FDZ Datenreport
Documentation of labour market data

No. 6/2009 (EN)

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

Volume I: Introduction and Overview

Wave 2 (2007/2008)

Daniel Gebhardt, Gerrit Müller, Arne Bethmann, Mark Trappmann, Bernhard Christoph, Christine Gayer, Bettina Müller, Anita Tisch, Bettina Siflinger, Hans Kiesl, Bernadette Huyer-May, Juliane Achatz, Claudia Wenzig, Helmut Rudolph, Tobias Graf, Anika Biedermann

Bundesagentur für Arbeit

Table of Contents

Data availability

The dataset described in this document is available for use by professional researchers. Further information can be found under “Individual Data / Household Data on the website: http://fdz.iab.de/.

List of diagrams

List of tables

1. Introduction

1.1 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), is a new dataset for labour market, welfare state and poverty research in Germany, creating a new empirical basis for the scientific community and for policy advice. This Datenreport provides an overview of the second survey wave, for which 12,487 individuals were interviewed in 8,429 households between December 2007 and July 2008. 10,114 individuals and 7,342 households were interviewed for the second time in the context of PASS.

The study is carried out as part of the IAB's research into the German Social Code Book II[1] (SGB II). The IAB has the statutory mandate to study the effects of benefits and services under SGB II aimed at integration into the labour market and subsistence benefits. However, due to its complex sample design, the study also enables researchers to answer questions far beyond these issues. Five core questions influenced the development of the new study, which are detailed in Achatz et al. (2007):

1. What options are there for regaining independence from Unemployment Benefit II?
2. In what ways does the social situation of a household change when it receives benefits?
3. How do the individuals concerned cope with their situation? Does their attitude towards action necessary to improve their situation change over time?
4. In what form does contact between benefit recipients and institutions providing basic social security take place? What are the actual institutional procedures applied in practice?
5. What employment history patterns or household dynamics lead to receipt of Unemployment Benefit II?

[1] Social Code Book II – Basic Social Security for Jobseekers (Sozialgesetzbuch (SGB) Zweites Buch (II) - Grundsicherung für Arbeitsuchende).

1.2 Additions to the existing data

German labour market, poverty and welfare state research already has access to various micro datasets. In particular, there are a number of longitudinal datasets available which already cover relatively long survey periods. A particularly important source in the field of survey data is the German Socio-Economic Panel Study (SOEP) (Wagner et al. 2007), which provides annual data at the individual and household level dating back to 1984. In addition, administrative data from the Federal Employment Agency (BA) is processed at the IAB and provided for research use by the Research Data Centre (FDZ) of the BA at the IAB, for example in the form of the Integrated Employment Biographies (IEBS), the IAB Employment Samples (IABS) or the Linked Employer-Employee Dataset (LIAB).

The spectrum of questions and the design of PASS are intended to close gaps in the existing stock of data. PASS has three main characteristics that extend analysis potential beyond that of the Federal Employment Agency's administrative data:

1. The panel takes the household context into account – including the situation before and after receipt of Unemployment Benefit II.

2. The panel is complete in that it covers all groups of persons and all employment biographies, not only people in dependent employment, unemployed people and those in need of assistance. The dataset also provides information on the status during phases of economic inactivity, self-employment or employment as civil servants.

3. The panel collects additional or significantly more detailed data on relevant characteristics such as attitudes, employment potential or job-search behaviour.

Compared to the existing surveys of individuals or households, PASS aims to improve the data situation in particular with regard to the following points:

1. The high case numbers of Unemployment Benefit II recipients (in wave 2: 4,557 respondents in 3,382 households receiving such benefits at the time when the survey was conducted) make it possible to conduct more detailed analyses – for example on the impact of Social Code Book II on certain target groups – and to ob-

tain more precise estimations of statistics and model coefficients than datasets in which benefit recipients are only included in proportion to their share of the population.

2. Collecting additional characteristics such as the intensity and type of contact to institutions providing basic social security, or participation in employment and training measures makes it possible to analyse the significance of institutional assistance for the population below the poverty line.

