Research Data Centre (FDZ)
of the German Federal
Employment Agency (BA)
at the Institute for
Employment Research (IAB)

FDZ-Datenreport

Documentation of labour market data

02/2014
EN

Codebook and Documentation of the Panel Study 'Labour Market and Social Security' (PASS)

Datenreport Wave 7

Marco Berg,
Ralph Cramer,
Christian Dickmann,
Reiner Gilberg,
Birgit Jesske,
Martin Kleudgen,
Arne Bethmann,
Benjamin Fuchs,
Martina Huber,
Mark Trappmann

Bundesagentur für Arbeit

Codebook and Documentation of the Panel Study
‘Labour Market and Social Security’ (PASS)

Datenreport Wave 7

Marco Berg, Ralph Cramer, Christian Dickmann, Reiner Gilberg, Birgit Jesske, Martin Kleudgen, infas Institut für angewandte Sozialwissenschaft GmbH - Arne Bethmann, Benjamin Fuchs, Mark Trappmann, Martina Huber, Institut für Arbeitsmarkt- und Berufsforschung (Institute for Employment Research – IAB)

FDZ-Datenreporte (FDZ data reports) describe FDZ data in detail. As a result, this series of reports has a dual function: on the one hand, users of the reports can ascertain whether the data offered is suitable for their research task, on the other hand, the data can be used to prepare evaluations. This data report documents the data preparation of the seventh PASS wave and is based upon the sixth wave’s data report: Marco Berg, Ralph Cramer, Christian Dickmann, Reiner Gilberg, Birgit Jesske, Martin Kleudgen, (all infas Institut für angewandte Sozialwissenschaft GmbH), Arne Bethmann, Benjamin Fuchs, Mark Trappmann, Anja Wurdack, Martina Huber (all Institut für Arbeitsmarkt- und Berufsforschung (IAB)): Codebuch und Dokumentation des ‚Panel Arbeitsmarkt und soziale Sicherung‘ (PASS) volume I: Datenreport Welle 6, FDZ Datenreport, 06/2013 (de), Nuremberg.

Contents

List of Tables

List of Figures

Data availability

The dataset described in this document is available for use by professional researchers. For further information, please refer to http://fdz.iab.de/.

1 Introduction

1.1 The objectives and research questions of the panel study 'Labour Market and Social Security'

The panel study 'Labour Market and Social Security' (PASS), established by the Institute for Employment Research (IAB), creates a new empirical dataset for labour market, welfare state and poverty research and policy counseling in Germany.

This study is conducted as part of IAB research on German Social Code Book II (SGB II)[1]. The IAB must fulfill a statutory mandate to study the effects of the benefits and services provided under SGB II, which are aimed at labour-market integration and subsistence benefits. However, due to its complex sampling design, this study also enables researchers to examine additional issues. The following five core questions, which are detailed in Achatz, Hirseland and Promberger (2007), influenced the development of this study:

1. What are the options for regaining financial independence from Unemployment Benefit (UB) II (Arbeitslosengeld II)?
2. How does a household's social situation change when it receives benefits?
3. How do individuals who receive benefits cope with their situations? Do recipient attitudes toward the actions required to improve their situations change over time?
4. How does contact between benefit recipients and institutions that provide basic social security take place? What actual institutional procedures are applied in practice?
5. What employment history patterns or household dynamics lead to receiving Unemployment Benefit II?

This data report provides an overview of the seventh survey wave, for which 14,449 individuals in 9,509 households[2] were interviewed between February 2013 and September 2013. This sample included 12,644 individuals and 8,379 households that had previously been interviewed for PASS.

This data report[3] of wave 7 documents the wave-specific aspects of the study. Chapter 1.2 contains a short description of the instruments and the survey program. Following a

---

1 Social Code Book II - basic security for job-seekers (Sozialgesetzbuch (SGB) Zweites Buch (II) - Grundsicherung für Arbeitsuchende).

2 These figures include evaluable interviews only. Additionally, repeatedly interviewed households were considered even if only a household interview but no personal or senior citizen interview could be conducted.

3 These reports were divided into the following two components for the first time in the wave 3 documentation: a wave-specific data report (including a codebook) and a cross-wave User Guide. The PASS project team at the IAB is responsible for creating the cross-wave User Guide. As of wave 3, infas has created the documentation for the wave-specific data report, which is based on the wave 2 data report. The cross-wave User Guide documents the entire

short overview of the innovations and characteristics of wave 7 (Chapter 1.3), the data report provides key figures on the wave's sample and response rates (Chapter 2). The data preparation process is described (Chapter 5), and an overview of the variables generated is presented (Chapter 4). Additionally, the weighting procedure is presented (Chapter 6). Separate tables list the frequencies of all of the variables included in the scientific use file that were recorded in wave 7 by their respective datasets (Volumes II through V).

1.2 Instruments and interview program

The information in PASS is collected using separate questionnaires for the household and individual levels. First, a household interview is conducted. This interview gathers information about the entire household. The target person for this household interview[4] was selected during the contact phase preceding the interviews. Personal interviews of the household members follow the household interview. The aim is to conduct a personal interview of each individual living in the household who is 15 years of age or older. Household members who are 65 or older receive a shortened version of the questionnaire (the senior citizens' questionnaire), which excludes questions that are irrelevant to that age group.

The survey instruments and interview program for wave 7 are based on those used in wave 6. However, individual questions and modules have been revised or newly developed (see Chapter 1.3 for an overview).

The PASS survey instruments are designed to allow not only repeat interviews of individuals and households but also first-time interviews[5].

Since wave 3, dependent interviewing has been used for certain questions to update information that the respondent had previously provided to avoid seam effects[6] in the repeat interviews and to increase data quality. Information about constant characteristics was generally not gathered again. Additionally, since wave 4, an integrated questionnaire for

---

study, details the objectives and design of PASS and presents the contents and instruments of the survey. Moreover, it describes the structure of the scientific use file and the concept of the variable types and their names.

4 The target person for the household interview should know as much as possible about general household issues, and target selection was based on the rules documented in the methods reports (Jesske & Quandt, 2011; Jesske & Schulz 2012; Jesske & Schulz 2013).

5 First-time interviewed households include the following groups: (1) households from the refreshment and replenishment samples of the current wave; and (2) households that split off from households interviewed during previous waves (split-off households). (For further explanation, please see the wave 4 methods report (Jesske & Quandt, 2011).)

6 In a panel data, the number of changes observed at the interface (seam) between interviews conducted in sequential panel waves is often considerably higher than the number of changes observed within an interview (see Jäckle 2008).

repeatedly interviewed households (HHalt) and first-time interviewed households (HHneu) has been used[7].

The cross-wave PASS User Guide elaborates the individual instruments and interview program. The following section reviews the characteristics and innovations of wave 7.

1.3 Characteristics and innovations of wave 7

At this point we outline the characteristics of the seventh wave for users who are already familiar with the data from previous PASS waves.

The characteristics and innovations of wave 7 affect the questions asked in the household and personal questionnaires (e.g., change of reference periods, modification of individual questions and new question modules)[8], sample and data preparation.

1.3.1 Personal questionnaire

The personal questionnaire updates the employment history information gathered since wave 2[9]. Wave 7 maintains the chronological retrospective surveying introduced in wave 4 (see section 1.3.1 in Berg et al., FDZ Datenreport 08/2011).

In wave 7, the following six new modules were added to the personal questionnaire:

- Estimation of annual income in 2012 (PBIOBLK06-PBIOBLK07);
- Attitude (children's leisure) (PEO1500*);
- Stigma awareness (PSV0100-PSV0200) ;
- Duty of search (minimum program of the provider module) (PSUBLK01-PSU0200*) ; and
- Quality of employment (PBIOBLK08-ET4900)

The "**quality of employment**" module contains questions regarding an evaluation of the present employment, job requirements, occupational chances and burdens of a job, as well as occupational perspectives. In this module's questionnaire, a separate variable for supply, "actual main employment" (Haupt_ET), and a variable to show the associated job title (Beruf_Haupt_ET) were necessary. These variables were generated by using either the employment episodes of the actual wave or the preload of the previous wave.

---

7 In this survey, split-off households are treated like new households.

8 Not all of the minor changes to the questionnaire (adding, modifying or deleting individual questions) are listed.

9 This information is gathered using the so-called dependent interviewing method. In dependent interviewing, information that was provided during previous interview waves is included in the interview text of the current interview to determine whether the information must be updated.

− Memory I (PME0100) and Memory II (PME0200)

The “**Memory I**” and “**Memory II**” modules involved a test of memory power and were integrated into the personal interviews. The respondents had to repeat a given list of words (10 terms) during the interview after a definite period. The respondents had one minute to answer the memory tests.

In the personal questionnaire in wave 7, the following additional adjustments were made.

Both small and large revisions were made to the “**health**” module. In particular, new filters were integrated so that the respondents could refuse to answer health questions. Furthermore, a new question regarding presentism was inserted for all surveyed persons (PG1400). The question concerning the frequency of sporting activities was matched to the intensity of the previous wave to ask for specific reasons for reduced or increased sporting activities (PG1236 – PG1239). The questions involving health risks (PG1205, PG1210 and PG1215), a medical certificate of more than 6 weeks during the last 12 months and consequent restrictions on employment (PG1220, PG1225 and PG1230), were removed. All questions on alcohol and smoking habits (PG1240 – PG1280), body size and weight (PG1285 and PG1290) were also deleted.

The “**sport**” module was differentiated between repeat respondents and new respondents in wave 7. Different filtering processes followed regarding the type of sport in childhood or adolescence. Furthermore, an additional question concerning the commute was integrated (PSB1310).

Because of the current legal situation, a differentiated collimation of the name of the agency no longer appeared in the “**measure**“ module, but it kept the name “job-centre”. Throughout, versions of the text were combined so that no distinction of cases with known or unknown names of agencies was necessary.

In the introduction to the job search task, the respondents were told that the name of the agency would be discussed as a job-centre.

In addition to the aforementioned extension and adaption, the following deletions were made from the personal questionnaire:

- Both parts I and II of the “**justice**“ module were completely deleted (justice I (PGR0100*) and justice II (PPGBLK01-PPG0300*)); and
- The “**agency contacts**” module (PTKBL01-PTK1600*) was intermitted for wave 7. Only the aspects of the job search task were surveyed including an adjusted introduction. For wave 8, an integration of a new revised module regarding the contact of agencies is planned.

1.3.2 Household questionnaire

In the household questionnaire in wave 7, only the „**education and participation package**” module was revised. All other modifications address only general wave-specific changes.

The “education and participation package” module was continued using dependent interviewing in wave 7. Therefore, different types of filter control and variants of the text for the new and returning respondents were created. Open questions in that module were removed (HBT1000 and HBTBLK05) and replaced by an item battery (HBT1200*).

Further deletions were not made in the household questionnaire.

1.3.3 Sample and data preparation

In wave 7, as in previous waves, a refreshment sample was drawn from the Federal Employment Agency (BA) subsample.[10] The aims are to guarantee the representativeness of the BA sample in the cross-section and to observe enough new transitions into benefits, that is, into UB II, over time. For the refreshment sample, benefit units were drawn receiving UB II in July 2012 but not on the sampling date of the first, second, third, fourth, fifth and sixth waves (see Chapter 2.1 and, on the concept of the refreshment sample, Trappmann et al., 2009, page 11 ff.). All of the households that were surveyed for the first time during wave 7 can be identified via the sample indicator (*sample*).

The data preparation was performed in close cooperation with the IAB. Basic procedures, such as updating datasets and correcting problems in the household structures, were discussed during the preparation process. Final decisions were made by the IAB.

The integration of the spell datasets into the module “**employment**” and the necessary preparatory steps were discussed and determined in agreement with the IAB. That procedure is documented in Chapter 5.7.

2 Key figures

This chapter provides a brief overview of important figures in the study, such as sample sizes (gross and net) and response rates. The panel sample is represented over the course of the previous waves. Figures are reported not only for both the original and replenishment samples but also for the complete study.

- Subsample 1 (BA sample) refers to the sample of benefits recipients from the process data of the Federal Employment Agency.
- Subsample 2 (MICROM sample) refers to the stratified population sample.
- Refreshment sample 1 (BA sample) is the sample drawn from the SGB II inflow between waves 1 and 2.

---

[10] Wave 1 of PASS includes two subsamples: (1) a sample of households receiving UB II, which was drawn from the Federal Employment Agency (BA) process data; and (2) a general population sample, stratified by status, drawn from a database provided by the commercial provider MICROM.

- Refreshment sample 2 (BA sample) is the sample drawn from the SGB II inflow between waves 2 and 3.
- Refreshment sample 3 (BA sample) is the sample drawn from the SGB II inflow between waves 3 and 4.
- Refreshment sample 4 (BA sample) is the sample drawn from the SGB II inflow between waves 4 and 5.
- Panel replenishment/supplement 1 (municipal register sample) is the sample drawn from the registration office inflows in ten new postcode regions during wave 5.
- Panel replenishment/supplement 2 (BA sample) is the sample drawn from the SGB II inflows in 100 new postcode regions during wave 5.
- Refreshment sample 5 (BA sample) is the sample drawn from the SGB II inflow between waves 5 and 6.
- Refreshment sample 6 (BA sample) is the sample drawn from the SGB II inflow between waves 6 and 7.

2.1 Sample size

Each sample in a panel begins with the interviewed households from the first survey wave. In PASS, the gross panel sample contains the interviewed households from wave 1 and the HHneu from the refreshment samples in waves 2 to 6. Only those households being interviewed for the first time that are willing to participate in the panel and are available for repeat interviews are considered.[11] Agreement to participate in the panel is only recorded during the first interview. Confirmation of these households' willingness in subsequent waves is not required. In addition to confirming willingness, access to the panel is induced during the first interview by general willingness to participate, that is, by providing an interview. Measures to ensure the best possible selection-free access to the panel as part of PASS are described in detail in the methods and field reports of waves 1 to 7.[12]

Wave 1 of PASS included 12,794 household interviews, of which 12,000 households agreed to participate in the panel. These wave-1 households constitute the sample for the beginning of the first tracking survey.

The panel concept in PASS assumes that new or split-off households emerge as individuals move out of panel households, which are considered separate households as soon as a household interview is conducted.

---

[11] Willingness to participate in the panel is confirmed by the household reference person and is thus valid for all household members. Households that were willing to participate in the panel have allowed their addresses to be stored for the purposes of this study's repeat interviews.

[12] See Hartmann et al. (2008); Büngeler et al. (2009); Büngeler et al. (2010); Jesske & Quandt (2011); Jesske & Schulz (2012).

This design results in a higher number of households compared to the original sample. Details about the procedures for the PASS panel concept can be found under “split-off households.” In addition to the expansion of the panel, loss of households can occur due to panel mortality. Households in which all respondents passed away or moved abroad are removed from the gross panel in subsequent waves. Moreover, panel losses may occur if no household interview could be conducted for a household for two consecutive waves. This situation arose for the first time at the end of wave 3 and affected the gross panel in waves 4,[13] 5, 6 and 7. The gross sample used for wave 7 included 10,763 panel households. That includes additionally HHneu from the refreshment sample (n=3,020) and newly formed split-off households in wave 6 (n=292) and wave 7 (n=458).

The case numbers for the gross sample size of the respective survey waves and subsamples are reported in the following table. In wave 7, at least one interview could be conducted for 8,379 households in the panel sample. In addition, 949 first-time household interviews were conducted from the refreshment sample, of which 914 were willing to participate in the panel. In addition, the households interviewed for the first time in wave 7 include 181 split-off households that arose because of the subsamples in waves 1 – 6.

[13] The survey institute change also influenced the panel gross in wave 4 because transmitting participant addresses from the IAB to infas required the target person's permission. For details on this procedure and its results, please refer to the methods report for wave 4 (Jesske & Quandt, 2011).

Table 1: Panel sample at the household level by wave and subsample[14]

		Sample										
	n	BA	Microm	BA-refreshment 1	BA-refreshment 2	BA-refreshment 3	BA-refreshment 4	EWO supplement	BA supplement	BA-refreshment 5	BA-refreshment 6	Total
Wave 1	HH-interview realised	6.804	5.990									12.794
	of this: HH willing to participate in panel	6.452	5.548									12.000
Wave 2	Panel-HH gross	6.520	5.611									12.131
	HH-interview realised	3.491	3.897	1.041								8.429
	of this: HH willing to participate in panel	3.360	3.766	1.003								8.129
Wave 3	Panel-HH gross	5.851	5.150	1.010								12.011
	HH-interview realised	3.754	3.901	694	1.186							9.535
	of this: HH willing to participate in panel	3.576	3.777	669	1.145							9.167
Wave 4*	Panel-HH gross	3.926	3.628	863	1.069							9.486
	HH-interview realised	2.815	2.977	563	745	748						7.848
	of this: HH willing to participate in panelt	2.754	2.933	554	727	723						7.691
Wave 5**	Panel-HH gross	3.392	3.334	679	960	727						9.092
	HH-interview realised	2.382	2.680	464	608	517	753	1.510	1.321			10.235
	of this: HH willing to participate in panel	2.347	2.633	456	598	512	702	1.415	1.257			9.920
Wave 6**	Panel-HH gross	2.902	3.021	576	768	687	653	1.324	1.185	961		12.077
	HH-interview realised	2.109	2.539	398	532	466	497	1.103	908	961		9.513
	of this: HH willing to participate in panel	2.078	2.503	389	519	460	492	1.087	890	919		9.337
Wave 7**	Panel-HH gross	2.523	2.789	479	656	552	622	1.264	1.134	925		10.944
	HH-interview realised	1.984	2.409	359	505	414	413	996	798	682	949	9.509
	of this: HH willing to participate in panel	1.954	2.383	357	502	412	407	969	783	671	914	9.352

Source: HH-Register and PENDDAT; Scientific Use File IAB

* Reduction of the gross sample due to objection procedures

** Expansion of the gross sample by supplementation

[14] The scientific use file's register files always comprise the net sample of realised interviews of the respective waves. In the case of split-off households it is possible that there is a subsequent expansion of the panel household gross of the previous wave if the split-off household was identified in the previous wave but could not be realised yet.

The 9,509 household interviews conducted in wave 7 correspond to 14,449 personal interviews. The following table lists the distribution of respondents across subsamples and survey waves.

Table 2: Panel sample size at the individual level by wave and subsample

	Personal interview realised	**Wave 1**	**Wave 2**	**Wave 3**	**Wave 4***	**Wave 5****	**Wave 6**	**Wave 7**
		abs.	abs.	abs.	abs.	abs.	abs.	abs.
Sample	**BA**	9,386	4,753	4,913	3,958	3,394	3,048	2,862
	Microm	9,568	6,392	6,207	5,016	4,511	4,245	4,001
	BA-Refreshment 1		1,342	898	786	653	558	505
	BA-Refreshment 2			1,421	983	822	719	688
	BA-Refreshment 3				1,025	760	679	590
	BA-Refreshment 4					1,019	716	599
	EWO supplement					2,589	1,990	1,784
	BA supplement					1,859	1,350	1,182
	BA-Refreshment 5						1,314	975
	BA-Refreshment 6						1,314	1,264
	Total	18,954	12,487	13,439	11,768	15,607	14,619	14,449

Source: P_Register; Scientific Use File IAB

* Reduction of the gross sample due to objection procedures

** Expansion of the gross sample by supplementation

For people without sufficient knowledge of German, interviews were offered in Turkish and Russian. Table 3 indicates how many households or persons were interviewed in these additional survey languages.

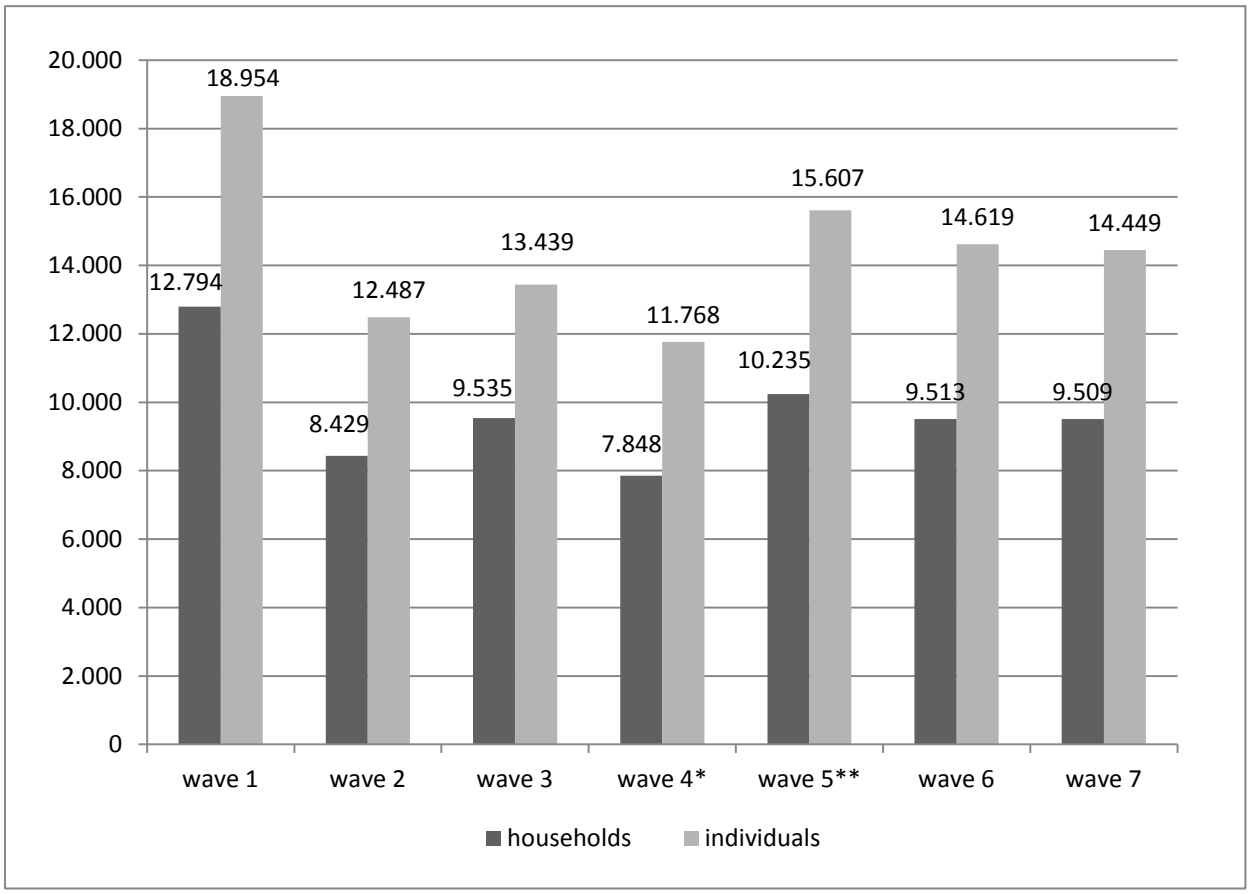
Table 3: Panel sample size of foreign-language interviews by wave

		Russian abs.	Turkish abs.
Wave 1	Households	275	163
	Individuals	432	305
Wave 2	Households	156	39
	Individuals	219	31
Wave 3	Households	210	69
	Individuals	330	109
Wave 4	Households	179	42
	Individuals	285	78
Wave 5	Households	159	36
	Individuals	259	58
Wave 6	Households	146	25
	Individuals	242	40
Wave 7	Households	145	29
	Individuals	245	43

Source: PENDDAT; Scientific Use File IAB

For the overall data pool of the realised panel sample, the following figure outlines households and individuals over the seven survey waves.

Figure 1: Realised panel sample for households and individuals by survey wave



* Reduction of the gross sample due to objection procedures

** Expansion of the gross sample by supplementation

2.2 Response rates

The response rate is calculated according to AAPOR standards (AAPOR, 2006). The response rate (RR1) is reported, which includes all cases of unknown eligibility in the denominator and therefore provides the minimum value of all response rates.[15] The response rate at the household level is calculated from the share of usable household in-

[15] This issue is addressed in very different ways in Germany. Frequently, a large number of individuals or households that were not interviewed are considered ineligible and are removed from the denominator when the response rate is calculated. When a sample is drawn from registers, neither a household that is not living at the expected address nor a household that claims not to belong to the target group may be considered to have provided a neutral non-response. Moreover, the population of PASS is not restricted to German-speaking respondents or individuals who can be interviewed; therefore, the non-response reasons "does not speak German" or "respondent is sick/unable to be interviewed" cannot be considered cases of neutral non-response.

terviews as a proportion of the total usable household interviews and non-neutral non-responses. Only households in which all members have passed away or moved abroad permanently are considered cases of neutral non-response. Households are considered usable if at least one complete household interview is available. New households are considered usable if both the household interview and at least one complete personal interview are available. The following response rates were obtained at the household level for wave 7:

Table 4: Response rate for wave 7 at the household level by subsample

Sample

Wave 7		BA	Microm	BA-refreshment 1	BA-refreshment 2	BA-refreshment 3	BA-refreshment 4	EW supplement	BA supplement	BA-refreshment 5	BA refreshment 6	Total
HH gross	abs.	2.523	2.789	479	656	552	622	1.264	1.134	925	3.020	13.964
	%	*100,0*	*100,0*	*100,0*	*100,0*	*100,0*	*100,0*	*100,0*	*100,0*	*100,0*	*100,0*	*100,0*
neutral nonresponses	abs.	15	15	2	4	10	2	8	7	6	30	99
	%	*0,6*	*0,5*	*0,4*	*0,6*	*1,8*	*0,3*	*0,6*	0,6	0,6	*101*	*0,7*
HH gross corrected*		2.508	2.774	477	652	542	620	1.256	1.127	919	2.990	13.865
	%	*100,0*	*100,0*	*100,0*	*100,0*	*100,0*	*100,0*	*100,0*	*100,0*	*100,0*	*100,0*	*100,0*
HH-interview realised	abs.	1.984	1.409	359	505	414	413	996	798	682	949	9.509
	%	*79,1*	*86,8*	*75,3*	*77,5*	*76,4*	*66,6*	*79,3*	*70,8*	*74,2*	*32,1*	*68,7*
of this: HH willing to participate in panel	abs.										914	
	%										*30,6*	

* HH gross - neutral non-responses

Source: HH-Register; Scientific Use File IAB - for BA refreshment 5 and supplementary samples: methodological research dataset by infas

In a household survey, one can distinguish between the response rates at the household level and within the household.

The response rate within households indicates the average proportion of household members aged 15 or older within evaluable households for whom a complete personal interview is available.

On average, the following response rates were obtained within interviewed households:

Table 5: Average response rate among interviewed households by wave and subsample

		Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7
		%	%	%	%	%	%	%
Sample	**BA**	85.6	85.5	83.1	88.4	88.7	89.3	89,2
	Microm	84.2	85.1	83.6	88.0	88.3	88.6	88,4
	BA-Refreshment 1		86.2	84.3	90.2	89.5	88.5	90,1
	BA-Refreshment 2			84.2	88.3	89.3	88.5	88,8
	BA-Refreshment 3				89.6	91.2	91.4	89,8
	BA-Refreshment 4					88.9	92.0	90,6
	EWO supplement					84.4	89.1	89,1
	BA supplement					90.0	91.5	92,0
	BA-Refreshment 5						89.9	90,7
	BA-Refreshment 6						89,9	90,1
	Total	84.9	85.4	83.5	88.5	88.3	89.5	89,5

Source: P_Register; Scientific Use File IAB

In addition to the between- and within-household response rates, the following table provides the repeat interview rate at the individual level. This value is the proportion of individuals willing to participate in the panel with whom an interview could be conducted in the subsequent wave.

Table 6: Proportion of personal interviews in waves 2 through 7 with respondents who were willing to participate in the panel by subsample

			Sample									
			BA	**Microm**	**BA-Refr. 1**	**BA-Refr. 2**	**BA-Refr. 3**	**BA-Refr. 4**	**EWO suppl.**	**BA suppl.**	**BA-Refr. 5**	**Total**
Wave 2	individuals willing to participate in the panel W 1	abs.	8,925	8,938								17,863
	re-interviewed individuals in W 2	abs.	4,274	5,829								10,103
	Share	*%*	*47.9*	*65.2*								*56.6*
Wave 3	individuals willing to participate in the panel W 2	abs.	4,686	6,292	1,298							12,276
	re-interviewed individuals in W 3	abs.	3,365	4,956	820							9,141
	Share	**%**	*71.8*	*78.8*	*63.2*							*74.5*
Wave 4*	individuals willing to participate in the panel W 3	abs.	4,844	6,100	894	1,380						13,218
	re-interviewed individuals in W 4	abs.	3,287	4,347	626	854						9,114
	Share	**%**	*67.9*	*71.3*	*70.0*	*61.9*						*69.0*
Wave 5	individuals willing to participate in the panel W 4	abs.	3,946	5,004	785	979	993					11,707
	re-interviewed individuals in W 5	abs.	2,972	4,151	570	714	702					9,109
	Share	**%**	*75.3*	*83.0*	*72.6*	*72.9*	*70.7*					*77.8*
Wave 6	individuals willing to participate in the panel W 5	abs.	3,378	4,468	645	819	756	957	2,439	1,786		15,248
	re-interviewed individuals in W 6	abs.	2,653	3,864	486	606	563	660	1,861	1,255		11,948
	Share	**%**	*78.2*	*85.7*	*74.4*	*73.7*	*74.1*	*64.8*	*71.9*	*67.5*		*76.6*
Wave 7	individuals willing to participate in the panel W 6	**abs.**	*3.034*	*4.216*	*555*	*711*	*667*	*712*	*1.973*	*1.337*	*1.264*	*14.469*
	re-interviewed individuals in W 7	**abs.**	*2.486*	*3.706*	*434*	*590*	*523*	*523*	*1.633*	*1.040*	*900*	*11.835*
	Share	**%**	*81,9*	*87,9*	*78,2*	*83,0*	*78,4*	*73,5*	*82,8*	*77,8*	*71,2*	*75,6*

Source: PENDDAT; Scientific Use File IAB

* Reduction of the gross sample due to objection procedures between wave 3 and wave 4

2.3 Panel participation agreements, merging data and linking with process data

Respondent consent is always required to store addresses for repeat interviews in a subsequent wave and to merge survey data with the process data obtained from the Federal Employment Agency.

Panel participation agreement was explained in detail in Chapter 2.1. HHneu[16] consent to participate in the panel is illustrated as follows:

Table 7: First-time interviewed households* consent to participate in the panel by wave**

	Realised HH interviews with first-time interviewed HH	Realised HH interviews with first-time interviewed HH willing to participate in panel	Share willing to participate in panel
	abs.	abs.	%
Wave 1	**12,794**	**12,000**	**93.8**
Wave 2	**1,086**	**1,048**	**96.5**
Wave 3	**1,327**	**1,285**	**96.8**
Wave 4*	**903**	**866**	**95.9**
Wave 5**	**3,688**	**3,476**	**94.3**
Wave 6	**1,112**	**1,068**	**96.0**
Wave 7	**1.130**	**1.089**	**96.4**

* Reduction of the gross sample due to objection procedures

** Expansion of the gross sample by supplementation

*** First-time interviewed HH from refreshment, supplement and split

Source: PENDDAT und HH_Register; Scientific Use File IAB

The consent to participate in the panel is recorded following the first personal interview in a new household during each wave. The information provided by that individual is assumed to apply to the household. That is, if the individual consents to participate in the panel, the household is considered willing to participate in the panel and if the individual

[16] All households in wave 1 are HHneu. Subsequently, only households from the refreshment samples and split-off households participating for the first time are considered HHneu. Therefore, since wave 2, households interviewed for the first time have been in the minority - the majority of household interviews conducted in these waves were conducted previously.

does not agree to participate in the panel, the household is considered unwilling to participate in the panel (see also Chapter 2.1).[17]

In contrast, permission to merge process data from the Federal Employment Agency with the survey data was obtained for each respondent who was interviewed using the personal questionnaire. This question does not apply to individuals aged 65 and over because it is not included in the senior citizens questionnaire. Consent to merging of these data is not obtained again in each wave.[18]

Table 8 provides an overview of obtained consent to merge data in each wave. Only interviews in which consent to merge data was requested in that wave as part of the personal questionnaire are listed.

[17] One individual confirms household willingness to participate in the panel. The information available on the household level was integrated into the individual dataset (*PENDDAT*) during data preparation. The individual respondents in the household were assigned the corresponding information available for that household. The same procedure was applied during wave 2. In wave 1; however, consent was recorded after each individual and senior citizen interview; therefore, data could vary within a household. Households with at least one individual willing to participate in the panel were considered willing to participate in the panel.

As part of updating address information after the first personal interview in re-interviewed households, it was explained that an interview would be conducted again the following year. If the respondent did not explicitly object to this notification, the household was considered to agree to participate in the panel and the panel variable in the individual dataset (*PENDDAT*) was updated accordingly.

[18] Due to filtering modifications, there were cases in which permission to merge data was raised again in waves 2 and 3 if the respondent had not previously agreed to that during the previous waves.

Table 8: Consent to merge data in personal interviews (respondents aged 15-65 years) obtained by wave

	Realised personal interviews from the wave in which the merging question was posed	Realised personal interviews from the wave in which consent to merging was granted	Share with granted consent to merging
	abs.	abs.	%
Wave 1	**17,249**	**13,766**	**79.8**
Wave 2	**3,358**	**2,560**	**76.2**
Wave 3	**2,656**	**2,128**	**80.1**
Wave 4*	**2,032**	**1,774**	**87.3**
Wave 5**	**5,145**	**4,414**	**85.8**
Wave 6	**2,482**	**2,002**	**80.7**
Wave 7	**1,973**	**1,613**	**81.8**

* Reduction of the gross sample due to objection procedures

** Expansion of the gross sample by supplementation

Basis: individuals 15 to 64 years of age

Source: PENDDAT; Scientific Use File IAB

2.4 Split-off households

PASS is designed as a dynamic panel. Individuals who join or are born into the household are interviewed if they are at least 15 years old. Individuals who move out of sample households for one year or more should continue to be interviewed; however, these individuals are considered new, split-off households. These split-off households also become sample households in PASS. All individuals 15 years of age or more living in these households become target persons for personal interviews. If part of this split-off household in turn splits off in subsequent waves, then this new split-off household also becomes a PASS sample household regardless of whether that new household contains anyone from the original sample (see infinite degree contagion model, Rendtel & Harms 2009, 267). However, individuals who have moved abroad are removed from the survey because they no longer belong to this population and research questions specific to SGB II no longer apply. Individuals who leave the household for less than one year continue to be considered household members.

There are 779 split-off households from waves 1 to 7, of which 484 could be interviewed during wave 7, including 126 newly split-off households from wave 7 and 55 HHneu that could be identified in wave 6. Please refer to the methods report for wave 7 for further information about split-off households (Jesske & Schulz, forthcoming).

The interviewed split-off households can be identified in the datasets by comparing the current household number (*hnr*) with the original household number (*uhnr*), which differs in these cases. The original household number (*uhnr*) contains the household number of the panel household from which the new household has separated. Split-off households assume the sample indicator (*sample*), sampling year (*jahrsamp*), primary sampling unit (*psu*) and stratification (*strpsu*) of their original household.

3 Dataset structure

The usual structure for editing a panel dataset - for example, the German Socio-Economic Panel (GSOEP) or the British Household Panel Survey (BHPS) - involves storing individual and household information in annual individual datasets. If required, these individual datasets can be supplemented with specific datasets, which might have a cross-wave data structure, such as register or spell data.

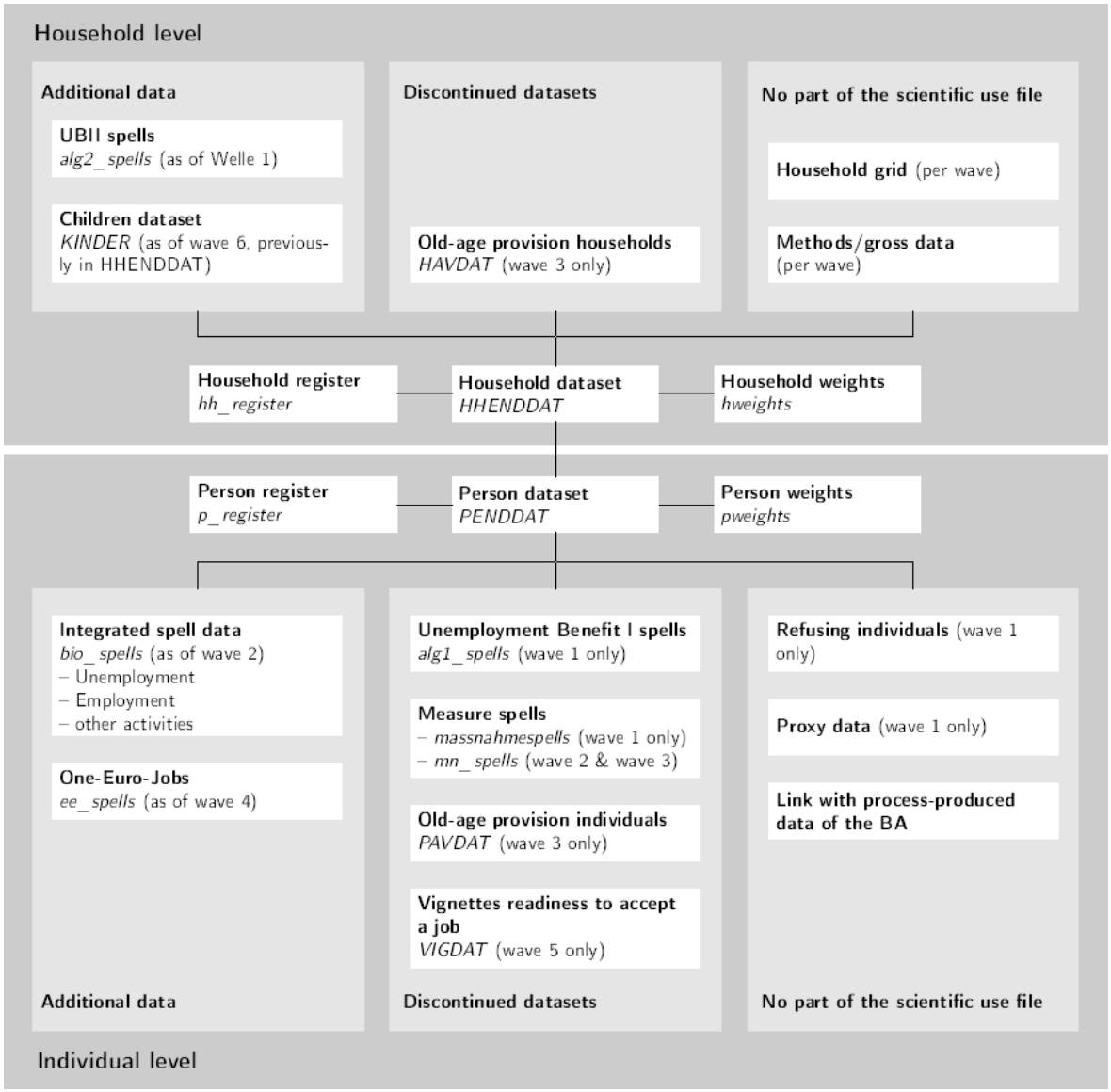
This data structure allows the information to be stored using relatively little storage space. The variables for each year can be identified immediately when examining the datasets. Identifying the merged additional information via key variables, such as household or personal identification numbers, is also quite simple. However, this common panel data structure increases the difficulty of working with these datasets. If analyses are conducted not only cross-sectionally but also longitudinally, then first, all of the relevant variables from each wave dataset must be integrated into a common dataset and care must be taken to ensure that the constructs are comparable for each year. For typical longitudinal analyses, the cross-wave dataset created in this way then must be reshaped into the so-called long format. Unlike the wide format, which contains a data matrix with one row per observation unit (e.g., the household or individual) and several datasets for each survey wave, in the long format, all of the waves assigned to an observation unit are arranged below one another. Rather than arranging information in wave-specific variables in the same row, in long format, the information is assigned to the same variable in each case in wave-specific rows for the observation units.

Reshaping the data into long format has both advantages and disadvantages. The decisive advantage of this variant is that this data structure is required for many longitudinal analyses (such as event history analyses). It is no longer necessary to invest additional time and effort creating a cross-wave file. The switch from long format to wide format is also quite easy to perform. STATA, for example, provides an option to switch between formats with little effort using the “reshape” command. Until a few years ago, the central argument against using this type of data structure was the significantly larger storage space required because even variables recorded in only one or a small number of survey waves require a complete column across all of the waves in the dataset. In addition, these long files become quite large with the increasing duration of the panel because all annual waves are appended, which significantly increases the storage space required and time needed to perform individual operations. The current wide availability of fast processors and large storage capacities even on simple desktop computers render this objection irrelevant. Another disadvantage occurs when merging additional data sources. Unlike datasets prepared in wide format, an additional variable is now required to identify an

observation clearly. This variable may be a wave identifier in the household or individual datasets or the spell number in the spell datasets, which are also available in long format. Furthermore, it is not immediately apparent which variables were included in each wave because all variables are present in the dataset. These variables are assigned a special code (-9) to identify waves during which they were not surveyed.

When the advantages and disadvantages of long format are weighed, the advantages of the long format clearly outweigh the disadvantages. Accordingly, household and individual PASS datasets (*HHENDDAT; PENDDAT*), corresponding weighting data (*hweights; pweights*) and a new dataset since wave 6 on children (KINDER) were prepared in long format.

Figure 2: Dataset structure of PASS in wave 7



At the household level, the scientific use file contains the data on household receipt of Unemployment Benefit II in spell form (*alg2_spells*). Since wave 4, the individual level has

contained an integrated biographic spell dataset (bio_spells) that integrates and replaces the previous spell datasets et_spells, al_spells and lu_spells. Furthermore, a one Euro spell dataset (ee_spells) was introduced during wave 4. The household and person registers (*hh_register; p_register*) are available in wide format. During wave 5, the scientific use file was extended at the individual level by one dataset for the vignette module (*VIGDAT*) and was complemented by a dataset on resident children (KINDER), which includes household information. For further information on the structure of each dataset, please refer to the PASS User Guide (Bethmann & Gebhardt, 2011).

