire about psychological concepts for L2 in 2009 are also provided at the SOEP-BASE project homepage[10].

Table 6: Topics of SOEP-BASE 2012 household questionnaire by categories

Category	**Topics in 2012**
Housing and expenses for house or flat	Housing status
	Size of dwelling
	Features of dwelling
	Owning and renting
Revenues and credits	Public welfare benefits
	Monthly earnings
Does your household have...?	People in need of care
	children

[8] Another specific feature of the 2012 SOEP-BASE questionnaire is the biographic part which includes questions about experiences during World War II. These special data will not be free for public use before 2015.
[9] http://www.diw.de/en/diw_02.c.222729.en/questionnaires.html
[10] http://www.diw.de/sixcms/detail.php?id=diw_01.c.395666.en

Table 7: Topics of SOEP-BASE 2012 individual questionnaire by categories

Category	Topics in 2012
Your current life situation	Satisfaction, Mobility, Financial situation
Your past occupation	
Your current occupation	
Income and other earnings	
Second Job	
Pension	
Health and illness	Sleeping, Nutrition, Smoking
	Aircraft noise
	Height, weight
	Physiological disorder, diseases
	Consultations, amount of hospital stays
	Health insurance
Sports	Current sport activities
	Past sport activities
Personality	Trust, Risk aversion
	Big Five, Life Orientation Test
	Loneliness
	Social desirability, Need to evaluate
	Religion
Attitude and opinions	Politics and parties
	Concerns
	Family situation and friends
	Partnership, Children
	Financial support
Experienced discriminations	
Biography: childhood and origin	Birth Date, Birth weight
	Breast feeding
	Country of birth, immigration
	Siblings
Biography: parents	Parents' country of birth and nationality
	Parents' school leaving certificate
Biography: education	School leaving certificate
	Professional education
Biography: GDR	Experienced limitations
	Attitudes towards past GDR
Biography: unemployment and former employers	Unemployment after unification
	Unemployment during last 10 years
	Number of jobs
Biography: Living	Residential history
	Satisfaction

5 Data Access

The BASE-II study, as a joint project, is an infrastructure for the entire research community. The SOEP-BASE data are available as scientific use data for the entire scientific community. Applications for data access require permission by the BASE-PIs.

To apply for the data access or for further information please contact the project coordinator, Dr. Katrin Schaar, at schaar@mpib-berlin.mpg.de.

References

Baltes, P. B. and K. U. Mayer (2001). The Berlin Aging Study: Aging from 70 to 100, Cambridge: University Press.

Bertram, L., Boeckenhoff, A., Demuth, I., Duezel, S., Eckhardt, R., Li, S.-C., Lindenberger, U., Pawelec, G., Siedler, T., Wagner, G. G., and E. Steinhagen-Thiessen (2013). Cohort Profile: The Berlin Aging Study II (BASE-II). International Journal of Epidemiology: 1-10.

Bohlender, A., and N. Siegel (2013). Datenerhebung in den Berliner Laborsamples 2012: Methodenbericht. Muenchen: TNS Infratest Sozialforschung.

Goebel, J., K. C. Spiess, Witte, N. R. J., and S. Gerstenberg (2007). Die Verknuepfung des Soep mit Microm-Indikatoren: Der Microm-Soep Datensatz. DIW Data Documentation 26.

Kroh, M. (2012). Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (Soep) (1984 until 2011). DIW Data Documentation 66.

Li, S.-C., Chicherio, C., Nyberg, L., von Oertzen, T., Nagel, I. E., Papenberg, G., Sander, T., Heekeren, H. R., Lindenberger, U., and L. Baeckman (2009). Ebbinghaus Revisited: Influences of the BDNF Val66Met Polymorphism on Backward Serial Recall Are Modulated by Human Aging. Journal of Cognitive Neuroscience 22(10): 2164–2173.

Lindenberger, U., J. Smith, Mayer, K. U., and P. B. Baltes (2009). Die Berliner Altersstudie. Berlin: Akademieverlag, Vol. 3.

Loevdén, M., Schaefer, S., Noack, H., Bodammer, N. C., ueKuehn, S., Heinze, H.-J., Duezel, E., Baeckman, L., and U. Lindenberger (2012). Spatial navigation training protects the hippocampus against age-related changes during early and late adulthood. Neurobiology of Aging 33: 620.e9–620.e22.

Richter, D., and J. Schupp (2012). SOEP Innovation Sample (SOEP-IS): Description, Structure and Documentation. SOEPpapers 463, Berlin.

Rowe, J. W., and R. L. Kahn (1987). Human aging: Usual and successful. Science, 237: 143–149.

Sassenroth, D., Kroh, M., and G. G. Wagner (2013). Selectivity Processes in and Weights for the Berlin Aging Study II (BASE-II). SOEPpapers (forthcoming).

Schmiedek, F., Loevdén, M., and U. Lindenberger (2010). Hundred days of cognitive training enhance broad cognitive abilities in adulthood: findings from the COGITO study. Frontiers in Aging Neuroscience 2(27): 1-10.

Siedler, T., J. Schupp, Spiess, K. C. and G. G. Wagner (2009). The German Socio-Economic Panel (SOEP) as Reference Data Set. Schmollers Jahrbuch 129(2): 367-374.

Siedler, T. and B. Sonnenberg (2010). Experiments, Surveys and the Use of Representative Samples as Reference Data. In: German Data Forum (RatSWD) (ed.). Building on Progress: Expanding the Research Infrastructure for the Social, Economic and Behavioral Sciences. Opladen & Farmington Hills, MI: Budrich UniPress, Vol. 1, 547-561.

Smith, J., Fleeson, W., Geiselmann, B., Settersten, R. A. and U. Kunzmann (2009). Wohlbefinden im hohen Alter: Vorhersagen aufgrund objektiver Lebensbedingungen und subjektiver Bewertungen. In: Lindenberger, U., J. Smith, Mayer, K. U., and P. B. Baltes (Ed.). Die Berliner Altersstudie. Berlin: Akademieverlag, Vol. 3, 521-549.

TNS Infratest Sozialforschung (2013). SOEP-RS BASE II 2008-2012 – Erhebungsinstrumente Berliner Altersstudie II. SOEP Survey Papers 137: Series A. Berlin: DIW/SOEP

Wagner, G. G., Frick, J. R., and J. Schupp (2007), The German Socio-Economic Panel Study (SOEP) –Scope, Evolution and Enhancements.Schmollers Jahrbuch, 127(1): 139-169.

Wenger, E., Schaefer, S., Noack, H., Kuehn, S., Mårtensson, J., Heinze, H.-J., Duezel, E., Baeckman, L., Lindenberger, U., and M. Loevdén (2012). Cortical thickness changes following spatial navigation training in adulthood and aging. NeuroImage 59: 3389–3397.

Wurm, M. (2013). „Verknuepfung von Fernerkundungsdaten und Survey-Daten (SOEP und BASE II) in staedtischen Raeumen fuer sozialwissenschaftliche Analysen“. Dissertation. Graz.