3. Linking the survey data with the process-generated data of the Federal Employment Agency makes it possible to validate the characteristics surveyed and also to conduct analyses in which the higher measurement precision of the process-generated data can be combined with further variables and the household context from the survey.

1.3 Special features and innovations of the 2nd wave

At this point we would like to provide a brief outline of the special features of the 2nd wave of PASS for users who have already worked with the data from the first panel wave.

The core topic in wave 2 is the collection of information regarding employment biographies (see Chapter 4.4). In the personal interviews, all periods of employment (with earnings in excess of € 400) and all periods of registered unemployment since January 2005 were recorded in spell form[2]. In addition, where there are gaps in the data, i.e. periods in which there is neither a spell of employment nor a spell of unemployment, the respondent is asked about the status of economic inactivity[3] if the period lasted longer than three months. Furthermore, forgotten periods of employment or unemployment are resurveyed and dating errors are corrected. When no spell of employment or unemployment is available for the time of the interview, the current status is determined using the gap module. These data are supplemented by pooled measures about the duration of periods of employment, unemployment and receipt of social assistance[4]

[2] See personal questionnaire of the 2nd wave: P38 – P63 (employment); P106 – P117 (unemployment)
[3] See personal questionnaire of the 2nd wave: P126 – P132
[4] See personal questionnaire of the 2nd wave: P82 – P104 (employment); P122 – P125 (unemployment); P261f. (social assistance)

during the course of life so far. For the respondents aged 15 to 64, the collection of data about employment biographies replaces the concept of the 1st wave, in which questions were only asked about the current status on the interview date (and possibly its start date). In contrast, employment biography data are not gathered in the 2nd wave either for senior citizens over the age of 65. The periods of employment, unemployment and economic inactivity recorded are prepared in the form of spell datasets[5], which are part of the scientific use file of PASS for the first time with the 2nd wave. The set of questions about job search, which was the core topic of the 1st wave, was reduced slightly in contrast.

Another new feature also results from the introduction of the employment biographies. In the 1st wave, periods during which the respondent had received Unemployment Benefit I were recorded as separate spells. In addition, it was ascertained whether the target person was registered as unemployed on the survey date. From wave 2 onwards, however, in the context of the employment biographies the periods of registered unemployment are recorded in spell form. Here, receipt of Unemployment Benefit I is surveyed as part of a spell of unemployment in order to avoid a dual response burden and discrepancies between the different statements. The spell dataset for Unemployment Benefit I which was part of the PASS scientific use file in the first wave is not continued in the second wave. Information about the periods in which the respondent drew Unemployment Benefit I is embedded in the unemployment spell dataset (*al_spells*) from the 2nd wave onwards and is therefore included in the spells of registered unemployment.

The recording of participation in employment and training measures was also changed in the 2nd wave (see Chapter 4.5). It had emerged that it was not possible to identify the measure types clearly by means of the questions asked in the 1st wave. As a result, the module of the personal questionnaire concerning participation in employment and training measures was thoroughly revised. The measure type is now identified at the start of the module by means of a multiple choice question. The changes in the survey instrument are reflected in the data structure. As it would not have been easy to continue the measure spell dataset recorded in the 1st wave, it was decided to create a new dataset for the measure spells recorded from the 2nd wave onwards (*mn_spells*).

[5] Employment spells: et_spells; unemployment spells: al_spells; spells during periods of economic inactivity: lu_spells

In the 1st wave of PASS, both at the household[6] and the individual level[7], questions were asked not only about the overall household income and the gross and net income, but also about individual income components. The set of questions, which was still incomplete in the 1st wave, was expanded considerably for the 2nd wave and now permits the generation of a component income. However, missing values for individual items (item non-response) and households which were not interviewed fully (unit non-response) make it difficult to produce the component income. In a project being conducted within the IAB an imputation concept is currently being developed. The aim is to make available a generated component income for wave 2 with the scientific use file of the 3rd wave.