4 Generated variables

4.1 Coding responses to open-ended survey questions

Some items of the survey were gathered as closed items with an open residual category or as open-ended items. In such cases, additional variables were usually generated,[19] which differed from the original variable only insofar as the information from the open-ended responses could not be coded to the corresponding categories. Moreover, in some cases, new categories were created based on the information obtained from open-ended questions. The name of these additional variables frequently differs from that of the original variable in the last digit only, where “0” is replaced by “1.” The items on country of birth, nationality and parent/grandparent country of residence before migration were anonymised and assigned variable names.[20] Tables 9 and 10 provide an overview of the open-ended survey questions that were coded for wave 7.[21]

[19] Other information from open-ended survey questions was not coded, such as the name of the institution providing basic social security (PTK0100).

[20] ogebland (country of birth); ostaatan (nationality); ozulanda to ozulandf (parent/grandparent country of residence before migration).

[21] Variables for which information was obtained via open-ended questions and coded in the previous waves but not in the current wave are not listed (with the exception of the spell dataset for Unemployment Benefit II). Observations in waves without obtaining information on these variables were coded -9 (item not asked in wave) and documented in the survey wave data report.

Table 9: Coding responses to open-ended questions at the household level in wave 7

Regular variable name	Coded to variable	Dataset	Name
HD1100a-o	*HD1101a-o*	*HHENDDAT*	Other Employment status of HH members, proxy information, if necessary
HW0880a-i	*HW0881a-j*	*HHENDDAT*	Other reason for moving out, not listed
HBT0200a-g	*HBT0201a-g*	*HHENDDAT*	Source of information on the educational package
HBT0900a-f	*HBT0901a-f*	*HHENDDAT*	Other reasons not to apply for services of the “Bildungs und Teilhabepaket”
HT0510a-g	*HT0511a-g*	*KINDER*	Other type of group or club that a child is member of
AL21300a-h – AL22100a-h	*AL21301a-h* *AL21401a-h* *AL21501a-h* *AL21601a-h* *AL21701a-h* *AL21801a-h* *AL21851a-h* *AL21901a-h* *AL22001a-h* *AL22101a-h* *AL22102a-h* *AL22103a-h*	*alg2_spells*	Other reason for benefit cut, not listed
AL22200a – AL22200h	*AL22201a-h*	*alg2_spells*	Other reason for discontinuation of receipt of UB II, not listed

Table 10: Coding responses to open-ended questions at the individual level in wave 7

Regular variable name	Coded to variable	Dataset	Name
PSM0200a-l	*PSM0201a-n*	*PENDDAT*	Other social network used in the last four weeks
PB0230 (code 6)	*PB0231*	*PENDDAT*	Other German school qualification, not listed (update)
PB0230 (code 7)	*PB0231*	*PENDDAT*	Other foreign school qualification, not listed (update)
PB0400 (code 9)	*PB0401*	*PENDDAT*	Other German school qualification, not listed (first survey or not reported in previous wave)
PB0400 (code 10)	*PB0401*	*PENDDAT*	Other foreign school qualification, not listed (first survey or not reported in previous wave)
PB1000	*PB1001*	*PENNDAT*	Other foreign school qualification, not listed (first survey or not reported in previous wave)
PB1600	*PB1601*	*PENDDAT*	Other qualification to which the foreign qualification corresponds, not listed
AL0600	*AL0601*	*bio_spells*	Other reason for no longer being registered as unemployed, not listed
BIO0100	*BIO0101*	*bio_spells*	Other type of activity, not listed
ET2400	*ET2401*	*bio_spells*	Other source to get notice of a job
ET2420	*ET2421*	*bio_spells*	Other social network as source to get notice of a job
PRD0300	*PRD0301*	*PENDDAT*	Other reason the rounded or estimated annual income of the previous year is mentioned
PSU0200a-g	*PSU0201a-g*	*PENDDAT*	Other reasons not to look for a job
EE0300a-h	*EE0301a-h*	*ee_spells*	Other reason for not participating in a one-euro job
EE1000a-e	*EE1001a-e*	*ee_spells*	Other reason why one-euro job was terminated prematurely
PEE0200a-d	*PEE0201a-e*	*PENDDAT*	Other source of information of one-euro jobs
PAS0900a-g	*PAS0901a-g PAS0901i*	*PENDDAT*	Other places where target pers. obtained information about job vacancies, not listed
PAS0920a-l	*PAS0921a-l*	*PENDDAT*	Other social network as source of information on job vacancies
PG0900a-f	*PG0901a-g*	*PENDDAT*	Other health problems, not listed
PG1300	*PG1301*	*PENDDAT*	Other health insurance, not listed
PG1500a-g	*PG1501a-i*	*PENDDAT*	Other reasons to play sports more often
PG1510a-g	*PG1511a-g*	*PENDDAT*	Other reasons to play sports less often
PSB0100	*PSB0101*	*PENDDAT*	Other sport done most often
PSB0700	*PSB0701*	*PENDDAT*	Other crucial factor to do this kind of sport
PSB0900	*PSB0901*	*PENDDAT*	Other sport done most often in childhood
PSB1200a-t	*PSB1201a-t*	*PENDDAT*	Other sport done most often in adolescence
PP1300a-e	*PP1301a-e*	*PENDDAT*	Other private caretaking activities
PMI0200	*ogebland*	*PENDDAT*	Other country of birth, not listed
PMI0500	*ostaatan*	*PENDDAT*	Other nationality, not listed
PMI1000a-f	*ozulanda-f*	*PENDDAT*	Other country of birth, not listed
			Country from which parent/grandparent mi-

			grated
PA1100[22]	*freiz1-3*	*PENDDAT*	First to third leisure time activity
PA1200[23]	*frwunsch*	*PENDDAT*	Desired leisure time activity
PA1300a-g	*PA1301a-g*	*PENDDAT*	Other reason for not pursuing the leisure time activity, not listed
PSH0200 (code 9)	*PSH0201*	*PENDDAT*	Other German school qualification of mother, not listed
PSH0200 (code 10)	*PSH0201*	*PENDDAT*	Other foreign school qualification of mother, not listed
PSH0300a-i (code 7)	*PSH0301a-i*	*PENDDAT*	Other German vocational qualification of mother, not listed
PSH0300a-i (code 8)	*PSH0301a-i*	*PENDDAT*	Other foreign vocational qualification of mother, not listed
PSH0500 (code 9)	*PSH0501*	*PENDDAT*	Other German school qualification of father, not listed
PSH0500 (code 10)	*PSH0501*	*PENDDAT*	Other foreign school qualification of father, not listed
PSH0600a-i (code 7)	*PSH0601a-i*	*PENDDAT*	Other German vocational qualification of father, not listed
PSH0600a-i (code 8)	*PSH0601a-i*	*PENDDAT*	Other foreign vocational qualification of father, not listed

4.2 Harmonisation

The survey instruments for some variables changed across waves. In particular, the integration of the module “employment biography” in wave 2 provided critical information on employment status, current main employment, status of economic inactivity and receipt of UB I in a different way than in wave 1. Since then, information has been collected not only for the date of the interview but also for particular periods.
To facilitate cross-wave analyses in such cases, variables are generated for important indicators, which are harmonised across waves. Harmonisation creates a special group within the generated variables (see Section 4.4) that is used to standardise indicators collected in different ways retrospectively.

Changes between the waves can affect the entire survey concept, categories and interviewed groups. Harmonised variables thus consider different source variables that result from changed survey concepts, categories or interviewed groups. This was an effort to standardise them across waves as much as possible before variables were generated.

[22] The variable PA1100 is not included in PENDDAT itself, since it does not include any additional information aside from the fact whether a target person has provided an open response or replied to the question with "don't know" or "details refused". Responses of "don't know" or "details refused" in PA1100 were included in the variables freiz1-3.

[23] The variable PA1200 is not included in PENDDAT itself, since it does not include any additional information aside from the fact whether a target person has provided an open response or replied to the question with "don't know" or "details refused". Responses of "don't know" or "details refused" in PA1200 were included in the variable frwunsch.

Thus far, the simple classification for occupational status (stibkz) has been harmonised; however, the need harmonisation is expected to increase with the duration of the panel.

Table 11: Harmonised variables in the individual dataset (*PENDDAT*)

Variable	Subject area	Name
stibkz	Employment	Current occupational status, simple classification, harmonised (anonymised)

Although explicitly harmonised variables also consider changes in categories and interviewed group across waves - in addition to changes in the survey concept - a second type of variable does not explicitly consider changes in the interviewed groups. These variables are generated for all waves but may contain information for different groups of respondents in each wave. These differences result from revisions to the filtering processes performed between waves and affect the source variables of generated variables.

Accordingly, cross-wave variables of this type apply in addition to harmonisations and standardise individual aspects across waves. In contrast to the harmonised variables, they are generated for each wave for all groups for which the corresponding source variables were collected. Thus, they can easily be used to evaluate the cross-section of a specific wave. However, in the longitudinal section, these differences must be considered before statements about changes between the waves can be made.

Before working with cross-wave but not harmonised variables, it should be verified whether differences in the interviewed groups might cause problems in the evaluations, and it should be determined whether standardisation is necessary.[24]

Subsequent cross-wave variables are different for the group for which they are generated.

[24] For example, in wave 1, the groups of respondents that were questioned about their employment were different from those questioned in the waves that followed. Accordingly, the respective groups that provided information about occupational status, occupational activities, working hours, fixed-term employment, etc., varied.

Table 12: Variables in the individual dataset (*PENDDAT*) are generated across waves but not completely harmonised

Variable	Subject area	Name
isco88	Employment	ISCO 88 (ZUMA coding), current employment, gen.
kldb	Employment	Classification of occupations 1992, current employment
azhpt2	Employment	Current actual working hrs. main employment (without marginal employment, incl. cat. info.), gen.
azges2	Employment	Current total actual working hrs. (without marginal employment, incl. cat. info.), gen.
befrist	Employment	Current activity: limited contract? Generated (all waves)
mps	Employment	Magnitude Prestige Scale, current employment, gen.
siops	Employment	Standard International Occupational Prestige Scale, current employment, gen.
isei	Employment	International Socio-Economic Index, current employment, gen.
egp	Employment	Class scheme acc. to Erikson, Goldthorpe and Portocarrero (EGP), current occupation, gen.
esec	Employment	European Socio-economic Classification (ESeC), current occupation, gen.
stib	Employment	Occupational status, code number, current employment, gen.
netges	Employment	Current total net income (without marginal employment, incl. cat. info.), gen.
alg1abez	Benefit receipt	Current receipt of UB I, gen.
aktmassn	Participation in measures	Current participation in a programme funded/promoted by the employment agency, gen.

4.3 Dependent interviewing

At various times in both the household and personal interviews, information was gathered via dependent interviewing, i.e., interviews that were dependent on the responses provided during a previous wave. In this approach, data from the previous interview are used to control the filter questions or are integrated directly into the question text of the current interview.

Two main goals were pursued, utilising information from previous waves. First, changes that occurred since the previous wave were recorded, depending on the information available from the previous wave.[25] At those points, information from previous waves was used to control the filter. Second, the respondent should have received information. In places where changes since the previous wave were to be collected, the interview date of the previous wave was included in the question text to clarify the definition of the reporting period.[26] In other places, especially where spell information was updated[27], the previous re-

[25] For example, individuals were only asked about their highest school qualification once. Only qualifications obtained since the previous interview were reported in subsequent waves.

[26] For example, if only new school qualifications were to be reported, the following question was asked: "Have you obtained a general school qualification since our last interview on [interview date of previous wave]?"

sponse was integrated into the question text to remind the respondent and prevent incorrect changes in status. Such changes are artifacts of the open-ended survey question arising out of inaccurate memories or imprecise information.
If information from a single wave in the dataset is reviewed, information is incomplete for some respondents due to dependent interviewing, which only represents the changes between survey dates. For respondents who are interviewed for the first time about a certain topic, complete information might be information available for that wave[28].
During data preparation, the recorded changes are combined with information from the previous wave to create variables and datasets with complete information. The spells in the existing spell datasets are then updated. In the cross-section datasets (*HHENDDAT, PENDDAT*), however, generated variables are created in which the information from the previous wave is combined with the reported changes.
Table 13 and 14 provide a brief overview of the relevant updates to the questionnaires and indicate the variables for which updated information was obtained. Cases for which generated variables were updated or continued are listed in Chapter 4.4 of this data report.

---

[27] Examples include updates of UB II receipts since the previous wave in the household interview or employment or unemployment updates in the individual interview.

[28] Individuals who were asked about their school qualifications for the first time reported their highest school qualification. Therefore, complete information on the highest school qualification is available for this wave in the recorded variable. In the subsequent wave, only newly obtained school qualifications are recorded. For example, if a school qualification is recorded, it is not clear whether it represents the individual's highest school qualification. In that sense, the information obtained in the subsequent wave is incomplete in its reported variables.

Table 13: Updated information in wave 7, household questionnaire

Household questionnaire for re-interviewed households (HHalt)			
Construct	**Q. no.**	**Note**	**Update in variable**
Housing situation		Form of accommodation, type of tenancy and type of hostel/home/hall of residence updated during the interview	HHENDDAT: *HW0200 to HW0400*
Household structure		Household size updated during the interview	HHENDDAT: *HA0100*
		Sex of the individuals in the household corrected during the interview, if necessary	HHENDDAT: *HD0100a* to *HD0100o*
		Age of the individuals in the household updated during the interview	HHENDDAT: *HD0200a* to *HD0200o*
		Family relationships updated during the interview	not provided in the SUF
Size of dwelling in sqm	HW1000	Updated in generated variable	HHENDDAT: *wohnfl*
Receipt of Unemployment Benefit II	Module "Unemployment Benefit II"	Updated in Unemployment Benefit II spell dataset	alg2_spells: Variables of the Unemployment Benefit II spell dataset
		Information on the HH's current receipt of Unemployment Benefit II	HHENDDAT: alg2abez
		Information on the benefit units's Unemployment Benefit II receipt	p_register: bgbezs7; bgbezb7

Table 14: Updated information since wave 7, personal questionnaire

Personal questionnaire			
Construct	**Q. no.**	**Note**	**Update in variable**
Highest general school qualification	PB0220-PB1100	Updated in generated variable	PENDDAT: *schul1* (without responses to open-ended questions) *schul2* (with responses to open-ended questions)
Year in which highest school qual. was gained	PB0410	Updated in generated variable	PENDDAT: *schulabj*
Vocational qualification	PB1200-PB1600	Highest vocational qualification, updated in generated variable	PENDDAT: *beruf1* (without responses to open-ended questions) *beruf2* (with responses to open-ended questions)
Year of vocational qualification	PB1310a-k	Updated in generated variable	*berabj*
Periods of updated activities in the BIO spell dataset	BIO0600z1, BIO0600z2, BIO0400z, BIO0500z	Updated in the BIO spell dataset for attached spells	bio_spells *BIO0400, BIO0500, BIO0600*
		Updated in the BIO spell dataset for attached spells Information on current employment, updated in generated variables	bio_spells: ET2300, ET2700 PENDDAT: *isco88; kldb; stib; stibkz; arbzeit; befrist; mps; siops; isei; egp; esec*
		Information on current economic inactivity/employment status, updated in generated variables	PENDDAT: *etakt; alakt; statakt*
Periods of receipt of Unemployment Benefit I in updated unemployment spells		Information on current receipt of Unemployment Benefit I	bio_spells: AL0700, AL0800, AL0900, *AL1000, AL1100, AL1200*
		Updated in the BIO spell dataset for attached spells	bio_spells: AL0600, AL0601 PENDDAT: *alg1abez*
Periods of updated activities in the EE spell dataset			*ee_spells:* *EE0800a, EE0800b*
Information regarding premature end in the EE spell dataset			*ee_spells:* *EE0900,* *EE1000a-EE1000e, EE1001a-EE1001e*

A distinction must be drawn between characteristics for which previously collected information is updated with information on changes between the survey dates and so-called

constant characteristics that are not expected to change over time. Therefore, these characteristics are recorded only once in PASS, but in some cases, corrections are possible. Because information on these characteristics is usually only available for the surveyed variables during the first interview, they are subsequently provided in the form of generated variables (see Chapter 4.4, Bethmann & Gebhardt, 2011).

4.4 Simple generated variables

Simple generated variables include variables for which different items in a construct are surveyed separately for technical reasons and then aggregated. Alternatively, information from the current wave is combined with information from the previous wave (see Chapter 4.3), such as the highest educational qualification (see Chapter 4.3). Important information can also be obtained by merging partial datasets (e.g., indicators for current receipt of UB I or II).

The simple generated variables for households and individuals who are interviewed on a topic for the first time can always be generated based on information from the current wave. Households and individuals who provided information on a topic during a previous wave can be differentiated in the cross-section datasets (*HHENDDAT; PENDDAT*) to indicate the origin of the variables necessary for variable generation. The three different types of simple generated variables are provided in table 15.

Table 15: Simple generated variables in the cross-section datasets (HHENDDAT; PENDDAT) for households and individuals who previously provided information on the topic

Type	Generation based on source data from		Description
	wave of the first survey of the topic for HH/individual	current wave	
unveränderlich (uv)	yes	no	Information gathered in the first survey is generally adopted in the subsequent wave – unless input errors were corrected in the current wave. Example: *zpsex* (sex)
fortgeschrieben (fs)	yes	yes	Information that was current in the previous wave is combined with information of the current wave and updated, if necessary. Example: *schul1* (highest school qualification)
unabhängig neu (neu)	no	yes	The variable is newly generated from the data of the current wave in each wave, regardless of the information from the previous wave. Example: *hhincome* (net income of household)

Explanations that are more detailed must be provided on the type “unveränderlich (uv)” simple generated variables for *PENDDAT*. A first-time survey of a topic with an individual does not always take place during the first wave in which the individual provides an interview. Two groups of individuals are considered first-time interview respondents even if they provide a repeat interview.

The first group is individuals moving back into a household. Individuals who move out of their previous household to form a split-off household (see Chapter 2.4) take their preload information with them. Thus, they can be treated correctly as either first-time interviews or repeated interviews. However, if an individual returns from a split-off household into a panel household in which he/she lived during a previous wave, the preload of this individual is not transferred from the split-off household to the original household. Individuals returning home are treated as first-time interviewees. This situation has occurred since wave 3. The first move-outs of HHalt occurred during wave 2, and returns may occur by wave 3.

An individual preload for dependent interviewing is created for an individual (see Chapter 4.3) only if he/she provided an interview during one of the two preceding waves. The context for this rule is that there is a point in time until which an individual is expected to re-

member the response in spell form. Individuals who last provided a personal or senior citizen interview during the third wave or earlier had passed this point. To reduce respondent stress and protect the validity of the information provided, which is presumably severely threatened beyond this limit, individuals whose reference date for information about spell results is before the relevant date are treated as first-time respondents.[29] This situation first occurred in wave 4 because that wave was the first time that a previous personal interview could have taken place more than two waves previously.

The information on which these generated variables are based is collected again for these two groups (e.g., in the module "social origin") because they are treated as first-time interviews. Data preparation treats this survey information identically to the information from individuals engaged in actual first-time interviews within the PASS framework. These generated variables, e.g., the status of the mother and father, are thus based on information from the current wave. No transfer of information from previous waves takes place, and there is no attempt to make the data fit plausibly with previous information. We assume that the information provided by the target person, which is processed to become generated variables, is consistent with previous information in a repeated survey. However, deviations from previously obtained information in the previous waves cannot be generally excluded. Individuals included in either group are flagged in *PENDDAT* by the variable *altbefr* as first-time respondents (code "0" or "-9" for wave 1).

These simple generated variables are provided in tables 16 to 21. The tables include short descriptions of each variable. Furthermore, the source variables to generate the variable are indicated.[30] For the cross-section datasets (*HHENDDAT and PENDDAT*), additional information identifies the type of simple generated variable shown in Table 16 (uv; fs; neu). This division is not used for spell datasets because there are no wave-specific observations. Instead, variables are newly generated at the spell level if the spell was newly included in the wave or was updated with information obtained in the current wave. In addition, register datasets follow a different logic, and no further differentiation was made.

---

29 Excluding previously granted consent to the merging of data. This preload information is generated regardless of when the previous personal interview was provided to avoid individuals negating question *RegP0100* and de facto withdrawing their consent. The option to withdraw consent to the merging of data remains unaffected by this decision.

30 The data report documents how the variables in the cross-section datasets (*HHENDDAT; PENDDAT*) were generated for observations in previous waves. The documentation for specific waves also describes the generation of wave-specific variables in the register datasets. The generated variables in the spell datasets were always generated in the updated datasets. If a spell was not updated, the generated variables remain unchanged (with the exception that a special code was used in the censoring indicator if the spell could not be continued for technical reasons). If a spell was updated, then the most current information was used, i.e. the variables provided with information from the current wave or cross-section variables in the spells relevant for the current wave.

Table 16: Wave 7 simple generated variables in the household (HHENDDAT) and KINDER datasets (in alphabetical order)

Variable	Variable label and description	Source var. for generated var. in wave 7
alg2abez	*Current receipt of UB II of the HH, generated* Indicator for the household's current receipt of Unemployment Benefit II	*zensiert; AL20300; AL20400; AL20500* (*alg2_spells*); information on further receipts of Unemployment Benefit II (*AL22700*); *hintjahr (HHENDDAT)*
anzgeschw	*Number of siblings in the household* *Indicator of an individual's number of siblings* *Parenthood and sibling status are surveyed separately. Individuals may share one parent but not call themselves siblings. Therefore, anzgeschw is not equivalent to sibling status, which can be generated through the parent indicator variable in p_register.*	
bik	*BIK region size classes (GKBIK10), generated* The information on region size was generated by infas by converting the postcode from the address to *GKBIK10* (neu).	Supplied by survey institute
blneualt	*Western German States or Eastern German States, generated* Divides the German states into the western states of the former FRG (excluding Berlin) and the eastern states of the former GDR (with Berlin). Infas determined the state based on the postcodes from the address data (neu).	Information generated and supplied by the survey institute on the federal state in which the household is resident at the survey date.
butaber	*Eligibility for education package at point of interview* *This variable indicates that a household is eligible to draw benefits from the education and participation package if he draw one of the benefits like UB II, children`s allowance, housing or social benefit since January of the year before the actual year of the survey (neu).*	AL20200; AL20400; AL20500 (alg2_spells); HA0250a-b; HW1800; HW1950; HEK0100; HEK0115; HEK1630; HEK1645 (HHENDDAT)

Variable	Variable label and description	Source var. for generated var. in wave 7
hhinckat	*Categorised household income per month (in EUR), gen.* Categorised information on the household's income aggregated from several survey items into one variable (neu)	*HEK0700; HEK0800; HEK0900; HEK1000; HEK1100* (*HHENDDAT*)
hhincome	*Household income per month (in EUR) incl. categorised information, gen.* This generated variable integrates information from categorised and open-ended survey questions on net household income (neu).	*HEK0600; HEK0700; HEK0800; HEK0900; HEK1000; HEK1100* (*HHENDDAT*)
hintdat	*Date of household interview* This generated variable indicates the date on which the household interview was conducted in the format YYMMDD (neu)	*hintjahr; hintmon; hinttag* (*HHENDDAT*)
hintnum	*interviewer in household interviews* The artificial identifier indicates the interviewer who conducted the interview. This information is consistent between PENDDAT and HHENDDAT as well as across waves. A definite characteristic of the label always identifies the same interviewer (neu).	*information that is generated and supplied by the survey institute*
kindu4	*Control variable: child under the age of 4 in the HH* A variable indicating that at least one individual in the household is under the age of four in the wave. As the generated variable is based only on the age details in the household dataset, it is irrelevant whether this individual aged four is actually the child of another individual living in the household (neu).	*HD0200a - HD0200o (HHENDDAT)*
kindu13	*Control variable child under the age of 13 in the HH* A variable indicating that at least one individual in the household is under the age of 13 in the wave. As the generated variable is based only on the age details in the household dataset, it is irrelevant whether this individual aged 13 is actually the child of another individual living in the household (neu).	*HD0200a - HD0200o* (*HHENDDAT*)

Variable	Variable label and description	Source var. for generated var. in wave 7
kindu15	*Control variable: child under the age of 15 in the HH* A variable indicating that at least one individual in the household is under the age of 15 in the wave. As the generated variable is based only on the age details in the household dataset, it is irrelevant whether this individual aged 15 is actually the child of another individual living in the household. If the response to the open-ended question on age was missing, the categorical follow-up question about the age groups was also used to generate the variable (neu).	*HD0200a - HD0200o*; categorical follow-up question about age group (in cases of no response in *HD0200*) (*HHENDDAT*)
kindu25	*Control variable: child under the age of 18 or pupils under the age of 25 in the HH.* A variable indicating whether at least one individual in the household is under the age of 18 or that at least one individual is between the age of 18 and 25 and pupil. As the generated variable is based only on the age details in the household dataset, it is irrelevant whether this individual of the age group is actually the child of another individual living in the household. If the response to the open-ended question on age was missing, the categorical follow-up question about the age groups was used to generate the variable as well (neu).	
wohnfl	*Living space in sqm, gen.* Information on the size of the living space in the household's current dwelling. In the case of re-interviewed households, the size of the living space was only asked as of the second wave if the household had moved house or if the house/apartment had changed since the previous wave (fs).	*For first survey:* *HW1000 (HHENDDAT)* For repeated survey: *wohnfl from previous wave; HW1000;* (*HHENDDAT*)

Table 17: Simple generated variables for wave 7 in the individual dataset (PENDDAT) (in alphabetical order)

Variable	Variable label and description	Source var. for generated var. in wave 7
akt1euro	*Current part. in one-euro job, generated* Indicator: respondent is participating in a one-euro job program at the time of the interview (neu).	*zensiert (ee_spells*)
alakt	*Currently reported as unemployed, generated (as of wave 2)* Indicator: the TP was unemployed at the date of the personal interview of that wave (neu).	*zensiert; spintegr; BIO0101 (bio_spells)*
alg1abez	*Current receipt of UB I, generated* Indicator: respondent is receiving Unemployment Benefit I at the interview date. In wave 6, the periods since January 2010 during which the respondent was unemployed were surveyed. For each spell, additional questions about whether and when the respondent received UB I (neu).	*AL0700; AL1000; AL1100; AL1200 (bio_spells)*
apartner	*Control variable: unmarried partner living in HH* Indicator: respondent has a cohabitee or partner whose status is not specified in the household (neu).	Information on relationships between household members (household grid); *PD0500 - PD0800* (*PENDDAT*)
azhpt1	*Current contractual working hrs. main employment (without marginal employment), gen* Weekly contractual working hours provide the respondent's primary employment at the time of the interview. Generated from open-ended questions about working hours.	*ET2005 (bio_spells)*
azges1	Akt. vertragl. Arb.zeit insgesamt (ohne Minijobs), gen. Wöchentliche, vertraglich vereinbarte Arbeitszeit aller Erwerbstätigkeiten, die der Befragte zum Interviewzeitpunkt ausübt, generiert aus offenen Angaben zur Arbeitszeit (neu)	*ET2005 (bio_spells)*

azges2	*Current total actual working hrs. (without marginal employment, incl. cat. info.), gen.* Actual weekly working hours for all positions held by the respondent at the time of the interview. Generated from responses to open-ended questions on working hours and a categorical follow-up question in which irregular working hours were reported (neu).	*ET2105; ET2205 (bio_spells)*
befrist	*Current employment: limited contract?* *Generated (all waves)* Indicator: the employment position held by the respondent at the interview date is on a limited contract (neu).	*PET2510a; PET2510b* (*PENDDAT*)
begjeewt	*Start year of first employment, generated* The first year during which the respondent was employed in a regular position. To generate this variable, information about the first regular position was combined with information from the employment spells if the respondent had previously reported his/her first regular employment since January 2010 (uv).	<u>*For first survey:*</u> *bjahr* (*bio_spells*); *PET3200b* (*PENDDAT*) <u>After first survey:</u> *begjeewt* from previous wave (*PENDDAT*)
begmeewt	*Start month of first employment, generated* The month during which the respondent first held regular employment (generated, see begjeewt) (uv).	<u>*For first survey:*</u> *bmonat* (*bio_spells*); *PET3200a* (*PENDDAT*) <u>After first survey:</u> *begmeewt* from previous wave *(PENDDAT)*
berabj	*Year of the highest vocational qualification* The year in which the respondent obtained his/her highest vocational qualification at the interview date (fs). <u>Note:</u> The year in which the reported vocational qualifications reported in wave 1 but asked in wave 2.	<u>For first survey:</u> *PB1310aj-kj* (*PENDDAT*) <u>For repeated survey:</u> *berabj* from previous wave; *PB1310aj-kj* (*PENDDAT*)

beruf1	*Highest vocational qual., excl. foreign qual and open info., generated* Identifies the highest vocational qualification obtained by the interview date by ranking the vocational qualifications cited by the respondents, excl. information from open-ended questions (fs).	<u>For first survey:</u> *PB0100; PB0200; PB0300; PB1200b; PB1200c; PB1300a-j;* (*PENDDAT*) <u>For repeated survey:</u> *beruf1* from previous wave; *PB0100; PB0200; PB1200a; PB1300a-j* (*PENDDAT*)
beruf2	*Highest vocational qual., incl. foreign qual and open info., generated* Defined as in beruf1 with the following differences: 1. Inclusion of responses to open-ended questions; 2. Inclusion of foreign qualifications; and 3. Degrees are not distinguished by type of institution (e.g., university or other institution of higher education) but by level (Bachelor's degree; Master's degree; Ph.D.) (fs).	<u>For first survey:</u> *PB0200; PB1301a-j; PB1500a; PB1500b; PB1500c; PB1601* (*PENDDAT*) <u>For repeated survey:</u> *beruf2* from previous wave; *PB0200; PB1301a-j; PB1500a; PB1500b; PB1500c; PB1601* (*PENDDAT*)
brges	*Current total gross income (without marginal employment, incl. cat. info.), gen.* Contains the cumulative information on gross income from all employment (> EUR 400). Generated from the answers provided in open-ended questions on gross income and categorical follow-up question when the "don't know" or "details refused" answers were provided to open-ended questions (neu)	*ET2802; ET2902; ET3002; ET3102; ET3202; ET3302 (bio_spells)*
brutto	*Gross income from the current main employment incl. categorised information, generated* A generated variable integrating information from categorised and open-ended survey questions on gross income (neu).	*ET2802; ET2902; ET3002; ET3102; ET3202; ET3302 (bio_spells)*
bruttokat	*Categorised gross income from the current main employment, generated* This variable aggregates the categorised information on gross income for a specific variable, which combines several items on income categories (neu).	*ET2802; ET2902; ET3002; ET3102; ET3202; ET3302 (bio_spells)*

ejhrlewt	*Time when last employment ended (year)* Last year in which the respondent was in employment. To generate this variable, information from the employment spells was combined with information on the last employment if the respondent had been out of work since January 2010 (fs).	<u>*For first survey:*</u> *PET1200b* (*PENDDAT*); *ejahr; emonat (bio_spells*) <u>For repeated survey:</u> *ejhrlewt from previous wave* (*PENDDAT*); *ejahr; emonat (bio_spells*)
ekin1517	*Control variable: own child aged between 15 and 17 in the household* A variable indicating whether the respondent has a natural child, a stepchild/adopted child or a child of non-specified status aged between 15 and 17 in the household (neu).	Information on relationships between household members (household grid)
ekind	*Control variable: own child in HH* A variable indicating whether the respondent has a natural child, a stepchild/adopted child or a child of non-specified status of any age in the household (neu). It can occur in rare household constellations that according to *ekind*, an individual has children living in the household, but their *pnr* does not appear in the pointers *zmhh* and *zvhh* of *p_register*. This can occur in case of same-sex relationships with children or if both the current and the former partner live in the household.	Information on relationships between household members (household grid)
ekin614	*Control variable: own child aged between 6 and 14 in the household* A variable indicating whether the respondent has a natural child, a stepchild/adopted child or a child of non-specified status aged between 6 and 14 in the household (neu).	Information on relationships between household members (household grid)
ekinu15	*Control variable: own child under the age of 15 in HH* A variable indicating whether the respondent has a natural child, a stepchild/adopted child or a child of non-specified status under the age of 15 in the household (neu).	Information on relationships between household members (household grid)
ekinu18	*Control variable: own child under the age of 18 in HH* A variable indicating whether the respondent has a natural child, a stepchild/adopted child or a child of non-specified status under the age of 18 in the household (neu).	Information on relationships between household members (household grid)
epartner	*Control variable: spouse or registered partner in HH* A variable indicating whether the respondent has a spouse or a same-sex registered partner in the household (neu).	Information on relationships between household members (household grid)

etakt	*Currently employed (>EUR 400 per month), gen. (as of wave 2)* A variable indicating whether the TP had an ongoing spell of employment at the time of the personal interview of the respective wave (i.e. employment earning > EUR 400) (neu).	*zensiert, spintegr, BIO0101 (bio_spells)*
famstand	*Marital status, gen.* Generation of a marital status variable integrating information from the personal questionnaire and the control variable *epartner* generated from the household dataset (neu).	*epartner; PD0500; PD0700* (*PENDDAT*)
gebhalbj	*Half-year of birth, gen.* A variable indicating whether the date of birth is in the first or second half of the year of birth (neu).	Information on month of birth
kindzges	*Total number of own children (living in and outside the household), gen.* Total number of the respondent's children including the children living in his/her household and the children living outside the household (neu).	Information on relationships between household members (household grid); *PD0900; PD1000; PD1100* (*PENDDAT*)
kindzihh	*Number of own children in the household, gen.* Variable generated on the basis of the responses in the household questionnaire concerning the number of children that an individual in the household has (total number of individuals in the household (half) matrix who count as children of the respondent plus the number of individuals in the household (half) matrix for whom the respondent is classified as being a parent) (neu). *Note: When using this variable it should be borne in mind that it relates to each individual person. This means that a child who lives in a household together with his/her parents is counted as a "child in the household" for both the father and the mother. Aggregating this variable across the household members will therefore not produce any meaningful results.*	Information on relationships between household members (household grid)

mberuf1	*Highest vocational qualification attained by the mother, incl. mother in the HH, excl. information from open-ended survey questions, gen.* In wave 1, the question about the mother's vocational qualification was asked only if the mother was not living in the survey household. If she was living in the household, this information was obtained from her personal interview. As of wave 2, the question regarding the mother's vocational qualification has been posed to all newly interviewed individuals regardless of whether the mother was living in the household. After wave 2, for respondents taking part in a repeated interview, the values were transferred from the generated variable *mberuf1* from the previous wave (uv).	<u>For first survey:</u> *PSH0300a-i* (*PENDDAT*) <u>After first survey:</u> *mberuf1* from previous wave (*PENDDAT*)
mberuf2	Defined as in *mberuf1* except that responses to open-ended questions were also considered to generate *mberuf2* (uv).	<u>For first survey:</u> *PSH0301a-i* (*PENDDAT*) <u>After first survey:</u> *mberuf2* from previous wave (*PENDDAT*)
mhh	*Control variable: mother living in HH* A variable indicating whether the respondent's biological mother, stepmother, adoptive mother or mother of non-specified status lives in the household (neu).	Information on relationships between household members (household grid)
migration	*Respondent's migration background, generated* The following four categories were included in a generated variable for migration background: no migration background; personal migration (first generation); migration of at least one parent but no personal migration (second generation); migration of at least one grandparent but not the respondent or either parent (third generation) (uv). <u>*Note:*</u> *The concept for generating this variable has been revised as of wave 2.* Previously, only the information on whether the respondent was born in Germany and which ancestor moved to Germany was collected. Now, information on whether an ancestor was born outside Germany and if applicable, which ancestor, is included. To guarantee consistency across waves, the variable for wave 1 was regenerated.	<u>*For first survey:*</u> *PMI0100; PMI0700; PMI0800a-f; PMI0900a-f* (*PENDDAT*) <u>After first survey:</u> *migration* from previous wave (*PENDDAT*)

mschul2	*Highest general school qualification attained by the mother, incl. mother in HH, incl. information from open-ended questions, gen.* Same as *mschul1*, apart from the fact that responses to open-ended questions were also taken into account for the generation of *mberuf2* (uv).	For first survey: *PSH0201* (*PENDDAT*) After first survey: *mschul2* from previous wave (*PENDDAT*)
mschul1	*Highest general school qualification attained by the mother, incl. mother in HH, incl. information from open-ended questions, gen.* In wave 1, the mother's highest academic qualification was inquired about only if the mother was not living within the survey household. If she was living in the household, this information was obtained from her personal interview (uv). As of wave 2, the mother's highest academic qualification has been asked of all newly interviewed individuals regardless of whether the mother was living in the survey household.	For first survey: PSH0200 (PENDDAT) After first survey: mschul1 from previous wave (PENDDAT)
mstib	*Mother's occupational status, code number, gen.* The detailed occupational status of the mother was generated from the individual variables (uv).	For first survey: *PSH0320; PSH0330; PSH0340; PSH0360; PSH0370; PSH0380* (*PENDDAT*) After first survey: *mstib (PENDDAT)*
netges	*Current total net income (without marginal employment, incl. cat. info.), gen.* This variable contains the accumulated information on net income from all employment positions (> EUR 400), which is generated from the answers to open-ended questions on net income and a categorical follow-up question when respondents provided "don't know" or "details refused" answers to open-ended questions (neu).	*ET3402; ET3502; ET3602; ET3702; ET3802; ET3902 (bio_spells)*
netto	*Net income of the current main employment incl. categorised information, gen.* A generated variable integrating information from categorised and open-ended survey questions on net income (neu).	*ET3402; ET3502; ET3602; ET3702; ET3802; ET3902 (bio_spells)*

nettokat	*Categorised net income from the current main employment, gen.* This variable aggregates the categorised information on net income for a specific variable, which combines several items on income categories (neu).	*ET3402; ET3502; ET3602; ET3702; ET3802; ET3902 (bio_spells)*
palter	*Age (from PD0100), gen.* The respondent's age is generated from the date of birth and date of the current personal interview (neu).	*PD0100; pintjahr, pintmon, pinttag* (*PENDDAT*)
panel	*Willingness to participate in the panel* (neu)	Information supplied by the survey institute regarding the households' willingness to participate in the panel.
pintdat	*Date of personal interview* This generated variable indicates the date on which the personal interview was conducted in the format YYMMDD (neu).	*pintjahr, pintmon, pinttag* (*PENDDAT*)
pintnum	*interviewer in personal interview* The artificial identifier indicates the interviewer who conducted the interview. This information is consistent between PENDDAT and HHENDDAT as well as across waves. A definite characteristic of the label always identifies the same interviewer (neu).	Information that is generated and supplied by the survey institute.
schul1	*Highest school qualification, excl. foreign qualifications and information from open-ended survey questions* This variable records the highest academic qualification. Equivalent Eastern and Western German qualifications were combined (e.g., EOS and Abitur), but information from open-ended questions was excluded (fs).	<u>For first survey:</u> *PB0200; PB0220; PB0230; PB0300; PB0400* (*PENDDAT*) <u>For repeated survey:</u> *schul1 from previous wave*; *PB0200; PB0220; PB0230; PB0300; PB0400* (*PENDDAT*)

schul2	*Highest school qualification, incl. foreign qualifications and information from open-ended survey questions* Defined as in schul1 with the following differences: 1. inclusion of responses to open-ended questions; and 2. inclusion of information about foreign qualifications (fs).	For first survey: *PB0200; PB0220; PB0231; PB0300; PB0401* (*PENDDAT*) For repeated survey: *schul2 from previous wave; PB0200; PB0220; PB0231; PB0300; PB0401* (*PENDDAT*)
schulabj	*Year in which highest school qual. was attained* Year in which the respondent attained his/her highest *academic* qualification (fs). Note: Re-interviewed respondents for whom information regarding the highest school qualification was already available from a previous wave were not asked in the current wave about the year when this qualification was attained if they had attained a new qualification since the previous wave. In this case, the year in which the qualification was attained was estimated depending on the month and year of the interview. If the interview in wave 6 was conducted before May 2012, it was assumed that the qualification was gained in 2011, if the interview was conducted later than May, the qualification was assumed to have been gained in 2012.	For first survey: *PB0220; PB0230; PB0410; pintjahr; pintmon* (*PENDDAT*) For repeated survey: *schulabj* from previous wave; *PB0220; PB0230; PB0410; pintjahr; pintmon* (*PENDDAT*)
statakt	*Current main status, generated (as of wave 2)* Indicates which main status the TP had at the date of the personal interview of the respective wave (neu).	*zensiert; spintegr; BIO0101; az2ges (bio_spells)*
stib	*Occupational status, code number, generated* A generated of the detailed code number for occupational status from the individual variables. A generated variable using information from the module “employment” (*ET060*-ET120**). If there was more than one ongoing employment spell, the one with the most hours of work was selected. If there was more than one ongoing spell with exactly the same amounts of hours, the one that started first was selected (neu).	*ET0605; ET0705; ET0805; ET0905; ET1005; ET1105; ET1205* (*bio_spells*)

stibeewt	*Occupational status, first employment, code number, generated*Detailed code number of the occupational status in the respondent's first regular employment. To generate the variable, information regarding the first regular employment was combined with information from the employment spells if the respondent had already reported his/her first regular employment during the questions on employment spells since January 2010 (uv).	<u>For first survey:</u> *PET3300; PET3400; PET3500; PET3600; PET3700; PET3800; PET3900 (PENDDAT); ET0605; ET0705; ET0805; ET0905; ET1005; ET1105; ET1205 (bio_spells)* <u>After first survey:</u> *stibeewt* from previous wave (*PENDDAT*)
stiblewt	*Occupational status, last employment, code number, generated* Detailed code number of the occupational status in the respondent's last employment. Information from the employment spells were combined with information on the last employment for the generation if the respondent has been unemployed since January 2010 (fs).	<u>For first survey:</u> *PET1210; PET1220; PET1230; PET1240; PET1250; PET1260; PET1270 (PENDDAT); ET0605; ET0705; ET0805; ET0905; ET1005; ET1105; ET1205 (bio_spells)* <u>For repeated survey:</u> *stiblewt* from previous wave *(PENDDAT); ET0605; ET0705; ET0805; ET0905; ET1005; ET1105; ET1205 (bio_spells)*
vberuf1	*Highest vocational qualification attained by the father, incl. father in the HH, excl. open info., gen.* A generated variable for father's highest vocational qualification analogous to *mberuf1* (uv).	<u>For first survey:</u> *PSH0600a-i* (*PENDDAT*) <u>After first survey:</u> *vberuf1 from previous wave* (*PENDDAT*)
vberuf2	*Highest vocational qualification attained by the father, incl. father in the HH, incl. open info., gen.* A generated variable for father's highest vocational qualification (incl. information from open-ended survey questions) analogous to *mberuf2* (uv).	<u>For first survey:</u> *PSH0601a-i* (*PENDDAT*) <u>After first survey:</u> *vberuf2 from previous wave* (*PENDDAT*)
vhh	*Control variable: father living in HH* Variable indicating that the respondent's natural father, stepfather, adoptive father or father of non-specified status is living in the household (neu).	Information on relationships between household members (household grid)
vschul1	*Highest general school qualification attained by the father, incl. father in HH, excl. information from open-ended questions, gen.* A generated variable for father's highest general *academic* qualification analogous to *mschul1* (uv).	<u>For first survey:</u> *PSH0500* (*PENDDAT*) <u>After first survey:</u> *vschul1 from previous wave* (*PENDDAT*)

vschul2	*Highest general school qualification attained by the father, incl. father in household, incl. open info., gen.* This generated variable records the father's highest general academic qualification (including information from open-ended survey questions) and is analogous to mschul2 (uv).	For first survey: *PSH0501* (*PENDDAT*) After first survey: *vschul2 from previous wave* (*PENDDAT*)
vstib	*Father's occupational status, code number, generated* The detailed occupational status of father is generated from individual variables (uv).	*For first survey:* *PSH0620; PSH0630; PSH0640; PSH0660; PSH0670; PSH0680* (*PENDDAT*) After first survey: *vstib* from previous wave *(PENDDAT)*