Furthermore in the 2nd wave a so-called refreshment sample was drawn for the BA subsample[8]. The aim of this is to guarantee the representativeness of the BA sample in the cross-section, too, and also to be able to observe sufficient new transitions into receipt of Unemployment Benefit II over time. For the refreshment sample, households are drawn which were in receipt of Unemployment Benefit II on 1 July 2007 but not on the sampling date for the 1st wave (see Chapter 0 and, on the concept of the refreshment sample, Trappmann et al 2009: 11ff.). These households which were surveyed for the first time in the 2nd wave can be identified via the sample indicator (*sample*).

The household and individual datasets of PASS are prepared in the form of cross-wave files in so-called long format. There is a separate row in the data matrix for every survey wave in which a household or a person was interviewed. In addition to the identifier for the household or the person, the wave indicator (*welle*) is also required in order to identify an observation clearly (see Chapter 10.1). The two register datasets (*hh_register; p_register*) provide information for example about the waves in which a household or a person was interviewed and about the reason why there is perhaps no interview available for individual waves

[6] See household questionnaire of the 2nd wave for households surveyed for the first time (HHneu): HH20 - HH23; HH63 – HH 85; household questionnaire of the 2nd wave for re-interviewed households (HHalt): HH81 – HH84; HH105 – HH129

[7] See personal questionnaire of the 2nd wave: P32 - P37; P64 – P79; P249 – P260P249 – P260; senior citizens' questionnaire of the 2nd wave: P20 – P28

[8] The 1st wave of PASS consists of two subsamples: (1) A sample of households in receipt of Unemployment Benefit II drawn from the administrative data of the Federal Employment Agency (BA), and (2) a general population sample, stratified by status, drawn from a database provided by the commercial provider MICROM.

Certain information at the household or individual level which was collected in the 1st wave (e.g. the highest general school qualification in the personal interview) is simply continued or updated in the follow-up survey in the 2nd wave, i.e. only changes since the time of the last interview are recorded. In these cases the variables of the 2nd wave do not contain the complete information about this household or this person, but only the changes that occurred between the two interviews. The complete information is generally made available in generated variables (e.g. for the highest general school qualification attained: *schul1*; *schul2*) (see Chapter 7.3).

With the delivery of the scientific use file of the 2nd wave a problem in the weighting datasets is also rectified (see Chapter 9.2). According to the scientific use file of the 1st PASS wave, which was published in summer 2008, there were 3.3 million foreigners in Germany. This figure is unrealistically low. Such divergences with regard to key target variables should not occur, as the weights are calibrated to benchmark statistics of the microcensus. However, in this case the Federal Statistical Office provided implausible benchmark figures. In order to rectify the error the calibration was repeated using the correct benchmark figures. The projection factors of the total sample and of the Microm sample[9] are affected by the changes, but the weights of the BA sample are not. Larger divergences in the results can be expected in analyses concerning people with a migration background.

[9] The following weights are affected by the changes: *qwhh*, *wqmihh* (*hweights*) and *wqp* and *wqmip* (*pweights*)

2. Design

2.1 Survey design

By establishing PASS, the IAB is setting up a new data base that creates a new empirical basis for research into the labour market, the welfare state and poverty in Germany. The survey pays particular attention to the dynamics of households in receipt of benefits in accordance with the Social Code Book II (SGB II) (see Chapter 1.1 on the objectives and questions of PASS, and in more detail Achatz et al 2007: 17ff.).

When developing the design of the study, the diverse requirements as regards contents and the challenges associated with the specific composition of the sample made it necessary to devise corresponding strategies (see Schnell 2007, Rudolph & Trappmann 2007). The following paragraphs describe some of the main challenges and the resulting special features of the survey design.

Analyses of inflows into receipt of Unemployment Benefit II, comparisons of households in receipt of benefits with households not receiving benefits, the investigation into hidden poverty and the formation of control groups require a comparison of benefit recipients with the rest of the population (see Promberger 2007). For this reason PASS combines subsamples from two sources: benefit recipient households drawn from the administrative data of the Federal Employment Agency (BA sample) and a sample of the general population disproportionately stratified according to status (Microm sample), which was drawn from the address database of a commercial provider.

In most of the cases longitudinal data are required in order to examine the key research questions of PASS. PASS is therefore structured as a panel study in which the same individuals and their households are to be re-interviewed in each wave. Only in this way is it possible not only to analyse inflows into and outflows from benefit receipt but also, for example, to examine changes in attitudes, action taken or the material situation before and after beginning benefit receipt.

In order to be able to analyse inflows into receipt of Unemployment Benefit II already after a short time and to guarantee the representativeness of the BA sample in the cross-section, a refreshment sample was drawn for the BA sample in the 2nd wave (on the concept of the refreshment sample see Trappmann et al 2009: 11ff.). For this,

households were drawn which were not in receipt of Unemployment Benefit II on the sampling date of the 1st wave but were receiving this benefit on 1 July 2007.

When examining research questions in the context of the Social Code Book II (SGB II) the respondents' action context and in particular here their household context is of importance for two different reasons: first, because the individuals always make decisions against the background of their household-specific circumstances. Second, because the SGB II also always examines the household context when activating benefit recipients, in the context of “support and demand“ (see Achatz et al. 2007). This is one of the reasons why PASS is designed as a household survey: within a household all members aged 15 or above are to be interviewed with a person-level questionnaire. The personal interviews are always preceded by a household interview in which general, household-related information is gathered.

In addition to the content-related demands on the study there are also methodical challenges which result from the special basic conditions and which are taken into account by special aspects of the study design. One of the challenges is being able to contact the survey population by telephone. A large proportion of the benefit recipients can not, or not easily, be contacted by phone, as there can be expected to be not only a large number of new telephone numbers but also a large share of unlisted numbers, for example, of mobile phones (see Schnell 2007, Rudolph & Trappmann 2007). In PASS this problem is addressed using a combination of telephone interviews (CATI) and personal interviews (CAPI) in a mixed-mode procedure. In order to cover the expected large proportion of people in the target population who have little or no knowledge of the German language, the survey instruments are translated into other languages and are used in the field by interviewers who are native-speakers of the respective language. In the 1st wave the questionnaire was translated into both Turkish and Russian and also into English and was used in the fieldwork in these versions. As only very few interviews were conducted in English in the 1st wave, however, the survey instrument was no longer translated into English for the 2nd wave.

2.2 Sampling design and sampling procedure

The sample in the 1st wave of PASS consisted of two subsamples. These two otherwise independent samples are connected in the first sampling stage via the selection of identical primary sampling units (for detailed information about the sampling design of the 1st wave see Rudolph & Trappmann 2007: 65ff.). The first subsample (BA sample) is a random sample of so-called 'Bedarfsgemeinschaften'[10], which can be roughly translated as 'benefit communities', in which at least one person was receiving Unemployment Benefit II in July 2006. This sample was drawn from the administrative data of the Federal Employment Agency (BA). As PASS is a household survey, the entire household in which a benefit community was living was recorded in each case. The second subsample is a sample of private households in Germany (Microm sample). For this a random sample of addresses was drawn from the MOSAIC database of addresses held by the commercial provider Microm. The sample was stratified disproportionately by status in such a way that households with a low social status and thus a greater risk of entry into benefit receipt had a higher probability of inclusion (on the results of the stratification see Trappmann et al. 2007).