Table 18: Wave 7 simple generated variables included in the spell dataset for Unemployment Benefit II (alg2_spells) (provided in the same order as in the dataset)

Variable	Variable label and description	Source var. for generated var. in wave 7
bmonat	*Spell of UB II: start month, generated* The month in which the spell of receiving Unemployment Benefit II began. If information was only available on the season when a spell began, the season was converted into a month to generate the variable. Note: The generated date variables were both checked for plausibility and corrected when necessary. The dates originally reported by the respondent have been included in the source variables as of wave 2. The season in which the spell began were recoded into months as follows: 21 beginning of year/winter → January 24 spring/Easter → April 27 middle of year/summer → July 30 autumn → October 32 end of year → December	*AL20100* (*alg2_spells*)
bjahr	*Spell of UB II: start year, generated* The year during which the spell of receiving Unemployment Benefit II ended. Note: see *bmonat*	*AL20200* (*alg2_spells*)
emonat	*Spell of UB II: end month, generated* The month during which the spell of UB II receipts ended. To generate this variable, information about the season was converted into a month. For right-censored spells (i.e., spells that were ongoing when the household was interviewed), the interview month was entered. Note: see *bmonat*	*AL20300* (*alg2_spells*); *hintmon (HHENDDAT)*
ejahr	*Spell of UB II: end year, generated* The year during which the spell of Unemployment Benefit II ended. In the case of right-censored spells (i.e., spells that were ongoing when the household was interviewed), the interview year was entered. Note: see *bmonat*	*AL20400* (*alg2_spells*); *hintjahr (HHENDDAT)*

alg2kbma - alg2kbmh	*UB II: 1st cut: start month, generated* The month during which Unemployment Benefit II was reduced. To generate this variable, information about the season was converted into a month. Note: These UB II reductions are embedded in spells of UB II receipts. Information on an individual benefit reduction can be distinguished via the indicator at the end of the respective variable (a - h). The generated date variables were checked for plausibility and corrected if necessary. The dates originally reported by the respondent have been included in the source variables since wave 2.	1st benefit cut: *AL21000a* (*alg2_spells*) to 8th benefit cut: *AL21000h* (*alg2_spells*)
alg2kbja- alg2kbjh	*UB II: 1st benefit cut: start year, generated* The year during which the Unemployment Benefit II reduction began. Note: see *alg2kma - alg2kbmf*	1st benefit cut: *AL21100a* (*alg2_spells*)to 8th benefit cut:*AL21100h* (*alg2_spells*)
alg2kema - alg2kemh	*UB II: 1st benefit cut: end month, generated* The month during which the Unemployment Benefit II reduction ended. To generate this variable, information on the season was converted into a month. If the respondent reported the duration of the benefit reduction, this information was used to calculate the end date of the benefit cut based on the generated start date. Note: see *alg2kma - alg2kbmf*	1st benefit cut: *alg2kbma; alg2kbja; AL21200a; AL21201a; AL21202a* (*alg2_spells*) to 8th cut:*alg2kbmh; alg2kbjh; AL21200h; AL21201h; AL21202h* (*alg2_spells*)
alg2keja - alg2kejh	*UB II: 1st benefit cut: end year, generated* Year in which the Unemployment Benefit II cut ended. If the respondent reported a duration for the benefit cut, this information was used to calculate the end date of the benefit cut based on the generated start date. Note: see *alg2kma - alg2kbmf*	1st benefit cut: *alg2kbma; alg2kbja; AL21200a; AL21201a; AL21202a* (*alg2_spells*) to 8th benefit cut: *alg2kbmh; alg2kbjh; AL21200h; AL21201h; AL21202h* (*alg2_spells*)

AL22150a *to* *AL22150h*	*UB II: benefit cut: which HH member's benefit was cut, gen.* This variable records which household members experienced reductions in Unemployment Benefit II. This is a string variable with 15 positions. Starting from the left, each position in this variable represents the position of one individual on the household grid. The first position of the variable, for example, indicates whether Unemployment Benefit II was cut for the first individual in the household during the particular benefit reduction spell, the second position indicates whether the second individual's benefit was reduced, etc. Because source information for the generated variable was collected from wave 2 to wave 4, all 15 positions are coded "I" (i.e., item not asked in wave) for all benefit cuts reported during the first wave and since wave 5 (see below). Each of the 15 positions of this variable, which represent one of a maximum of 15 individuals in the household, is assigned one of the following codes indicating each individual' benefit status. Codes: 1 – the household member's UB II was cut 2 - the household member's UB II was not cut W – don't know K – not specified T – not applicable (filter) F – question mistakenly not asked U – implausible value I – item not recorded in wave.	Information which household member's benefit was cut in the respective benefit cut spell (only surveyed until wave 4).
zensiert	*Spell of UB II: spell ongoing at time of last HH interview (right-censored.), generated* The censoring indicator shows whether a spell was still ongoing at the time of the last household interview. Note: A spell is regarded as censored if one of the following conditions is met: (a) It is a censored spell of a household from one of the previous waves that had not been re-interviewed in the subsequent waves up to the current wave. (b) A household surveyed in wave 6 reports that a spell of UB II is still ongoing on the interview date in wave 7, or an end date is reported that is identical to the interview date in wave 7 and it is confirmed in the follow-up question that the benefit receipt is still currently ongoing. Code -5 was given if the household reference person of the previous wave was no longer living in the household in wave 7 and was not interviewed in wave 7.	*AL20300; AL20400, AL20500* (*alg2_spells*)

Table 19: Simple generated variables for wave 7 in the BIO spell dataset (bio_spells) (in the same order presented in the dataset)

Variable	Variable label and description	Source var. for generated var. in wave 7
bmonat	*Employment: start month, generated* The month during which the employment spell began. To generate the variable information on the season was converted into a month. Note: The generated date variables were checked for plausibility and corrected if necessary. The dates originally reported by the respondent are included in the source variables. Details regarding the season in which the spell began were recoded into months as follows: 21 beginning of year/winter → January 24 spring/Easter → April 27 middle of year/summer → July 30 autumn → October 32 end of year → December	*BIO0200 (bio_spells)*
bjahr	*Employment: start year, generated* The year during which the employment spell began. Note: see *bmonat*	*BIO0300 (bio_spells)*
emonat	*Employment: end month, generated* The month during which the employment spell ended. To generate the variable information on the season was converted into a month and for right-censored spells (i.e., spells that were ongoing when the individual was interviewed), the interview month was entered. Note: see *bmonat*	*BIO0400, BIO0600 (bio_spells)*; *pintmon (PENDDAT)*
ejahr	*Employment: end year, generated* The year during which the employment spell ended. For right-censored spells (i.e., spells that were ongoing when the individual was interviewed), the interview month was entered. Note: see *bmonat*	*BIO0500, BIO0600* (*bio_spells*); *pintjahr* (*PENDDAT*)

zensiert	*Employment: spell still currently ongoing (right censoring)* The censoring indicator shows whether a spell was ongoing at the time of the personal interview in the previous wave, i.e., whether it is a right-censored spell. Note: A spell is considered censored if one of the following conditions is met: the individual reports an end date of the BIO spell that the employment is ongoing on the interview date. Alternatively, when a reported end date is identical to the interview date, the follow-up question confirms that the activity is ongoing.	*BIO0400; BIO0500; BIO0600* (*bio_spells*)
stib	*Occupational status, code number, generated* A detailed code for individual occupational status is generated from the individual variables.	*Collection of spell information in wave 7* *ET0605; ET0705; ET0805; ET0905; ET1005; ET1105; ET1205 (bio_spells*) Otherwise, the value from the previous wave remains
az1	*Weekly contractual working hours*	*Collection of spell information in wave 7* *ET2005 (bio_spells*) Otherwise, the value from the previous wave remains
az2	*Weekly working hours incl. details in the case of irregular working hours, gen.* An integrated variable on weekly hours worked in the position held by the respondent, combining responses to open-ended questions on working hours and a categorical follow-up question. For the closed categories, the follow-up question utilised the mean values for the categories. For the open-ended category, the median of the weekly working hours reported (40 hours or more) was used.	*Collection of spell information in wave 7* *ET2105; ET2205 (bio_spells*) Otherwise, the value from the previous wave remains

alg1bm	*Receipt of UB I: start month, generated* The month during which the spell of Unemployment Benefit I began. To generate this variable, information on the season was converted into a month. Note: Periods during which Unemployment Benefit I is received are embedded in the spells of registered unemployment. An individual can receive a maximum of one period of UB I per period of registered unemployment. The generated date variables were checked for plausibility and corrected if necessary. The dates originally reported by the respondent are included in the source variables. For conversion to months, see *bmonat*.	*AL0800 (bio_spells*)
alg1bj	*Receipt of UB I: start year, generated* The year during which the spell of Unemployment Benefit I began. Note: see *alg1bm*	*AL0900 (bio_spells*)
alg1em	*Receipt of UB I: end month, generated* The month during which the spell of Unemployment Benefit I ended. To generate the variable information, the season was converted into a month. For right-censored spells (i.e., spells that were ongoing at the time of the interview), the interview date was entered. Note: see *alg2kma - alg2kbme*	*AL1000; AL1200 (bio_spells*); *pintmon* (*PENDDAT*)
alg1ej	*Receipt of UB I: end year, generated* The year during which the spell of receiving Unemployment Benefit I ended. In right-censored spells (i.e., spells that were ongoing at the time of the interview), the interview date was entered. Note: see *alg2kma - alg2kbme*	*AL1100; AL1200 (bio_spells*); *pintjahr* (*PENDDAT*)
alg1akt	*Receipt of UB I: spell still currently ongoing (right censoring)* This variable indicates whether the spell of receiving Unemployment Benefit I was ongoing at the time of the personal interview during the previous wave, i.e., whether it is right-censored. Note: A spell is considered censored if one of the following conditions is met: the individual reports an end date for receiving Unemployment Benefit I that indicates that the benefits are ongoing. Alternatively, an end date identical to the interview date is reported. The follow-up question confirms that benefits are ongoing. This variable is generated based on generated date variables, which have been checked for plausibility.	*emonat; ejahr; AL1000;* *AL1100; AL1200 (bio_spells)*

br	*Gross income (incl. categorised info.), gen.* This variable is generated for spells that are ongoing during wave 7 using wave 7 data. For spells that ended or have not been updated in wave 7, information from wave 6 is used to calculate the variable.	*ET2802; ET2902; ET3002; ET3102; ET3202; ET3302* *ET2801; ET2901; ET3001; ET3101; ET3201; ET3301 (bio_spells)*
net	*Net income (incl. categorised info.), gen.* For ongoing spells during wave 7, this variable is generated using wave data. For spells that ended or have not been updated in wave 7, the information from wave 6 is used to calculate the variable.	*ET3402; ET3502; ET3602; ET3702; ET3802; ET3902; ET3401; ET3501; ET3601; ET3400; ET3500; ET3600; ET3700; ET3800; ET3900; ET3701; ET3801; ET3901; (bio_spells)*

Table 20: Wave 7 simple generated variables included in the one - euro spell dataset (ee_spells) (in the same order presented in the dataset)

Variable	Variable label and description	Source var. for generated var. in wave 7
bmonat	*Measure: start month, generated* The month during which the active labour market policy spell began. To generate this variable, information about the season was converted into a month. Note: The generated date variables were checked for plausibility and corrected if necessary. The dates reported by the respondent (excluding identified implausible values) are included in the source variables. Seasons during which the spell began were recoded into months as follows: 21 beginning of year/winter → January 24 spring/Easter → April 27 middle of year/summer → July 30 autumn → October 32 end of year → December	*EE0600a (ee_spells)*
bjahr	*Measure: start year, generated* The year during which the active labour market policy spell began. Note: see *bmonat*	*EE0600b (ee_spells)*
emonat	*Measure: end month, generated* The month during which the active labour market policy ended. To generate the variable, information about the season was converted into a month. For right-censored spells (i.e., spells that were ongoing at the time of the interview), the interview date was entered. Note: see *bmonat*	*EE0600a; EE0600b; EE0700; EE0800a; EE0800b (ee_spells); pintmon, pintjahr (PENDDAT)*
ejahr	*Measure: end year, generated* The year during which the active labour market policy spell ended. For right-censored spells (i.e., spells that were ongoing when the individual was interviewed), the interview date was entered. Note: see *bmonat*	*EE0600a; EE0600b; EE0700; EE0800a; EE0800b (ee_spells); pintjahr; pintjahr (PENDDAT)*
zensiert	*Measure: spell still currently ongoing (right censoring)* The censoring indicator records whether a spell was ongoing at the time of the personal interview during the previous wave, i.e., whether this is a right-censored spell.	*EE0700 (ee_spells)*

Table 21: Wave 7 simple generated variables included in the person register dataset (p_register) (in alphabetical order)

Variable	Variable label and description	Source variable(s) for wave 7 generated variables
alter7	*Age of individual in wave 7 (2013)* A variable contains the best available information about an individual's age. This is either (a) the age calculated from the date of birth reported in wave 7 or (b) the age reported in the household interview if no date of birth is available from wave 7. The information from alter7 is transferred to the household dataset, which corresponds to the information in *HD0200a* to *HD0200o*. This procedure is consistent with conventions in the field. Even during the fieldwork, age was populated using the best available information. During fieldwork, the age variable is first populated using the age information obtained from the household interview. If a personal interview is conducted, this variable is overwritten in the database using the age calculated from the details obtained in the personal interview (date of birth, date of personal interview). The age information provided in the household and individual datasets are based on this variable. The best age information included in the household dataset for wave 7 was considered during the plausibility checks as well as generating the benefit unit and household type.	*PD0100; pintjahr; pintmon; pinttag (PENDDAT); HD0200a* to *HD0200o (HHENDDAT)*
erwprox7	*Employment status according to HH interview in wave 7 (2013)* This variable is transferred unchanged as HD1101* from the current wave from the HHENDDAT dataset.	*HD1101**
kinddat7	*Person included in the KINDER dataset in wave 7 (2013)* *This variable indicates whether an individual is included in the KINDER dataset*.	*pnr (KINDER)*
korrsex	*Info. on sex was corrected between survey waves* For individuals who belonged to a sample HH in more than one wave, this variable indicates whether their sex was corrected in the household interview.	*HD0100a* to *HD0100o* of all waves *(HHENDDAT)*
lastint	*Survey wave of last interview at individual level* This variable indicates the wave in which the last individual interview was conducted (personal or	Personal interviews from all waves *(PENDDAT)*

	senior citizen interview).	
neuj7	*Year in which individual joined current HH, reported in wave 7 (2013)* This variable indicates the year during which an individual joined the current household of which he/she is a member reported during wave 7. Note: The wave 7 interview with the re-interviewed household provides that date when the individual moved or was born into the household since the previous wave.	Information on the date since which an individual has belonged to a household. Surveyed in the household grid
neum7	*Month in which individual joined current HH, reported in wave 7 (2013)* This variable indicates the month that the individual joined the household of which he/she is a current member. Note: see *neuj7*	Date an individual joined a household. Surveyed in the household grid.
wegj7	*Year since which individual has no longer been living in previous HH, reported in wave 7 (2013)* This variable indicates the year that the individual ceased to be a member of the household of the previous wave. Note: Information on the date comes from the wave 7 interview with the household in which the individual was living in the previous wave.	Date an individual ceased to belong to a household. Surveyed in the household grid
wegm7	*Month since which individual has no longer been living in previous HH, reported in wave 7 (2013)* This variable indicates the month that the individual ceased to be a member of the household of the previous wave. Note: see *wegj7*	Date an individual ceased to belong to a household. Surveyed in the household grid
zdub7	*Pointer: Personal identification no. of the individual doubled by the TP in wave 7 (2013)* *Indicates that an individual from an original HH currently lives in a split-off HH without the original HH having reported the move of this individual.* Note: For matchings with the *p_register* via the personal identification number, one must first generate a match variable equalling *zdub**, if it exceeds 0, or otherwise equalling *pnr*. Chapter 5.4.1.2 of the data report for wave 5 of PASS provides a detailed explanation on the reasons for the introduction of this variable.	Information on all original household members of an original household and all of its split-off households are included in the household grid of the current and the previous waves.
zmhh7	*Pointer: Personal ID number of target person's mother in HH in wave 7 (2013)*	Relationships between household members (household grid).

	Contains the personal identification number of the mother if she is living in the household. Biological mothers, stepmothers, adoptive or foster mothers and mothers whose status is not specified are considered mothers.	
zparthh7	*Pointer: personal ID number of target person's partner in HH in wave 7 (2013)* Contains the personal identification number of a partner living in the household. Spouses, registered partners, cohabitees and partners whose status is not specified are considered partners.	Relationships between household members (household grid).
zupanel	*Survey wave in which individual joined panel* This variable indicates the wave in which the individual was a member of a sample household for the first time.	The individuals living in a household across waves (household grid).
zvhh7	*Pointer: Personal ID number of target person's father in HH in wave 7 (2013)* *Contains the personal identification number of the father if he lives in the household. Biological fathers, stepfathers, adoptive or foster fathers and fathers whose status is not specified are considered fathers.*	Relationships between household members (household grid).

The individual-level datasets contain a multitude of generated and constructed variables, including variables (e.g., occupational status) that are recorded in more than one dataset. Figure 3 provides an overview of both the simple and complex generated variables at the individual level.

Figure 3: Overview of generated variables for wave 7 at the individual level

	PENDDAT					BIO-Spells	EE_Spells
	Current status	Employment history		Social origin		Employment and unem-ploy-ment bi-ography	One-euro job participation
		Last employ-ment	First employ-ment	Mother	Father		
Education	berabj						
	beruf1			mberuf1	vberuf1		
	beruf2			mberuf2	vberuf2		
	schulabj						
	schul1			mschul1	vschul1		
	schul2			mschul2	vschul2		
Education classi-fication	casmin			mcasmin	vcasmin		
	isced97			misced97	visced97		
	bilzeit			mbilzeit	vbilzeit		
Information on current status	akt1euro						
	alakt						
	etakt						
	statakt					spelltyp	
Socio-economic position	egp	egplewt	egpeewt	megp	vegp	egp	
	esec	eseclewt	eseceewt	mesec	vesec	esec	
	isei	iseilewt	iseieewt	misei	visei	isei	
	mps	mpslewt	mpseewt	mmps	vmps	mps	
	siops	siopslewt	siopseetw	msiops	vsiops	siops	
Occupational sta-tus	stip	stiblewt	stibeewt	mstib	vstib	stib	
	stibkz						
Date of employ-ment			begmeewt			bmonat	bmonat
			begjeewt			bjahr	bjahr
		emonlewt				emonat	emonat
		ejhrlewt				ejahr	ejahr
Date of unem-ployment						alg1bm	
						alg1bj	
						alg1em	
						alg1ej	
Information on employment	befrist						
	azhpt1					az1	
	azhpt2					az2	
	azges1						
	azges2						
Occupation	isco88	iscolewt	iscoeewt	misco	visco	isco88	
	kldb	kldblewt	kldbeewt	mkldb	vkldb	kldb	
Employed in which industry	branche					branche	
Income	netges						
	brges						
	netto						
	nettokat						
	brutto						
	bruttokat						
Benefit receipt	alg1abez					alg1akt	
	hhalg2						
Houshold context and civil status	hhgr						
	famstand						
	vhh						
	mhh						
	apartner						
	epartner						
	ekind						
	ekin614						
	ekinu15						
	ekinu18						
	ekin1517						
	kindzges						
	kindzihh						
Migration backround	ogebland						
	ostaatan						
	ozulanda						
	ozulandb						
	ozulandc						
	ozulandd						
	ozulande						

	ozulandf						
	migration						
Information on individual	gebhalbj						
	palter						
	zplathh						
	zpsex						
General	altbefr						
	fb_vers						
	panel						
	pintdat						
	RegP0100						
	sample						
Leisure time behaviour	freiz1						
	freiz2						
	freiz3						
	frwunsch						

4.5 Constructed variables

Constructed variables are generated variables that require more extensive coding or recoding. In most cases, these variables have been empirically tested elsewhere and are based on theoretical concepts. At least some of these are standardized instruments used in social sciences or economics, such as the European Socio-economic Classification (ESeC), the International Standard Classification of Education (ISCED) or equivalised household income. This chapter provides detailed descriptions of the constructed variables made available in the PASS data, along with a short overview of the theoretical background and the most important references.

Individual level

Education in years

Variable name	*bilzeit*
Variable label	Duration of school education and vocational training in years, generated
Source variables	*schul2; beruf2*
Type / dataset	Education / individual-level data
Prepared by	Bernhard Christoph
Explanation	For many statistical models, a linear variable for education and training is more appropriate than a categorical variable. For school qualifications, it is easy to convert categorical data to linear data. The linear value simply corresponds to the time spent in school until attainment of the final qualification. Care must be taken to ensure that equivalent qualifications are assigned identical durations. An upper secondary school certificate, for example, should always be labeled with the same duration regardless of whether it was obtained after twelve or thirteen years of education. Final qualifications were assigned the following durations:
	Lower secondary school certificate, lower secondary school certificate from the former GDR (POS) after completion of grade 8: 8 years
	Other degree: 9 years
	Intermediate secondary school certificate;
	intermediate secondary school certificate from the former GDR (POS) after completion of grade 10: 10 years
	Entrance qualification for university for applied sciences: 12 years
	General qualification for university or subject-specific higher education entrance (including EOS—similar qualification in the former GDR): 13 years
	Vocational qualifications differ because of their numerous, different requirements and potentially large differences in income even for qualifications with similar training duration. The training duration may not be subjected to a simple one-to-one conversion process. This problem can be avoided by attempting to operationalise the growth in human capital related to a particular vocational qualification (see e.g., Helberger, 1988).This study adopts a similar approach. Only the respondent's highest vocational qualification was considered, and the years estimated to represent the human capital growth resulting from this qualification were added to the years of education.

Training as a semi-skilled worker:	+1 year
Apprenticeship, vocational school, school for health care occupations:	+1.5 years
Master craftsman certificate:	+3 years
Vocational academy:	+3 years
Applied sciences/Bachelor's degree:	+3 years
University/Master's degree:	+5 years
Ph.D.:	+8 years
Other German qualification:	+1.5 years
Other foreign qualification:	+1.5 years

Literature: Helberger (1988)

Education in years, mother

Variable name	*mbilzeit*
Variable label	Duration of school education and vocational training of mother in years, generated
Source variables	*mschul2; mberuf2*
Category / dataset	Education / individual-level data
Prepared by	Bernhard Christoph
Explanation	General description: see “Education in years”
	When generating the parents’ years of education and training variables, the values added for vocational qualifications differ from those used to construct the corresponding variable for the respondents because information on vocational education/training was collected in less detail for parents (especially for tertiary education). The following values are assigned to particular courses of education/training:
	Training as a semi-skilled worker: +1 year
	Apprenticeship, vocational school,
	Health care occupations: +1.5 years
	Master craftsman certificate: +3 years
	Vocational academy: +3 years
	University, applied sciences: +3 years
	University: +5 years
	Other German qualification: +1.5 years
	Other foreign qualification: +1.5 years
Literature:	Helberger (1988)

Education in years, father

Variable name	*vbilzeit*
Variable label	Duration of school education and vocational training of father in years, generated
Source variables	*vschul2; vberuf2*
Category / dataset	Education / individual-level data
Prepared by	Bernhard Christoph
Explanation	General description: see Education in years (above). When generating the parents' years of education and training variables, the values added for vocational qualifications differ from those used to construct the corresponding variable for the respondents because information on vocational education/training was collected in less detail for parents (especially for tertiary education). The following values are assigned to particular courses of education/training:

Training as a semi-skilled worker:	+1 year
Apprenticeship, vocational school,	
Health care occupations:	+1.5 years
Master craftsman certificate:	+3 years
Vocational academy:	+3 years
University, applied sciences:	+3 years
University:	+5 years
Other German qualification:	+1.5 years
Other foreign qualification:	+1.5 years

Literature:	Helberger (1988)

CASMIN

Variable name	*casmin*
Variable label	Education classified acc. to CASMIN, updated version, generated
Source variables	*schul2; beruf2*
Category / dataset	Education / individual-level data
Prepared by	Bernhard Christoph
Explanation	The CASMIN educational classification was developed within the framework of the CASMIN project (Comparative Analysis of Social Mobility in Industrial Nations) in order to compare academic and vocational qualifications internationally (König, Lüttinger & Müller,. 1987). An updated version is now available (Brauns & Steinmann, 1999). The procedures applied in the panel to recode qualifications according to the CASMIN classification, especially for problematic cases, follow the procedures described in Lechert, Schroedter and Lüttinger (2006) and Granato (2000). The slightly differing category values of the education variable in this dataset are considered. Details are presented in the table below. Cells containing valid CASMIN combinations are highlighted in light gray, whereas those containing missing values are dark grey.

School / Occup.	Not surv.	Pupil	Not asked	NA	No details	Don't know	No qual.	Special needs school	Lower sec. school	Interm. Sec. school	Entrance qual. for uni. of app. Sci.	Upper sec. leaving cert.	Other Ger. qual.	Other foreign qual.
Not surv.	-10	-	-	-	-	-	-	-	-	-	-	-	-	-
Implaus. value	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
Pupil	-	-5	-	-	-	-	-	-	-	-	-	-	-	-
Not asked	-	-	-4	-	-	-	-	-	-	-	-	-	-	-
NA	-	-	-	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
No details	-	-	-	-3	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
Don't know	-	-	-	-3	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1
No qual.	-	-	-	-3	-2	-1	1a	1a	1b	2b	2c_gen	2c_gen	1b	1b
Semi-skilled	-	-	-	-3	-2	-1	1a	1a	1b	2b	2c_gen	2c_gen	1b	1b
Apprenticeship	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
Voc. school	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
Health care school	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
Master craftsman	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
Vocational academy	-	-	-	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a
UAS/ Bachelor's	-	-	-	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a
Uni./Master's	-	-	-	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b
PhD	-	-	-	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b
Other Ger. qual.	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
Other foreign qual.	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c

Literature: Brauns et al. (1999); Granato (2000); König et al. (1987); Lechert et al. (2006)

MCASMIN

Variable name	*mcasmin*
Variable label	Education of mother classified acc. to CASMIN, updated version, generated
Source variables	*mschul2; mberuf2*
Category / dataset	Education / individual-level data
Prepared by	Bernhard Christoph
Explanation	General description: see CASMIN (above). Because the education variable has different category values for respondents and their parents, the coding pattern for *mcasmin* and *vcasmin* differs slightly from the pattern used in *casmin*. The following table details the differences.

School / Occup.	Not surv.	Personal inter-view missing	Parent un-known	Not asked	NA	No de-tails	Don't know	No qual.	Special needs school	Lower sec. school	Interm. Sec. school	En-trance qual. for uni. of app. Sci.	Upper sec. leaving cert.	Other Ger. qual.	Other foreign qual.
Not surv.	-10	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Implaus. value	-	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
Personal inter-view missing	-	-6	-	-	-	-	-	-	-	-	-	-	-	-	-
Parent unknown	-	-	-5	-	-	-	-	-	-	-	-	-	-	-	-
Not asked	-	-	-	-4	-	-	-	-	-	-	-	-	-	-	-
NA	-	-	-	-	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
No details	-	-	-	-	-3	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
Don't know	-	-	-	-	-3	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1
No qual.	-	-	-	-	-3	-2	-1	1a	1a	1b	2b	2c_gen	2c_gen	1b	1b
Semi-skilled	-	-	-	-	-3	-2	-1	1a	1a	1b	2b	2c_gen	2c_gen	1b	1b
Apprentice-ship	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
Master craftsman	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
Vocational acad-emy	-	-	-	-	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a
UAS	-	-	-	-	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a
Uni.	-	-	-	-	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b
Other Ger. qual.	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
Other foreign qual.	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c

Literature: Brauns et al. (1999); Granato (2000); König et al. (1987); Lechert et al. (2006)

VCASMIN

Variable name	*vcasmin*
Variable label	Education of father classified acc. to CASMIN, updated version, generated
Source variables	*vschul2; vberuf2*
Category / dataset	Education / individual-level data
Prepared by	Bernhard Christoph
Explanation	General description: see CASMIN (above). Because the education variable has different category values for respondents and their parents, the coding pattern for *mcasmin* and *vcasmin* differs slightly from the pattern used in *casmin*. The following table details the differences.

School / Occup.	Not surv.	Personal inter-view missing	Parent un-known	Not asked	NA	No de-tails	Don't know	No qual.	Special needs school	Lower sec. school	Interm. Sec. school	En-trance qual. for uni. of app. Sci.	Upper sec. leaving cert.	Other Ger. qual.	Other foreign qual.
Not surv.	-10	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Implaus. value	-	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
Personal inter-view missing	-	-6	-	-	-	-	-	-	-	-	-	-	-	-	-
Parent unknown	-	-	-5	-	-	-	-	-	-	-	-	-	-	-	-
Not asked	-	-	-	-4	-	-	-	-	-	-	-	-	-	-	-
NA	-	-	-	-	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
No details	-	-	-	-	-3	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
Don't know	-	-	-	-	-3	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1
No qual.	-	-	-	-	-3	-2	-1	1a	1a	1b	2b	2c_gen	2c_gen	1b	1b
Semi-skilled	-	-	-	-	-3	-2	-1	1a	1a	1b	2b	2c_gen	2c_gen	1b	1b
Apprenticeship	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
Master craftsman	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
Vocational acad-emy	-	-	-	-	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a
UAS	-	-	-	-	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a
Uni.	-	-	-	-	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b
Other Ger. qual.	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
Other foreign qual.	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c

Literature: Brauns et al. (1999); Granato (2000); König et al. (1987); Lechert et al. (2006)

ISCED 97

Variable name	*isced97*
Variable label	Education classified acc. to isced97, updated version, generated
Source variables	*schul2; beruf2*
Category / dataset	Education / individual-level data
Prepared by	Bernhard Christoph
Explanation	The ISCED-97, (International Standard Classification of Education) developed by the OECD (OECD 1999; for an outline, see also BMBF, 2003), is an education classification alternative to CASMIN. Note that the coding for the ISCED-97 classification includes categories that cannot reasonably be assigned to these data. The ISCED values “0” (pre-primary education/kindergarten) and “1” (primary education) do not apply because the respondents are at least 15 years old. Instead, a separate group was created for individuals with an education below ISCED level 2 (ISCED 2 = lower or intermediate secondary school certificate). Therefore, only ISCED levels 2 to 6 are coded in this dataset.

Coding details are shown in the table below. Cells containing valid combinations according to ISCED are highlighted in light grey, those containing defined missing values are dark grey.

School / **Occup.**	Not surv.	Pupil	Not asked	NA	No details	Don't know	No qual.	Special needs school	Lower sec. school	Interm. Sec. school	Entrance qual. for uni. of app. Sci.	Upper sec. leav-ing cert.	Other Ger. qual.	Other foreign qual.
Not surv.	-10													
Implaus. value					-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
Pupil		-5												
Not asked			-4											
NA				-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
No details				-3	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
Don't know				-3	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1
No qual.				-3	-2	-1	1	1	2	2	3a	3a	2	2
Semi-skilled				-3	-2	-1	2	2	2	2	3a	3a	2	2
Apprenticeship				-3	-2	-1	3b	3b	3b	3b	4a	4a	3b	3b
Voc. school				-3	-2	-1	3b	3b	3b	3b	4a	4a	3b	3b
Health care school				5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
Master craftsman				5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
Vocational acade-my				5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
UAS/Bachelor's				5a	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a
Uni./Master's				5a	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a
PhD				6	6	6	6	6	6	6	6	6	6	6
Other Ger. qual.				-3	-2	-1	2	2	2	2	3a	3a	2	2
Other foreign qual.				-3	-2	-1	2	2	2	2	3a	3a	2	2

Literature: BMBF (2003); OECD (1999)

MISCED 97

Variable name	*misced97*
Variable label	Education of mother classified acc. to isced97, updated version, generated
Source variables	*mschul2; mberuf2*
Category / dataset	Education / individual-level data
Prepared by	Bernhard Christoph
Explanation	For the theoretical background and variable generation details, see ISCED-97. In contrast to the ISCED-97 coding applied to respondent education, it is not possible to generate 6 ISCED levels for parents because data on the corresponding qualifications (i.e., Ph.D. or equivalent) were not collected for parents. Therefore, only ISCED levels 2 to 5 are coded in this dataset. The following table provides the coding details.

Occup. \ School	Not surv.	Personal interview missing	Parent unknown	Not asked	NA	No details	Don't know	No qual.	Special needs school	Lower sec. school	Interm. Sec. school	Entrance qual. for uni. of app. Sci.	Upper sec. leaving cert.	Other Ger. qual.	Other foreign qual.
Not surv.	-10	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Implaus. value	-	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
Personal interview missing	-	-6	-	-	-	-	-	-	-	-	-	-	-	-	-
Parent unknown	-	-	-5	-	-	-	-	-	-	-	-	-	-	-	-
Not asked	-	-	-	-4	-	-	-	-	-	-	-	-	-	-	-
NA	-	-	-	-	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
No details	-	-	-	-	-3	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
Don't know	-	-	-	-	-3	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1
No qual.	-	-	-	-	-3	-2	-1	1	1	2	2	3a	3a	2	2
Semi-skilled	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2
Apprenticeship	-	-	-	-	-3	-2	-1	3b	3b	3b	3b	4a	4a	3b	3b
Master craftsman	-	-	-	-	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
Vocational academy	-	-	-	-	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
UAS	-	-	-	-	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a
Uni.	-	-	-	-	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a
Other Ger. qual.	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2
Other foreign qual.	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2

Literature: BMBF (2003); OECD (1999)

VISCED 97

Variable name	*visced97*
Variable label	Education of father classified acc. to isced97, updated version, generated
Source variables	*vschul2; vberuf2*
Category / dataset	Education / individual-level data
Prepared by	Bernhard Christoph
Explanation	For the theoretical background and variable generation details, see ISCED-97. In contrast to the ISCED-97 coding applied to respondent education, it is not possible to generate 6 ISCED levels for parents because data on the corresponding qualifications (i.e., Ph.D. or equivalent) were not collected for parents. Therefore, only ISCED levels 2 through 5 are coded in this dataset. The following table provides the coding details.

School / **Occup.**	Not surv.	Personal interview missing	Parent unknown	Not asked	NA	No details	Don't know	No qual.	Special needs school	Lower sec. school	Interm. Sec. school	Entrance qual. for uni. of app. Sci.	Upper sec. leaving cert.	Other Ger. qual.	Other foreign qual.
Not surv.	-10	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Implaus. value	-	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
Personal interview missing	-	-6	-	-	-	-	-	-	-	-	-	-	-	-	-
Parent unknown	-	-	-5	-	-	-	-	-	-	-	-	-	-	-	-
Not asked	-	-	-	-4	-	-	-	-	-	-	-	-	-	-	-
NA	-	-	-	-	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
No details	-	-	-	-	-3	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
Don't know	-	-	-	-	-3	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1
No qual.	-	-	-	-	-3	-2	-1	1	1	2	2	3a	3a	2	2
Semi-skilled	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2
Apprenticeship	-	-	-	-	-3	-2	-1	3b	3b	3b	3b	4a	4a	3b	3b
Master craftsman	-	-	-	-	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
Vocational academy	-	-	-	-	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
UAS	-	-	-	-	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a
Uni.	-	-	-	-	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a
Other Ger. qual.	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2
Other foreign qual.	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2

Literature: BMBF (2003); OECD (1999)

International Standard Classification of Occupations 1988 (ISCO-88); ZUMA coding

Generated	Employment	Variable name	Source variables
	current	*isco88*	ET2500
	Spell data (*bio_spells*)	*isco88*	ET2500
	first	*iscoeewt*	ET2500, PET1280, PET3950
	last	*iscolewt*	ET2500, PET1280
	of father	*visco*	PSH0800
	of mother	*misco*	PSH0700
Variable label	Current empl.: ISCO-88 (ZUMA coding), generated		
	Spell data (*bio_spells*): ISCO-88 (ZUMA coding), generated		
	First empl.: ISCO-88 (ZUMA coding), first employment, generated		
	Last empl.: ISCO 88 (ZUMA coding), last employment, generated		
	Father: ISCO-88 (ZUMA coding) of the father, generated		
	Mother: ISCO-88 (ZUMA coding) of the mother, generated		
Category / dataset	Occupation / individual-level data		
Contact person	Bernhard Christoph		
Explanation	The International Standard Classification of Occupations (ISCO) was developed by the International Labour Organization (ILO) to allow international comparison. An advantage of the ISCO-88 is that in addition to the employment, the qualification level generally necessary to perform the job is also considered when assigning an occupation to a particular occupational code. This constitutes a major difference from the Classification of Occupations provided by the German Federal Statistical Office (KldB), which is also provided in this dataset.		
Literature:	ILO (1990)		

Classification of Occupations 1992 (KldB92)

Generated	Employment	Variable name	Source variables
	current	*kldb_it*	ET2500
	Spell data (*bio_spells*)	*kldb*	ET2500
	first	*kldbeewt*	ET2500, PET1280, PET3950
	last	*kldblewt*	ET2500, PET1280
	of father	*vkldb*	PSH0800
	of mother	*mkldb*	PSH0700
Variable label	Current empl.: Classification of Occupations 1992, current employment		
	Spell data (*bio_spells*): Classification of Occupations 1992, generated		
	First empl.: Classification of Occup. 1992, first empl., gen.		
	Last empl.: Classification of Occupations 1992, last empl., gen.		
	Father: Classification of Occupations 1992 of father, generated		
	Mother: Classification of Occupations 1992 of mother, generated		
Category / dataset	Occupation / individual-level data		
Contact person	Bernhard Christoph		
Explanation	The KldB92 is the current version of the Classification of Occupations published by the German Federal Statistical Office (Statistisches Bundesamt). This classification system was developed to match the German occupational structure, which is based solely on employment.		
Literature:	StBA (1992)		
Explanation	The KldB92 is the current version of the Classification of Occupations published by the German Federal Statistical Office (Statistisches Bundesamt). This classification system was developed to match the German occupational structure, which is based solely on employment.		

Erikson, Goldthorpe and Portocarrero (EGP) Class Scheme

Generated	Employment	Variable name	Source variables
	current	*egp*	*isco88, stib*
	Spell data (*bio_spells*)	*egp*	*isco88, stib*
	first	*egpeewt*	*iscoeewt, stibeewt*
	last	*egplewt*	*iscolewt, stiblewt*
	of father	*vegp*	*visco, vstib*
	of mother	*megp*	*misco, mstib*
Variable label	Current empl.: Class scheme acc. to Erikson, Goldthorpe & Portocarrero (EGP), current occupation, generated		
	Spell data (*bio_spells*): Class scheme acc. to Erikson, Goldthorpe & Portocarrero (EGP), gen.		
	First empl.: Class scheme acc. to Erikson, Goldthorpe & Portocarrero (EGP), first employment, gen.		
	Last empl.: Class scheme acc. to Erikson, Goldthorpe & Portocarrero (EGP), last employment, gen.		
	Father: Class scheme acc. to Erikson, Goldthorpe & Portocarrero (EGP), occupation of father, gen.		
	Mother: Class scheme acc. to Erikson, Goldthorpe & Portocarrero (EGP), occupation of mother, gen.		
Category / dataset	socio-economic position / individual-level data		
Prepared by	Bernhard Christoph		
Explanation	The class scheme developed by Erikson, Goldthorpe and Portocarrero (Erikson et al., 1979, 1982; Erikson & Goldthorpe, 1992) is among the most common instruments for operationalising class.		
	For this variable, data are coded by ISCO-88 occupational classification and occupational status. The coding procedure is based on an earlier approach elaborated by Christoph et al. (2005), who provide a detailed description of the procedure. Here, in contrast, unpaid family workers were not coded as self-employed but as individuals in dependent employment consistent with the coding applied in the European Socio-Economic Classification (ESeC), which is described in the next section.		
	One difference between the EGP coding applied here and the ESeC coding is that in the EGP coding procedure, cases are “missing” (-7) in which the occupational activity seemed incompatible with occupational status (e.g., “directors and chief executives” [ISCO=1210] who reported that they were “employees performing simple duties” [StiB=51]). To ensure compatibility with the standardised coding procedure we adopted, we did not apply a comparable revision procedure using the EseC codes.		
Literature:	Christoph (2005); Erikson and Goldthorpe (1992); Erikson et al. (1982); Erikson et al. (1979).		