In the first sampling stage 300 postcodes were drawn from the postcode register. These postcodes serve as primary sampling units in PASS (on the selection of the primary sampling units see Rudolph & Trappmann 2007: 77ff.). The selection probability of a postcode sector was dependent on the number of households in the particular sector according to the MOSAIC database (probability proportional to size). Within each sampling point benefit communities (BA sample) or addresses (Microm sample) were drawn. The number of benefit communities to be drawn for the BA sample depended on the rate of benefit recipients (number of benefit communities in the sampling point

[10] A so-called 'Bedarfsgemeinschaft', or 'benefit community', includes all individuals in a household who receive benefits jointly (i.e. as a joint payment). In the majority of cases the 'benefit community' and the household will be the same, which in particular applies in the case of (married or unmarried) couples and parents with children under the age of 25. However, under specific circumstances the 'benefit community' might not include all household members, or a household in which everybody receives benefit payments might be made up of more than one 'benefit community'. An example of the former is if a grown-up child lives with his/her parents and earns just enough to make his/her own living but has insufficient means to support his/her mother and father – in this case the 'benefit community' will only include the parents. An example of the latter is a three- (or more) generation household: since a 'benefit community' may only consist of two generations, this type of household will be made up of two such 'benefit communities', one consisting of the grandparent(s) and one of the parent(s) and child(ren). (Trappmann et al. 2007)

according to BA process data divided by the number of households in the sampling point according to the MOSAIC database). On average 20 ‘benefit communities' were selected per sampling point. As the number of selected ‘benefit communities' is proportional to the benefit recipient rate in the sampling points, a uniform selection probability is also guaranteed in the BA sample (Rudolph & Trappmann 2007: 78ff.). All members of each household in which a ‘benefit community' was living were surveyed.

For the Microm sample 100 addresses were drawn within each sampling point. In order to obtain an over-representation of the lower status classes, addresses of lower status classes had a higher inclusion probability. The addresses drawn in this way were visited by employees of the field institute conducting the survey, who wrote down all of the names that were on the doorbell panels. At the field institute a random selection of these doorbells was made. If a doorbell panel had more than one name on it, one of these names was selected. Each selected person’s entire household was surveyed.

All of the households in the two samples of the 1st wave were to be re-interviewed in the 2nd wave (see Chapter 3.2 regarding 2nd wave response rates). In addition to this, in the 2nd wave households that had split off from the households interviewed in the 1st wave were also surveyed (see Chapter 3.4). They were each assigned to the subsample from which their original household had been drawn in the 1st wave.

In addition, in the 2nd wave a refreshment sample was drawn of benefit communities that had begun receipt of Unemployment Benefit II. These are benefit communities which were receiving Unemployment Benefit II in July 2007 but not in July 2006 (the time when the sample for the 1st wave was drawn). The sample was drawn in the postcode sectors that had already been selected for the 1st wave (primary sampling units) following the procedure used in the first wave. The households in the refreshment sample (with benefit receipt in July 2007) and those households from the BA subsample of the 1st wave which were still in receipt of Unemployment Benefit II in July 2007, taken together, can be projected to all households with at least one recipient of Unemployment Benefit II in Germany at that time.

3. Key statistics

This chapter provides a brief overview of key statistics of the study, such as sample sizes and response rates. These figures refer to the 2nd wave and are reported separately both for the two original subsamples and the refreshment sample and for the study as a whole.

- Subsample 1 (BA sample) refers in the following to the sample of benefit recipients from the process data of the Federal Employment Agency.
- Subsample 2 (Microm sample) refers to the stratified population sample.
- Refreshment sample (BA sample) is the name of the sample drawn from the inflow of benefit recipients who began to receive Unemployment Benefit II between wave 1 and wave 2.

3.1 Sample size

The sample of the 2nd wave of PASS includes a total of 8,429 households. Table 1 shows the distribution of these households across the three subsamples and the modes of data collection.

Table 1: Sample size at the household level

n	CATI	CAPI	Total
Subsample 1 (BA)	2,301	1,190	3,491
Subsample 2 (Microm)	2,299	1,598	3,897
Refreshment sample (BA)	778	263	1,041
Total	5,378	3,051	8,429

At the individual level these 8,429 households consist of 12,487 respondents. Table 2 depicts the corresponding distribution of these individuals across the samples and the modes of data collection.