European Socio-economic Classification (ESeC)

Generated	Employment	Variable name	Source variables
	current	*esec*	*isco88, stib, PET2000, PET2700*
	Spell data (*bio_spells*)	*esec*	*isco88, stib, ET1100, ET1101, ET1102, ET1300, ET1301, ET1302,*
	first	*eseceewt*	*iscoeewt, stibeewt, PET1261*
	last	*eseclewt*	*iscolewt, stiblewt, PET3801*
	of father	*vesec*	*visco, vstib, PSH0670*
	of mother	*mesec*	*misco, mstib, PSH0370*
Variable label	Current empl.: European Socio-economic Classification (ESeC), current occupation, gen.		
	Spell data (*bio_spells*): European Socio-economic Classification (ESeC), gen.		
	First empl.: European Socio-economic Classification (ESeC), first employment, gen.		
	Last empl.: European Socio-economic Classification (ESeC), last employment, gen.		
	Father: European Socio-economic Classification (ESeC), occupation of father, gen.		
	Mother: European Socio-economic Classification (ESeC), occupation of mother, gen.		
Category / dataset	socio-economic position / individual-level data		
Prepared by	Bernhard Christoph		
Explanation	The European Socio-economic Classification is largely based on the EGP class scheme. Unlike the latter, great importance was attached to international comparability of the operationalisation and validation of the classification (for a general description, see Rose & Harrison, 2007; for Germany, see Müller et al. 2006, 2007).		
	The Stata do-file required to generate the ESeC was kindly provided by Heike Wirth from GESIS-ZUMA (Fischer & Wirth 2007). We simply adjusted the file to meet the requirements of this study. This do-file, originally written in standard SPSS syntax by Harrison and Rose (2006) as a standard program to generate the ESeC, was converted into Stata.		
Literature:	Fischer and Wirth (2007); Harrison Rose (2006); Müller et al. (2006, 2007); Rose and Harrison (2007)		

Magnitude Prestige Scale (MPS)

Generated	Employment	Variable name	Source variables
	current	*mps*	*isco88*
	Spell data (*bio_spells*)	*mps*	*isco88*
	first	*mpseewt*	*iscoeewt*
	last	*mpslewt*	*iscolewt*
	of father	*vmps*	*visco*
	of mother	*mmps*	*misco*
Variable label	Current empl.: Magnitude Prestige Scale, current occupation, gen.		
	Spell data (*bio_spells*): Magnitude Prestige Scale, generated		
	First empl.: Magnitude Prestige Scale, first employment, gen.		
	Last empl.: Magnitude Prestige Scale, last employment, gen.		
	Father: Magnitude Prestige Scale, occupation of father, gen.		
	Mother: Magnitude Prestige Scale, occupation of mother, gen.		
Category / dataset	socio-economic position / individual-level data		
Contact person	Bernhard Christoph		
Explanation	The MPS (Wegener, 1985, 1988) is the only Germany-specific instrument available to operationalize social prestige based on detailed occupation information. The scale was originally developed for the 1968 version of the International Standard Classification of Occupations (ISCO-68). Because occupation codes in this study were based on the more recent ISCO-88 classification and the Classification of Occupations (KldB) developed by the Federal Statistical Office, a variant of the scale adapted to the ISCO-88 was used (Christoph 2005). Infas merged the data as part of the occupational coding procedure.		
Literature:	Christoph (2005); Wegener (1985, 1988)		

Standard International Occupational Prestige Scale (SIOPS/Treiman Scale)

Generated	Employment	Variable name	Source variables
	current	*siops*	*isco88*
	Spell data (*bio_spells*)	*siops*	*isco88*
	first	*siopseewt*	*iscoeewt*
	last	*siopslewt*	*iscolewt*
	of father	*vsiops*	*visco*
	of mother	*msiops*	*misco*
Variable label	Current empl.: Standard International Occupational Prestige Scale, current occupation, gen.		
	Spell data (*bio_spells*): Standard International Occupational Prestige Scale, generated		
	First empl.: Standard International Occupational Prestige Scale, first employment, gen.		
	Last empl.: Standard International Occupational Prestige Scale, last employment, gen.		
	Father: Standard International Occupational Prestige Scale, occupation of father, gen.		
	Mother: Standard International Occupational Prestige Scale, occupation of mother, gen.		
Category / dataset	socio-economic position / individual-level data		
Contact person	Bernhard Christoph		
Explanation	The Treiman Prestige Scale, which was originally constructed by Treiman (1977) for ISCO-68, is the first and only prestige scale available for international comparative research on occupations. Since its adaptation to the ISCO-88 (Ganzeboom & Treiman, 1996, 2003), the scale has commonly been called the “Standard International Occupational Prestige Scale”. Infas merged the data as part of the occupational coding procedure.		
Literature:	Ganzeboom and Treiman (1996, 2003); Treiman (1977)		

International Socio-Economic Index (ISEI)

Generated	Employment	Variable name	Source variables
	current	*isei*	*isco88*
	Spell data (*bio_spells*)	*isei*	*isco88*
	first	*iseieewt*	*iscoeewt*
	last	*iseilewt*	*iscolewt*
	of father	*visei*	*visco*
	of mother	*misei*	*misco*
Variable label	Current empl.: International Socio-Economic Index, current employment, gen.		
	Spell data (*bio_spells*): International Socio-Economic Index, generated		
	First empl.: International Socio-Economic Index, first employment, gen.		
	Last empl.: International Socio-Economic Index, last employment, gen.		
	Father: International Socio-Economic Index, occupation of father, gen.		
	Mother: International Socio-Economic Index, occupation of mother, gen.		
Category / dataset	socio-economic position / individual-level data		
Contact person	Bernhard Christoph		
Explanation	The ISEI is among the most common indices of this kind, in part, due to the fact that, unlike most other SEIs, the ISEI is based on an original theoretical concept that considers the occupation and its socio-economic status as an intervening variable in the relationship between education and income. The ISEI was developed for the ISCO-68 (Ganzeboom, De Graaf & Treiman, 1992); it was later adapted to the ISCO-88 (Ganzeboom & Treiman, 1996, 2003). Infas merged the data as part of the occupational coding procedure.		
Literature:	Ganzeboom et al. (1992); Ganzeboom and Treiman (1996, 2003)		

Classification of Economic Activities 2003 (Klassifikation der Wirtschaftszweige 2003 (WZ2003))

Generated	Employment	Variable name	Source variables
	current	*branche*	ET2600
	Spell data (*bio_spells*)	*branche*	ET2600
Variable label	Current empl.: Current activity: economic sector/industry (WZ2003)		
	Spell data (*bio_spells*): economic sector/industry (WZ2003), generated		
Category / dataset	socio-economic position / individual-level data		
Contact person	Bernhard Christoph		
Explanation	The information obtained from the open-ended survey question about the sector/industry in which the respondent is employed was coded using the 2-digit Classification of Economic Activities of the Federal Statistical Office (WZ2003) code. At the two-digit level, this classification largely corresponds to the European Nomenclature générale des Activités économiques dans les Communautés Européennes (NACE) in revision 1.1.		
Literature:	StaBA (2002); EG (2002)		

Leisure activities pursued and desired by young people

Variable name	*freiz1, freiz2, freiz3, frwunsch*
Variable label	freiz1: leisure time activity 1, pursued freiz2: leisure time activity 2, pursued freiz3: leisure time activity 3, pursued frwunsch: leisure time activity, desired
Source variables	*PA1100 (for freiz1-freiz3); PA1200 (for frwunsch)*
Category / dataset	leisure time / individual-level data
Prepared by	Johanna Eckert (DJI), Arne Bethmann, Claudia Wenzig

Explanation

Explanation:

The variables freiz1, freiz2, freiz3 and frwunsch are based on newly developed categories for youth leisure activities. This scheme originates in the three most popular (PA1100) and desired (PA1200) leisure activities obtained through open-ended questions. The most popular leisure activities were converted into three individual variables according to the question text. Only one desired leisure activity was considered. Additional responses were not included in the coding.

The scheme was developed inductively based on corrected information. To achieve comparability among waves, the new scheme includes all leisure activities that were asked in restricted questions during previous waves. Furthermore, the scheme is designed to allow expansion, if necessary, over subsequent waves with new (sub)categories.

The scheme includes not only 16 main categories but also categories for no leisure activities and information that could not be assigned. The ranking of the 16 main categories results from the frequency with which they were mentioned. The main categories can be differentiated into 77 subcategories.

	Main category / variable characteristic	**Number of subcategories**
1000	Sports and exercise	31
2000	Spending time with family and friends	4
3000	Computer, games and communication	5
4000	Making / listening to music	6
5000	Reading	-
6000	Culture, cinema, television and events	8
7000	Creative hobbies, crafts, cooking and baking	11
8000	Going out, partying, nightlife	3
9000	Hanging out, relaxing	-
10000	Shopping	-
11000	Traveling, trips, tours and being mobile	3
12000	Spending time with pets	-
13000	Volunteer work	4
14000	Learning and education	-
15000	Games and mental exercise	2
16000	Side job	-
99998	No leisure activity	-
99999	Information cannot be assigned	-

Literature:

Johanna Eckert, Arne Bethmann, Claudia Wenzig (planned): Manual coding “Pursued and desired leisure time activities by young people”. PASS wave 5 (2011).

Household or benefit unit level

Equivalised household income, previous OECD weighting

Variable name	*oecdinca*
Variable label	equivalised household income, old OECD weighting (rounded)
Source variables	*HD0200a-HD0200o; HA0100; hhincome*
Category / dataset	socio-economic position / household-level data
Prepared by	Bernhard Christoph
Explanation	Equivalised household income considers the savings achievable through joint housekeeping in multi-individual households compared to single households. The per-capita income of the household is not divided by the actual number of individuals but by a divisor, which is usually less than this figure, and is calculated based on the assumed needs of household members (equivalised household size). According to the previous OECD scale, only the first household member (15 or older) is assigned a weighting factor of 1.0. Household members at least 15 years of age are assigned a weighting factor of 0.7, and children up to age 14 are assigned a weighting factor of 0.5 to calculate equivalised household size.
Literature:	Hauser (1996); OECD (1982)

Equivalised household income, modified OECD weighting

Variable name	*oecdincn*
Variable label	equivalised household income, modified OECD weighting (rounded) .
Source variables	*HD0200a-HD0200o; HA0100; hhincome*
Category / dataset	socio-economic position / household-level data
Prepared by	Bernhard Christoph
Explanation	General description: see “Equivalised household income, previous OECD weighting (above). The modified OECD equivalence scale assumes a weighting factor of 1.0 only for the first household member (15 or older). Household members at least 15 years old are assigned a weighting factor of 0.5, and children up to age 14 are assigned a weighting factor of 0.3 to calculate household size. For more information on the modified OECD scale, see Hagenaars, de Vos, and Zaidi (1994).
Literature:	Hagenaars et al. (1994)

Deprivation index, unweighted

Variable name	*depindug2*
Variable label	All waves: deprivation index, unweighted (item total: 23) .
Source variables	*HLS0100a-HLS0400a; HLS0100b-HLS0400b; HLS0600a-HLS1200a; HLS0600b-HLS1200b; HLS1400a-HLS2500a; HLS1400b-HLS2500b;*
Category / dataset	material situation / household-level data
Prepared by	Bernhard Christoph
Explanation	Following Ringen (1988), poverty researchers usually distinguish between direct and indirect measures of poverty. Indirect measurement focuses on the resources available to attain a particular standard of living, especially (equivalised household) income. This method is also called the resource-based approach to measuring poverty. In contrast, direct measurement attempts to record the household's ownership of goods and to determine the extent to which the households cannot afford certain goods or activities that are considered relevant. This method is also called the deprivation approach (see, e.g., Halleröd 1995). Previous scientific research suggests that the population classified as poor by the resource-based approach is not always identical to that identified by the deprivation approach. To define with precision who is to be considered poor, combining measures of resource poverty and deprivation is often been suggested—i.e., to classify as poor only those individuals identified by both approaches (see Halleröd 1995; Nolan & Whelan 1996; Andreß & Lipsmeier 2001). The deprivation index is based on a list of 23 goods or activities. The surveyed households are asked to indicate whether they possessed these goods or participated in the activities mentioned. The unweighted index simply adds the number of items that respondents indicated they did not possess or in which they did not participate. However, only items that are missing for financial reasons are counted to prevent consumer preferences (e.g., a household choosing not to own a car or television) from being misinterpreted as a reduced standard of living. Additionally, an item was only accepted as missing for financial reasons if explicitly confirmed in the answers to both questions. "Don't know" or "details refused" answers were considered available goods or missing for a non-financial reason. This assumption does not apply to all cases. Alternatively, an index value for households that failed to answer a question for (at least) one particular good could be excluded (through listwise deletion). Of the 23 goods and activities surveyed, however, this method would quickly lead to a large number of missing index values. Therefore, the first method described was selected. Nevertheless, compared to the listwise deletion procedure, there is a risk that the number of goods missing for financial reasons is underestimated by this method. For waves 1 through 4, the variable *depindug* provides a version of the unweighted deprivation index based on 26 items, i.e., adding to the items mentioned above HLS0500*, HLS1300* and HLS2600*. These three items have not been asked since wave 5. Thus, *depindug2* was newly integrated into the dataset and has been generated retroactively since wave 1.
Literature:	Andreß and Lipsmeier (2001); Halleröd (1995); Nolan and Whelan (1996); Ringen (1988)

Deprivation Index, weighted

Variable name	*depindg2*
Variable label	Deprivation index, weighted (items not missing for financial reasons; total of weighted items: 13,14)
Source variables	*HLS0100a-HLS0400a; HLS0100b-HLS0400b; HLS0600a-HLS1200a; HLS0600b-HLS1200b; HLS1400a-HLS2500a; HLS1400b-HLS2500b; PLS0100-PLS0400; PLS0600-PLS1200; PLS1400-PLS2500;*
Category / dataset	All waves: Deprivation Index, weighted (item total: 11.08)
Prepared by	Bernhard Christoph
Explanation	For a general description: see deprivation index, unweighted (above). Unweighted indices, such as the one described above, are often criticised for assigning all items included identical weightings. For example, the difference in asking whether a dwelling has an indoor toilet or whether there is a VCR/DVD player in the household immediately reveals the vast difference in the reduction of household's standard of living caused by the lack of an item. It therefore seems reasonable to weight the items. However, empirical research indicates that in most cases, weighted and unweighted index variants do not yield significantly different results (see Lipsmeier, 1999). For this survey, we weighted items according to the proportion of respondents who considered a particular item as necessary. We selected this procedure not only because it is conceptually convincing and commonly used (applied by Halleröd 1995, for example) but also because it can be implemented without unreasonable costs. The deprivation weightings determined for the individual questionnaire items are assumed highly stable over time, and these items only need to be administered once or in long intervals. Moreover, the large PASS sample allowed us to split the sample into several randomly selected subsamples, each of which classified only some items. Alternative weighting methods, such as restricting the indices to items that are considered necessary by a minimum proportion of the respondents (e.g., Andreß & Lipsmeier 1995, Andreß et al. 1996) or theoretically restricting the indices to a few fundamental items (e.g., Nolan & Whelan 1996), were not utilised in this survey but can be generated, if necessary, from the data provided. A discussion of the different methods of index weighting can be found in Andreß and Lipsmeier (2001, esp. p. 28 ff.). For waves 1 through 4, the variable *depindg* provides a version of the weighted deprivation index based on 26 rather than 23 items, i.e., in addition to the items mentioned above, it includes the following items: HLS0500*; HLS1300* and HLS2600*; and PLS0500, PLS1300 and PLS2600. These three HLS items have not been asked since wave 5. Thus, *depindg2* is newly integrated into the dataset and has been generated retroactively since wave 1.
Literature:	Andreß and Lipsmeier (1995, 2001); Andreß et al. (1996); Halleröd (1995); Lipsmeier (1999); Nolan and Whelan (1996)

Household typology

Variable name	*hhtyp*
Variable label	Household type, generated
Source variables	Household information on age and relationships between household members.
Category / dataset	Household structure / household data
Prepared by	Daniel Gebhardt
Explanation	Various household typologies exist (see, e.g., Lengerer, Bohr & Jansen, 2005 for the Microcensus household typology; Porst (1984) and Beckmann & Trometer 1991 for the ALLBUS typology; and Frick, Göbel & Krause (n.d.) for the SOEP). The household typology used in PASS follows the latter typology. The decisive differentiation criteria are existing partnerships, number and age of children and existing generational relationships. Whereas the SOEP typology is based on the relationship of the household members to the head of the household, PASS uses information on the relationships among all household members. The PASS typology includes the ages of household members as indicated in the household interview and household size.

Definition of relationships for generating the household type:

- Couples: married couples, registered partnerships, nonmarried partnerships and partnerships whose status is not specified (missing value for the follow-up question about the type of partnership).
- Child of an individual: biological child, stepchild, adopted/foster child or child whose status is not specified (missing value for the follow-up question about type of relationship to the child).
- Parent of an individual: biological parent, stepparent, adoptive/foster parent or parent whose status is not specified (missing value in follow-up question about type of parenthood).

Definition of household type:

- One-person household: A household consisting of only one individual.
- Couple without children: A household consisting of two individuals living as a couple.
- One-parent household: A household consisting solely of one parent and his/her children. No restrictions apply to children's ages.
- Couple with children under the age of 16: A household consisting of two individuals living as a couple and their respective and/or mutual children. All of the children are younger than 16.
- Couple with children aged 16 or over: A household consisting of two individuals living as a couple and their respective and/or mutual children. All of the children are aged 16 or over.
- Couple with children both under and over 16: A household consisting of two individuals living as a couple and their respective and/or mutual children. Some children living in the household are younger than 16 and others are older than 16.
- Multigeneration household: A household consisting of members of at least three generations in linear succession. The core of the household is multigenerational, i.e., at least one individual in the household is both a child and a parent of another member of the household. Other people living in the household include parents, children, siblings, the central member's partner or a partner's siblings.
- Other household: A household that could not be assigned to another household type.
- Generation not possible (missing values): All households with at least one missing value (-1, -2, -4) or implausible value (-8) in the main category of a relationship or age variable (except for households with three or fewer members in unambiguous relationship constellations for which the household type was generated even if ages were missing).

Literature: Beckmann and Trometer (1991); Frick et al. (n.d.); Lengerer et al. (2005); Porst (1984)

Wave 7 benefit unit ID

Variable name	*bgnr7*
Variable label	Benefit unit ID in wave 7
Source variables	Household information on age and relationships between household members
Category / dataset	Benefit unit / person register
Prepared by	Gerrit Müller
Explanation	The *bgnr7* variable is created at the individual level. It assigns an identification number to each household member that indicates the individual's relationship to a particular benefit unit. Consequently, household members with the same identification number constitute a benefit unit. The *bgnr7* variable is composed of the known household number and a two-digit indicator to identify the benefit unit within the household. The identification of a household member's relationship to a benefit unit is based solely on information about the relationships between household members from the household grid along with the ages obtained from the household interview. Therefore, the benefit units identified in this way are considered synthetic benefit units. The identification process does not consider information about actual benefits received, individual members' ability to work or qualification status, but it does identify groups of individuals in the same household who are or would be considered benefit units in jointly receiving benefits according to the provisions of Book II of the German Social Code in the event that such benefits are needed. This artificial allocation procedure is necessary because information about the existence of a benefit unit and the identification of individuals affiliated with that unit cannot be collected directly in the context of an interview. The allocation of an individual to a benefit unit is based on the latest version of the German Social Code, Book II, Section 7, Subsection 3 (last amended on 21 March 2013). Each individual ages 25- 65 constitutes a separate benefit unit unless he or she is living in a partnership and/or has a child/children younger than 25 who has/have no partner/children of their own. In the latter case, the benefit unit consists of the individual, his/her partner and child(ren). If two individuals live in the same household with a mutual child but do not indicate that they are living in a partnership, a partnership is nevertheless assumed to exist according to Section 7, Subsection 3a. The corresponding individuals and their child(ren) are assigned to the same benefit unit. Individuals who are between the ages of 15 and 25 are generally assigned to their parents unless they are already living with a partner (or a child of their own) in a joint household. Individuals between the ages of 15 and 25 who live without their parents, partner or children constitute a separate benefit unit. Individuals older than 65 are not covered by Book II of the German Social Code and are therefore not considered members of a benefit unit (coded 0) unless they live with a partner who is under 65 (or a child under 25). Likewise, children who have not reached age 15 who live in a household without their parents are not considered members of a benefit unit (code 0) because they are covered by the provisions of German Social Code Book XII. Benefit units were not assigned to households with missing information on relationships or the age of certain household members. Instead, all members of these households were assigned code 99. By approximation, such households are interpreted as households consisting of only one benefit unit.
Literature:	German Social Code Book II – basic security for job-seekers (Sozialgesetzbuch, Zweites Buch - Grundsicherung für Arbeitssuchende (SGB II))

Wave 7 benefit unit typology

Variable name	*bgtyp7*
Variable label	Type of benefit unit in wave 7
Source variables	Household information on age and relationships between household members.
Category / dataset	Benefit unit / person register
Prepared by	Gerrit Müller
Explanation	The benefit unit typology is based on the same concept as the synthetic benefit unit used for variable *bgnr7*. Until age 25, children are considered members of their parents' benefit unit unless they themselves have a partner or child. BA statistics typologies are often still established based on reaching legal age (the 18th birthday). For example, according to our typology, households in which the youngest child is between 18 and 24 years old and that are classified as one-parent benefit units are considered single households in BA statistics. This difference must be noted when comparing PASS data with figures from the official statistics. Code 0, no benefit unit, was assigned to households in which one or more member(s) were not covered by Social Code Book II (see also code 0 for *bgnr7*). Code 5, generation impossible (missing values), was assigned to households with missing information on relationships or the ages of individual household members (see code 99 for *bgnr7*).
Literature:	–

Benefit unit receiving Unemployment Benefit II on the wave 7 sampling date

Variable name	*bgbezs7*
Variable label	Benefit unit in receipt of UB II on the sampling date in wave 7 (2013)
Source variables	*HA0250*, HA0300, AL20100, AL20200, AL20300, AL20400, AL20606, AL20706*, HA0400, sample, hnr, bgnr7, hhgr*
Category / dataset	Benefit unit / person register
Prepared by	Mark Trappmann
Explanation	For each benefit unit that was identified according to the procedure described for variable *bgnr7*, this variable indicates whether the benefit unit was actually receiving Unemployment Benefit II on the sampling date of wave 7.
Literature:	–

Benefit unit receiving Unemployment Benefit II on the wave 7 survey date

Variable name	*bgbezb7*
Variable label	Benefit unit in receipt of UB II on the survey date in wave 7 (2013)
Source variables	*AL20606, AL20706a-o, zensiert* (*alg2_spells*), *sample, hhgr, bgnr7*
Category / dataset	Benefit unit / person register
Prepared by	Daniel Gebhardt
Explanation	For each benefit unit that was identified according to the procedure described for variable *bgnr7,* this variable indicates whether the benefit unit was actually receiving Unemployment Benefit II on the wave 7 survey date.
Literature:	–

Number of benefit units within the household

Variable name	*anzbg*
Variable label	Number of synthetic benefit units in the HH, generated
Source variables	*bgnr7, hnr*
Category / dataset	Benefit unit / household dataset
Prepared by	Daniel Gebhardt
Explanation	This variable indicates the number of benefit units existing in the household. The benefit units were identified according to the procedure to generate the variable *bgnr7*.
Literature:	–

Number of benefit units in the household receiving benefits on the sampling date

Variable name	*nbgbezug*
Variable label	Number of benefit units in the HH receiving benefits on the sampling date
Source variables	*bgbezs7, bgnr7, hnr*
Category / dataset	Benefit unit / household dataset
Prepared by	Daniel Gebhardt
Explanation	This variable indicates the number of benefit units within a household that were receiving benefits according to Social Code Book II on the sampling date. The value was calculated via the household number by aggregating the benefit units within a household that were actually receiving benefits according to variable *bgbezs7* from the person register.
Literature:	–

5 Data preparation

Since wave 3, infas, not the IAB, has been responsible for preparing the data. To guarantee consistent data preparation in the longitudinal section, infas was provided with the relevant syntax files for data preparation from wave 2, necessary sources, intermediary datasets and documentation of individual operations. Important decisions, such as the correction of structural problems in participating households or the development of the *bio_spells* dataset, which was first developed in wave 4, were made with the IAB. The IAB was also available for questions during data preparation.

The information gathered in the wave 7 interviews is available from infas as ASCII data. First, infas prepared the following datasets from the raw data[31]:

- Household dataset for the cross-section, including the spell-reshaped questions for the modules “childcare”, “social participation” and “educational package”
- Household dataset for the longitudinal section (module “Unemployment Benefit II”)
- Dataset updating household composition (matrix)
- Dataset updating family relationships in the household (relationship matrix)
- Individual/senior citizen dataset for the cross-section
- Individual dataset for longitudinal section I (module “employment biography [spells]”)
- Individual dataset for longitudinal section II (module “measures”)
- Dataset for open texts (across household, personal and senior citizen interviews)

Second, a more detailed, formal and content-oriented verification of the data was performed. These data were then prepared as the scientific use file. Furthermore, infas provides a gross dataset along with special datasets that are not derived directly from the actual survey instruments.

The data checks conducted at infas can be divided into three steps, which are detailed in the following sections. First, the household structure of the re-interviewed households was reviewed and when necessary, corrected. If serious problems were identified in the structure, the corresponding interviews were removed (see Chapter 5.1 on this issue). This step was followed by a detailed review of the filter questions (applying corrections if necessary). Filter errors were marked and specific codes were set for missing values (see Chapter 5.2 on this issue). Next, selected items were verified for plausibility. Clearly implausible or contradictory responses were marked by a specific missing code. However, such data corrections were limited.

The following table reviews the steps of the data preparation:

[31] The software packages Stata (versions 11 and 12) and PASW (version 18) were used for data preparation.

Table 22: Overview of the steps to prepare the wave 7 PASS data

	Procedure
1	Import the raw data into working datasets
2	Check the household structure (see Chapter 5.1)
3	Remove problematic interviews (household and/or individual levels) (see Chapter 5.1)
4	Integrate individual and senior citizen datasets
5	Correct the household structure of re-interviewed households (see Chapter 5.1)
6	Filter checks at the household level (see Chapter 5.2)
7	Construct a household grid dataset and perform plausibility checks (see Chapter 5.3)
8	Generate synthetic benefit units (see description of variables, Chapter 4.5)
9	Generate new control variables based on the household data after filter checks, household grid dataset and plausibility checks
10	Filter checks at the individual level (see Chapter 5.2)
11	Code information from open-ended survey questions (see Chapter 4.1)
12	Plausibility checks of household and individual-level data (excluding spell data) (see Chapter 5.3)
13	Prepare, plausibility check and construct spell datasets (see Chapters 5.6 to 5.8 and Chapter 5.3)
14	Simple generated variables (see Chapter 4.4)
15	Complex generated variables (see Chapter 4.5)
16	Generation of the data structure for the scientific use file (household, individual and register datasets)
17	Anonymisation (see Chapter 5.5)

5.1 Structure checks and removing interviews

A structure check was conducted before the filter checks. Here, interviews that were not considered successful were to be identified and if necessary, removed from the datasets. In addition, the structure of re-interviewed households was compared with the structure reported during the previous wave to identify and if necessary, to correct implausible or problematic changes in household composition and errors in the allocation of the personal interviews to their respective positions in the household. To observe households in the longitudinal section, it is essential that the individuals be assigned consistently to their position in the household and the respondents can be identified clearly across waves. A personal identification number must not be assigned to different individuals in different waves. If the correct household composition was unclear, all of the interviews conducted with this household in wave 7 were removed from the dataset. If a personal interview was conducted with the wrong individual without further problems in household composition, then only the personal interview was removed.

Different processes identified problematic cases. The relevant cases were discussed as part of a formal procedure between infas and the IAB. The final decision on how to proceed with these cases was made by the IAB. The following specifies the extent of the checks conducted. Not every check in every wave identifies problems. The result of a check is usually that an issue occurs in few cases. Furthermore, known error sources are

absorbed during the interviews. For example, the intention of the survey instrument is that not all known target persons can move out of a panel household at the same time and that at least one remaining individual is at least 15 years old.

- By comparing the first names reported in the current and previous waves, changes in household composition that had not been recorded correctly were identified. Instead of recording moves into and out of a household in the relevant places during the household interview, interviewers sometimes renamed household members or changed their age or sex. All cases in which a first name had been changed that could not be attributed to correcting the spelling and for which the year of birth reported in the previous wave differed by more than one year from that reported in the current wave were reviewed individually. A decision was made as to whether the interviewer made a simple change requiring correction of the first name, age or sex or an inadmissible change to the household structure.
- Furthermore, whether more than one individual with the same date of birth was living in the household was reviewed. Whether these cases were plausible was decided in the context of the household, using two waves. The remaining cases then underwent another review. Households in which a date of birth was reported in the current and previous waves by individuals in different positions in the household structure were identified. Here, it seemed reasonable to suspect that a different individual provided the personal interview in the current wave. In the context of the household and individual-level data of the current and previous wave, individual decisions were made for each household and personal interview.
- In general, the date of birth from the personal/senior citizen interview of the current wave displaces all other age information on that individual, e.g., from the household grid, and is the basis for all generated variables utilising age. The date of birth is corrected in *PD0100*. If an individual's year of birth changes significantly according to *PD0100* but the day and month stay the same, the previously known date of birth has never changed according to *PD0100*, and at least two pieces of information about the date of birth from *PD0100* are available from previous waves, then the year of birth is reset to the value from the previous waves considering the whole household. Consider a hypothetical individual whose date of birth is recorded as February 1, 1972 in at least two previous waves and whose date of birth is now recorded as February 1, 1992. This date of birth would make this individual younger than the other children in the household. Without a correction, such an arrangement leads to an implausible relationship structure, which would consequently mean that synthetic benefit units could not be generated. Hence, in the example above, the date is corrected to February 1, 1972 in the current wave.
- To identify households that are considered not successfully surveyed, the datasets at the household and individual level are merged. Personal interviews without a full

household interview and household interviews for which no individual interview was available were marked.[32]

- Moves into and out of a household are another important factor. Panel households with reported move-outs were generally inspected and correlated with the split-off households. Evaluations were made as to whether the remaining household of the panel household is plausible. Interviews from panel households in which all household members leave except individual children under 15 years old were discarded for the panel and split-off households. If more than one individual moved, whether these individuals formed a joint split-off or several different households was considered and whether this is plausible was determined. For instance, cases in which one partner left the panel household with young children but the children formed several split-off households were considered implausible. In cases of a non-realised split-off household, move-outs were considered plausible, but all individuals who moved out were remerged into one joint split-off household.
- Individual cases occurred in which the panel household indicates that individuals formed a split-off household, but all members could be identified in the split-off household. Alternatively, not all members of the panel household live in the split-off household, and at least one member of the panel household was not reported as having moved out or moved to a split-off household other than the one observed. Decisions were made as to which reported move-outs were considered valid and which were discarded as implausible. If a reported move-out was retroactively discarded as implausible, the individual who had allegedly moved out was retroactively re-integrated into the household panel.
- In split-off households, individuals who are not known from the panel household but who join PASS through the split-off household might still originate from the panel household. Two situations promote these cases. The first situation arises when a panel household reports several individuals moving out and the split-off individuals formed more than one household. In that case, a dynamic preload is created for the current file for all split-off households identified through the panel household. If, however, individuals who, according to the panel household, live in various split-off households are actually sharing a split-off household, those individuals who were not assigned to this split-off household by the panel household but to another split-off household do not have a preload and are included as new individuals.
- It is possible that individuals from a panel household move out of or into a household that was formed as split-off household during a previous wave and that was successfully surveyed at that time. Thus, there is another move from the original panel household into this split-off household after the separation of the split-off household. Regardless of whether the panel household from which the split-off household emerged was successfully surveyed during the wave of the move, such cases cannot

[32] New sample households for which a household interview but no valid personal interview was available were removed from the dataset following the procedure used in wave 1. In contrast, the household interviews of re-interviewed households and split-off households were retained.

be controlled in the field. To do so, the split-off household would have to be provided with the personal information of all individuals from the panel household (and possibly all individuals in other split-offs from this panel household) as a preload. The few cases in which such a situation might occur do not justify such efforts in the field. Instead, these cases must be found during the structure checks. Note that in this context, split-off households must be considered in the waves following their first successful survey even if they are considered panel households in field control. In both cases, the personal identification numbers of the individuals in the split-off household are corrected retrospectively. It must also be considered that these individuals are treated as new respondents in the personal/senior citizen interview although they might have already participated in an interview. This deviation is generally not corrected (see also Chapter 4.4).

- In panel households that reported a move-out as of wave 2, a return to the household can also occur as of wave 3. Recognising these individuals as moving back in and assigning them their former household position instead of a new household position is a function of the household grid. Whether these requirements were met in the field in all cases was also evaluated. For individuals who were identified in the current wave as moving back in by comparing the first name, age and sex with the members who previously moved out of the household, the household structure must be changed. These changes led to retroactive changes of the personal identification number of the individual and the individual information in the household interview - e.g., information about childcare or the reasons for a cut in Unemployment Benefit II - to the correct position within the structural check. Whether an individual who is marked in the field as moving back in is the same individual who moved out during a previous wave was also verified. If not, this change represents an individual who is new to PASS. Changes to the household structure are also made in this case.
- In case of moves back into a household, whether the split-off household in which the individual lived was successfully surveyed during the current wave and whether the split-off household reported that the individual moved out were verified. In addition, the status of individuals who moved back into their panel household during a previous wave must continue to be verified with the split-off household provided the split-off household is part of the current panel sample. If an individual who moves back in is still considered a current household member in his/her split-off household, a decision was made as to whether this was plausible or whether either household structure should be corrected.
- Returns are not the only cases of individuals being considered current household members of several households. This situation can also occur when a member of a split-off household is not recorded as having moved out of the panel household. Individual cases can be acknowledged as plausible after examination of both household structures. These cases are documented in the *zdub** variables in the person register. For further explanation, please refer to Chapters 4.4 and 5.4.1.2 of the data report for Wave 5 of PASS (Berg et. al., 2012).
- Other issues concerning the relationship of a panel household and its split-off households can also arise. Individuals who joined PASS via a split-off household might move to the panel household. Another possibility is that individuals move from one

split-off household to another. Generally, all individuals in a panel household and all of its split-off households must be considered a network. The structure checks are designed so that individual moves among the households of such a network are detected regardless of the direction in which an individual moves.

- Household structure verification generally evaluates the changes between waves, not the plausibility of the structure. Therefore, the household structure first-time interviews can only be verified to a limited extent. For first-time households, information concerning first name, age and sex is reviewed to determine whether individual household members are listed multiple times. In this case, only the initially reported household position is maintained. This situation might lead to other changes in the household structure. If, for example, in a household interviewed for the first time, there are four individuals and the individuals in positions 2 and 3 are identical, individual 3 is removed and individual 4 is retroactively moved to position 3. As a rule, in a household interviewed for the first time with X household members, positions 1 to X are to be filled without gaps. Someone retroactively recognised as moving back through a subsequent change in his or her personal identification number also makes it necessary to move the individual information in the household interview.
- Thanks to feedback provided by a field interviewer, a household that was included twice in the panel sample during wave 4 was detected. Household 10015439 had been included in the sample as the identical household 15044862 since wave 1. Both households were successfully surveyed during waves 1 and 3 and not surveyed during wave 2. In wave 4, household 10015439 was successfully surveyed. This duplicate was detected because “both” households were assigned to the CAPI interviewer for that point. The household composition remained the same across all waves. Household 15044862, which was not surveyed in wave 4, will be deleted from the sample for wave 5. There will be no retroactive removal of the duplicate from waves 1 to 3 because to do so would affect weighting. The duplicate household is coded 26 in the *hnettod4* variable in *hh_register*, which identifies the reason for non-surveying. All household members of the duplicate household are coded 56 in the *pnettod4* variable in *p_register*.
- Individual decisions were also made to address cases that proved to be problematic during the structure checks. Here, the seriousness of the particular problem was significant. In cases in which the correct household composition in wave 7 was unclear, all of the interviews from wave 7 were removed. In wave 8, these households will be treated as households that did not participate in wave 7. If in retroactively removed household interviews moves-out were reported, the split-off households were discarded. This removal affected both the interviews conducted in the current wave in these split-off households and the sample of the subsequent wave. Split-off households that developed from a discarded interview of a panel household are retroactively classified as not having been conducted and do not contribute to the panel sample of the subsequent wave. If there was merely a problem in assigning individuals to their respective positions in the household, i.e., if it was suspected that a personal interview had been conducted with the wrong individual in wave 7, then only that personal or senior citizen interview was removed. Structural problems with no serious consequences that could be solved, for example, by removing a personal interview, first name, age and

sex were made at the household level. The incorrect information concerned was replaced with the last valid value from the previous wave or the value from the previous wave added to the number of years since the last valid interview.

In addition, all interviews with individuals for households with no complete household interview were removed. In the opposite case, i.e., households for which no individual-level interview was available, a distinction was made between re-interviewed households and households from the refreshment sample. Households from the refreshment sample that were not successfully surveyed were removed following the procedure used in the previous waves. In the case of re-interviewed households without interviews at the individual level, however, the household interview was not deleted.

The *netto* variables (*hnettok7*, *hnettod7*, *pnettok7*, *pnettod7*) in the household and person register datasets indicate removed interviews. Through the corresponding variables in the household register, it is possible to trace the re-interviewed households whose household interviews were later removed. Net variables in the person register allow for tracing the cases in which only single individual-level interviews or all of the interviews in the household were deleted. In the case of households from the refreshment sample of wave 7 without at least one valid household and personal interview, it is not possible to trace deleted interviews in the register datasets because these households were not included in the datasets.

5.2 Filter checks

During the filter checks, the correct operation of the filter questions in the instruments was verified using a statistical program. If certain questions were asked when the value of the relevant filter variable would have required something else (for example, if detailed information was requested about vocational training although the respondent had stated that he/she did not have any vocational qualification), these variables were set to missing code "-3" (not applicable), which they would also have received through correct use of the filters.[33] Moreover, some items were not asked in individual cases when those questions would have been necessary according to the filter (e.g., if no further information was recorded about vocational training although the respondent had stated that he/she had undergone such training). In these cases, the missing code "-4" (question mistakenly not asked) was assigned. An assignment of code "-4" can also be based on the household structure evaluation described in Chapter 5.1. If an individual's move-out is retroactively discarded as implausible and the individual is retroactively classified as belonging to his or her former household, then individual information about these individuals in the household interview must be coded retroactively as mistakenly not surveyed. Thus, the code "-4" does not always refer to a problem in the survey instrument. If code "-4" is assigned to a question that is relevant for filtering subsequent questions, then the subsequent questions are also coded "-4" in case these subsequent questions are not asked. If these questions

[33] As is customary in such cases, the filter checks were conducted beginning with the items that were asked first.

were asked because, for instance, several filter questions linked to this subsequent question and another filter question triggered the question correctly, the value recorded there remains.

In an additional step, the missing codes assigned by the field institute and system missing codes were replaced by standard values for all variables. Table 23 provides an overview of the assigned values. Codes “-1” and “-2” are the standard “don’t know” and “details refused” answers recorded during the survey, respectively. Code “-3” is the general “not applicable” code for questions not asked due to filters. As described above, code “-4” was assigned if a question was not asked because of a filter error. Codes “-5” through “-7” are question-specific codes. These can be either specific missing codes (e.g., “Not applicable, not available for the labour market”) or special categories for valid values (e.g., a category for an income of greater than € 99,999 in the open question on income). These codes were only assigned as required.

Table 23: Overview of the missing codes used

	Explanation
-1	“don’t know”
-2	“details refused”
-3	“not applicable (filter)” (question not asked due to filter)
-4	“question mistakenly not asked” (question should have been asked)
-5	question-specific code number 1, only assigned as required
-6	question-specific code number 2, only assigned as required
-7	question-specific code number 3, only assigned as required
-8	“implausible value”
-9	“item not surveyed in wave”
-10	“item not surveyed in questionnaire version”[34]

The value “-8” is a specific missing code assigned during the plausibility checks (see Chapter 5.3 on plausibility checks). The missing code “-9” became necessary for the first time in wave 2. It is assigned if an item was not asked during a specific wave.

Because the dataset is prepared in long format, as was described above, variables that were no longer asked in any version of the questionnaire as of wave 2 are coded “-9” for the observations in this wave. Variables included for the first time after wave 1 are retroactively coded “-9” for observations of waves in which they were not surveyed. Code “-10” can be used to consider differences between questionnaires, that is, between the person-

[34] As of wave 4, code "-10" has only been used to differentiate between personal and senior citizen questionnaires. Up to and including wave 3, there was an additional differentiation at the household level between first-time and repeatedly interviewed households. The differentiation at the household level is not continued in wave 4 due to the merger of the questionnaire versions into one comprehensive household questionnaire.

al questionnaire and senior citizen questionnaire or between two versions of the household questionnaire until wave 3.

5.3 Plausibility checks

For the plausibility checks, an extensive list of theoretically possible contradictions in the respondents' statements was checked. The checks conducted during the previous waves were adapted and extended for the current wave. Furthermore, the household structure and spell data were checked for plausibility - especially for inadmissible overlaps within the individual spell types. Generally, only the data gathered in the cross-section of wave 7 were verified. No checks were conducted in the longitudinal section, that is, to compare the information provided in the current wave with that provided in the previous wave.

In detail, the following steps were conducted:

1. Contradiction check: In general, contradictions were only corrected either if the implausibility could be defined as particularly serious and/or if the alteration was considered minor. The latter applied, for example, if only a small number of cases were affected or if one missing code (e.g., "-3") was replaced by another (e.g., "-8"). Two strategies were used to filter implausible statements. Either the implausible responses were corrected directly, or they were assigned a specific missing code.