Table 2: Sample size at the individual level

n	CATI	CAPI	Total
Subsample 1 (BA)	3,105	1,648	4,753
Subsample 2 (Microm)	3,795	2,597	6,392
Refreshment sample (BA)	988	354	1,342
Total	7,888	4,599	12,487

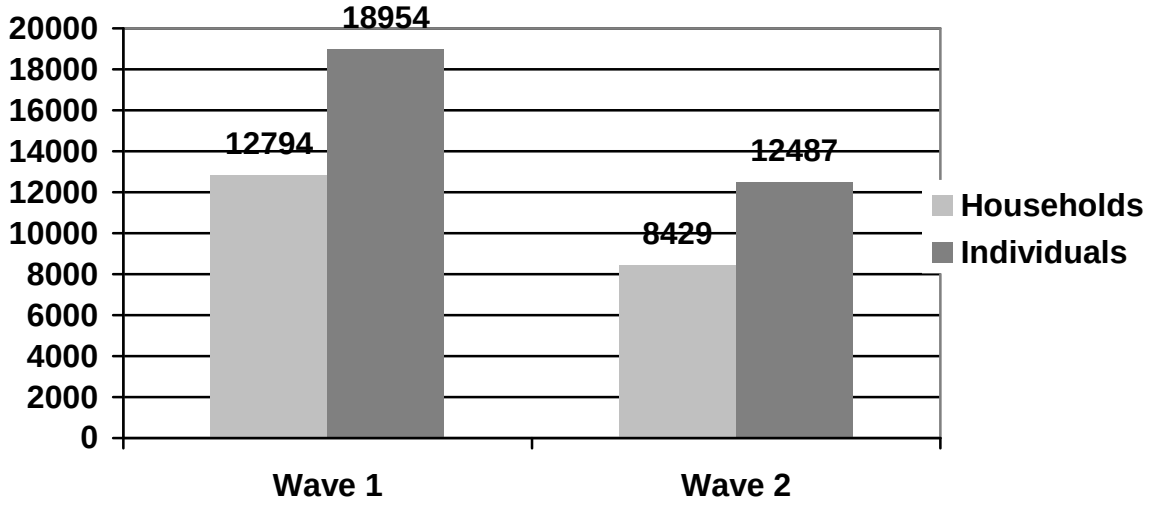
As explained in Chapter 2.1, respondents without sufficient knowledge of the German language had the option of being interviewed in another language. The alternative interview languages on offer were Turkish and Russian. In the 2nd wave an interview in English was no longer offered. Table 3 shows how many households or individuals were interviewed in the respective languages.

Table 3: Sample size of interviews in languages other than German

n	Households	Individuals
Russian	156	219
Turkish	39	31
Total	195	250

As a result of panel mortality it is not always possible to re-interview all of the households and individuals in the subsequent waves. On the other hand, new survey units also enter the study via refreshment samples and split households. Figure 1 provides an overview of the sample sizes at the household and the individual level in the different waves.

Figure 1: Samples sizes of the different waves



3.2 Response rates

In a household survey one can distinguish between the response rate at the household level and the response rate within households. The “response rate at the household level” means the share of usable households as a proportion of the total of all usable households plus the cases of non-neutral non-response. Only households in which all members have died and households which have moved abroad permanently are regarded as cases of neutral non-response. The response rate RR1 in accordance with AAPOR standards (AAPOR 2006) is reported, in which all cases of unknown eligibility are also included in the denominator and which therefore assumes the lowest value of all response rates[11]. Re-interviewed households considered usable if at least one complete household interview was conducted. New households are only considered usable if not only the household interview but also at least one complete personal interview is available. The “response rate within the households” is used to denote the average proportion of all household members aged 15 or above within usable households for whom a complete personal interview is available.