2. Implausible responses were only corrected if it was highly probable that the interviewer had entered information incorrectly: for example, if the interviewer entered a monthly total rent of EUR 9,998.-. Here, it was assumed in the plausibility check that the five-digit missing code "99998" (don't know) was entered incorrectly. This response and other similar responses were recoded to the corresponding missing categories. If the recoded missing categories triggered a filter in subsequent questions, as is the case for the categorical question of income, then the categorical questions were retroactively set to code "-4" (question mistakenly not asked).

    - However, it was rarely the case that a value could be recognised as an incorrect entry with certainty. In most cases, it was only possible to establish a contradiction between two statements but not to identify specific incorrect entries that had led to the implausible statement. Therefore, in these cases, no corrections were made, and the specific missing value code "-8" was assigned instead. It was decided on an individual basis whether the code was assigned to one of the two variables involved in the contradiction or to both of them.

3. Plausibility check of the household structure: This check was conducted based on the information collected in the household interview about family relationships between household members, age, sex and first name. Prior to this check, information about relationships in the household was supplemented by information about partnerships reported in the personal interview.

    - To identify implausible household structures, the information on relationships was first combined with the demographic information for individual household

members. For the households that were identified as implausible during these checks, individual decisions were made considering overall household structure and other information gathered during the interviews (e.g., on marital status in the personal interview). Implausible relationships were marked as such ("-8") or corrected based on additional information on the household context if it was highly probable that an error had occurred. For example, in the case of two people of the same sex who were both biological parents of a third member of the household, the sex was corrected based on the first name.

If the first names also indicated two people were of the same sex and if there was no other relevant information available, then the relationship was marked as implausible based on the household structure.

- In a second step, checks were conducted comparing sets of three family relationships for plausibility. The following provides an example of a relationship structure that would be classified as implausible: individual A is individual B's spouse. Individual A is the biological parent of individual C. Individual C is a sibling of individual B. If such a combination or similarly implausible combination of relationships was identified, an attempt was made to make the relationship plausible based on the household context. In the case described, the relationship data were corrected by coding individual C as a child of individual B, whose status was not specified. The aim was to correct as many of the implausible entries as possible because a plausible and complete set of relationships is necessary to generate the benefit unit.

4. In addition, the spell datasets were subjected to a number of plausibility checks, as detailed in Chapters 5.6 through 5.8.

5.4 Retroactive changes in waves 1 to 6

During the data preparation process for the scientific use file for wave 7, some changes were also made to the waves that had already been delivered. These changes included corrections of errors that were detected after the completion of the scientific use file of wave 6. The corrected data can now be used in the SUF datasets of the current wave, wave 7. Tables 24 through 28 provide an overview of the retroactive changes to the delivered waves of PASS[35].

[35] Adjustments to value or variable labels are only considered here if this changes the interpretation of variables or values.

Table 24: Overview of retroactive changes to the household dataset (HHENDDAT)

Altered variable	Dataset concerned	Altered wave	Type of alteration	Description of the alteration
-	-	-	-	-

Table 25: Overview of retrospective alterations in the individual dataset (PENDDAT)

Altered variable	Dataset concerned	Altered wave	Type of alteration	Description of the alteration
-	-	-	-	-

Table 26: Overview of retroactive corrections to spell datasets (bio_spells, alg2_spells, and ee_spells)

Altered variable	Dataset concerned	Altered wave	Type of alteration	Description of the alteration
-	-	-	-	-

Table 27: Overview of retrospective alterations to the register datasets (hh_register; p_register)

Altered variable	Dataset concerned	Altered wave	Type of alteration	Description of the alteration
-	-	-	-	-

Table 28: Overview of retrospective alterations to the weighting datasets (hweights; pweights)

Altered variable	Dataset concerned	Altered wave	Type of alteration	Description of the alteration
-	-	-	-	-

5.5 Anonymisation

All data obtained by the IAB, a special department of the Federal Employment Agency (BA), are social data, which places high demands on data protection. It was therefore necessary to include some of the variables in the scientific use file in simplified form. These variables are generally labeled with the flag “anonymised” in the variable label.

For the same reason, it was also necessary to exclude available regional information, excluding the German states and information about East/West Germany. To protect the data, neither family relationships in the household nor the first names of the household members are part of the scientific use file. References to the household structure are provided, however, by generated variables. For example, the household and benefit unit type (*hhtyp*[36], *bgtyp*[37]), indicator variables on partners in the household (*apartner; epartner*[38]), indicator variables pointing to parents, partners in the household (*zmhh; zvhh; zparthh*[39]) and various indicator variables for parents (*mhh; vhh*[40]) or children of the target person (e.g., *ekind*[41]) living in the household are provided.
Table 29 provides an overview of the variables concerned and the process of anonymisation[42] in each dataset. Table 30 provides the anonymised variables for the employment spell dataset.

---

36 Contained in the household dataset (*HHENDDAT*), see Chapter 4.5.2

37 Wave-specific variables contained in the person register (*p_register*), see Chapter 4.4.

38 Contained in the individual dataset (*PENDDAT*), see Chapter 4.4.

39 Wave-specific variables contained in the person register (*p_register*), see Chapter 4.4.

40 Contained in the individual dataset (*PENDDAT*), see Chapter 4.4.

41 Contained in the individual dataset (*PENDDAT*), see Chapter 4.4.

42 If non-anonymised versions of one or several variables are indispensable for your research, please contact the Forschungsdatenzentrum (Research Data Center) to determine the possibility of obtaining access to the data. The form of this access will depend on the research project and the variables necessary.

Table 29: Overview of the anonymised variables in the individual dataset (PENDDAT) in wave 7

Varname	Variable label	Procedure
PD0100	Year of birth (date of birth, anon.)	The precise date of birth was shortened to year of birth.
gebhalbj	Half-year of birth, gen.	The precise date of birth was shortened to an indicator for the first or second half of the year.
PET1210	Last occupational status, simple classification (anon.)	For technical reasons, professional and regular soldiers were recorded separately. Due to the few case numbers and because this group is not usually asked about occupational status, this group was merged with civil servants and judges.
PET1250	Last occup. status civil servant: detailed info., incl. soldiers (anon.)	This variable contains additional cases. The professional and regular soldiers from PET1240 were added to the corresponding civil servants category. The variable for professional and regular soldiers PET1240 is not supplied.
PET1211	Last occup. status, simple class. (incl. spell info.) (anon.), gen.	Procedure as for *PET1210*.
PET1251	Last occup. status civil servant: detailed info., incl. soldiers (incl. spell info.) (anon.), gen.	Procedure as for *PET1250*. The variable for professional and regular soldiers PET1240 is not supplied.
stiblewt	Occupational status, last employment, code number, gen.	When generating the occupational status variable, professional and regular soldiers were assigned to the corresponding civil servant category.
PET1510	Current occup. status, simple classification, surv. as of wave 2 (anon.)	Procedure as for *PET1210*.
PET1900	Current occup. status civil servant: detailed info., incl. soldiers (anon.)	Procedure as for *PET1250*. The variable for professional and regular soldiers PET1800 surveyed in the senior citizens' interviews is not supplied. For the personal interviews, no generated variable for professional and regular soldiers is incorporated into the individual dataset from the employment spells ET090*.
stibkz	Current occupational status, simple classification, harmonised (anon.)	When generating the occupational status variable, professional and regular soldiers are assigned to the corresponding civil servants category.
stib	Occupational status, code number, gen.	Procedure as for *stiblewt*.

PET3300	First occup. status, simple classification (anon.)	Procedure as for *PET1210*.
PET3700	First occup. status civil servant: detailed info., incl. soldiers	Procedure as for *PET1250*. The variable for professional and regular soldiers PET3600 is not supplied.
PET3301	First occup. status, simple class. (merged, incl. spell info.) (anon.), gen.	Procedure as for *PET1210*.
PET3701	First occup. status civil servant: detailed info., incl. soldiers, (merged, incl. spell info) (anon.), gen.	Procedure as for *PET1250*. The variable for professional and regular soldiers PET3600 is not supplied.
stibeewt	Occupational status, first employment, code number, gen.	Procedure as for *stiblewt*.
PSH0320	Mother's occup. status at that time, simple classification (anon.)	Procedure as for *PET1210*.
PSH0360	Mother's occup. status at that time, civil servant, incl. soldiers: detailed info. (anon.)	Procedure as for *PET1250*. The variable for professional and regular soldiers PSH0350 is not supplied.
mstib	Mother's occupational status, code number, gen.	Procedure as for *stiblewt*.
PSH0620	Father's occup. status at that time, simple classification (anon.)	Procedure as for *PET1210.*
PSH0660	Father's occup. status at that time, civil servant, incl. soldiers: detailed info. (anon.)	Procedure as for *PET1250*. The variable for professional and regular soldiers PSH0650 is not supplied.
vstib	Father's occupational status, code number, gen.	Procedure as for *stiblewt*.
PMI0200	Not born in Germany: country of birth	Countries with very low case numbers were grouped into larger categories.
ogebland	Country of birth, incl. open info., categories (anon.)	Procedure as for *PMI0200*.
PMI0500	No German nationality: which nationality? (anon.)	Nationalities of countries with very low case numbers were grouped into larger categories.

ostaatan	Nationality, incl. open info., categories (anon.)	Procedure as for *PMI0500*.
PMI1000a	Father: country of res. before migration (anon.)	Countries of residence before migration with very low case numbers were grouped into larger categories.
PMI1000b	Mother: country of residence before migration (anon.)	Procedure as for *PMI1000a*.
PMI1000c	Father's father: country of residence before migration (anon.)	Procedure as for *PMI1000a*.
PMI1000d	Father's mother: country of res. before migration (anon.)	Procedure as for *PMI1000a*.
PMI1000e	Mother's father: country of residence before migration (anon.)	Procedure as for *PMI1000a*.
PMI1000f	Mother's mother: country of residence before migration (anon.)	Procedure as for *PMI1000a*.
ozulanda	Father: country of residence before migration, incl. open info., categories (anon.)	Procedure as for *PMI1000a*.
ozulandb	Mother: country of residence before migration, incl. open info., categories (anon.)	Procedure as for *PMI1000a*.
ozulandc	Father's father: country of residence before migration, incl. open info., categories (anon.)	Procedure as for *PMI1000a*.
ozulandd	Father's mother: country of residence before migration, incl. open info., categories (anon.)	Procedure as for *PMI1000a*.
ozulande	Mother's father: country of residence before migration, incl. open info., categories (anon.)	Procedure as for *PMI1000a*.
ozulandf	Mother's mother: country of residence before migration, incl. open info., categories (anon.)	Procedure as for *PMI1000a*.

Table 30: Overview of the anonymised variables in the BIO spell dataset (bio_spells) in wave 7

Varname	Variable label	Procedure
ET0605	Wave 7, Occup. status, simple classification (anon.)	Procedure as for *PET1210*.
ET1005	Wave 7, Occ. status: civil servant / judge / soldier, detailed information (anonymised) (anon.)	Procedure as for *PET1250*. The variable for professional and regular soldiers is not supplied.
stib	Occ. status, code number, gen.	Procedure as for *stiblewt*.

5.6 Receipt of Unemployment Benefit II

UB II is recorded at the household level in spell form in waves 1 to 6. This concept was continued in wave 7 but with a slightly revised set of questions.

5.6.1 Concept for updating the spells of Unemployment Benefit II receipt that were ongoing in the previous wave

To update spells for which UB II was ongoing during the previous wave and therefore were right-censored in the spell dataset, dependent interviewing questions are included. Households with ongoing spells from the previous wave start here again with the interview.

The households from the refreshment sample that were interviewed for the first time in wave 7 were asked about their receipt of UB II during the period since the last change in the household composition. If this change was before January 2011 or if no information was provided about changes in the household, then the household's receipt of UB II from January 2011 on was recorded.

5.6.2 Structure of the Unemployment Benefit II spell dataset

The structure and contents of the spell dataset on UB II change due to the integration of the spells of UB II reported in wave 7. Here, it is necessary to distinguish among (1) new variables that refer to a particular wave, (2) new variables that do not refer to a particular wave and (3) variables that are no longer asked in wave 7.

1. Additionally, in wave 7, new wave-specific, cross-sectional variables were included in the UB II spell dataset. These variables include *AL20606, AL20706a to AL20706o, AL20806 and AL20906*. These variables refer to the interview date in wave 7. Cross-sectional variables also exist for the interview dates of the previous waves that contain the analogous information referring to the respective wave. Table 31 provides an overview of the cross-sectional information contained in the UB II spell dataset.

Table 31: Cross-sectional variables in the UB II spell dataset (alg2_spells)

	Cross-sectional variable with information referring to				
	Wave 1:	**Wave 2:**	**Wave 3:**	**…**	**Wave 7:**
Does the HH receive UB II for all HH members?	*AL20600*	*AL20601*	*AL20602*	…	*AL20606*
Does the HH receive UB II for individuals 1 to 15?	*AL20700a* to *AL20700o*	*AL20701a* to *AL20701o*	*AL20702a* to *AL20702o*	…	*AL20706a* to *AL20706o*
Amount of monthly UB II receipt?	*AL20800*	*AL20801*	*AL20802*	…	*AL20806*
Has a cut of UB II begun?	*AL20900*	*AL20901*	*AL20902*	…	*AL20906*

2. Not available in wave 7 compared to wave 6.
3. Not available in wave 7 compared to wave 6.

5.6.3 Plausibility checks and corrections to the Unemployment Benefit II spell dataset

As in waves 1 to 6, the information on UB II was also subjected to a number of plausibility checks in wave 7. Inadmissible overlaps and dates of spells of UB II or benefit cuts were corrected when necessary. In principle, changes were only made to the generated date variables (*bmonat; bjahr; emonat; ejahr*) of the spell of UB II receipt, the spells of benefit cuts (*alg2kbm*; alg2kbj*; alg2kem*; alg2kej**) and the censoring indicator of the spell of UB II receipt (*zensiert*). If it was not possible to remove implausible data by correcting the dates, then in a small number of cases, spells of UB II receipt or cuts were merged or deleted.

5.6.4 Updating the Unemployment Benefit II spell dataset

After the spells of Unemployment Benefit II reported in wave 7 had been converted into spell format, and after inadmissible overlaps and implausible dates were corrected following the plausibility checks and corrections, the spells of UB II that were ongoing at the time of the interview in the previous wave were updated using the information gathered in wave 7. Two variants are to be distinguished here. In the first (1), only the censoring indicator *zensiert* is changed. The second variant (2) is an update of the spell that was censored during the previous wave using information gathered in wave 7. Here, the censoring indicator is integrated into the spell of receiving UB II, which was ongoing during the previous wave, as are the generated and recorded end dates, wave-specific cross-sectional information (see above) and new spells of benefit cuts. In addition to updating spells that were censored during the previous wave, new spells that were reported in wave 7 are merged with the spell dataset (3). These three variants are outlined briefly below:

1. Cases in which the household in wave 7 contradicts an ongoing spell of receiving UB II at the interview date in the previous wave.

   If the household contradicted an ongoing spell of receiving UB II at the time of the previous wave, either explicitly or implicitly (by reporting an end date that preceded the interview date in the previous wave) in the update question, then *zensiert* was set to "2" (no). The information provided in the interview of the previous wave is assumed correct. Because it is not possible to make reliable statements about the continued duration of the benefit receipt beyond the date of the interview in the previous wave, it is assumed that the benefit receipt ended during the month of the interview in the previous wave. The reported and generated variables for the end date of the spell (*AL20300, AL20400* and *emonat, ejahr*), along with the question of whether a spell continues (*AL20500*), remain unchanged.[43] The generated end date of the UB II spell (*emonat; ejahr*) had been set to the interview date of the previous wave in the previous wave.
2. Cases in which the household reports the end date of a spell of benefit receipt that was ongoing in the previous wave.

   If information about the end date of a spell of UB II receipt that was censored in the previous wave is available in wave 7, then the spell that was censored in the previous wave was updated using the current information. First, the recorded end date (*AL20300; AL20400*), the generated end date (*emonat; ejahr*), the follow-up question as to whether the receipt of UB II is ongoing (*AL20500*) and the censoring indicator (zensiert) are overwritten with the information gathered in the previous wave. Furthermore, the spells of benefit cuts reported in wave 7 and the cross-sectional data referring to wave 7 (*AL20606; AL20706a to AL20706o, AL20806, AL20906*) were included.
3. Spells of UB II receipt reported for the first time during wave 7 that do not update any spells that were censored in the previous wave.

   Spells reported for the first time during wave 7 were added to the UB II spell dataset. Next, the spell counter was generated anew to create a variable *spellnr* without gaps.

5.7 Employment biographies

Employment, unemployment and gap periods at the individual level were recorded in spell form in waves 2 and 3. This concept of a modular spell survey was changed to an integrated survey of the employment biography in wave 4. For individuals who were asked for their employment biography for the first time in wave 7, the reference date for the start of the retrospective interval was adjusted. In wave 7, all spells of employment and unemployment since January 2011 were to be reported here. Individuals who were interviewed about their employment biography during the previous wave, however, should report all new spells since the date of the last interview.

---

[43] The same applies here. Only the censoring indicator is changed. The reported end date, the question for continuing spells and the generated end date remain unchanged.

5.7.1 Concept for updating the spells that were ongoing in the previous wave

Continuing ET, AL and gap spells were updated in wave 7. To update the spells that were ongoing during the previous wave and were therefore right-censored in the spell dataset, dependent interviewing questions are included in the personal questionnaires.

5.7.2 Structure of the BIO spell dataset

With respect to its structure, the BIO spell dataset has oriented itself on the modular ET, AL and LU spell datasets of waves 2 to 3 since wave 4. ET-specific variables kept their names in the BIO spell dataset compared to the ET SUF of wave 3, analogous to the AL- and LU-specific variables. Variables which are the same in ET, AL and LU have been standardised (BIO0100, BIO0101, BIO0200, BIO0300, BIO0400, BIO0500, BIO0600) as of wave 4 or were already standardised in the original datasets of the SUF wave 3 (bmonat, bjahr, emonat, ejahr, zensiert). Furthermore, variables for type of activity (spelltyp), spell integration (spintegr) and comprehensive spell number (spellnr) are available.

Due to the integration of the employment and unemployment spells reported in wave 7 into the BIO spell dataset, new ET- and AL-specific variables are added. Here, it is necessary to distinguish between (1) new variables that refer to a particular wave and (2) new variables that do not refer to a particular wave.

1. The ET-specific variables in the BIO spell dataset *ET0600 to ET2200* are considered wave-specific, cross-section information that refer to wave 2; variables ET0601 to ET2201 refer to wave 3, *ET0552 to ET2202* refer to wave 4, ET0553 to ET2203 refer to wave 5, ET0554 to ET2204 refer to wave 6, and ET0555 to ET2205 are cross-section information that refers to wave 7. Table 32 provides an overview of the ET-specific cross-section information in the BIO spell dataset.

Table 32: ET-specific cross-section variables in the BIO spell dataset (bio_spells)

	Cross-sectional variable with information referring to...				
	Wave 2:	**Wave 3:**	**Wave 4:**	**Wave 5:**	**Wave 7:**
Occupational status (simple and detailed classification)	*ET0600* *ET0700* *ET0800* *ET1000* *ET1100* *ET1200*	*ET0601* *ET0701* *ET0801* *ET1001* *ET1101* *ET1201*	*ET0552* *ET0602* *ET0702* *ET0802* *ET1002* *ET1102* *ET1202*	*ET0553* *ET0603* *ET0703* *ET0803* *ET1003* *ET1103* *ET1203*	*ET0555* *ET0605* *ET0705* *ET0805* *ET1005* *ET1105* *ET1205*
Supervisory function; number of employees supervised	*ET1300* *ET1400*	*ET1301* *ET1401*	*ET1302* *ET1402*	*ET1303* *ET1403*	*ET1305* *ET1405*
Cancellation of limitation of an initially limited employment	*ET1700*	*ET1701*	*ET1702*	*ET1703* *ET1753a* *ET1753b*	*ET1705* *ET1755a* *ET1755b*
Working hours (contracted; actual; average for irregular working hours)	*ET2000* *ET2100* *ET2200*	*ET2001* *ET2101* *ET2201*	*ET1952* *ET2002* *ET2102* *ET2202*	*ET1953* *ET2003* *ET2103* *ET2203*	*ET1955* *ET2005* *ET2105* *ET2205*
Income for current ongoing spells				*ET2600-* *ET3900*	*ET2802-* *ET3902*

The BIO spell dataset also includes an AL-specific variable which is understood as wave-specific cross-sectional information (AL1300 for wave 2; AL1301 for wave 3, AL1302 for wave 4, AL1303 for wave 5, AL1304 for wave 6 and AL1305 for wave 7). Table 33 gives an overview of the cross-sectional information contained in the spell dataset.

Table 33: AL-specific cross-section variables in the BIO spell dataset (bio_spells)

	Cross-sectional variable with information referring to …					
	Wave 2:	**Wave 3:**	**Wave 4:**	**Wave 5:**	**Wave 6:**	**Wave 7:**
Amount of monthly UB I receipt?	*AL1300*	*AL1301*	*AL1302*	*AL1303*	*AL1304*	*AL1305*

2. Does not exist in wave 7 compared with wave 6.

5.7.3 Plausibility checks and corrections of the spell datasets

At the individual level, the plausibility checks and corrections orient themselves by wave 2 to wave 4. As in wave 4, checks were made only within one spell type. Cross-spell type checks were not conducted. As with the spell data on receiving UB II, correction and re-coding were only conducted for the generated date variables. Here, details on seasons were recoded into months, "-8" values were set for implausible responses and date information was replaced or rendered plausible. Because only the generated date variables were edited, the original information gathered in the survey is available to the user in the date variables BIO0200-BIO0500 and AL0800-AL1100, thus permitting the user to conduct his/her own checks and corrections.

In addition, in some cases it was necessary to delete entire spells. For example, spells that were obviously recorded twice were removed. Spells that are completely outside the survey period but for which data were nonetheless collected were also deleted.

5.7.4 Update of spell datasets

After the spells reported in wave 7 had been converted into spell format, plausibility checks and corrections for inadmissible overlaps and spells with implausible dates were corrected. The spells that were ongoing at the time of the previous interview wave were updated using the information recorded in wave 7.

Three variants are to be distinguished here. In the first (1), only the censoring indicator *zensiert* is changed. The second variant (2) is an update of the spell that was censored in the previous wave using information gathered in wave 7 in the narrow sense. Here, the censoring indicator is integrated into the spell that was ongoing during the previous wave, as are the generated and recorded end dates and wave-specific cross-sectional information (see above).

In addition to updating spells that were censored during the previous wave, new spells reported in wave 7 are merged with the spell dataset (3). These three variants are outlined briefly below:

1. Cases in which the individual in wave 7 contradicts an ongoing spell on the interview date in the previous wave.

If the individual contradicted the information that there was an ongoing spell at the time of the previous wave, either explicitly or implicitly (by reporting an end date that preceded the interview date in the previous wave) in the update question, then the censoring indicator *zensiert* was set to “2” (no). The information provided in the interview of the previous wave is assumed correct. Because it is not possible to make any reliable statements about the continued duration of the spell beyond the date of the interview in the previous wave, it is assumed that the spell ended during the month of the interview in the previous wave. The reported and generated variables on the end date of the spell (BIO0400*, BIO0500 and emonat, ejahr*), along with the question of whether a spell continues (BIO0600) remain unchanged[44]. The generated end date of the spell (*emonat; ejahr*) was already set to the interview date of the previous wave in the previous wave.

2. Cases in which the individual reports the end date of a spell that was ongoing in the previous wave.

If information about the end date of a spell that was censored during the previous wave is available in wave 7, then the spell that was censored was updated using the current information. For ET spells, the recorded end date (BIO0400*; BIO0500*), the generated end date (*emonat; ejahr*), the follow-up question as to whether the spell was ongoing (BIO0600), the reason for the cancellation of a work contract (*ET2300*), the generated variables on occupational status and weekly working hours (*stib, az1, az2*) and the censoring indicator (*zensiert*) were overwritten with the information gathered in wave 7. Furthermore, the cross-sectional data referring to wave 7 (*ET0555 to ET2205*) were included.

For AL spells, the recorded end date (BIO0400*; BIO0500*), the generated end date (*emonat; ejahr*), the follow-up question as to whether the spell was ongoing (BIO0600), the reason for the end of unemployment (*AL0600, AL0601*) and the censoring indicator (zensiert) were overwritten with the information gathered in wave 7. Furthermore, the cross-sectional data referring to wave 7 (*AL1305*) were included. AL spell data, moreover, feature the exception that the spell of UB I (receipt of UB I) is recorded within an AL spell. Which information is updated depends on whether UB I was already received during this spell of unemployment and whether this benefit was ongoing during the previous wave.

- If, in the previous wave, there was also an ongoing receipt of UB I in the AL spell to be updated, then the recorded end date of the receipt (*AL1000, AL1100*), the indicator as to whether the spell is ongoing (*AL1200*), the generated end date of

[44] Thus, the reported end date remains completed with the interview date of the wave in which the spell was censored or the special code "0" for continuing spells. In addition, the question about whether the spell continued (for the case that the end date corresponds with the interview date) is not changed. The generated date variables continue to contain the last valid information, which here is the interview date for the wave in which the spell was censored.

the receipt (*alg1em, alg1ej*) and the censoring indicator of the receipt (*alg1akt*) were overwritten with the information obtained in wave 7.

- If no UB I was received in previous waves in the AL spell to be updated, then the information on UB I receipt was overwritten with the information obtained in wave 7. In addition to the indicator as to whether UB I was received in the AL spell (*AL0700*), the reported start and end date (*AL0800, AL0900, AL1000, AL1100*), the indicator for ongoing receipt (*AL1200*) and the respective generated variables (*alg1bm, alg1bj, alg1em, alg1ej, alg1akt*) were replaced with the newly recorded information.
- If there was UB I receipt in the AL spell to be updated in the past but that ended in the previous wave, no changes were made to these spells.

3. Spells reported for the first time in wave 7 that do not update any spells that were censored in the previous wave.

Spells reported for the first time in wave 7 were added to the BIO spell dataset. Next, the spell counter was generated anew to create a variable *spellnr* without gaps.

Updating the spell datasets does not affect the spell numbers of the previous wave's SUF. Spells already included in the wave 6 SUF (*spellnret, spellnral, spellnrlu, spellnr*) maintain their spell number. The new spells from wave 7 are added to the respective dataset and the spell numbers are updated.

5.8 One-Euro job spell dataset (ee_spells)

In wave 4, the concept for surveying participation in employment and training measures was thoroughly revised. The MN spell dataset has been replaced by the one Euro spell dataset (ee_spells) as of wave 4. This was updated in wave 7. The reference date as of which to consider one-Euro jobs was January 2012 for wave 7.

5.8.1 Concept for updating the spells that were ongoing in the previous wave

Continuing ee_spells were updated in wave 7. To update the spells that were ongoing in the previous wave and were therefore right-censored in the spell dataset, dependent interviewing questions are included in the personal questionnaires.

5.8.2 Structure of the EE spell dataset

By integrating the one-Euro jobs (OEJ) reported in wave 7 in the OEJ spell dataset (ee_spells), new variables are added that refer to a specific wave. Table 34 gives an overview of the cross-sectional information contained in the EE spell dataset.

Table 34: Cross-sectional variables in the EE spell dataset (ee_spells)

Cross-sectional variable with information referring to...	Wave 4:	Wave 5:	Wave 6:	Wave 7:
Weekly working hours in the OEJ	*EE1100*	*EE1101*	*EE1102*	*EE1103*
OEJ is the same work permanent co-workers do	*EE1200*	*EE1201*	*EE1202*	*EE1203*
Which kind of training necessary for OEJ	*EE1300*	*EE1301*	*EE1302*	*EE1303*
Only work or also training/classes?	*EE1400*	*EE1401*	*EE1402*	*EE1403*
Assessment OEJ	*EE1500a-EE1500h*	*EE1501a-EE1501h*	*EE1501a-EE1501h*	*EE1503a-EE1503h*

For the OEJ spell dataset, it must be considered that there are also spells if the OEJ was not performed, i.e., if there was no participation.

5.8.3 Plausibility checks and corrections in the EEJ spell dataset

The OEJ spell dataset on the participation in OEJ was both checked for plausibility and corrected. The plausibility checks contained checks for dates, for the reference date for the newly integrated spells in wave 7 (January 2012) and for logical inconsistencies in cases of respondents with several OEJ spells.

Only the generated date variables (bmonat, bjahr, emonat, ejahr) were corrected and re-coded. Details on seasons were recoded into months, “-8” values were assigned for implausible responses and date information was replaced or rendered plausible. Next, a spell counter *spellnr* was generated. The variable generation was performed analogously to the chronological counters in the BIO spell datasets. Non-participating spells were not included in the sorting and thus kept their original *position within the survey wave*. Spells from wave 6 maintained their spell number for the wave 7 SUF.

6. Weighting wave 7

The weighting concept for wave 7 generally follows the concepts developed in previous waves (see Berg et al., 2013). The starting point for the wave 7 weighting procedure and for the longitudinal section from wave 6 to wave 7 were the cross-sectional weights from wave 6 for households and individuals. The two weights for each household and two weights for each individual were updated. This chapter of the data report documents the technical details and exact models used to generate the weights for wave 7. An overview of the weighting concept used in PASS can be found in chapter 8 (Trappmann, 2013a) of the PASS User Guide (Bethmann, Fuchs, and Wurdack, 2013). Examples of how to use the weights can be found in Chapter 12 (Trappmann, 2013b).

6.1 Design weights for the panel households in wave 7

New "household design weights" were generated for wave 7 from the cross-sectional weights for households of wave 6, taking into account people moving into households from within Germany. This step was performed by using the weight share procedure as described in wave 2 (see Gebhardt et al., 06/2009). Births, deaths or move-outs from households have no influence on weight; moves into households from within Germany, however, increase the inclusion probability of a household because the individuals who moved into the household also had the opportunity to be included in the sample in waves 1 to 6. The new design weight for subsample i dw_ihh_7 is therefore calculated from the old cross-sectional weight wq_ihh_6:

$$1/dw_ihh_7=1/wq_ihh_6 + (n_{sample\ i}/n_{population\ i})$$

The new design weight is only an intermediate step and therefore is not included in the data.

6.2 Design weights for the refreshment sample in wave 7

In wave 7, the panel was refreshed by sampling new households from new inflows to benefit receipt. All households that were receiving benefits in July 2012 but had had no probability of being selected for the register data sample in the same month in 2011, 2010, 2009, 2008, 2007 and 2006 had a likelihood of being selected. This refreshment could be achieved by selecting only benefit units in which no member was receiving benefits in July of the previous years. The refreshment sample was drawn from the 300 points of the first wave and the 100 replenishment points of wave 5. Analogous to the special pps procedure used to draw the first register data sample, which is described in Rudolph and Trappmann (2007), the sample size was proportional to the share of new benefit recipients in the population in the sampling point (at the time when the sampling points were selected). The calculation of the design weights is also described in the same article. For cases with *sample = 10*, the design weight of the refreshment sample is included in the variable *dw_ba*.

6.3 Propensity to participate again - households

In this step, again similar to the procedure in wave 6, the probability of re-participation in wave 7 was estimated for each household that participated in wave 6 based on logit models for willingness to participate in the panel, availability and participation. Additionally, households that participated in wave 5 but not in wave 6 (temporary non-responses) were considered in the modeling for wave 7. In addition to variables from the household and personal interviews with the head of the household conducted during the previous wave, other fieldwork variables were included, e.g., number of contact attempts. The estimated propensities of all three models were multiplied. The reciprocal value of this product can be found in the variable *hpbleib* for each wave. The longitudinal weight for a household from one of the samples of wave 1 for the total period possible [t_1, t_2, t_3, t_4, t_5, t_6, t_7] across all six waves can be obtained as the product of the cross-sectional weight to t_1, *hpbleib* (wave 1 to wave 2) and *hpbleib* (wave 2 to wave 3, etc.) (see also the PASS User Guide section 12 (Trappmann, 2013b)).

Table 35: Variable overview, codes and reference categories for logit models of re-participating households

Variable code and reference category	Explanation
alter_1	Household reference person (HRP) younger than 30 years
alter_2	HRP 30-39 years of age
alter_4	HRP 50-64 years of age
alter_5	HRP 65 years and older
Reference category	HRP 40-49 years of age
sex_1	HRP male
Reference category	HRP female
nichtdeutsch	HRP nationality other than German
Reference category	HRP German nationality or missing information
schulbil_1	School qualification HRP: no qualification
schulbil_2	School qualification HRP: lower secondary school
schulbil_4	School qualification HRP: college/university qualification
Reference category	School qualification HRP: intermediate secondary school/pupil
gesundheit_3	Subjective evaluation of the health state of the HRP: satisfactory
gesundheit_4	Subjective evaluation of the health state of the HRP: not so good
gesundheit_5	Subjective evaluation of the health state of the HRP: bad
Reference category	Subjective evaluation of the health state of the HRP: very good to good
zufrieden_1	General life satisfaction HRP: scale value 0-2
zufrieden_2	General life satisfaction HRP: scale value 3-5
zufrieden_3	General life satisfaction HRP: scale value 6-8
Reference category	General life satisfaction HRP: scale value 9-10
anz_0_3	Number of individuals in the household aged 0-3 years
anz_4_6	Number of individuals in the household aged 4-6 years
anz_7_14	Number of individuals in the household aged 7-14 years
anz_65	Number of individuals in the household aged 65 years and older
Reference category	Number of individuals in the household aged 15-64 years
eigentum	Type of residential property: proprietor
Reference category	Type of residential property: tenant, missing information
wnka_1	Number of “don’t know” and “details refused” responses in household and personal interviews of the HRP: none
wnka_3	Number of “don’t know” and “details refused” responses in household and personal interviews of the HRP: 11 and more
Reference category	Number of “don’t know” and “details refused” responses in household and personal interviews of the HRP: 1-10
hhincome_1	Household income: up to EUR 870
hhincome_2	Household income: EUR 871-1,400
hhincome_4	Household income: more than EUR 2,200
Reference category	Household income: EUR 1,401-2,200
alg2_1	UB II receipt of the household: current receipt of UB II
Reference category	UB II receipt of the household: no current receipt of UB II
stichprobe1	BA sample
stichprobe3	Refreshment sample (BA) wave 2
stichprobe4	Refreshment sample (BA) wave 3
stichprobe5	Refreshment sample (BA) wave 4
stichprobe6	Replenishment sample (EWO) wave 5
stichprobe7	Replenishment sample (BA) wave 5
stichprobe8	Refreshment sample (BA) wave 5
Stichprobe9	Refreshment sample (BA) wave 6
Reference category	Microm sample
anzkon_1	Number of contact attempts CATI/CAPI: 1 contact attempt
anzkon_3	Number of contact attempts CATI/CAPI: 4-9 contact attempts
anzkon_4	Number of contact attempts CATI/CAPI: 10 and more contact attempts
Reference category	Number of contact attempts CATI/CAPI: 2-3 contact attempts

Variable code and reference category	Explanation
blneualt_2	New federal states
Reference category	Old federal states
bundesld_1	Federal state: Schleswig-Holstein
bundesld_2	Federal state: Hamburg
bundesld_3	Federal state: Lower-Saxony
bundesld_4	Federal state: Bremen
bundesld_6	Federal state: Hesse
bundesld_7	Federal state: Rhineland-Palatinate
bundesld_8	Federal state: Baden-Wuerttemberg
bundesld_9	Federal state: Bavaria
bundesld_10	Federal state: Saarland
bundesld_11	Federal state: Berlin
bundesld_12	Federal state: Brandenburg
bundesld_13	Federal state: Mecklenburg-Vorpommern
bundesld_14	Federal state: Saxony
bundesld_15	Federal state: Saxony-Anhalt
bundesld_16	Federal state: Thuringia
Reference category	Federal state: North Rhine-Westphalia
bik_1	BIK size class of municipality: population of less than 2,000
bik_2	BIK size class of municipality: population of 2,000 to under 5,000
bik_3	BIK size class of municipality: population of 5,000 to under 20,000
bik_4	BIK size class of municipality: population of 20,000 to under 50,000
bik_5	BIK size class of municipality: population of 50,000 to under 100,000 STYP 2/3/4
bik_6	BIK size class of municipality: population of 50,000 to under 100,000 STYP 1
bik_7	BIK size class of municipality: population of 100,000 to under 500,000 STYP 2/ 3/ 4
bik_8	BIK size class of municipality: population of 100,000 to under 500,000 STYP 1
bik_9	BIK size class of municipality: population of 500,000 and more STYP 2/ 3/ 4
Reference category	BIK size class of municipality: population of 500,000 and more STYP 1

Table 36: Logit models on re-participation for willingness to participate in a panel, availability and participation

	Willingness to participate in the panel		Contact		Participation	
	Coef.	p	Coef.	p	Coef.	p
alter_1	.1653619	0.552	-.6149784	0.003	-.5200979	0.000
alter_2	.2861791	0.298	.0773882	0.730	-.1870612	0.056
alter_4	-.1077226	0.628	.3186933	0.134	.3645157	0.000
alter_5	.2560294	0.617	.4089901	0.559	.0802022	0.729
sex_1	-.040139	0.805	-.203618	0.160	-.0748417	0.253
nichtdeutsch	-.0029554	0.992	-.2017413	0.389	-.4289519	0.000
schulbil_1	.3807655	0.334	.0926146	0.782	-.4037096	0.003
schulbil_2	.2088833	0.305	-.2961028	0.088	-.0183351	0.821
schulbil_4	-.1124137	0.577	.0524427	0.788	.0818881	0.336
gesundheit_3	-.0364798	0.846	-.0695437	0.685	.0979859	0.201
gesundheit_4	.2056602	0.400	-.0748888	0.719	.0843137	0.370
gesundheit_5	-.1486741	0.642	-.20786	0.461	-.0319324	0.814
zufrieden_1	-.0205211	0.962	-.2333184	0.520	-.3313231	0.069
zufrieden_2	.2198598	0.430	.1580884	0.514	-.0950522	0.394
zufrieden_3	-.0063743	0.977	.1384873	0.489	-.0337765	0.709

	Willingness to participate in the panel		Contact		Participation	
	Coef.	p	Coef.	p	Coef.	p
anz_0_3	-.0705367	0.784	-.0580309	0.761	.1253476	0.172
anz_4_6	.3018359	0.396	.1449624	0.570	-.0411012	0.683
anz_7_14	-.242717	0.111	.014824	0.919	-.0232759	0.694
anz_65	-.3774934	0.180	.3183305	0.464	.1429383	0.285
eigentum	-.0469357	0.837	1.41043	0.000	.3759917	0.000
wnka_1	.442847	0.014	.0470867	0.759	.1311077	0.058
wnka_3	-.3161653	0.150	-.0095407	0.964	-.1289798	0.195
hhincome_1	-.6506646	0.012	-.4166859	0.052	.0428092	0.680
hhincome_2	-.4876026	0.036	-.0995725	0.613	-.0610413	0.494
hhincome_4	.2306914	0.385	.5225021	0.035	.0911237	0.343
alg2_1	.0943682	0.641	.1404639	0.405	.1622797	0.047
stichprobe1	.461706	0.124	-.2998475	0.229	-.1636364	0.142
stichprobe3	.2861768	0.567	-.3101952	0.377	-.4081909	0.016
stichprobe4	-.0709671	0.855	.0116141	0.973	-.1090852	0.499
stichprobe5	.1960883	0.657	-.2408514	0.494	-.2097362	0.201
stichprobe6	-.1036861	0.724	-.9199709	0.003	-.5208163	0.000
stichprobe7	-.0175324	0.959	-.0342918	0.915	-.4815003	0.000
stichprobe8	.219095	0.621	-.036684	0.917	-.5998574	0.000
Stichprobe9	-.8121316	0.005	-.4485471	0.106	-.6516609	0.000
bundesld_1	.4577384	0.385	-.8594901	0.007	.0509883	0.788
bundesld_2	-.7797506	0.039	.4072801	0.455	.2731528	0.239
bundesld_3	.5189818	0.102	-.5402043	0.023	.3144291	0.011
bundesld_4	-.9526218	0.054	.2554677	0.736	.17505	0.599
bundesld_6	.2028897	0.583	-.074883	0.830	.2168033	0.161
bundesld_7	.3052425	0.459	-.2196181	0.538	-.1174112	0.476
bundesld_8	.2411724	0.413	-.0763242	0.776	-.1104821	0.339
bundesld_9	.5319902	0.086	.1626458	0.542	.1844909	0.104
bundesld_10	-.6692604	0.169	-.0205511	0.978	.3172153	0.292
bundesld_11	-.1972091	0.519	.6408913	0.066	.3144364	0.043
bundesld_12	.5560753	0.210	-.0793074	0.830	.1746551	0.297
bundesld_13	.5401473	0.374	.1388815	0.799	.055332	0.802
bundesld_14	.7587425	0.087	.5491007	0.135	.4138011	0.011
bundesld_15	.9777823	0.065	.4162121	0.360	.4789828	0.008
bundesld_16	.8032965	0.130	.1394068	0.735	-.0264719	0.879
bik_1			.5075231	0.525	.3964418	0.272
bik_2			.3039467	0.545	.1567093	0.449
bik_3			.8261366	0.019	.0440056	0.745
bik_4			.2196707	0.421	.3184644	0.015
bik_5			-.105345	0.706	.1467814	0.296
bik_6			-.0189519	0.958	.7325545	0.000
bik_7			.3000301	0.236	-.0445475	0.698
bik_8			.5977443	0.010	.143903	0.168
bik_9			.258165	0.391	.1133093	0.411
anzkon_1			-1.394604	0.000		
anzkon_3			-.7207082	0.002		
anzkon_4			-2.210193	0.000		
cons	3.905161	0.000	4.770719	0.000	1.806383	0.000
n	9,510		9,341		9,112	
Log likelihood	-800.20325		-897.59042		-3400.0963	
PseudoR2	0.0570		0.1653		0.0537	

6.4 Propensity to participate - first-time interviewed split-off households

This step calculated the propensities to participate for new split-off households, i.e., households that are included in the panel due to the relocation of one individual of the panel sample in a new household. Here, only split-off households that had not been interviewed in the previous waves were considered. This condition means that the participation propensities for first-time participating split-off households were modeled separately following the criterion of originating in wave 6 (split-off W6 households) or originating in wave 7 (split-off W7 households). The probability of re-participation was estimated via logit models for availability and participation. Missing time-stable information on the household reference person (HRP) was added from the previous wave when necessary. The estimated propensities of the two models were multiplied. The reciprocal value of the product for the split-off households can also be found in the variable *hpbleib*.