Response rate at the household level

The following response rates were obtained at the household level for the 2nd wave [12]

Table 4: Response rates at the household level

Response rate	CATI	CAPI	Total
Re-interviewed households (BA)	48.9 %	58.5 %	51.1 %
Re-interviewed households (Microm)	66.9 %	61.9 %	64.7 %
New households	27.0 %	24.5 %	26.3 %
Split-off households	13.3 %	14.8 %	13.4 %
Total	41.0 %	57.4 %	45.0 %

Response rate within the households

The following response rates within the households are reported:

Table 5: Response rates within the households

Response rate	CATI	CAPI	Total
Re-interviewed households (BA)	84.0 %	89.8 %	85.5 %
Re-interviewed households (Microm)	83.8 %	87.1 %	85.1 %
New households	85.8 %	90.1 %	86.2 %
Split-off households	88.3 %	87.5 %	88.3 %
Total	84.2 %	88.2 %	85.4 %

Repeat interview rate at the individual level

In addition to the response rates at the household level and within the households, the following table shows the repeat interview rate at the individual level. This reports the

[11] This is dealt with in very different ways in Germany. Frequently a large number of individuals or households that were not interviewed are counted as “ineligible” and are removed from the denominator when the response rate is calculated. When a sample is drawn from registers, however, 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 counted as a case of neutral non-response. Moreover, the population of PASS is not restricted to German-speaking respondents or to people who are able to be interviewed, so the non-response reasons “does not speak German” or “respondent is sick / unable to be interviewed” can not be regarded as cases of neutral non-response either.

[12] The cases are assigned to the data collection mode which was used at the beginning of the fieldwork. Subsequent changes from CATI to CAPI or vice versa are not recorded in the statistics. The figures for new respondents are taken from the Infratest field report for the second wave of PASS (Büngeler et al. 2009).

proportion of individuals interviewed in wave 1 with whom it was also possible to conduct an interview in wave 2.

Table 6: Repeat interview rates at the individual level

Repeat interview rate	
Subsample 1 (BA)	48.0 %
Subsample 2 (Microm)	65.2 %
Total	56.6 %

3.3 Agreement to panel participation and mergers, linking with process data

The respondents' consent is always required for saving addresses for the purpose of repeat interviews in the next wave and for merging the survey data with the process data of the Federal Employment Agency. In both cases high consent rates were obtained from the respondents taking part for the first time in wave 2:

Willingness to participate in panel: 97.7 %

Consent to merging of process data: 84.2 %

Cases that could actually be linked: 85.4 %

3.4 Split households

PASS is designed as a dynamic panel. People who move into or are born into sample households are also interviewed as long as they are aged 15 or over. People who move out of sample households or do not live in the household for a year or longer should continue to be interviewed, however. These individuals' new households are seen as split-offs from the original sample households. These split-off parts of the household (or split-off households) themselves become sample households of PASS. All of the people over the age of 15 living in these households become target persons for personal interviews. Should it occur in one of the subsequent waves that part of this split-off household in turn splits off, then this new split-off household, too, becomes a PASS sample household, irrespective of whether there is still anyone from one of the original samples living there ("infinite degree contagion model", Rendtel and Harms 2009, 267). Individuals who have moved abroad, on the other hand, cease to be included in the survey as they no longer belong to the population and because the research questions specific to SGB II no longer apply. People who do not live in the

household for less than a year continue to be counted as household members and do not constitute a new PASS household.

Between the survey dates of the 1st and 2nd waves a total of 344 households split off from the households already included in the first wave of the survey. It was possible to interview 46 of these split-off households during the fieldwork period of the 2nd wave. The split-off households that were not surveyed will be contacted again in the 3rd wave as long as they have not definitely refused to participate. The split-off households can be identified in the datasets by comparing the current household number (*hnr*) and the original household number (*uhnr*), which differ 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 take over from their original household the sample indicator (*sample*), the information as to the sampling year (jahrsamp), the primary sampling unit (*psu*) and its stratification (*strpsu*).

4. Instruments and interview programme

4.1 Overview and innovations

In PASS information is collected by means of separate questionnaires at the household and the individual level. First a household interview is conducted with each household. In this interview information referring to the entire household is gathered (see Chapter 4.2). The target person for this household interview[13] is already selected during the contact phase which precedes the actual interviews. The household interview is followed by personal interviews (see Chapter 4.3) with the individual household members. The aim is to hold a personal interview with all of the persons living in the household who are aged 15 or over – household members who are 65 or older receive a short version of the questionnaire (senior citizens' questionnaire) which does not include questions that are irrelevant for this age group.