Table 37: Variable overview, codes and reference categories for the logit models of the split-off households participating for the first time (waves 6 and 7)

Variable code and reference category	Explanation
alter_1 (Split W6) alter_2 (Split W6) alter_4 (Split W6) alter_5 (Split W6) Reference category (Split W6)	Household reference person (HRP) younger than 30 years HRP 30-39 years of age HRP 50-64 years of age HRP 65 years and older HRP 40-49 years of age
alter_1 (Split W7) alter_2 (Split W7) alter_3 (Split W7) Reference category (Split W7)	Household reference person (HRP) younger than 30 years HRP 30-39 years of age HRP 40-49 years of age HRP 50 years or older
sex_1 Reference category	HRP male HRP female
nichtdeutsch Reference category	HRP has nationality other than German HRP has German nationality or missing information
schulbil_1 (Split W6) schulbil_2 (Split W6) Reference category (Split W6)	School qualification HRP: no qualification, lower secondary school School qualification HRP: college/university qualification School qualification HRP: intermediate secondary school/still pupil
schulbil_1 (Split W7) schulbil_2 (Split W7) schulbil_4 (Split W7) Reference category (Split W7)	School qualification HRP: no qualification School qualification HRP: lower secondary school School qualification HRP: college/university qualification School qualification HRP: intermediate secondary school/still pupil
stichprobe_ba Reference category	BA samples (incl. BA refreshment samples and BA replenishment sample) Microm sample (incl. EWO replenishment sample)
anzkon_1 anzkon_3 anzkon_4 Reference category	Number of contact attempts CATI/CAPI: 1 contact attempt Number of contact attempts CATI/CAPI: 4-9 contact attempts Number of contact attempts CATI/CAPI: 10 and more contact attempts Number of contact attempts CATI/CAPI: 2-3 contact attempts

Table 38: Logit models on the first participation of split-off wave 6 households for availability and participation

	Contact		Participation	
	Coef.	**p**	**Coef.**	**p**
alter_1	.3681974	0.510	.5909438	0.205
alter_2	.6776091	0.289	.0174486	0.975
alter_4	.2260042	0.612	.2111833	0.607
alter_5	1.777198	0.106	.2666991	0.652
sex_1	-.9182511	0.019	-.0227698	0.947
nichtdeutsch	-.4533548	0.509	-.0324933	0.957
schulbil_1	.3088436	0.466	.394357	0.287
schulbil_3	.5333129	0.305	.3871969	0.350
anzkon_1	.3370282	0.578		
anzkon_3	.4077702	0.397		
anzkon_4	-.2761052	0.572		
stichprobe_ba	.1083233	0.779	-.4856419	0.141
cons	1.654118	0.001	-1.473673	0.000
n	292		256	
Log likelihood	-104.10601		-130.43528	
Pseudo R^2	0.0452		0.0207	

Table 39: Logit models on the first participation of split-off wave 7 households for availability and participation

	Contact		Participation	
	Coef.	**p**	**Coef.**	**p**
alter_1	.3472081	0.490	-.6535865	0.092
alter_2	.6295513	0.298	-1.130708	0.012
alter_3	.4326617	0.292	.0270757	0.914
sex_1	.3960823	0.290	.5039242	0.031
nichtdeutsch	.874886	0.413	-.783173	0.240
schulbil_1	.2355631	0.776	.8487859	0.129
schulbil_2	-.3201862	0.427	.0572184	0.832
schulbil_4	-.2179491	0.631	.3431401	0.238
anzkon_1	-1.637894	0.001		
anzkon_3	1.122442	0.305		
anzkon_4	-.3865559	0.567		
stichprobe_ba	-.8553874	0.042	-.5271163	0.024
cons	3.518005	0.000	-.6417301	0.014
n	459		417	
Log likelihood	-123.89058		-241.22498	
Pseudo R^2	0.1179		0.0558	

6.5 Nonresponse weighting for households from the BA refreshment sample and the BA panel replenishment sample of wave 7

Again, a two-stage nonresponse modeling for the households from the refreshment sample of BA new inflows into UB II receipt (sample = 10) was performed (availability and participation) similar to the wave 6 refreshment sample. The participation probability derived from this procedure can be found in variable *prop_t0*.

Table 40: Variable overview, codes and reference categories for the logit models of the BA refreshment sample of wave 7

Variable code and reference category	Explanation
alter_1	Household reference person (HRP) younger than 30 years
alter_2	HRP 30-39 years of age
alter_4	HRP 50-64 years of age
Reference category	HRP 40-49 years of age
sex_1	HRP male
Reference category	HRP female
nichtdeutsch	HRP has nationality other than German
Reference category	HRP has German nationality or missing information
schulbil_1	School qualification HRP: no qualification
schulbil_2	School qualification HRP: lower secondary school
schulbil_4	School qualification HRP: college/university qualification
schulbil_5	School qualification HRP: Details refused
Reference category	School qualification HRP: intermediate secondary school/still pupil
anz_persBG_2	Number of individuals in the benefit unit: 2 individuals
anz_persBG_3	Number of individuals in the benefit unit: 3 and more individuals
Reference category	Number of individuals in the benefit unit: 1 individual
anz_verwfBG_1	Number of individuals capable of work in the benefit unit: none
anz_verwfBG_3	Number of individuals capable of work in the benefit unit: 2 and more individuals
Reference category	Number of individuals capable of work in the benefit unit: 1 individual
BG_typ_2	Type of benefit unit: single parent
BG_typ_3	Type of benefit unit: couple without children
BG_typ_4	Type of benefit unit: couple with children under the age of 18
BG_typ_5	Type of benefit unit: other benefit unit
Reference category	Type of benefit unit: single
famstand_2	Marital status: married/ widowed
famstand_3	Marital status: widowed
famstand_4	Marital status: divorced
famstand_5	Marital status: separated
Reference category	Marital status: single
blneualt_2	Neue Bundesländer
Reference category	Alte Bundesländer
bundesld_1	Federal state: Schleswig-Holstein
bundesld_2	Federal state: Hamburg
bundesld_3	Federal state: Lower-Saxony
bundesld_4	Federal state: Bremen
bundesld_6	Federal state: Hesse
bundesld_7	Federal state: Rhineland-Palatinate
bundesld_8	Federal state: Baden-Wuerttemberg
bundesld_9	Federal state: Bavaria
bundesld_10	Federal state: Saarland
bundesld_11	Federal state: Berlin
bundesld_12	Federal state: Brandenburg
bundesld_13	Federal state: Mecklenburg-Vorpommern
bundesld_14	Federal state: Saxony
bundesld_15	Federal state: Saxony-Anhalt
bundesld_16	Federal state: Thuringia
Reference category	Federal state: North Rhine-Westphalia
bik_1	BIK size class of municipality: population of less than 2,000 to under 5,000 (BIK-Region size classes 1 and 2 combined)
bik_2	BIK size class of municipality: population of 5,000 to under 20,000
bik_3	BIK size class of municipality: population of 20,000 to under 50,000

bik_4	BIK size class of municipality: population of 50,000 to under 100,000 STYP 2/3/4
bik_5	BIK size class of municipality: population of 50,000 to under 100,000 STYP 1
bik_6	BIK size class of municipality: population of 100,000 to under 500,000 STYP 2/3/4
bik_7	BIK size class of municipality: population of 100,000 to under 500,000 STYP 1
bik_8	BIK size class of municipality: population of 500,000 and more STYP 2/ 3/ 4
Reference category	BIK size class of municipality: population of 500,000 and more STYP 1
anzkon_1	Number of contact attempts CATI/CAPI: 1 contact attempt
anzkon_3	Number of contact attempts CATI/CAPI: 4-9 contact attempts
anzkon_4	Number of contact attempts CATI/CAPI: 10 and more contact attempts
Reference category	Number of contact attempts CATI/CAPI: 2-3 contact attempts

Table 41: Logit models on the first participation for availability and participation of the BA refreshment sample and BA replenishment sample of wave 7

	Kontakt		Teilnahme	
	Coef.	p	Coef.	p
alter_1	.0524774	0.869	-.0393177	0.761
alter_2	.7726509	0.035	.0497856	0.704
alter_4	.69421	0.086	.2389005	0.070
sex_1	-.2146974	0.403	.0598636	0.509
nichtdeutsch	-.6031051	0.019	-.2447335	0.028
schulbil_1	.3224202	0.553	-.2902533	0.162
schulbil_2	.273708	0.490	-.1795233	0.198
schulbil_4	.5215917	0.291	.3374379	0.026
schulbil_5	-.0812563	0.811	-.2757269	0.024
anz_persBG_2	-.2865387	0.717	-.1084502	0.718
anz_persBG_3	-.6634868	0.449	-.2963462	0.395
anz_verwfBG_1	-.6156957	0.473	1.097764	0.004
anz_verwfBG_3	1.124908	0.094	-.1321264	0.537
BG_typ_2	.4954853	0.551	.4028605	0.203
BG_typ_3	-.5248671	0.598	.0334093	0.921
BG_typ_4	-.0947281	0.929	.5711397	0.119
BG_typ_5	.0086258	0.990	.0515203	0.847
famstand_2	-.2287179	0.629	.0809372	0.636
famstand_3	.2996569	0.525	.3105824	0.029
famstand_4	-.5577763	0.147	-.0171178	0.914
famstand_5	.3426494	0.686	.3900423	0.109
blneualt_2	.7359803	0.040		
bundesld_1			-.1092373	0.638
bundesld_2			-.0205502	0.938
bundesld_3			-.1159136	0.446
bundesld_4			-1.175796	0.031
bundesld_6			-.1908747	0.270
bundesld_7			-.0425545	0.845
bundesld_8			.0865975	0.577
bundesld_9			.1330312	0.354
bundesld_10			-.2107162	0.560
bundesld_11			.078331	0.655
bundesld_12			-.4259874	0.092
bundesld_13			-.373682	0.272

bundesld_14			.2784301	0.200
bundesld_15			.043926	0.865
bundesld_16			-.0934828	0.714
bik10_1	.6865896	0.513	.7437453	0.009
bik10_2	.2841948	0.544	.2765113	0.109
bik10_3	.4744491	0.393	.0736344	0.677
bik10_4	.1482443	0.772	.2405833	0.181
bik10_5	-.515828	0.290	-.1574245	0.548
bik10_6	.5717225	0.189	-.0975493	0.529
bik10_7	.0977069	0.745	.1867681	0.148
bik10_8	.3234643	0.490	.2219957	0.172
anzkon_1	-.9326998	0.116		
anzkon_3	.2509391	0.654		
anzkon_4	-2.562142	0.000		
cons	4.552326	0.000	-.8105567	0.000
n	3,020		2,926	
Log likelihood	-335.64027		-1795.3685	
Pseudo R^2	0.1983		0.0262	

6.6 Propensity to participate again - individuals

The decisive longitudinal weight is not the household but the individual-level weight because these units are stable over time. The propensities to participate again for individuals in wave 7 were estimated using additional personal characteristics via logit models for willingness to participate in the panel, availability and participation. The dependence of the personal sample conveyed via the household context and correction of the estimation of standard errors made necessary by it were considered in these models by clustering the error terms at the household level. The predicted propensities of the models were multiplied. The reciprocal value of this product can be found in variable *ppbleib*. The longitudinal weight for an individual for the period [t_1; t_2; t_3; t_4; t_5; t_6, t_7] across all six waves can be obtained as the product of the cross-sectional weight to t_1, *ppbleib* (wave 1 to wave 2) and *ppbleib* (wave 2 to wave 3, etc.).

Table 42: Variable overview, codes and reference categories for the logit models of re-participating individuals

Variable code and reference category	Explanation
alter_1	Individual younger than 30 years
alter_2	Individual 30-39 years of age
alter_4	Individual 50-64 years of age
alter_5	Individual 65 years and older
Reference category	Individual 40-49 years of age
sex_1	Individual male
Reference category	Individual female
nichtdeutsch	Individual has nationality other than German
Reference category	Individual has German nationality or missing information
schulbil_1	School qualification individual: no qualification
schulbil_2	School qualification individual: lower secondary school
schulbil_4	School qualification individual: college/university qualification
Reference category	School qualification individual: intermediate secondary school/still pupil
gesundheit_3	Subjective evaluation of the health state of the individual: satisfactory
gesundheit_4	Subjective evaluation of the health state of the individual: not so good
gesundheit_5	Subjective evaluation of the health state of the individual: bad
Reference category	Subjective evaluation of the health state of the individual: very good to good
zufrieden_1	General life satisfaction of the individual: scale value 0-2
zufrieden_2	General life satisfaction of the individual: scale value 3-5
zufrieden_3	General life satisfaction of the individual: scale value 6-8
Reference category	General life satisfaction of the individual: scale value 9-10
anz_0_3	Number of individuals in the household aged 0-3 years
anz_4_6	Number of individuals in the household aged 4-6 years
anz_7_14	Number of individuals in the household aged 7-14 years
anz_65	Number of individuals in the household aged 65 years and older
Reference category	Number of individuals in the household aged 15-64 years
eigentum	Type of residential property: proprietor
Reference category	Type of residential property: tenant, missing information
wnka_1	Number of “don’t know” and “details refused” responses in household and personal interviews of the individual: none
wnka_3	Number of “don’t know” and “details refused” responses in household and personal interviews of the individual: 11 and more
Reference category	Number of “don’t know” and “details refused” responses in household and personal interviews of the individual: 1-10
hhincome_1	Household income: up to EUR 870
hhincome_2	Household income: EUR 871-1,400
hhincome_4	Household income: more than EUR 2,200
Reference category	Household income: EUR 1,401-2,200
alg2_1	UB II receipt of the household: current receipt of UB II
Reference category	UB II receipt of the household: no current receipt of UB II
stichprobe1	BA sample
stichprobe3	Refreshment sample (BA) wave 2
stichprobe4	Refreshment sample (BA) wave 3
stichprobe5	Refreshment sample (BA) wave 4
stichprobe6	Replenishment sample (EWO) wave 5
stichprobe7	Replenishment sample (BA) wave 5
stichprobe8	Refreshment sample (BA) wave 5
stichprobe9	Refreshment sample (BA) wave 6
Reference category	Microm sample
anzkon_1	Number of contact attempts CATI/CAPI: 1 contact attempt
anzkon_3	Number of contact attempts CATI/CAPI: 4-9 contact attempts
anzkon_4	Number of contact attempts CATI/CAPI: 10 and more contact attempts

Reference category	Number of contact attempts CATI/CAPI: 2-3 contact attempts
blneualt_2	New federal states
Reference category	Old federal states
bundesld_1	Federal state: Schleswig-Holstein
bundesld_2	Federal state: Hamburg
bundesld_3	Federal state: Lower-Saxony
bundesld_4	Federal state: Bremen
bundesld_6	Federal state: Hesse
bundesld_7	Federal state: Rhineland-Palatinate
bundesld_8	Federal state: Baden-Wuerttemberg
bundesld_9	Federal state: Bavaria
bundesld_10	Federal state: Saarland
bundesld_11	Federal state: Berlin
bundesld_12	Federal state: Brandenburg
bundesld_13	Federal state: Mecklenburg-Vorpommern
bundesld_14	Federal state: Saxony
bundesld_15	Federal state: Saxony-Anhalt
bundesld_16	Federal state: Thuringia
Reference category	Federal state: North Rhine-Westphalia
bik_1	BIK size class of municipality: population of less than 2,000
bik_2	BIK size class of municipality: population of 2,000 to under 5,000
bik_3	BIK size class of municipality: population of 5,000 to under 20,000
bik_4	BIK size class of municipality: population of 20,000 to under 50,000
bik_5	BIK size class of municipality: population of 50,000 to under 100,000 STYP 2/3/4
bik_6	BIK size class of municipality: population of 50,000 to under 100,000 STYP 1
bik_7	BIK size class of municipality: population of 100,000 to under 500,000 STYP 2/3/4
bik_8	BIK size class of municipality: population of 100,000 to under 500,000 STYP 1
bik_9	BIK size class of municipality: population of 500,000 and more STYP 2/ 3/ 4
Reference category	BIK size class of municipality: population of 500,000 and more STYP 1

Table 43: Logit models on re-participation for willingness to participate in a panel, availability and participation

	Willingness to participate in the panel		Contact		Participation	
	Coef.	p	Coef.	p	Coef.	p
alter_1	.5387252	0.011	-.3278182	0.059	-.5817339	0.000
alter_2	.2969068	0.241	.0504754	0.804	-.1311099	0.121
alter_4	-.014256	0.946	.4099556	0.044	.3215936	0.000
alter_5	.4157218	0.411	-.2043568	0.676	.1033921	0.529
sex_1	.0264776	0.822	-.0314522	0.766	-.0642579	0.121
nichtdeutsch	-.1599025	0.549	-.2698596	0.292	-.4227938	0.000
schulbil_1	.0655585	0.839	-.1702726	0.559	-.4426275	0.000
schulbil_2	.359111	0.060	-.2020069	0.200	-.0909565	0.147
schulbil_4	-.094674	0.631	-.0014163	0.994	-.0157743	0.813
gesundheit_3	.0676779	0.695	.0641359	0.693	.0721477	0.203
gesundheit_4	.4135375	0.073	-.0649727	0.719	.1627175	0.029
gesundheit_5	-.0363564	0.904	-.2252076	0.405	.03563	0.750
zufrieden_1	.0534566	0.904	-.3381872	0.342	-.3371156	0.030
zufrieden_2	.0974879	0.716	-.0362204	0.876	-.1653499	0.064
zufrieden_3	-.2556492	0.241	-.1365815	0.485	-.0220317	0.742
alg2_1	-.0736926	0.752	.0730625	0.691	.0384422	0.622

eigentum	.2298709	0.340	1.543306	0.000	.3581379	0.000
anz_0_3	-.0544564	0.825	-.0793356	0.649	.1612564	0.051
anz_4_6	.5209505	0.156	.3410723	0.240	-.0809624	0.343
anz_7_14	-.302593	0.071	.132697	0.462	-.0146615	0.777
anz_65	-.4775012	0.123	.7654526	0.044	.0708246	0.473
wnka_1	.1963329	0.251	-.0266136	0.847	.2103417	0.000
wnka_3	-.2400594	0.258	-.0411512	0.824	-.0817277	0.295
hhincome_1	-.9115488	0.001	-.4480155	0.042	.2947801	0.003
hhincome_2	-.7901132	0.001	-.1340456	0.523	.0623063	0.461
hhincome_4	.1366573	0.665	.5314825	0.063	.0224828	0.792
stichprobe1	.5247436	0.083	-.4712072	0.067	-.2121834	0.037
stichprobe3	.2140845	0.684	-.3452794	0.341	-.4768145	0.002
stichprobe4	-.1152575	0.790	-.2505105	0.521	-.0643513	0.666
stichprobe5	-.3644285	0.502	-.2649503	0.469	-.3807125	0.010
stichprobe6	-.150185	0.637	-1.280175	0.000	-.4148683	0.000
stichprobe7	.0379638	0.918	-.1350946	0.684	-.463734	0.000
stichprobe8	.0796245	0.867	-.232847	0.535	-.6372529	0.000
stichprobe9	-.9121166	0.002	-.8027169	0.004	-.6023137	0.000
blneualt_2	.344365	0.106	.4374013	0.026		
anzkon_1			-1.809857	0.000		
anzkon_3			-.7552675	0.005		
anzkon_4			-2.20168	0.000		
bundesld_1					.3516475	0.054
bundesld_2					.1259233	0.594
bundesld_3					.3252159	0.004
bundesld_4					.2843329	0.344
bundesld_6					.1475666	0.286
bundesld_7					.0801895	0.618
bundesld_8					-.0703863	0.505
bundesld_9					.2622488	0.012
bundesld_10					.4087855	0.126
bundesld_11					.3189031	0.039
bundesld_12					.3231011	0.041
bundesld_13					.2209192	0.284
bundesld_14					.3734952	0.007
bundesld_15					.4425552	0.008
bundesld_16					.2111018	0.190
bik_1			-.0115073	0.988	.7868304	0.017
bik_2			.2035832	0.689	.0208119	0.910
bik_3			.5210306	0.147	-.0226748	0.850
bik_4			.0817178	0.776	.1445965	0.198
bik_5			-.3282069	0.285	.0829867	0.518
bik_6			-.0391836	0.910	.6088737	0.001
bik_7			-.0719944	0.780	-.1055074	0.322
bik_8			.4418322	0.056	.1958517	0.053
bik_9			-.0684688	0.827	.0901113	0.467
cons	4.452159	0.000	5.139258	0.000	1.630923	0.000
n	14,619		14,411		14,117	
Log likelihood	-1030.2564		-1198.5188		-5894.6493	
Pseudo R^2	0.0557		0.1650		0.0561	

Note: The correction of standard errors was made by means of an estimation clustered across households.

6.7 Integration of the weights to yield the total weight before calibration

This step again involved combining the household weights of the new replenishment and panel household samples (including the refreshments from waves 2 to 6) that were modified by the non-response modeling. The multiple selection probability of a sampled benefit recipient living in the same household as a benefit recipient in previous years without being a member of the benefit unit himself/herself was ignored. The new design weights of the benefit recipient sample are projected in the cross-section to all individuals who were living in a household that included at least one benefit unit in either July 2006, in July 2007, in July 2008, in July 2009, in July 2010, in July 2011, or in July 2012. It is only when calculating new weights for the total sample that it becomes necessary to adjust the weights for all households receiving benefits in July 2012. For this adjustment, the inclusion probability in the other sample was estimated for cases from the Microm sample (wave 1), EWO replenishment sample (wave 5) and new refreshment sample (wave 7). For cases from the refreshment sample, the mean wave 1 selection probability in the Microm sample respectively, the mean wave 5 selection probability of EWO refreshment in the respective postcode area and the average participation probability (for waves 1 to 7) in that sample were assumed. For cases from the Microm sample, if they are (according to survey data) new recipients of UB II who first received the benefit between the last five sampling dates (waves 2, 3, 4, 5, 6 and 7), the mean selection probability of a household in the refreshment sample in the respective postcode area and the average participation probability in that sample were assumed. The two weights were then integrated to form a new total weight.

6.8 Integration of temporary non-responses (households)

Households that skipped one wave - i.e., did not participate (temporary non-responses) - could participate again in wave 7, as was possible in wave 6. No longitudinal weights are calculated for these households, i.e., (weighted) longitudinal evaluations can only be made with participants across all waves in question. Non-participation of a household can only occur in one wave; if a household skips two consecutive waves, it will no longer be contacted. To calculate mutual cross-sectional weights including the temporary non-responses, there was a convex combination of the modified household weights of the temporary non-responses and the modified household weights of the panel household sample (not of the refreshment sample) before calibration. Thus, the convex combination of the household weights was made before calibration; the calibration was then made with the new combined household weights.

Although the household weights modified by non-response modeling already serve as projection factors for the panel and refreshment sample, it was necessary to calculate such modified household weights as an estimator for the respective population again for the temporary non-responses. The starting point was the calibrated household weights of wave 5 (wave 6 is the temporary non-response).

For temporary non-responses, the probability of non-participation in wave 6 in case of participation in wave 5 (non-participation propensities wave 6) and the probability of participa-

tion in wave 7 in case of a non-participation in wave 6 (participation propensities wave 7) was determined. The probability of non-participation in wave 6 is calculated from 1– participation probability in wave 6.

The described propensities for participation and non-participation were estimated via logit models. The estimated probabilities of the respective models were multiplied. The modified household weight of the temporary non-responses was then calculated by multiplying the calibrated household weights of wave 5 by the reciprocal value of this product.

Table 44: Variable overview, codes and reference categories for the logit models of the temporary non-responses

Variable code and reference category	Explanation
alter_1	Household reference person (HRP) younger than 30 years
alter_2	HRP 30-39 years of age
alter_4	HRP 50-64 years of age
alter_5	HRP 65 years and older
Reference category	HRP 40-49 years of age
sex_1	HRP male
Reference category	HRP female
nichtdeutsch	HRP has nationality other than German
Reference category	HRP has German nationality or missing information
schulbil_1	School qualification HRP: no qualification
schulbil_2	School qualification HRP: lower secondary school
schulbil_4	School qualification HRP: college/university qualification
Reference category	School qualification HRP: intermediate secondary school/still pupil
gesundheit_3	Subjective evaluation of the health state of the HRP: satisfactory
gesundheit_4	Subjective evaluation of the health state of the HRP: not so good
gesundheit_5	Subjective evaluation of the health state of the HRP: bad
Reference category	Subjective evaluation of the health state of the HRP: very good to good
zufrieden_1	General life satisfaction HRP: scale value 0-2
zufrieden_2	General life satisfaction HRP: scale value 3-5
zufrieden_3	General life satisfaction HRP: scale value 6-8
Reference category	General life satisfaction HRP: scale value 9-10
anz_0_3	Number of individuals in the household aged 0-3 years
anz_4_6	Number of individuals in the household aged 4-6 years
anz_7_14	Number of individuals in the household aged 7-14 years
anz_65	Number of individuals in the household aged 65 years and older
DinvalidAge	Age responses that cannot be evaluated
Reference category	Number of individuals in the household aged 15-64 years
eigentum	Type of residential property: proprietor
Reference category	Type of residential property: tenant, missing information
wnka_1	Number of “don’t know” and “details refused” responses in household and personal interviews of the HRP: none
wnka_3	Number of “don’t know” and “details refused” responses in household and personal interviews of the HRP: 11 and more
Reference category	Number of “don’t know” and “details refused” responses in household and personal interviews of the HRP: 1-10
hhincome_1	Household income: up to EUR 870
hhincome_2	Household income: EUR 871-1,400
hhincome_4	Household income: more than EUR 2,200
Reference category	Household income: EUR 1,401-2,200
alg2_1	UB II receipt of the household: current receipt of UB II

Reference category	UB II receipt of the household: no current receipt of UB II
bundesld_1	Federal state: Schleswig-Holstein
bundesld_2	Federal state: Hamburg
bundesld_3	Federal state: Lower-Saxony
bundesld_4	Federal state: Bremen
bundesld_6	Federal state: Hesse
bundesld_7	Federal state: Rhineland-Palatinate
bundesld_8	Federal state: Baden-Wuerttemberg
bundesld_9	Federal state: Bavaria
bundesld_10	Federal state: Saarland
bundesld_11	Federal state: Berlin
bundesld_12	Federal state: Brandenburg
bundesld_13	Federal state: Mecklenburg-Vorpommern
bundesld_14	Federal state: Saxony
bundesld_15	Federal state: Saxony-Anhalt
bundesld_16	Federal state: Thuringia
Reference category	Federal state: North Rhine-Westphalia
bik_1	BIK size class of municipality: population of less than 2,000
bik_2	BIK size class of municipality: population of 2,000 to under 5,000
bik_3	BIK size class of municipality: population of 5,000 to under 20,000
bik_4	BIK size class of municipality: population of 20,000 to under 50,000
bik_5	BIK size class of municipality: population of 50,000 to under 100,000 STYP 2/3/4
bik_6	BIK size class of municipality: population of 50,000 to under 100,000 STYP 1
bik_7	BIK size class of municipality: population of 100,000 to under 500,000 STYP 2/3/4
bik_8	BIK size class of municipality: population of 100,000 to under 500,000 STYP 1
bik_9	BIK size class of municipality: population of 500,000 and more STYP 2/ 3/ 4
Reference category	BIK size class of municipality: population of 500,000 and more STYP 1

Table 45: Logit models of temporary non-responses

	Re-participation in wave 6 to determine the W6 non-participation probability (1-participation probability W6)		Re-participation in wave 7 in case of non-participation in wave 6	
	Coef.	**p**	**Coef.**	**p**
alter_1	-.4743669	0.000	.2456239	0.145
alter_2	-.3288472	0.000	.1417846	0.379
alter_4	.2784932	0.000	.1256505	0.424
alter_5	.3598132	0.040	-.1809721	0.678
sex_1	-.0121581	0.809	-.2014488	0.066
not German	-.2051992	0.025	-.4252032	0.038
schulbil_1	-.4059292	0.000	-.471519	0.051
schulbil_2	-.114012	0.063	-.2872954	0.029
schulbil_4	.0378661	0.570	-.165429	0.249
gesundheit_3	.0880068	0.129	.1988218	0.107
gesundheit_4	.2645713	0.000	.0172278	0.917
gesundheit_5	.1007179	0.339	-.0734729	0.755
zufrieden_1	-.2049899	0.143	.1089547	0.727
zufrieden_2	-.1130124	0.185	.3371563	0.078
zufrieden_3	.0971773	0.170	.3666297	0.027
anz_0_3	.0817801	0.256	.01226	0.935
anz_4_6	.0408614	0.602	.1361519	0.360
anz_7_14	.1446098	0.003	.2017501	0.045
anz_65	.1139695	0.254	-.1805836	0.478
eigentum	.0257483	0.711	-.3977567	0.016
wnka_1	.1650295	0.002	.1639676	0.165
wnka_3	-.3501697	0.000	-.3164336	0.041
hhincome_2	.035456	0.622	.2841929	0.067
hhincome_4	.1627327	0.029	.1539175	0.365
alg2_1	.0488758	0.444	.1025195	0.447
bundesld_1	-.2516076	0.064	.3974806	0.153
bundesld_2	-.0911938	0.582	-.2929243	0.456
bundesld_3	.0287209	0.759	.1998363	0.348
bundesld_4	-.0517554	0.835	.5663817	0.264
bundesld_6	-.0961587	0.393	.4309436	0.070
bundesld_7	-.3104453	0.016	.4521949	0.086
bundesld_8	.2183348	0.033	.5228321	0.019
bundesld_9	.0257511	0.770	.3711294	0.058
bundesld_10	-.3460677	0.085	.2707379	0.519
bundesld_11	.0244938	0.837	.0426199	0.871
bundesld_12	.0618555	0.620	.1180042	0.673
bundesld_13	.0742838	0.661	.8754297	0.008
bundesld_14	.1267097	0.271	.4842854	0.041
bundesld_15	.4178095	0.003	.2245659	0.462
bundesld_16	.2451664	0.082	.1467476	0.640
bik_1	.5718375	0.023	-.4980492	0.396
bik_2	.1767954	0.292	.4842978	0.167
bik_3	.4054323	0.000	-.051608	0.840
bik_4	.1320543	0.177	-.0805754	0.712
bik_5	.2219597	0.039	.1367682	0.556
bik_6	.1770547	0.206	.506376	0.067
bik_7	.1411398	0.125	.2816695	0.149
bik_8	.1408248	0.085	.0505897	0.772

bik_9	.1183753	0.258	-.3528978	0.153
cons	.9612614	0.000	-1.909141	0.000
n	10,235		2,263	
Log likelihood	-5205.7308		-1125.6561	
Pseudo R^2	0.0373		0.0480	

The convex combination of the weights of the participants across all waves (panel household sample) and the temporary non-responses was made for the weights of all three subsamples i (Microm, BA and total) by multiplying the respective modified household weights by the share of the panel household sample or the temporary non-responses from the total sample, i.e., the sum of the panel household sample and temporary non-responses:

$dw_i hh_{temp.Ausfall} * (n_{temp.Ausfall\ i} / (n_{temp.Ausfall\ i} + n_{Bestand\ i}))$ for temporary non-responses and

$dw_i hh_{Bestand} * (n_{Bestand\ i} / (n_{temp.Ausfall\ i} + n_{Bestand\ i}))$ for the panel household sample.

6.9 Calibration to the household weight, wave 7, cross-section

Another calibration of the modified design weights, including the non-response weighting at the household level using the GREG procedure to the benchmark values of the Federal Statistical Office for 2012, followed. For households receiving benefits the weights were adjusted to the statistics of the Federal Employment Agency for July 2012. As in the previous year, the increase in UB II receipt since the previous year at the level of benefit units (280,269) was also included as an additional benchmark value in the total sample. Cases in the previous samples from waves 1 to 7 that, according to wave 7 of the survey, were receiving UB II in July 2012, will be projected to the benchmark statistics of the Federal Employment Agency on UB II.

The main objective of weighting is to balance distortions arising from the sample design (with different selection probabilities) and through selective participation or non-participation. By using the weights, population values from the sample can be estimated in an unbiased way. If the weights show a high variance, a large variance of the estimation functions can result. This is the trade-off between bias and variance so typical for statistics. The weighting reduces the bias; however, a too-severe increase in the variance caused by weighting is also to be avoided. Therefore, attempts are made to avoid very large weighting factors (and subsequently, very small factors) whenever possible and to make appropriate corrections to the weights if necessary. Within the framework of the calibration at hand, these corrections are made at two points:

- The input weights for the calibration (the modified design weights after considering non-response analyses) were trimmed before calibration, i.e., they were replaced by new input weights. The maximum and minimum of the trimmed design weights were determined by using particular percentiles of the distribution depending on the distribution of the design weights.

- In addition, the interval of weights was limited during calibration, i.e., a maximum and a minimum limit for weights was determined. Here, the total width of the weights was determined; the range of the pure calibration weights can be calculated from the relation of original weights to the trimmed input weight. Notably, narrower limits for the weights result in less variance of the weights and thus less variance of the estimations; too-narrow limits can, however, make the calibration of all benchmark values impossible.

To evaluate the weights, in addition to the average value and the standard deviation, the efficiency measure (E) is described as follows. The efficiency measure E is based on the variance of the weighting factor. The efficiency measure indicates the size of the effective case number of a passive characteristic that does not correlate with active characteristics when using the weight. The effective case number is the number of respondents who would have produced the same sample error in an unlimited random sample given the variance of the characteristic in the sample. The efficiency measure expresses the relation of n to n‘ as percentage.

6.10 Calibration of the BA sample

The population of the cumulated BA sample of all six waves consists of all of the households in Germany with at least one benefit unit receiving benefits according to SGB II at one of the (until now) seven drawing dates (July 2006, July 2007, July 2008, July 2009, July 2010, July 2011 or July 2012). In wave 7, only the benchmark values of the BA statistics from July 2012 are calibrated. The calibration thus only influences the weights of the households from the BA sample in which at least one benefit unit receiving benefits according to SGB II was living in July 2012. The starting points for the calibration were modified design weights, including the non-response weighting. The modified design weights were trimmed at the fifth and ninety-fifth percentiles of their distribution and then rescaled so that they totaled the untrimmed design weights. The projection factors of the trimmed design weights range from 182.89 to 3,457.02. The relation between the total projection factors after calibration and the trimmed design weights was limited downwards to 0.2 and upwards to 5.0. Thus, the total projection factors after calibration lie between a minimum of 36.58 and a maximum of 4,365.07.

A calibration was made for the following characteristics:

Benefit unit basis BA statistics:

- Increase in BU UB II recipients
- Number of BCs receiving benefits according to SGB II by federal states
- Number of BCs receiving benefits according to SGB II by number of individuals under 65 years of age in the benefit unit and by west/east
- Number of BCs receiving benefits according to SGB II by number of children under 15 years of age in the benefit unit and by west/east

- Number of BCs receiving benefits according to SGB II consisting of a single parent with child(ren), by west/east

As in the previous year, an additional benchmark was included. This is the increase in UB II recipients since the previous year at the level of benefit units (280,269).

For the calibration, the benchmark variable for each household must have a valid value. Therefore, the very low nonresponse item was imputed before calibration. The imputation was made by means of the average value and the modal value of the respective variable.

Because the imputation only serves the feasibility of the calibration, the imputed values were set back to missing values after the calibration. A projection with the calibrated weights without considering the nonresponse item thus leads to slight deviations from the values as presented in the following.

Table 46: Nominal distributions and distributions after calibration (BA sample, households)

Benchmark figure	Characteristics benchmark figure from BA statistics	Unweighted distribution	Nominal values from BA statistics	Distribution with calibrated weights
Number BCs receiving benefits in accordance with SGB II by federal states (16 categories)	Number BCs Schleswig-Holstein	118	117.580	117.580
	Number BCs Hamburg	103	100.580	100.580
	Number BCs Lower-Saxony	351	307.054	307.054
	Number BCs Bremen	43	50.422	50.422
	Number BCs North Rhine-Westphalia	920	817.347	817.347
	Number BCs Hesse	192	204.784	204.784
	Number BCs Rhineland-Palatinate	134	113.839	113.839
	Number BCs Baden-Wuerttemberg	274	226.754	226.754
	Number BCs Bavaria	341	231.483	231.483
	Number BCs Saarland	54	41.052	41.052
	Number BCs Berlin	278	320.929	320.929
	Number BCs Brandenburg	199	152.943	152.943
	Number BCs Mecklenburg-Vorpommern	90	114.812	114.812
	Number BCs Saxony	254	247.132	247.132
	Number BCs Saxony-Anhalt	209	170.404	170.404

	Number BCs Thuringia	133	111.744	111.744
Number of BCs receiving benefits in accordance with SGB II by number of individuals under 65 years of age in the benefit unit (1, 2, 3, 4, and “5 or more”) and by west/east (10 categories)	Number BCs with 1 individual under 65 (west)	1.072	1.214.722	1.214.722
	Number BCs with 2 individuals under 65 (west)	654	448.390	448.390
	Number BCs with 3 individuals under 65 (west)	437	269.869	269.869
	Number BCs with 4 individuals under 65 (west)	199	162.758	162.758
	Number BCs with 5 or more individuals under 65 (west)	168	115.156	115.156
	Number BCs with 1 individual under 65 (east)	549	666.329	666.329
	Number BCs with 2 individuals under 65 (east)	328	235.221	235.221
	Number BCs with 3 individuals under 65 (east)	168	117.707	117.707
	Number BCs with 4 individuals under 65 (east)	70	62.506	62.506
	Number BCs with 5 or more individuals under 65 (east)	48	36.201	36.201
Number of BCs receiving benefits in accordance with SGB II by number of individuals under 15 years of age in the benefit unit (0, 1, 2, 3, “4 or more”) and by west/east (10 categories)	Number BCs without children under 15 years of age (west)	1.691	1.517.424	1.517.424
	Number BCs with 1 child under 15 years of age (west)	495	377.065	377.065
	Number BCs with 2 children under 15 years of age (west)	227	211.713	211.713
	Number BCs with 3 children under 15 years of age (west)	83	75.420	75.420
	Number BCs with 4 or more children under 15 years of age (west)	34	29.273	29.273
	Number BCs without children under 15 years of age (east)	862	826.113	826.113
	Number BCs with 1 child under 15 years of age (east)	183	171.066	171.066
	Number BCs with 2 children under 15 years of age (east)	87	84.802	84.802
	Number BCs with 3 children under 15 years of age (east)	20	25.792	25.792
	Number BCs with 4 or more children under 15 years of age (east)	11	10.191	10.191
Number BCs receiving benefits in accordance with SGB II consisting of a single parent with children by west/east (4 categories)	Number BCs with a single parent (west)	595	442.433	442.433
	Rest BCs without a single parent (west)	1.935	1.768.462	1.768.462
	Number BCs with a single parent (east)	194	184.821	184.821
	Rest BCs without a single parent (east)	969	933.143	933.143

Table 47: Parameters of distribution of weights

1% percentile	113,1369
5% percentile	157,8378
10% percentile	185,2142
25% percentile	270,791
50% percentile	496,3308
75% percentile	1290,896
90% percentile	2324,095
95% percentile	3127,467
99% percentile	3598,883
Mean	916,3456
Standard deviation	915,9244
Minimum	36,57862
Maximum	4265,073
Number of observations	3571
Efficiency measure	**49,48%**

6.11 Population sample

All private households in Germany form the population. The starting points for the calibration were modified design weights, including the nonresponse weighting. The modified design weights were trimmed at the fifth and ninety-fifth percentiles of their distribution and after that rescaled so that they totaled the untrimmed design weights. The projection factors of the trimmed design weights range from 2,421.07 to 34,552.18. The relation between the total projection factors after calibration and the trimmed design weights was limited downwards to 0.2 and upwards to 4.0. Thus, the total projection factors after calibration lie between minimal 484.21 and maximal 58,343.09.

A calibration was made for the following characteristics:

Benefit units based on BA statistics:

- Number of BCs receiving benefits according to SGB II by federal states
- Number of BCs receiving benefits according to SGB II by number of individuals under 65 years of age in the benefit unit and by west/east
- Number of BCs receiving benefits according to SGB II by number of children under 15 years of age in the benefit unit and by west/east
- Number of BCs receiving benefits according to SGB II consisting of a single parent with child(ren), by west/east

Households based on Mikrozensus 2012:

- Number of households by federal state and BIK type
- Number of households by household size and west/east
- Number of households by “children under 15 years of age in the household yes/no” and west/east

For the calibration, each benchmark variable for each household must have a valid value. Therefore, the very low nonresponse item was imputed before calibration. The imputation was made by means of the average value and the modal value of the respective variable.

Because the imputation only serves the feasibility of the calibration, the imputed values were set back to missing values after the calibration. A projection with the calibrated weights without considering the nonresponse item thus leads to slight deviations from the values as presented in the following.

Table 48: Nominal distributions and distributions after calibration (population sample, households)

Benchmark figure	Characteristics benchmark figure from BA statistics and Mikrozensus 2012	Unweighted distribution	Nominal values	Distribution with calibrated weights
Number of BCs receiving benefits in accordance with SGB II by west/east	Number BGs west	137	2.210.895	2.210.181
	Number BGs east	54	1.117.964	1.118.677
Number of BCs receiving benefits in accordance with SGB II by number of individuals under 65 years of age in the benefit unit	Number BCs with 1 individual under 65 (west)	65	1.881.051	1.881.764
	Number BCs with 2 individuals under 65 (west)	52	683.611	683.136
	Number BCs with 3 individuals under 65 (west)	39	387.576	387.338
	Number BCs with 4 or more individuals under 65 (west)	35	376.621	376.621
Number of BCs receiving benefits in accordance with SGB II by number of individuals under 15 years of age in the benefit unit	Number BCs without children under 15 years of age (west)	130	2.343.537	2.343.774
	Number BCs with 1 child or more under 15 years of age (west)	61	985.322	985.084

Number BCs receiving benefits in accordance with SGB II consisting of a single parent with children	Number BCs with a single parent	42	627.254	627.016
	Rest BCs without a single parent	149	2.701.605	2.701.842
Number of households by federal state and BIK type (spelling: “Federal state. BIK type”)	1.1 bis 1.6	29	457.000	457.000
	1.7 bis 1.10	70	924.000	924.000
	2.10	41	989.000	989.000
	3.1 bis 3.5	119	1.408.000	1.408.000
	3.7 bis 3.8	159	1.389.000	1.389.000
	3.9 bis 3.10	98	1.030.000	1.030.000
	4.8 bis 4.10	24	361.000	361.000
	5.2 bis 5.4	115	1.129.000	1.129.000
	5.5 bis 5.6	98	993.000	993.000
	5.7 bis 5.8	223	2.945.000	2.945.000
	5.9 bis 5.10	297	3.554.000	3.554.000
	6.1 bis 6.4	79	633.000	633.000
	6.5 bis 6.8	84	1.111.000	1.111.000
	6.9 bis 6.10	92	1.197.000	1.197.000
	7.1 bis 7.6	76	909.000	909.000
	7.7 bis 7.10	86	962.000	962.000
	8.1 bis 8.4	122	1.227.000	1.227.000
	8.5 bis 8.8	140	2.168.000	2.168.000
	8.9 bis 8.10	122	1.621.000	1.621.000
	9.1 bis 9.4	151	1.502.000	1.502.000
	9.5 bis 9.7	180	1.505.000	1.505.000
	9.8 bis 9.9	123	1.414.000	1.414.000
	9.10	108	1.560.000	1.560.000

	10.3 bis 10.8	40	492.000	492.000
	11.10	113	2.015.000	2.015.000
	12.1 bis 12.4	37	462.000	462.000
	12.5 bis 12.7	35	271.000	271.000
	12.9 bis 12.10	38	518.000	518.000
	13.1 bis 13.6	26	518.000	518.000
	13.7 bis 13.9	27	327.000	327.000
	14.1 bis 14.4	67	592.000	592.000
	14.5 bis 14.8	37	671.000	671.000
	14.9 bis 14.10	67	933.000	933.000
	15.1 bis 15.4	42	377.000	377.000
	15.5 bis 15.7	59	533.000	533.000
	15.8 bis 15.9	28	265.000	265.000
	16.1 bis 16.4	75	518.000	518.000
	16.5 bis 16.8	78	594.000	594.000
Number of households by household size (1,2,3,4,“5 and more individuals“) and west/east (10 categories)	Number households with 1 individual (west)	695	12.308.000	12.308.000
	Number households with 2 individuals (west)	1.047	10.809.000	10.809.000
	Number households with 3 individuals (west)	422	3.989.000	3.989.000
	Number households with 4 individuals (west)	352	3.193.000	3.193.000
	Number households with 5 or more individuals (west)	160	1.181.000	1.181.000
	Number households with 1 individual (east)	213	3.730.000	3.730.000
	Number households with 2 individuals (east)	322	3.118.000	3.118.000
	Number households with 3 individuals (east)	109	1.057.000	1.057.000
	Number households with 4 individuals (east)	57	538.000	538.000
	Number households with 5 and more individuals (east)	28	151.000	151000
Number of households by “children under 15 years of age in the house-	Number households with children under 15 (west)	556	5.515.000	5.515.000
	Number households without children under 15 (west)	2.120	25.965.000	25.965.000

hold yes/no" and west/east	Number households with children under 15 (east)	120	1.265.000	1.265.000
	Number households without children under 15 (east)	609	7.329.000	7.329.000

Table 49: Parameters of distribution of weights

1% percentile	2022,851
5% percentile	2554,923
10% percentile	3083,755
25% percentile	4964,543
50% percentile	8437,925
75% percentile	15062,21
90% percentile	26920,13
95% percentile	32953,87
99% percentile	41051,79
Mean	11769,16
Standard deviation	9581,802
Minimum	484,2139
Maximum	58343,09
Number of observations	3405
Efficiency measure	**60,1%**

6.12 Total sample

All of the private households in Germany form the population. The starting points for the calibration were modified design weights, including the non-response weighting. The modified design weights were trimmed at the fifth and ninety-fifth percentiles of their distribution and after that rescaled so that they totaled the untrimmed design weights. The projection factors of the trimmed design weights range from 148.27 to 22,500.0. The relation between the total projection factors after calibration and the trimmed design weights was limited downwards to 0.3 and upwards to 3.0. Thus, the total projection factors after calibration lie between min. 44.48 and max. 34,347.43.

A calibration was made for the following characteristics:

Benefit unit basis BA statistics:

- Number of BCs receiving benefits according to SGB II by federal states

- Number of BCs receiving benefits according to SGB II by number of individuals under 65 years of age in the benefit unit and by west/east
- Number of BCs receiving benefits according to SGB II by number of children under 15 years of age in the benefit unit and by west/east
- Number of BCs receiving benefits according to SGB II consisting of a single parent with child(ren), by west/east

Household basis Mikrozensus 2012:

- Number of households by federal state and BIK type
- Number of households by household size and west/east
- Number of households by “children under 15 years of age in the household yes/no” and west/east

In addition, the increase in UB II recipients since the previous year at the level of benefit units (280,269) was included as an additional benchmark value in the total sample.

For the calibration, each benchmark variable for each household must have a valid value. Therefore, the very low non-response item was imputed before calibration. The imputation was made by means of the average value and the modal value of the respective variable.

Because the imputation only serves the feasibility of the calibration, the imputed values were set back to missing values after the calibration. A projection with the calibrated weights without considering the non-response item thus leads to slight deviations from the values as presented below.

Table 50: Nominal distributions and distributions after calibration (total sample, households)

Benchmark figure	Characteristics benchmark figure from BA statistics and Mikrozensus 2012	Unweighted distribution	Nominal values	Distribution with calibrated weights
Number BCs receiving benefits in accordance with SGB II by federal states (16 categories)	Number BCs Schleswig-Holstein	124	117.580	117.579
	Number BCs Hamburg	105	100.580	100.578
	Number BCs Lower-Saxony	379	307.054	307.048
	Number BCs Bremen	47	50.422	50.421
	Number BCs North Rhine-Westphalia	982	817.347	817.317
	Number BCs Hesse	198	204.784	204.776

	Number BCs Rhineland-Palatinate	144	113.839	113.835
	Number BCs Baden-Wuerttemberg	279	226.754	226.749
	Number BCs Bavaria	352	231.483	231.478
	Number BCs Saarland	57	41.052	41.051
	Number BCs Berlin	287	320.929	320.921
	Number BCs Brandenburg	210	152.943	152.940
	Number BCs Mecklenburg-Vorpommern	93	114.812	114.810
	Number BCs Saxony	263	247.132	247.133
	Number BCs Saxony-Anhalt	222	170.404	170.483
	Number BCs Thuringia	142	111.744	111.740
Number of BCs receiving benefits in accordance with SGB II by number of individuals under 65 years of age in the benefit unit (1, 2, 3, 4, and “5 or more”) and by west/east (10 categories)	Number BCs with 1 individual under 65 (west)	1.115	1.214.722	1.214.722
	Number BCs with 2 individuals under 65 (west)	691	448.390	448.365
	Number BCs with 3 individuals under 65 (west)	468	269.869	269.858
	Number BCs with 4 individuals under 65 (west)	214	162.758	162.743
	Number BCs with 5 or more individuals under 65 (west)	179	115.156	115.144
	Number BCs with 1 individual under 65 (east)	571	666.329	666.329
	Number BCs with 2 individuals under 65 (east)	343	235.221	235.295
	Number BCs with 3 individuals under 65 (east)	176	117.707	117.696
	Number BCs with 4 individuals under 65 (east)	76	62.506	62.506
	Number BCs with 5 or more individuals under 65 (east)	51	36.201	36.200
Number of BCs receiving benefits in accordance with SGB II by number of individuals under 15 years of age in the benefit unit (0, 1, 2, 3, “4 or more”) and by	Number BCs without children under 15 years of age (west)	1.780	1.517.424	1.517.386
	Number BCs with 1 child under 15 years of age (west)	517	377.065	377.050
	Number BCs with 2 children under 15 years of age (west)	244	211.713	211.707
	Number BCs with 3 children under 15 years of age (west)	87	75.420	75.420
	Number BCs with 4 or more children under 15 years of age (west)	39	29.273	29.270
	Number BCs without children under 15 years of age (east)	903	826.113	826.181

west/east (10 categories)	Number BCs with 1 child under 15 years of age (east)	187	171.066	171.062
	Number BCs with 2 children under 15 years of age (east)	93	84.802	84.801
	Number BCs with 3 children under 15 years of age (east)	22	25.792	25.791
	Number BCs with 4 or more children under 15 years of age (east)	12	10.191	10.191
Number BCs receiving benefits in accordance with SGB II consisting of a single parent with children by west/east (4 categories)	Number BCs with a single parent (west)	627	442.433	442.424
	Rest BCs without a single parent (west)	2.040	1.768.462	1.768.409
	Number BCs with a single parent (east)	204	184.821	184.818
	Rest BCs without a single parent (east)	1.013	933.143	933.209
Number of households by federal state and BIK type (spelling: “Federal state.BIK type”)	1.1 bis 1.6	86	457.000	457.000
	1.7 bis 1.10	222	924.000	924.000
	2.10	189	989.000	989.000
	3.1 bis 3.5	288	1.408.000	1.408.000
	3.7 bis 3.8	440	1.389.000	1.389.000
	3.9 bis 3.10	236	1.030.000	1.030.000
	4.8 bis 4.10	81	361.000	361.000
	5.2 bis 5.4	306	1.129.000	1.129.000
	5.5 bis 5.6	274	993.000	993.000
	5.7 bis 5.8	689	2.945.000	2.945.000
	5.9 bis 5.10	896	3.554.000	3.554.000
	6.1 bis 6.4	175	633.000	633.000
	6.5 bis 6.8	197	1.111.000	1.111.000
	6.9 bis 6.10	193	1.197.000	1.197.000
	7.1 bis 7.6	166	909.000	909.000
	7.7 bis 7.10	222	962.000	962.000
	8.1 bis 8.4	221	1.227.000	1.227.000
	8.5 bis 8.8	333	2.168.000	2.168.000

	8.9 bis 8.10	295	1.621.000	1.621.000
	9.1 bis 9.4	246	1.502.000	1.502.000
	9.5 bis 9.7	349	1.505.000	1.505.000
	9.8 bis 9.9	268	1.414.000	1.414.000
	9.10	295	1.560.000	1.560.000
	10.3 bis 10.8	123	492.000	492.000
	11.10	537	2.015.000	2.015.000
	12.1 bis 12.4	161	462.000	462.000
	12.5 bis 12.7	152	271.000	271.000
	12.9 bis 12.10	135	518.000	518.000
	13.1 bis 13.6	133	518.000	518.000
	13.7 bis 13.9	92	327.000	327.000
	14.1 bis 14.4	204	592.000	592.000
	14.5 bis 14.8	170	671.000	671.000
	14.9 bis 14.10	241	933.000	933.000
	15.1 bis 15.4	138	377.000	377.000
	15.5 bis 15.7	206	533.000	533.000
	15.8 bis 15.9	138	265.000	265.000
	16.1 bis 16.4	198	518.000	518.000
	16.5 bis 16.8	214	594.000	594.000
Number of households by household size (1,2,3,4,“5 and more individuals“) and west/east (10 categories)	Number households with 1 individual (west)	2.289	12.308.000	12.308.000
	Number households with 2 individuals (west)	2.233	10.809.000	10.809.000
	Number households with 3 individuals (west)	1.122	3.989.000	3.989.000
	Number households with 4 individuals (west)	733	3.193.000	3.193.000
	Number households with 5 or more individuals (west)	413	1.181.000	1.181.000
	Number households with 1 individual (east)	1.083	3.730.000	3.730.000
	Number households with 2 individuals (east)	937	3.118.000	3.118.000

	Number households with 3 individuals (east)	400	1.057.000	1.057.000
	Number households with 4 individuals (east)	199	538.000	538.000
	Number households with 5 and more individuals (east)	100	151.000	151000
Number of households by "children under 15 years of age in the household yes/no" and west/east	Number households with children under 15 (west)	1.841	5.515.000	5.515.000
	Number households without children under 15 (west)	4.949	25.965.000	25.965.000
	Number households with children under 15 (east)	601	1.265.000	1.265.000
	Number households without children under 15 (east)	2.118	7.329.000	7.329.000

Table 51: Parameters of distribution of weights

1% percentile	89,17792
5% percentile	139,7844
10% percentile	181,6879
25% percentile	334,2173
50% percentile	1089,224
75% percentile	5197,42
90% percentile	14164,47
95% percentile	21068,59
99% percentile	25247,92
Mean	4214,323
Standard deviation	6346,929
Minimum	44,47986
Maximum	34347,43
Number of observations	9509
Efficiency measure	**30,6%**

6.13 Calibration of the person weight, wave 7, cross-section

As in previous waves, the person weights were calibrated under the restriction that they differ as little as possible from the calibrated household weights. The calibrated household weights were quasi-inherited by the individual household members. These input weights were calibrated at the individual level.

As in the previous year, the increase in UB II recipients since the previous year at the level of individuals between 15 and 64 years (358,071) was also included as an additional

benchmark value in the total sample. Again, those cases in the previous samples from waves 1 to 6 of the survey who were receiving UB II in July 2012 are projected to the benchmark statistics of the Federal Employment Agency on receipt of UB II.

Before calibration, the calibrated household weights that formed the input weight were also trimmed. For the calibration of person weights, the range of weights was determined to a certain interval.

6.14 BA sample

The population of the cumulated BA sample of all six waves consists of all individuals aged 15 and over who are living in a household in which there was at least one benefit unit receiving benefits according to SGB II at one of the (until now) seven drawing dates (in July 2006, July 2007, July 2008, July 2009, July 2010, July 2011 or July 2012). Only those individuals aged 15 and over who were living in a benefit unit that received benefits according to SGB II in July 2012 were considered for calibration. Individuals living in a household that did not receive benefits and individuals living in a household with at least one benefit unit according to SGB II but who were not part of a benefit unit themselves were removed from the dataset for the calibration. The weighting of these individuals was calculated in a different way (see below).

The starting point for the calibration is the calibrated household weights of the BA sample. They were trimmed at the fifth and ninety-fifth percentiles of their distribution and then rescaled so that they totaled the untrimmed calibrated household weights. The trimmed projection factors range from 303.15 to 6,469.44. The relation between the total projection factors after calibration and the trimmed design weights was limited downwards to 0.2 and upwards to 3.0. Thus, the total projection factors after calibration lie between a minimum of 69.63 and a maximum of 8,733.51.

A calibration was made for the following characteristics:

Benefit recipients basis BA statistics:

- Number of individuals aged 15 and over in benefit units receiving benefits according to SGB II, by federal states
- Number of individuals in benefit units receiving benefits according to SGB II, by age (15-24 and 25-64)
- Number of individuals aged 15 and over in benefit units receiving benefits according to SGB II by sex and by west/east
- Number of individuals aged 15 and over in benefit units receiving benefits according to SGB II, by “single parent yes/no” and by west/east
- Number of individuals aged 15 and over in benefit units receiving benefits according to SGB II, by nationality (German/non-German)

As in the previous year, the increase in UB II recipients since the previous year at the level of individuals between 15 and 64 years (358,071) was included as an additional benchmark value in the total sample.

For the calibration, each benchmark variable for each individual must have a valid value. Therefore, the very low non-response item was imputed before calibration. The imputation was made by means of the average value and the modal value of the respective variable.

Because the imputation only serves the feasibility of the calibration, the imputed values were set back to missing values after the calibration. A projection with the calibrated weights without considering the nonresponse item thus leads to slight deviations from the values as presented below.

Table 52: Nominal distributions and distributions after calibration (BA sample, individuals)

Benchmark figure	Characteristics benchmark figure from BA statistics and Mikrozensus 2012	Unweighted distribution	Nominal values	Distribution with calibrated weights
Number of individuals aged 15 and over in benefit units receiving benefits in accordance with SGB II by federal states (16 categories)	Number individuals in BCs Schleswig-Holstein	148	160.475	160.475
	Number individuals in BCs Hamburg	128	134.323	134.323
	Number individuals in BCs Lower-Saxony	456	427.075	427.075
	Number individuals in BCs Bremen	57	67.894	67.894
	Number individuals in BCs North Rhine-Westphalia	1.202	1.154.436	1.154.436
	Number individuals in BCs Hesse	261	288.580	288.580
	Number individuals in BCs Rhineland-Palatinate	180	157.411	157.411
	Number individuals in BCs Baden-Wuerttemberg	357	306.673	306.673
	Number individuals in BCs Bavaria	416	302.790	302.790
	Number individuals in BCs Saarland	79	55.601	55.601
	Number individuals in BCs Berlin	344	429.338	429.338
	Number individuals in BCs Brandenburg	263	200.607	200.607
	Number individuals in BCs Mecklenburg-Vorpommern	113	150.999	150.999
	Number individuals in BCs Saxony	353	324.157	324.157
	Number individuals in BCs Saxony-Anhalt	277	226.369	226.369

	Number individuals in BCs Thuringia	166	145.816	145.816
Number of individuals in benefit units receiving benefits in accordance with SGB II by age (15-24 and 25-64; 2 categories)	Number individuals in BCs aged 15-24	718	810.116	810.116
	Number individuals in BCs aged 25-64	4.082	3.722.428	3.722.428
Number of individuals aged 15 and over in benefit units receiving benefits in accordance with SGB II by sex and west/east (4 categories)	Number men in BCs (west)	1.496	1.473.008	1.473.008
	Number women in BCs (west)	1.788	1.582.250	1.582.250
	Number men in BCs (east)	768	752.816	752.816
	Number women in BCs (east)	748	724.470	724.470
Number of individuals aged 15 and over in benefit units receiving benefits in accordance with SGB II by "single parent yes/no", sex and west/east (8 categories)	Number non single parents in BCs (west)	2.690	2.612.825	2.612.825
	Number single parents in BCs (west)	594	442.433	442.433
	Number non single parents in BCs (east)	1.317	1.292.465	1.292.465
	Number single parents in BCs (east)	199	184.821	184.821
Number of individuals aged 15 and over in benefit units receiving benefits in accordance with SGB II by nationality (German/non-German)	Number non-German individuals in BCs	639	941.503	941.503
	Number German individuals in BCs	4.161	3.591.041	3.591.041

Table 53: Parameters of distribution of weights

1% percentile	79,33319
5% percentile	116,4964
10% percentile	142,3242
25% percentile	247,7235
50% percentile	542,9979
75% percentile	1284,422
90% percentile	2382,249
95% percentile	3003,766
99% percentile	5376,827
Mean	944,8705
Standard deviation	1035,822
Minimum	60,63037
Maximum	8733,514
Number of observations	4797
Efficiency measure	**45,4%**

6.15 Population sample

All individuals over 14 years of age in private households in Germany form the basic population. The starting points for the calibration were calibrated household weights of the population sample. These weights were trimmed at the fifth and ninety-fifth percentiles of their distribution and after that rescaled so that they totaled the untrimmed calibrated household weights. The trimmed projection factors lie between a minimum of 2,876.83 and a maximum of 37,146.54. The relation between the total projection factors after calibration and the trimmed design weights was limited downwards to 0.2 and upwards to 3.0. Thus, the total projection factors after calibration lie between a minimum of 575.37 and a maximum of 111,439.6.

A calibration was made for the following characteristics:

Benefit recipients basis BA statistics:

- Number of individuals aged 15 and over in benefit units receiving benefits according to SGB II, by federal states
- Number of individuals in benefit communities receiving benefits according to SGB II, by age (15-24 and 25-64)

- Number of individuals aged 15 and over in benefit units receiving benefits according to SGB II by sex and by west/east
- Number of individuals aged 15 and over in benefit units receiving benefits according to SGB II, by “single parent yes/no” and by west/east
- Number of individuals aged 15 and over in benefit units receiving benefits according to SGB II, by nationality (German/non-German)

Population based on Mikrozensus 2012:

- Number of individuals aged 15 and over in private households by federal state
- Number of individuals aged 15 and over in private households, by age, sex and west/east region
- Number of individuals aged 15 and over in private households, by household size and west/east region
- Number of individuals aged 15 and over in private households, by academic qualifications and west/east region
- Number of individuals aged 15 and over in private households, by marital status and west/east region
- Number of individuals aged 15 and over in private households, by nationality
- Number of unemployed individuals including participants in measures, by west/east region
- Number of employees subject to social security, by west/east region

The source for the benchmark value of employment status was the BA statistics because the definition of unemployment and employment subject to social insurance in PASS does not correspond to the ILO.

For the calibration, each benchmark variable for each individual must have a valid value. Therefore, the very low nonresponse item was imputed before calibration. The imputation was made by means of the average value and the modal value of the respective variable.

Because the imputation only serves the feasibility of the calibration, the imputed values were set to missing values after the calibration. A projection with the calibrated weights without considering the nonresponse item therefore leads to slight deviations from the values as presented below.

Table 54: Nominal distributions and distributions after calibration (population sample, individuals)

Benchmark figure	Characteristics benchmark figure from BA statistics and Mikrozensus 2012	Unweighted distribution	Nominal values	Distribution with calibrated weights
Number of individuals in benefit units receiving benefits in accordance with SGB II by west/east (2 categories)	Number individuals in BCs west	196	3.055.258	3.055.258
	Number individuals in BCs east	81	1.477.286	1.477.286
Number of individuals in benefit units receiving benefits in accordance with SGB II by age (15-24 and 25-64; 2 categories)	Number individuals in BCs aged 15-24	46	810.116	810.116
	Number individuals in BCs aged 25-64	231	3.722.428	3.722.428
Number of individuals aged 15 and over in benefit units receiving benefits in accordance with SGB II by sex (2 categories)	Number men in BCs	121	2.225.824	2.225.824
	Number women in BCs	156	2.306.720	2.306.720
Number of individuals aged 15 and over in benefit units receiving benefits in accordance with SGB II by “single parent yes/no (2 categories)	Number non single parents in BCs	234	3.905.290	3.905.290
	Number single parents in BCs	43	627.254	627.254
Number of individuals aged 15 and over in benefit units receiving benefits in accordance with SGB II by nationality (German/non-German)	Number non-German individuals in BCs	37	941.503	941.503
	Number German individuals in BCs	240	3.591.041	3.591.041
Number of individuals aged 15 and over in private households by federal state (16 categories)	Number individuals in private households Schleswig-Holstein	166	2.422.000	2.422.000
	Number individuals in private households Hamburg	62	1.569.000	1.569.000
	Number individuals in private households Lower-Saxony	636	6.767.000	6.767.000
	Number individuals in private households Bremen	38	578.000	578.000
	Number individuals in private households North Rhine-Westphalia	1.222	15.344.000	15.344.000
	Number individuals in private households Hesse	454	5.255.000	5.255.000
	Number individuals in private households Rhineland-Palatinate	303	3.446.000	3.446.000

	Number individuals in private households Baden-Wuerttemberg	657	9.220.000	9.220.000
	Number individuals in private households Bavaria	985	10.782.000	10.782.000
	Number individuals in private households Saarland	71	890.000	890.000
	Number individuals in private households Berlin	157	3.077.000	3.077.000
	Number individuals in private households Brandenburg	188	2.185.000	2.185.000
	Number individuals in private households Mecklenburg-Vorpommern	91	1.432.000	1.432.000
	Number individuals in private households Saxony	269	3.618.000	3.618.000
	Number individuals in private households Saxony-Anhalt	226	2.011.000	2.011.000
	Number individuals in private households Thuringia	260	1.932.000	1.932.000
Number of individuals aged 15 and over in private households by age (in 5-year classes), gender and west/east (56 categories)	Number men in private households (west), 15-19 years	171	1.824.000	1.824.000
	Number men in private households (west), 20-24 years	148	2.028.000	2.028.000
	Number men in private households (west), 25-29 years	106	1.960.000	1.960.000
	Number men in private households (west), 30-34 years	88	1.986.000	1.986.000
	Number men in private households (west), 35-39 years	98	1.917.000	1.917.000
	Number men in private households (west), 40-44 years	177	2.555.000	2.555.000
	Number men in private households (west), 45-49 years	222	2.820.000	2.820.000
	Number men in private households (west), 50-54 years	233	2.563.000	2.563.000
	Number men in private households (west), 55-59 years	220	2.138.000	2.138.000
Number of individuals aged 15 and over in private households by age (in 5-year classes), gender and west/east (56 categories)	Number men in private households (west), 60-64 years	177	1.934.000	1.934.000
	Number men in private households (west), 65-69 years	172	1.539.000	1.539.000
	Number men in private households (west), 70-74 years	205	1.847.000	1.847.000
	Number men in private households (west), 75-79 years	122	1.251.000	1.251.000
	Number men in private households (west), 80+ years	77	1.147.000	1.147.000
	Number women in private households (west), 15-19 years	155	1.725.000	1.725.000
	Number women in private households (west), 20-24 years	120	1.921.000	1.921.000

	Number women in private households (west), 25-29 years	77	1.936.000	1.936.000
	Number women in private households (west), 30-34 years	96	1.987.000	1.987.000
	Number women in private households (west), 35-39 years	144	1.904.000	1.904.000
	Number women in private households (west), 40-44 years	202	2.506.000	2.506.000
	Number women in private households (west), 45-49 years	273	2.755.000	2.755.000
	Number women in private households (west), 50-54 years	276	2.531.000	2.531.000
	Number women in private households (west), 55-59 years	248	2.226.000	2.226.000
	Number women in private households (west), 60-64 years	231	2.023.000	2.023.000
	Number women in private households (west), 65-69 years	177	1.652.000	1.652.000
	Number women in private households (west), 70-74 years	204	2.092.000	2.092.000
	Number women in private households (west), 75-79 years	112	1.543.000	1.543.000
	Number women in private households (west), 80+ years	63	1.963.000	1.963.000
	Number men in private households (east), 15-19 years	27	287.000	287.000
	Number men in private households (east), 20-24 years	41	477.000	477.000
	Number men in private households (east), 25-29 years	30	558.000	558.000
	Number men in private households (east), 30-34 years	27	527.000	527.000
	Number men in private households (east), 35-39 years	30	465.000	465.000
	Number men in private households (east), 40-44 years	36	618.000	618.000
	Number men in private households (east), 45-49 years	55	697.000	697.000
	Number men in private households (east), 50-54 years	68	674.000	674.000
	Number men in private households (east), 55-59 years	59	606.000	606.000
	Number men in private households (east), 60-64 years	54	525.000	525.000
	Number men in private households (east), 65-69 years	46	418.000	418.000
	Number men in private households (east), 70-74 years	42	519.000	519.000
	Number men in private households (east), 75-79 years	40	347.000	347.000
	Number men in private households (east), 80+ years	18	283.000	283.000

Number of individuals aged 15 and over in private households by age (in 5-year classes), gender and west/east (56 categories)	Number women in private households (east), 15-19 years	22	266.000	266.000
	Number women in private households (east), 20-24 years	29	435.000	435.000
	Number women in private households (east), 25-29 years	33	518.000	518.000
	Number women in private households (east), 30-34 years	27	472.000	472.000
	Number women in private households (east), 35-39 years	28	436.000	436.000
	Number women in private households (east), 40-44 years	41	563.000	563.000
	Number women in private households (east), 45-49 years	57	661.000	661.000
	Number women in private households (east), 50-54 years	57	662.000	662.000
	Number women in private households (east), 55-59 years	74	629.000	629.000
	Number women in private households (east), 60-64 years	77	554.000	554.000
	Number women in private households (east), 65-69 years	47	458.000	458.000
	Number women in private households (east), 70-74 years	69	621.000	621.000
	Number women in private households (east), 75-79 years	32	456.000	456.000
	Number women in private households (east), 80+ years	25	523.000	523.000
Number of individuals aged 15 and over in private households by household size (1, 2, 3, 4, “5 or more individuals“) and west/east (10 categories)	Number individuals in private households with 1 individual (west)	693	12.308.000	12.308.000
	Number individuals in private households with 2 individuals (west)	1.779	21.157.000	21.157.000
	Number individuals in private households with 3 individuals (west)	864	9.884.000	9.884.000
	Number individuals in private households with 4 individuals (west)	806	8.933.000	8.933.000
	Number individuals in private households with 5 or more individuals (west)	452	3.991.000	3.991.000
	Number individuals in private households with 1 individual (east)	212	3.730.000	3.730.000
	Number individuals in private households with 2 individuals (east)	568	6.050.000	6.050.000
	Number individuals in private	215	2.582.000	2.582.000

	households with 3 individuals (east)			
	Number individuals in private households with 4 individuals (east)	116	1.421.000	1.421.000
	Number individuals in private households with 5 or more individuals (east)	80	472.000	472.000
Number of individuals aged 15 and over in private households by highest school qualification and west/east (12 categories)	Number individuals in private households with highest school qualification: still pupil (west)	211	2.270.000	2.270.000
	Number individuals in private households with highest school qualification: no qualification (west)	119	2.208.000	2.208.000
	Number individuals in private households with highest school qualification: lower secondary school (west)	1.482	22.018.000	22.018.000
	Number individuals in private households with highest school qualification: intermediate secondary school; intermediate secondary school in the former GDR (west)	1.318	14.011.000	14.011.000
	Number individuals in private households with highest school qualification: university (of applied sciences) qualification (west)	1.464	15.766.000	15.766.000
	Number individuals in private households with highest school qualification: still pupil (east)	29	372.000	372.000
	Number individuals in private households with highest school qualification: no qualification (east)	22	316.000	316.000
	Number individuals in private households with highest school qualification: lower secondary school (east)	276	3.071.000	3.071.000
	Number individuals in private households with highest school qualification: Intermediate secondary school; intermediate secondary school in the former GDR (east)	528	6.763.000	6.763.000
	Number individuals in private households with highest school qualification: university (of applied sciences) qualification (east)	336	3.733.000	3.733.000

Number of individuals aged 15 and over in private households by marital status and west/east (10 categories)	Number individuals in private households with marital status: single (west)	1.077	10.578.000	10.578.000
	Number individuals in private households with marital status: married, civil partnership (west)	2.910	35.809.000	35.809.000
	Number individuals in private households with marital status: divorced (west)	342	5.136.000	5.136.000
	Number individuals in private households with marital status: widowed (west)	265	4.750.000	4.750.000
	Number individuals in private households with marital status: single (east)	272	3.550.000	3.550.000
	Number individuals in private households with marital status: married, civil partnership (east)	725	7.967.000	7.967.000
	Number individuals in private households with marital status: divorced (east)	100	1.467.000	1.467.000
	Number individuals in private households with marital status: widowed (east)	94	1.271.000	1.271.000
Number of individuals aged 15 and over in private households by nationality	Number individuals in private households non-German	188	6.652.000	6.652.000
	Number individuals in private households German	5.597	63.876.000	63.876.000
Unemployed individuals incl. participants in measures west/east	Not unemployed west	4.409	53.596.402	53.596.402
	Unemployed individuals incl. participants in measures west	185	2.676.598	2.676.598
	Not unemployed east	1.119	13.104.561	13.104.561
	Unemployed individuals incl. participants in measures east	72	1.150.439	1.150.439
Employees subject to social security contributions west/east	Employees not subject to social security contributions west	2.785	32.588.081	32.588.081
	Employees subject to social security contributions west	1.809	23.684.919	23.684.919
	Employees not subject to social security contributions east	661	8.800.390	8.800.390
	Employees subject to social security contributions east	530	5.454.610	5.454.610

Table 55: Parameters of distribution of weights

1% percentile	1077,426
5% percentile	1976,18
10% percentile	2639,394
25% percentile	4159,173
50% percentile	7580,121
75% percentile	14561,65
90% percentile	27114,35
95% percentile	39315,35
99% percentile	68661,68
Mean	12191,53
Standard deviation	13561
Minimum	575,3656
Maximum	111439,6
Number of observations	5785
Efficiency measure	**44,7%**

6.16 Total sample

All individuals aged 15 and over in private households in Germany form the population. The starting point for the calibration was the calibrated household weight of the total sample. That weight was trimmed at the fifth and ninety-fifth percentiles of their distribution and then rescaled so that they totaled the untrimmed calibrated household weights. The trimmed projection factors range from 157.58 to 23,979.67. The relation between the total projection factors after calibration and the trimmed design weights was limited downwards to 0.3 and upwards to 3.5. Thus, the total projection factors after calibration lie between a minimum of 47.27 and a maximum of 83,928.84.

A calibration was made for the following characteristics:

Benefit recipients based on BA statistics:

- Number of individuals aged 15 and over in benefit units receiving benefits according to SGB II, by federal states
- Number of individuals in benefit units receiving benefits according to SGB II, by age (15-24 and 25-64)
- Number of individuals aged 15 and over in benefit units receiving benefits according to SGB II, by sex and by west/east
- Number of individuals aged 15 and over in benefit units receiving benefits according to SGB II, by “single parent yes/no” and by west/east

- Number of individuals aged 15 and over in benefit units receiving benefits according to SGB II, by nationality (German/non-German)

Population based on Mikrozensus 2012:

- Number of individuals aged 15 and over in private households, by federal state
- Number of individuals aged 15 and over in private households, by age, sex and west/east
- Number of individuals aged 15 and over in private households, by household size and west/east
- Number of individuals aged 15 and over in private households, by academic qualifications and west/east
- Number of individuals aged 15 and over in private households, by marital status and west/east
- Number of individuals aged 15 and over in private households, by nationality

Population based on BA statistics:

- Number of unemployed individuals including participants in measures, by west/east
- Number of employees subject to social security, by west/east

The source for the benchmark value of employment status was the BA statistics because the definition of unemployment and employment subject to social insurance in PASS does not correspond to the ILO concept.

In addition, the increase in UB II recipients since the previous year at the level of individuals between 15 and 64 years of age (358,071) was included as an additional benchmark value in the total sample.

For the calibration, each benchmark variable for each individual must have a valid value. Therefore, the very low non-response item was imputed before calibration. The imputation was made by means of the average value and the modal value of the respective variable.

Because the imputation is only required for the feasibility of the calibration, the imputed values were set back to missing values after the calibration. A projection with the calibrated weights without considering the non-response item therefore leads to slight deviations from the values, as presented below.

Table 56: Nominal distributions and distributions after calibration (total sample, individuals)

Benchmark figure	Characteristics benchmark figure from BA statistics and Mikrozensus 2012	Unweighted distribution	Nominal values	Distribution with calibrated weights
Number of individuals aged 15 and over in benefit units receiving benefits in accordance with SGB II by federal states (16 categories)	Number individuals in BCs Schleswig-Holstein	155	160.475	160.475
	Number individuals in BCs Hamburg	131	134.323	134.323
	Number individuals in BCs Lower-Saxony	492	427.075	427.075
	Number individuals in BCs Bremen	64	67.894	67.894
	Number individuals in BCs North Rhine-Westphalia	1.289	1.154.436	1.154.436
	Number individuals in BCs Hesse	269	288.580	288.580
	Number individuals in BCs Rhineland-Palatinate	197	157.411	157.411
	Number individuals in BCs Baden-Wuerttemberg	364	306.673	306.673
	Number individuals in BCs Bavaria	435	302.790	302.790
	Number individuals in BCs Saarland	84	55.601	55.601
	Number individuals in BCs Berlin	354	429.338	429.338
	Number individuals in BCs Brandenburg	278	200.607	200.607
	Number individuals in BCs Mecklenburg-Vorpommern	120	150.999	150.999
	Number individuals in BCs Saxony	365	324.157	324.157
	Number individuals in BCs Saxony-Anhalt	301	226.369	226.369
	Number individuals in BCs Thuringia	179	145.816	145.816
Number of individuals in benefit units receiving benefits in accordance with SGB II by age (15-24 and 25-64; 2 categories)	Number individuals in BCs aged 15-24	764	810.116	810.116
	Number individuals in BCs aged 25-64	4.313	3.722.428	3.722.428
Number of individuals aged 15 and over in benefit units receiving benefits in accordance with SGB II by sex and west/east (4 categories)	Number men in BCs (west)	1.579	1.473.008	1.473.008
	Number women in BCs (west)	1.901	1.582.250	1.582.250
	Number men in BCs (east)	806	752.816	752.816
	Number women in BCs (east)	791	724.470	724.470

Number of individuals aged 15 and over in benefit units receiving benefits in accordance with SGB II by "single parent yes/no" and west/east (4 categories)	Number non single parents in BCs (west)	2.853	2.612.825	2.612.825
	Number single parents in BCs (west)	627	442.433	442.433
	Number non single parents in BCs (east)	1.388	1.292.465	1.292.465
	Number single parents in BCs (east)	209	184.821	184.821
Number of individuals aged 15 and over in benefit units receiving benefits in accordance with SGB II by nationality (German/non-German)	Number non-German individuals in BCs	676	941.503	941.503
	Number German individuals in BCs	4.401	3.591.041	3.591.041
Number of individuals aged 15 and over in private households by federal state (16 categories)	Number individuals in private households Schleswig-Holstein	458	2.422.000	2.422.000
	Number individuals in private households Hamburg	256	1.569.000	1.569.000
	Number individuals in private households Lower-Saxony	1.462	6.767.000	6.767.000
	Number individuals in private households Bremen	117	578.000	578.000
	Number individuals in private households North Rhine-Westphalia	3.251	15.344.000	15.344.000
	Number individuals in private households Hesse	897	5.255.000	5.255.000
	Number individuals in private households Rhineland-Palatinate	633	3.446.000	3.446.000
	Number individuals in private households Baden-Wuerttemberg	1.343	9.220.000	9.220.000
	Number individuals in private households Bavaria	1.832	10.782.000	10.782.000
	Number individuals in private households Saarland	195	890.000	890.000
	Number individuals in private households Berlin	713	3.077.000	3.077.000
	Number individuals in private households Brandenburg	662	2.185.000	2.185.000
	Number individuals in private households Mecklenburg-Vorpommern	327	1.432.000	1.432.000
	Number individuals in private households Saxony	916	3.618.000	3.618.000
	Number individuals in private households Saxony-Anhalt	756	2.011.000	2.011.000
	Number individuals in private households Thuringia	631	1.932.000	1.932.000
Number of individuals aged 15 and over in private households by age (in 5-year classes), gen-	Number men in private households (west), 15-19 years	403	1.824.000	1.824.000
	Number men in private households (west), 20-24 years	335	2.028.000	2.028.000

der and west/east (56 categories)	Number men in private households (west), 25-29 years	361	1.960.000	1.960.000
	Number men in private households (west), 30-34 years	368	1.986.000	1.986.000
	Number men in private households (west), 35-39 years	313	1.917.000	1.917.000
	Number men in private households (west), 40-44 years	423	2.555.000	2.555.000
	Number men in private households (west), 45-49 years	506	2.820.000	2.820.000
Number of individuals aged 15 and over in private households by age (in 5-year classes), gender and west/east (56 categories)	Number men in private households (west), 50-54 years	543	2.563.000	2.563.000
	Number men in private households (west), 55-59 years	486	2.138.000	2.138.000
	Number men in private households (west), 60-64 years	426	1.934.000	1.934.000
	Number men in private households (west), 65-69 years	296	1.539.000	1.539.000
	Number men in private households (west), 70-74 years	234	1.847.000	1.847.000
	Number men in private households (west), 75-79 years	130	1.251.000	1.251.000
	Number men in private households (west), 80+ years	80	1.147.000	1.147.000
	Number women in private households (west), 15-19 years	373	1.725.000	1.725.000
	Number women in private households (west), 20-24 years	395	1.921.000	1.921.000
	Number women in private households (west), 25-29 years	391	1.936.000	1.936.000
	Number women in private households (west), 30-34 years	496	1.987.000	1.987.000
	Number women in private households (west), 35-39 years	447	1.904.000	1.904.000
	Number women in private households (west), 40-44 years	529	2.506.000	2.506.000
	Number women in private households (west), 45-49 years	583	2.755.000	2.755.000
	Number women in private households (west), 50-54 years	626	2.531.000	2.531.000
	Number women in private households (west), 55-59 years	550	2.226.000	2.226.000
	Number women in private households (west), 60-64 years	455	2.023.000	2.023.000
	Number women in private households (west), 65-69 years	269	1.652.000	1.652.000
	Number women in private households (west), 70-74 years	225	2.092.000	2.092.000
	Number women in private households (west), 75-79 years	122	1.543.000	1.543.000
	Number women in private households (west), 80+ years	79	1.963.000	1.963.000

	Number men in private households (east), 25-29 years	188	558.000	558.000
	Number men in private households (east), 30-34 years	179	527.000	527.000
	Number men in private households (east), 35-39 years	135	465.000	465.000
	Number men in private households (east), 40-44 years	130	618.000	618.000
	Number men in private households (east), 45-49 years	210	697.000	697.000
	Number men in private households (east), 50-54 years	227	674.000	674.000
Number of individuals aged 15 and over in private households by age (in 5-year classes), gender and west/east (56 categories)	Number men in private households (east), 55-59 years	228	606.000	606.000
	Number men in private households (east), 60-64 years	204	525.000	525.000
	Number men in private households (east), 65-69 years	94	418.000	418.000
	Number men in private households (east), 70-74 years	54	519.000	519.000
	Number men in private households (east), 75-79 years	46	347.000	347.000
	Number men in private households (east), 80+ years	20	283.000	283.000
	Number women in private households (east), 15-19 years	97	266.000	266.000
	Number women in private households (east), 20-24 years	125	435.000	435.000
	Number women in private households (east), 25-29 years	199	518.000	518.000
	Number women in private households (east), 30-34 years	171	472.000	472.000
	Number women in private households (east), 35-39 years	152	436.000	436.000
	Number women in private households (east), 40-44 years	166	563.000	563.000
	Number women in private households (east), 45-49 years	231	661.000	661.000
	Number women in private households (east), 50-54 years	232	662.000	662.000
	Number women in private households (east), 55-59 years	242	629.000	629.000
	Number women in private households (east), 60-64 years	204	554.000	554.000
	Number women in private households (east), 65-69 years	81	458.000	458.000
	Number women in private households (east), 70-74 years	80	621.000	621.000

	Number women in private households (east), 75-79 years	39	456.000	456.000
	Number women in private households (east), 80+ years	30	523.000	523.000
Number of individuals aged 15 and over in private households by household size (1, 2, 3, 4, “5 or more individuals“) and west/east (10 categories)	Number individuals in private households with 1 individual (west)	2.278	12.308.000	12.308.000
	Number individuals in private households with 2 individuals (west)	3.522	21.157.000	21.157.000
	Number individuals in private households with 3 individuals (west)	2.081	9.884.000	9.884.000
	Number individuals in private households with 4 individuals (west)	1.552	8.933.000	8.933.000
	Number individuals in private households with 5 or more individuals (west)	1.011	3.991.000	3.991.000
	Number individuals in private households with 1 individual (east)	1.079	3.730.000	3.730.000
	Number individuals in private households with 2 individuals (east)	1.519	6.050.000	6.050.000
	Number individuals in private households with 3 individuals (east)	766	2.582.000	2.582.000
	Number individuals in private households with 4 individuals (east)	400	1.421.000	1.421.000
	Number individuals in private households with 5 or more individuals (east)	241	472.000	472.000
Number of individuals aged 15 and over in private households by highest school qualification and west/east (12 categories)	Number individuals in private households with highest school qualification: still pupil (west)	499	2.270.000	2.270.000
	Number individuals in private households with highest school qualification: no qualification (west)	519	2.208.000	2.208.000
	Number individuals in private households with highest school qualification: lower secondary school (west)	3.699	22.018.000	22.018.000
	Number individuals in private households with highest school qualification: intermediate secondary school; intermediate secondary school in the former GDR (west)	2.949	14.011.000	14.011.000
	Number individuals in private households with highest school qualification: university (of applied sciences) qualification (west)	2.778	15.766.000	15.766.000

	Number individuals in private households with highest school qualification: still pupil (east)	116	372.000	372.000
	Number individuals in private households with highest school qualification: no qualification (east)	138	316.000	316.000
	Number individuals in private households with highest school qualification: lower secondary school (east)	927	3.071.000	3.071.000
	Number individuals in private households with highest school qualification: Intermediate secondary school; intermediate secondary school in the former GDR (east)	1.942	6.763.000	6.763.000
	Number individuals in private households with highest school qualification: university (of applied sciences) qualification (east)	882	3.733.000	3.733.000
Number of individuals aged 15 and over in private households by marital status and west/east (10 categories)	Number individuals in private households with marital status: single (west)	3.309	10.578.000	10.578.000
	Number individuals in private households with marital status: married, civil partnership (west)	5.199	35.809.000	35.809.000
	Number individuals in private households with marital status: divorced (west)	1.503	5.136.000	5.136.000
	Number individuals in private households with marital status: widowed (west)	433	4.750.000	4.750.000
	Number individuals in private households with marital status: single (east)	1.511	3.550.000	3.550.000
	Number individuals in private households with marital status: married, civil partnership (east)	1.706	7.967.000	7.967.000
	Number individuals in private households with marital status: divorced (east)	616	1.467.000	1.467.000
	Number individuals in private households with marital status: widowed (east)	172	1.271.000	1.271.000
Number of individuals aged 15 and over in private households by nationality	Number individuals in private households non-German	1.140	6.652.000	6.652.000
	Number individuals in private households German	13.309	63.876.000	63.876.000
Unemployed individuals incl. participants in measures west/east	Not unemployed west	8.395	53.596.402	53.596.402
	Unemployed individuals incl. participants in measures west	2.049	2.676.598	2.676.598

	Not unemployed east	2.924	13.104.561	13.104.561
	Unemployed individuals incl. participants in measures east	1.081	1.150.439	1.150.439
Employees subject to social security contributions west/east	Employees not subject to social security contributions west	6.641	32.588.081	32.588.081
	Employees subject to social security contributions west	3.803	23.684.919	23.684.919
	Employees not subject to social security contributions east	2.414	8.800.390	8.800.390
	Employees subject to social security contributions east	1.591	5.454.610	5.454.610

Table 57: Parameters of distribution of weights

1% percentile	53,96318
5% percentile	96,04667
10% percentile	144,5919
25% percentile	327,5365
50% percentile	1304,052
75% percentile	5833,994
90% percentile	14399,46
95% percentile	22033,35
99% percentile	38690,41
Mean	4881,168
Standard deviation	8325,565
Minimum	47,27327
Maximum	83928,84
Number of observations	14449
Efficiency measure	**25,6%**

6.17 Estimating the BA cross-sectional weights for households and individuals not receiving Unemployment Benefit II

Finally, in wave 7, some households and individuals remained that could not be assigned a BA cross-sectional household weight or a BA cross-sectional person weight by means of calibration. The number of these households is larger in wave 7 than in wave 6 because a larger part of the BA sample of waves 1 to 6 has withdrawn from benefits. These are the following three groups that were not receiving benefits in July 2012 but that belong to the population of the BA sample (households receiving UB II in July 2006, July 2007, July 2008, July 2009, July 2010, July 2011 or July 2012 and individuals in households re-

ceiving UB II in July 2006, July 2007, July 2008, July 2009, July 2010, July 2011 or July 2012).

- From the refreshment sample: Individuals in the household who are not members of a benefit unit: Here, the person weight was obtained from the BA household weight in wave 7 after calibration (*wqbahh*) by dividing it by the proportion of these individuals who gave a personal or senior citizen interview - provided that their household was participating.
- Panel households in which nobody continued to receive UB II in July 2012: The household retains the BA weight before calibration. Individuals in households with interviews in both waves were assigned a new BA person weight, which is obtained by multiplying their old BA person weight by the reciprocal re-participation probability *ppbleib*. Individuals in these households who did not provide a personal interview in wave 6 are assigned a new BA person weight calculated by dividing the BA household weight of their household for wave 7 by the proportion of such individuals who participate if their household is taking part.
- Individuals who are not members of a benefit unit in panel households that continued to receive UB II in July 2012: Individuals in these households with interviews in both waves were assigned a new BA person weight, which is obtained by multiplying their BA person weight from the previous wave by the reciprocal re-participation probability *ppbleib*.

The individuals and households were also adjusted to a benchmark figure for the individuals or benefit units that did not continue to receive UB II. The exact population of this group is unknown but can be approximated from the total of all cumulated BA subsamples minus the individuals or benefit units currently receiving benefits. The number of individuals who are no longer receiving UB II at wave 7 is 4,382,041. The number of benefit units that are no longer receiving UB II is 3,203,818.

7 Appendix: Brief description of the dataset

Content characteristics

Categories	Comments
Topics/ characteristics categories	**Socio-demographic characteristics:** artificial individual ID; sex; year of birth; age; marital status; number of children living in and outside the household; nationality; country of origin and migration background; school and vocational qualifications (incl. generated scales: CASMIN, ISCED-97, number of years of schooling and vocational training), parents' school and vocational qualifications; health indicators; religious denomination; social contacts; childcare and school attendance of children; household income (incl. individual components and equivalised household income); basic information on assets and liabilities; household equipment (deprivation index); housing and residential environment; detailed information on the topic of old age benefits (only wave 3); **Employment-related characteristics:** employment status/economic inactivity status; marginal employment; working hours; occupational status (detailed); employment (ISCO-88 and KldB-92); ISCO-based measures of occupational status and prestige (ISEI, SIOPS, MPS, EGP, ESeC); earned income (gross and net); employment biographies with employment/unemployment spells and periods of economic inactivity since January 2005 (from wave 2 onwards); limited-term employment; supervisory function; employer: public service/private industry; employer: number of employees; other employment; pooled information on the employment and unemployment history; detailed information on the subject of job-search; reservation wage; **Characteristics on receiving benefits:** Unemployment Benefit I: start and end dates of the spell(s) of benefit receipt since January 2005 (wave 1 only); information on periods of Unemployment Benefit I receipt in the context of registered unemployment since January 2005 (from wave 2 onwards); amount of benefit; reason for end; Unemployment Benefit II: start and end dates of the spell(s) of benefit receipt since January 2005; reason for start and end; identification of household members receiving benefits; amount of benefits received; benefit cuts (start date, duration, reasons, which household members' benefit cut); Measurement participation: type of measure; start and end dates of measure; indicator of dropout; reasons for dropout; type of access to measure; assessment of measure; working hours in measure; comparison to regular employment; economic sector/industry; from wave 4 onwards only, one-Euro job; Contacts with Unemployment Benefit II institutions: number and type of contacts; contents of discussion; offers; integration agreement; assessment of institution;

Categories	Comments
Topics/ characteristics categories (continued)	**Subjective indicators:** satisfaction; fears and problems; employment orientation; education aspiration; sex role orientation; subjective social position (top-bottom scale); subjective assessment of health state; personality scale “big five”
Data unit	Individuals and households receiving Unemployment Benefit II in July 2006 (sample I) Individuals and households in the resident population of Germany (sample II) Individuals and households receiving Unemployment Benefit II in July 2007 but without receipt in July 2006 (sample III; refreshment sample 1) Individuals and households receiving Unemployment Benefit II in July 2008 but without receipt in July 2006 or July 2007 (sample IV; refreshment sample 2) Individuals and households receiving Unemployment Benefit II in July 2009 but without receipt in July 2006, July 2007 or July 2008 (sample V; refreshment sample 3) Individuals and households receiving Unemployment Benefit II in July 2010 but without receipt in July 2006, July 2007, July 2008 or July 2009 (sample VIII; refreshment sample 4) Individuals and households of the resident German population (sample VI, panel refreshment/replenishment sample) Individuals and households receiving UB II in July 2010 (sample VII, panel refreshment/replenishment sample) Individuals and households receiving Unemployment Benefit II in July 2011 but without receipt in July 2006, July 2007, July 2008, July 2009 or July 2010 (sample IX; refreshment sample 5) Individuals and households receiving Unemployment Benefit II in July 2012 but without receipt in July 2006, July 2007, July 2008, July 2009, July 2010 or July 2011 (sample X; refreshment sample 6) Note: individuals aged 65 and over are interviewed using a shorter version of the questionnaire

Categories	Comments
Case numbers	Wave 1: Sample I: 9,386 individuals (living in 6,804 households) Sample II: 9,568 individuals (living in 5,990 households) Wave 2: sample I: 4,753 individuals (living in 3,491 households) Sample II: 6,392 individuals (living in 3,897 households) Sample III: 1,342 individuals (living in 1,041 households) Wave 3: sample I: 4,913 individuals (living in 3,754 households) Sample II: 6,207 individuals (living in 3,901 households) Sample III: 898 individuals (living in 694 households) Sample IV: 1,421 individuals (living in 1,186 households) Wave 4: sample I: 3,958 individuals (living in 2,815 households) Sample II: 5,016 individuals (living in 2,977 households) Sample III: 786 individuals (living in 563 households) Sample IV: 983 individuals (living in 745 households) Sample V: 1,025 individuals (living in 748 households) Wave 5: Sample I: 3,394 individuals (in 2,382 households) Sample II: 4,511 individuals (in 2,680 households) Sample III: 653 individuals (living in 464 households) Sample IV: 822 individuals (living in 608 households) Sample V: 760 individuals (in 517 households) Sample VI: 2,589 individuals (in 1,510 households) Sample VII: 1,859 individuals (in 1,321 households) Sample VIII: 1,019 individuals (living in 753 households) Wave 6: Sample I: 3,048 individuals (living in 2,109 households) Sample II: 4,245 individuals (living in 2,539 households) Sample III: 558 individuals (living in 398 households) Sample IV: 719 individuals (living in 532 households) Sample V: 679 individuals (living in 466 households) Sample VI: 1,990 individuals (living in 1,103 households) Sample VII: 1,350 individuals (living in 908 households) Sample VIII: 716 individuals (living in 497 households) Sample IX: 1,314 individuals (living in 961 households) Wave 7: Sample I: 2,861 individuals (living in 1,984 households) Sample II: 4,001 individuals (living in 2,409 households) Sample III: 505 individuals (living in 359 households) Sample IV: 688 individuals (living in 505 households) Sample V: 590 individuals (living in 414 households) Sample VI: 599 individuals (living in 413 households) Sample VII: 1,784 individuals (living in 996 households) Sample VIII: 1,182 individuals (living in 798 households) Sample IX: 975 individuals (living in 682 households) Sample X: 1,264 individuals (living in 949 households)

Categories	Comments
Data collection mode	CATI and CAPI CAPI interviews were conducted when a sample household could not be reached by telephone or when a personal interview was requested. Wave 1: N (CATI): 12,414 individuals (8,445 households) N (CAPI): 6,540 individuals (4,339 households) Wave 2: N (CATI): 7,888 individuals (5,378 households) N (CAPI): 4,599 individuals (3,051 households) Wave 3: N (CATI): 7,776 individuals (5,664 households) N (CAPI): 5,663 individuals (3,871 households) Wave 4: N (CATI): 6,913 individuals (4,669 households) N (CAPI): 4,855 individuals (3,179 households) Wave 5: N (CATI): 7,358 individuals (4,987 households) N (CAPI): 8,249 individuals (5,248 households) Wave 6: N (CATI): 6,069 individuals (4,058 households) N (CAPI): 8,550 individuals (5,455 households) Wave 7: N (CATI): 5,779 individuals (3,874 households) N (CAPI): 8,670 individuals (5,635 households)

Categories	Comments
Interview languages	Wave 1: German: 18,205 individuals (12,347 households) Russian: 432 individuals (275 households) Turkish: 305 individuals (163 households) English: 12 individuals (9 households) Wave 2: German: 12,237 individuals (8,234 households) Russian: 219 individuals (156 households) Turkish: 31 individuals (39 households) English: no longer offered in wave 2 due to the low case numbers in wave 1 Wave 3: German: 13,000 individuals (9,256 households) Russian: 330 individuals (210 households) Turkish: 109 individuals (69 households) Wave 4: German: 11,405 individuals (7,627 households) Russian: 285 individuals (179 households) Turkish: 78 individuals (42 households) Wave 5: German: 15,290 individuals (10,040 households) Russian: 259 individuals (159 households) Turkish: 58 individuals (36 households) Wave 6: German: 14,337 individuals (9,342 households) Russian: 242 individuals (146 households) Turkish: 40 individuals (25 households) Wave 7: German: 14,161 individuals (9,335 households) Russian: 245 individuals (145 households) Turkish: 43 individuals (29 households)

Categories	Comments
Response rates	Wave 1: Sample I: 35.1 % Sample II: 26.6 % Total: 30.5 % Wave 2: Sample I (HHs agreeing to participate only): 51.1 % Sample II (HHs agreeing to participate only): 64.7 % Sample III: 26.3 % Split-off households (from samples I and II): 13.4 % Total: 45.0 % Wave 3: Sample I (HHs agreeing to participate only): 64.5 % Sample II (HHs agreeing to participate only): 76.4 % Sample II (HHs agreeing to participate only): 69.0 % Sample IV: 31.2% Total: 60.6 % Wave 4: Sample I (HHs agreeing to participate only): 72.1 % Sample II (HHs agreeing to participate only): 82.4 % Sample III (HHs agreeing to participate only): 65.6 % Sample IV (HHs agreeing to participate only): 68.2 % Sample V: 30.9 % Total: 59.5 % Wave 5: Sample I (HHs agreeing to participate only): 71.1 % Sample II (HHs agreeing to participate only): 81.3 % Sample III (HHs agreeing to participate only): 69.2 % Sample IV (HHs agreeing to participate only): 63.7 % Sample V: (HHs agreeing to participate only): 71.5 % Sample VI: 24.5 % Sample VII: 24.5 % Sample VIII: 27.1 % Total: 43.9 % Wave 6: Sample I (HHs agreeing to participate only): 73.3 % Sample II (HHs agreeing to participate only): 85.1 % Sample III (HHs agreeing to participate only: 70.2 % Sample IV (HHs agreeing to participate only): 69.9 % Sample V (HHs agreeing to participate only): 68.4 % Sample VI (HHs agreeing to participate only): 78.4 % Sample VII (HHs agreeing to participate only): 84.1 % Sample VIII (HHs agreeing to participate only): 77.1 % Sample IX: 30.8 % Total: 67.4 % Wave 7: Sample I (HHs agreeing to participate only): 79.1 % Sample II (HHs agreeing to participate only): 86.8 % Sample III (HHs agreeing to participate only: 75.3 % Sample IV (HHs agreeing to participate only): 77.5 % Sample V (HHs agreeing to participate only): 76.4 % Sample VI (HHs agreeing to participate only): 66.6 % Sample VII (HHs agreeing to participate only): 79.3 % Sample VIII (HHs agreeing to participate only): 70.8 % Sample IX (HHs agreeing to participate only): 74.2 % Sample X: 32.1% Total: 68.7 %

Categories	Comments
Response rates within households	Stage 1: Sample I: 85.6 %, Sample II: 84.3 %, Total: 85.0 % Wave 2: Sample I (re-interviewed households only): 85.5 % Sample II (re-interviewed households only): 85.1 % Sample III: 86.2 % Split-off households (from samples I and II): 88.3 % Total: 85.4 % Wave 3: Sample I (re-interviewed households only): 83.1 % Sample I (re-interviewed households only): 83.6 % Sample III (re-interviewed households only): 84.3 % Sample IV: 84.2 % Split-off households (from samples I - III): 84.2 % Total: 83.5 % Wave 4: Sample I (re-interviewed households only): 88.4 % Sample I (re-interviewed households only): 88.0 % Sample III (re-interviewed households only): 90.2 % Sample IV (re-interviewed households only): 88.3 % Sample V: 89.6 % Split-off households (from samples I - IV): 86.4 % Total: 88.5 % Wave 5: Sample I (re-interviewed households only): 88.7 % Sample I (re-interviewed households only): 88.3 % Sample III (re-interviewed households only): 89.5 % Sample IV (re-interviewed households only): 89.3 % Sample V (re-interviewed households only): 91.2 % Sample VI: 84.4 % Sample VII: 90.0 % Sample VIII: 88.9 % Split-off households (from samples I - V): 89.9 % Total: 88.3 % Wave 6: Sample I (re-interviewed households only): 89.3 % Sample I (re-interviewed households only): 88.6 % Sample III (re-interviewed households only): 88.5 % Sample IV (re-interviewed households only): 88.5 % Sample V (re-interviewed households only): 91.4 % Sample VI (re-interviewed households only): 92.0 % Sample VII (re-interviewed households only): 89.1 % Sample VIII (re-interviewed households only): 91.5 % Sample IX: 89.9 % Split-off householdes (from samples I-VIII): 91.7 % Total: 89.5 % Wave 7: Sample I (re-interviewed households only): 89.2 % Sample I (re-interviewed households only): 88.4 % Sample III (re-interviewed households only): 90.1 % Sample IV (re-interviewed households only): 88.8 % Sample V (re-interviewed households only): 89.8 % Sample VI (re-interviewed households only): 92.6 % Sample VII (re-interviewed households only): 89.1 % Sample VIII (re-interviewed households only): 92.0 % Sample IX (re-interviewed households only): 90.7 % Sample X: 90.1% Split-off householdes (from samples I-IX): 90.3 % Total: 89.5 %

Categories	Comments
Fieldwork period	Wave 1: December 2006-June 2007 Wave 2: December 2007-July 2008 Wave 3: December 2008-August 2009 Wave 4: February 2010-September 2010 Wave 5: February 2011-September 2011 Wave 6: February 2012-September 2012 Wave 7: February 2013-September 2013
Period	Wave 1: fieldwork period and retrospective spell data as of January 2005 Wave 2: fieldwork period and retrospective spell data as of January 2005 or the respective reference period of the spell type Wave 3: fieldwork period and retrospective spell data as of 01/2006 or the respective reference period of the spell type Wave 4: fieldwork period and retrospective spell data as of 01/2008 or the respective reference period of the spell type Wave 5: fieldwork period and retrospective spell data as of 01/2009 or the respective reference period of the spell type Wave 6: fieldwork period and retrospective spell data as of 01/2010 or the respective reference period of the spell type Wave 7: fieldwork period and retrospective spell data as of 01/2011 or the respective reference period of the spell type
Time reference	Repeat interview (household panel)
Regional structure	German federal state, east/west Germany (Further regional information is available but is not contained in the scientific use file for data protection reasons. Detailed information is available on request.)
Territorial allocation	On the survey date

Methodological characteristics

Categories	Comments
Survey design	Original sample wave 1: two-stage random sample with two sub-populations Stage 1: selection of 300 postcode sectors as primary sampling units (PSU) for both subsamples. The sampling probability of the individual postcode areas depended on the particular size of the area in terms of the number of residents (probability proportional to size/pps). Stage 2, sample I: drawing of benefit units from the register data of the Federal Employment Agency. The number of the gross sample drawn per PSU depended on the PSU size in terms of the relative proportion of benefit recipients within the respective postcode sector (probability proportional to size/pps). The average size of the gross sample was N=100 per postcode area. Stage 2, sample II: for sample II, first a sample of residential buildings was drawn from a commercial database (Microm mosaic). This was then stratified using a stratification index contained in the database at a ratio of 4:2:1 for low-, medium- or high-status households, respectively. Interviewers from the surveying institute visited the selected buildings. In the event that a building accommodated several households, this fact was noted, and then one of the households was selected by the institute as the household to be interviewed. The gross sample comprised N=100 households per postcode area. Refreshment sample for sample I in wave 2: In addition to continuing sample I (which was drawn for wave 1) in the second wave, a refreshment sample was drawn from the register data of the Federal Employment Agency. Benefit units that received Unemployment Benefit II in July 2007 but not in July 2006 were selected, i.e., new recipients. The sample was drawn in the postcode areas selected for wave 1 following the procedure used in wave 1. Refreshment sample for sample I in wave 3: Also in wave 3, a refreshment sample for sample I was drawn from the register data of the Federal Employment Agency. To do so, benefit units that received Unemployment Benefit II in July 2008 but not in July 2006 or July 2007 were selected, i.e., new benefit recipients. The sample was drawn in the postcode sectors selected for wave 1 following the procedure used in wave 1. Refreshment sample 3 for sample I in wave 4 (sample V): Also in wave 4, a refreshment sample for sample I was drawn from the register data of the Federal Employment Agency. Benefit units that were receiving Unemployment Benefit II in July 2009 but not in July 2006, July 2007 or July 2008 were selected. These benefit units thus depict the inflows to benefit receipt. The sample was drawn in the postcode sectors selected for wave 1 following the procedure used in wave 1. Refreshment sample 4 for sample I in wave 5:

	Also in wave 5, a refreshment sample for sample I was drawn from the register data of the Federal Employment Agency. Benefit units that were receiving Unemployment Benefit II in July 2010 but not in July 2006, July 2007, July 2008 or July 2009 were selected. These benefit units thus depict the inflows to benefit receipt. The sample was drawn in the postcode sectors selected for wave 1 following the procedure used in wave 1. In wave 5, the panel of the original sample was refreshed with two replenishment samples based on a two-staged random sample with two subpopulations. Stage 1: selection of 100 postcode sectors as primary sampling units (PSU) for both subsamples. The sampling probability of the individual postcode sectors depended on the particular size of the sector in terms of the number of residents (probability proportional to size/pps). Stage 2, sample VII: drawing of benefit units from the register data of the Federal Employment Agency with sampling date July 2010. The number of benefit recipients to be selected per point was selected as the product of the permanent sample size (sample size individuals per point) in the population sample with the quotient from benefit recipient rate in the point and benefit recipient rate across Germany. Stage 2, sample VI: in sample VI, the individuals were drawn from the registration offices' registers. To do so, 96 municipalities were assigned to the 100 postcode areas. The drawing of the personal addresses from the possible choices in the municipalities was made by systematic random sampling (interval sampling). Sampling of addresses from the registration offices' registers was made for birth years of 1992 and earlier. One hundred forty-four addresses were drawn from the municipalities' registers in each sample point. Refreshment sample 5 for sample I in wave 6: In wave 6, a refreshment sample for sample I was again drawn from the register data of the Federal Employment Agency. Benefit units that were receiving Unemployment Benefit II in July 2011 but not in July 2006, July 2007, July 2008, July 2009 or July 2010 were selected, i.e., new benefit recipients. The sample was drawn in the postcode sectors selected for wave 1 following the procedure used in wave 1. Refreshment sample 6 for sample I in wave 7: In wave 7, a refreshment sample for sample I was again drawn from the register data of the Federal Employment Agency. Benefit units that were receiving Unemployment Benefit II in July 2012 but not in July 2006, July 2007, July 2008, July 2009, July 2010 or July 2011 were selected, i.e., new benefit recipients. The sample was drawn in the postcode sectors selected for wave 1 following the procedure used in wave 1.

Categories	Comments
Institutions involved in survey	Institute for Employment Research (IAB); TNS Infratest Sozialforschung (waves 1 to 3), infas Institut für angewandte Sozialwissenschaft GmbH (as of wave 4)
Frequency of data collection	Annually (panel)
File format and size	STATA, SPSS (several files)
File architecture	Household dataset: HHENDDAT.dta/.sav Individual dataset: PENDDAT.dta/.sav Spell data Unemployment Benefit I: alg1_spells.dta/.sav (wave 1 only) Spell data Unemployment Benefit II: alg2_spells.dta/.sav Spell data unemployment: al_spells.dta/.sav (waves 2 and 3) Spell data employment: et_spells.dta/.sav (waves 2 and 3) Spell data gaps: lu_spells.dta/.sav (waves 2 and 3) from wave 4 onwards: spell data on employment, unemployment and gaps integrated: bio_spells.dta/.sav Spell data measures: mn_spells.dta/.sav (from wave 2 onwards) Spell data participation in measures: massnahmespells.dta/.sav (wave 1 only) Register data on households: hh_register.dta/.sav Register data on individuals: p_register.dta/.sav Weighting data on households: hweights.dta/.sav Weighting data on individuals: pweights.dta/.sav Old-age provision household level: HAVDAT.dta/.sav (wave 3 only) Old-age provision individual level: PAVDAT.dta/.sav (wave 3 only) Vignette data: VIGDAT.dat/.sav (wave 5 only) Children data: KINDER.dat/.sav (from wave 6 onwards)

Categories	Comments
Data access	Scientific Use File (SUF)
Degree of anonymisation	Factually anonymised
Sensitive characteristics	None

Literature:

AAPOR (2006). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys.* 4th Edition. Lenexa: AAPOR.

Achatz, J., Hirseland, A. & Promberger, M. (2007). Rahmenkonzept für das IAB-Panel „Arbeitsmarkt und Soziale Sicherung". In M. Promberger (Ed.), Neue Daten für die Sozialstaatsforschung: Zur Konzeption der IAB-Panelerhebung „Arbeitsmarkt und Soziale Sicherung", *IAB-Forschungsbericht 12/2007* (pp. 11-32), Nuremberg.

Andreß H.-J., Burkatzki, E., Lipsmeier, G., Salentin, K., Schulte, K. & Strengmann-Kuhn, W. (1996). *Leben in Armut. Analysen der Verhaltensweisen armer Haushalte mit Umfragedaten.* Final report of the DFG project „Versorgungsstrategien privater Haushalte im unteren Einkommensbereich (VuE)". Bielefeld.

Andreß, H.-J. & Lipsmeier, G. (1995). Was gehört zum notwendigen Lebensstandard und wer kann ihn sich leisten? Ein neues Konzept zur Armutsmessung. *In Politik und Zeitgeschichte 31-32,* pp. 35-49.

Andreß, H.-J. & Lipsmeier, G. (2001). *Armut und Lebensstandard. Gutachten im Rahmen des Armuts- und Reichtumsberichts der Bundesregierung.* BMAS. Bonn.

Beckmann, P. & Trometer, R. (1991). Neue Dienstleistungen des ALLBUS: Haushalts- und Familientypologien, Klassenschema nach Goldthorpe. *ZUMA-Nachrichten 28,* pp. 7-17.

Berg, M., Cramer, R., Dickmann, C., Gilberg, R., Jesske, B., Marwinski, K., Gebhardt, D., Wenzig, C. & Wetzel, M. (2010). Codebuch und Dokumentation des ‚Panel Arbeitsmarkt und soziale Sicherung' (PASS). Vol. 1: Datenreport Welle 3. *FDZ Datenreport 06/2010.* Nuremberg.

Berg, M., Cramer, R., Dickmann, C., Gilberg, R., Jesske, B., Kleudgen, M., Bethmann, A., Fuchs, B., Trappmann, M., Wurdack, A. (2013). Codebuch und Dokumentation des „Panel Arbeitsmarkt und soziale Sicherung" (PASS). Bd. 1: Datenreport Welle 6. FDZ Datenreport 06/2013. Nürnberg.

Bethmann, A., Fuchs, B. & Wurdack, A. (Hrsg.)(2013). User Guide "Panel Labour Market and Social Security" (PASS). Wave 6. FDZ Datenreport 07/2013. Nürnberg.

Brauns, H. & Steinmann, S. (1999). Educational Reform in France, West-Germany and the United Kingdom. Updating the Casmin Classification. *ZUMA-Nachrichten 44*. pp. 7-45.

Bundesministerium für Bildung und Forschung [BMBF] (2003). Berufsausbildung sichtbar gemacht. Schaubildsammlung. 4th edition. Bonn: BMBF.

Büngeler, K., Gensicke, M., Hartmann, J., Jäckle, R. & Tschersich, N. (2009). IAB-Haushaltspanel im Niedrigeinkommensbereich Welle 2 (2007/2008). Methoden- und Feldbericht. *FDZ-Methodenreport 08/2009.* Nuremberg.

Büngeler, K., Gensicke, M., Hartmann, J., Jäckle, R. & Tschersich, N. (2010): IAB-Haushaltspanel im Niedrigeinkommensbereich Welle 3 (2008/2009). Methoden- und Feldbericht. *FDZ-Methodenreport 10/2010.* Nuremberg.

Christoph, B. (2005). Zur Messung des Berufsprestiges: Aktualisierung der Magnitude-Prestigeskala auf die Berufsklassifikation ISCO88. *ZUMA-Nachrichten 57.* pp. 79-127.

Europäische Gemeinschaften [EG] (2002). Verordnung (EG) Nr. 29/2002 der Kommission vom 19. Dezember 2001 zur Änderung der Verordnung (EWG) Nr. 3037/90 des Rates betreffend die statistische Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft. Amtsblatt der Europäischen Gemeinschaften L6/3-L6-33. Brussels.

Erikson, R. & Goldthorpe, J. (1992). *The Constant Flux. A Study of Class Mobility in Industrial Society.* Oxford: Clarendon Press.

Erikson, R., Goldthorpe, J. & Portocarero, L. (1979). Intergenerational Class Mobility in Three Western Societies: England, France and Sweden. *British Journal of Sociology 30.* pp. 415-441.

Erikson, R., Goldthorpe, J. & Portocarero, L. (1982): Social Fluidity in Industrial Nations: England, France and Sweden. *British Journal of Sociology 33.* pp. 1-34.

Fischer, A. & Wirth, H. (2007): *Constructing Version 4 of ESEC Classes from 3-digit ISCO (Stata-do file).* Mannheim: Gesis-ZUMA.

Frick, J., Göbel, J. & Krause, P. (n.d.). $HGEN: Generated Household Level Variables. [http://www.diw.de/documents/dokumentenarchiv/17/60053/hgen.pdf (8.11.2007)].

Ganzeboom, H. & Treiman, D. (1996). Internationally Comparable Measures for Occupational Status for the 1988 International Standard Classification of Occupations. *Social Science Research 25.* pp. 201-239.

Ganzeboom, H. & Treiman, D. (2003). Three Internationally Standardised Measures for Comparative Research on Occupational Status. In H. Jürgen, P. Hoffmeyer-Zlotnik & C. Wolf (Eds.), Advances in Cross-National Comparison. A European Working Book for Demographic and Socio-Economic Variables (pp. 159-193), New York: Kluwer Academic / Plenum Publishers.

Ganzeboom, H., De Graaf, P. & Treiman, D. (1992). A Standard International Socio-Economic Index of Occupational Status. *Social Science Research 21.* pp. 1-56.

Gebhardt, D., Müller, G., Bethmann, A., Trappmann, M., Christoph, B., Gayer, C., Müller, B., Tisch, A., Siflinger, B., Kiesl, H., Huyer-May, B., Achatz, J., Wenzig, C., Rudolph, H., Graf, T. & Biedermann, A. (2009). Codebuch und Dokumentation des ‚Panel Arbeitsmarkt und soziale Sicherung' (PASS). Datenreport Welle 2 (2007/2008). *FDZ Datenreport 06/2009.* Nuremberg.

Gebhardt, D., Müller, G., Bethmann, A., Trappmann, M., Christoph, B., Gayer, C., Müller, B., Tisch, A., Siflinger, B., Kiesl, H., Huyer-May, B., Achatz, J., Wenzig, C., Rudolph, H., Graf, T. & Biedermann, A. (2009). Codebuch und Dokumentation des „Panel Arbeitsmarkt

und soziale Sicherung“ (PASS). Datenreport Welle 2 (2007/2008). FDZ Datenreport 06/2009. Nürnberg.

Granato, N. (2000). Mikrodaten-Tools: CASMIN-Bildungsklassifikation. Eine Umsetzung mit dem Mikrozensus 1996. *ZUMA-Technischer Bericht 2000/12.* Mannheim.

Hagenaars, A., de Vos, K. & Zaidi, M. (1994). *Poverty Statistics in the Late 1980s: Research Based on Micro-data.* Luxembourg: Office for Official Publications of the European Communities.

Halleröd, B. (1995). The Truly Poor: Direct and Indirect Consensual Measurement of Poverty in Sweden. *Journal of European Social Policy 5(2).* pp. 111-129.

Harrison, E. & Rose, R. (2006). ESeC User Guide, Appendix 6 (SPSS-Syntax: Esec Full) [http://www.iser.essex.ac.uk/esec/guide/docs/ Appendix6.sps (31.11.2007)]

Hartmann, J., Brink, K., Jäckle, R. & Tschersich, N. (2008). IAB-Haushaltspanel im Niedrigeinkommensbereich. Methoden- und Feldbericht. *FDZ Methodenreport 07/2008.* Nuremberg.

Hauser, R. (1996). Zur Messung individueller Wohlfahrt und Ihrer Verteilung. In Statistisches Bundesamt (Ed.), *Wohlfahrtsmessung. Aufgabe der Statistik im gesellschaftlichen Wandel* (pp. 13-38), Stuttgart: Metzler-Poeschel.

Helberger, C. (1988). Eine Überprüfung der Linearitätsannahme der Humankapitaltheorie. In H.-J. Bodenhöfer (Hrsg.), *Bildung, Beruf, Arbeitsmarkt* (pp. 151-170), Berlin: Duncker & Humblot.

International Labour Office [ILO] (1990). *International Standard Classification of Occupations. ISCO-88.* Geneva: International Labour Office.

Jäckle, A. (2008). The Causes of Seam Effects in Panel Surveys. *ISEP Working Paper Series 2008-14.* Essex.

Jesske, B. & Quandt, S. (2011). Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 4. Erhebungswelle – 2010 (Haupterhebung). *FDZ-Methodenreport 08/2011.* Nuremberg.

Jesske, B. & Schulz, S. (2012). Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 5. Erhebungswelle – 2011 (Haupterhebung), *FDZ Methodenreport 11/2012*, Nuremberg.

Jesske, B. & Schulz, S. (2012). Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 6. Erhebungswelle – 2012 (Haupterhebung), FDZ Methodenreport 10/2013. Nürnberg.

König, W., Lüttinger, P. & Müller, W. (1987). *Eine vergleichende Analyse der Entwicklung und Struktur von Bildungssystemen. Methodologische Grundlagen und Konstruktion einer vergleichbaren Bildungsskala.* CASMIN-Projekt. Arbeitspapier Nr. 12. Mannheim.

Lechert, Y., Schroedter, J. & Lüttinger, P. (2006). Die Umsetzung der Bildungsklassifikation CASMIN für die Volkszählung 1970, die Mikrozensus- Zusatzerhebung 1971 und die Mikrozensen 1976-2004. *ZUMA-Methodenbericht 2006/12.* Mannheim.

Lengerer, A., Bohr, J. & Janßen, A. (2005). Haushalte, Familien und Lebensformen im Mikrozensus – Konzepte und Typisierungen. *ZUMA-Arbeitsbericht 2005/05.* Mannheim.

Lipsmeier, G. (1999). Die Bestimmung des notwendigen Lebensstandards – Einschätzungsunterschiede und Entscheidungsprobleme. *Zeitschrift für Soziologie 28(4).* pp. 281-300.

Müller, W., Wirth, H., Bauer, G., Pollak, R. & Weiss, F. (2006). ESeC – Kurzbericht zur Validierung und Operationalisierung einer europäischen sozioökonomischen Klassifikation. *ZUMA-Nachrichten 59.* pp. 111–119.

Müller, W., Wirth, H., Bauer, G., Pollak, R. & Weiss, F. (2007): Entwicklung einer europäischen sozioökonomischen Klassifikation. *Wirtschaft und Statistik 5/2007.* pp. 527-530.

Nolan, B. & Whelan, C. (1996). Measuring Poverty Using Income and Deprivation Indicators: Alternative Approaches. *Journal of European Social Policy 6(3).* pp. 225-240.

Organisation for Economic Co-Operation and Development [OECD] (Ed.) (1999). *Classifying Educational Programmes. Manual for ISCED-97 Implementation in OECD Countries.* 1999 Edition. Paris: OECD.

Organisation for Economic Co-Operation and Development [OECD] (Ed.) (1982): *The OECD List of Social Indicators.* Paris: OECD.

Porst, R. (1984). Haushalt und Familien 1982. Zur Erfassung und Beschreibung von Haushalts- und Familienstrukturen mit Hilfe repräsentativer Bevölkerungsumfragen. *Zeitschrift für Soziologie 13(2).* pp. 164-175.

Rammstedt, B. & John, O. (2005). Kurzversion des Big Five Inventory (BIF-K). Diagnostica 51(4). pp. 195-206.

Rendtel, U. & Harms, T. (2009). Weighting and calibration for household panels. In P. Lynn (Ed.), Methodology of Longitudinal Surveys (pp. 265-286), Chichester: Wiley.

Ringen, S. (1988). Direct and Indirect Measurement of Poverty. *Journal of Social Policy 17(3).* pp. 351-365.

Rose, R. & Harrison, E. (2007). The European Socio-Economic Classification: A New Social Class Schema for Comparative European Research. *European Societies 9(3).* pp. 459-490.

Rudolph, H. & Trappmann, M. (2007). Design und Stichprobe des Panels „Arbeitsmarkt und Soziale Sicherung“ (PASS). In M. Promberger (Hrsg.), Neue Daten für die Sozialstaatsforschung: Zur Konzeption der IAB-Panelerhebung „Arbeitsmarkt und Soziale Sicherung“, IAB-Forschungsbericht 12/2007 (S. 60-101), Nürnberg.

Sozialgesetzbuch Zweites Buch [SGB II]: Grundsicherung für Arbeitssuchende.

Spieß, M. & Rendtel, U. (2000). Combining an ongoing panel with a new cross-sectional sample. *DIW-Discussion Papers 198.* Berlin.

Statistisches Bundesamt [StBA] (1992). Klassifizierung der Berufe. Systematisches und alphabetisches Verzeichnis der Berufsbenennungen. Wiesbaden: Statistisches Bundesamt.

Statistisches Bundesamt [StBA] (2002). Klassifikation der Wirtschaftszweige, Edition 2003 (WZ2003). Wiesbaden: Statistisches Bundesamt.

Trappmann, M., Christoph, B., Achatz, J., Wenzig, C., Müller, G. & Gebhardt, D. (2009). Design and stratification of PASS. A New Panel Study for Research on Long Term Unemployment. *IAB-Discussion Paper 5/2009.* Nuremberg.

Trappmann, M. (2013a). Weighting. In: Bethmann, A., Fuchs, B. & Wurdack, A. (Hrsg.)(2013). User Guide "Panel Labour Market and Social Security" (PASS). Wave 6. FDZ Datenreport 07/2013 (S. 56-66). Nürnberg.

Trappmann, M. (2014b). Weights. In: Bethmann, A., Fuchs, B. & Wurdack, A. (Hrsg.)(2013). User Guide "Panel Labour Market and Social Security" (PASS). Wave 6. FDZ Datenreport 07/2013 (S. 81-99). Nürnberg.

Treiman, D. (1977). *Occupational Prestige in Comparative Perspective.* New York: Academic Press.

Wegener, B. (1985): Gibt es Sozialprestige? *Zeitschrift für Soziologie 14.* pp. 209-235.

Wegener, B. (1988): *Kritik des Prestiges.* Opladen: Westdeutscher Verlag.

Imprint

FDZ-Datenreport 2/2014 (EN)

Publisher
The Research Data Centre (FDZ)
of the Federal Employment Agency
in the Institute for Employment Research
Regensburger Str. 104
D-90478 Nuremberg

Editorial staff
Dr. Jörg Heining, Dagmar Theune

Technical production
Dagmar Theune

Download
http://doku.iab.de/fdz/reporte/2014/DR_02-14_EN.pdf

Internet
http://fdz.iab.de/

Corresponding author:
Forschungsdatenzentrum (FDZ) der
Bundesagentur für Arbeit im Institut für
Arbeitsmarkt- und Berufsforschung (IAB),
Regensburger Str. 100, 90478 Nürnberg,
Email: iab.fdz@iab.de

Arne Bethmann, Institut für Arbeitsmarkt-
und Berufsforschung (IAB),
Regensburger Str. 104, 90478 Nürnberg,
Tel.: +49 (0) 911/179-2307
Email: Arne.Bethmann@iab.de