The survey instruments and interview programme of the 2nd wave are based on those used in the 1st wave of PASS, but have been adapted and revised. The introduction of the employment biography module resulted in fundamental changes to the way in which information is collected about periods of employment, unemployment and economic inactivity at the individual level (see Chapter 4.4).

The instruments of the 1st wave were designed for the initial survey of the households and individuals in PASS. The instruments of the 2nd wave permit both initial inter-

[13] The target person for the household interview should know as much as possible about general issues regarding the household. In re-interviewed households this was the same person who had completed the questionnaire in the 1st wave. If this person was not available during the entire fieldwork period or was no longer a member of the household, then another adult who knows a lot about the household was selected. In the refreshment sample, which was drawn from the BA data, the person registered with the BA as the household reference person should answer the questions about the household. In the case of split-off households a person who used to be a member of the original household and is at least 15 years old should be selected as the target person. If the person with whom the household interview was conducted in the original household in the 1st wave now lives in the split-off part of the household, then this person should be selected as the target person of the household interview in the split-off household. Whenever a particular target person who was already known by name was not available during the fieldwork period, the interviewers tried to conduct the household interview with a person aged over 15 who knew as much as possible with regard to general household issues.

views[14] and repeat interviews with households and individuals who had already taken part in the 1st wave. In order to avoid seam effects[15] in the repeat interviews and to increase the data quality, dependent interviewing was used for certain questions to update information that the respondent had provided in the last interview. Furthermore, information about constant characteristics was generally not gathered again (see Chapter 10.2). This resulted in corresponding adaptations to the respective question modules, as respondents taking part for the first time and those participating in a repeat interview sometimes had to be asked different questions or questions had to be skipped. Owing to the complex updating of the household structure, at the household level a separate questionnaire was developed for re-interviewed households (HHalt), whilst only minor changes were made to the questionnaire for households taking part in the survey for the first time (HHneu).

In addition, the module on participation in measures of active labour market policy was thoroughly revised for the 2nd wave, after it had become clear in the 1st wave that in many cases it was not possible to identify the measure type reliably (see Chapter 4.5).

In contrast to the 1st wave, no proxy or refuser interviews were conducted in the 2nd wave. They were not included in the scientific use file in the 1st wave anyway.

The following sections are intended to provide an overview of the subject spectrum of the survey at the household level (Chapter 4.2) and the individual level (Chapter 4.3). The new employment biography module (Chapter 4.4) and the revised employment and training measure module (Chapter 4.5), which are part of the personal questionnaire (but not the senior citizens' questionnaire) in the 2nd wave, are presented separately.

[14] In the 2nd wave there were two types of household that were included in the survey for the first time: (1) households in the refreshment sample of the 2nd wave and (2) households which had split off from households that were involved in the 1st wave of the survey (split-off households). Furthermore, two types of individual were interviewed for the first time: (1) individuals who were members of a PASS household for the first time in the 2nd wave and (2) individuals who were already members of a PASS household in the 1st wave but for whom no interview from the 1st wave is available.

[15] In a panel data set the number of changes observed at the interface (seam) between one interview and the one conducted in the subsequent panel wave is often considerably higher than the number of changes observed within one interview (see Jäckle 2008)

4.2 Household level

In the 2nd wave two versions of the household questionnaire were used: (1) the household questionnaire for re-interviewed households and (2) the household questionnaire for households taking part in the survey for the first time, i.e. households from the refreshment sample of the 2nd wave or households which had split off from households that were part of the 1st wave of PASS. The main difference between the two versions of the questionnaire is in the collection of data regarding the current household composition as of the interview date. In the case of re-interviewed households, the household structure of the previous wave was updated. For this a) information about people who had moved into the household was incorporated, b) people who had moved out of the household since the previous wave were identified and c) changes in the relationships between the household members were recorded. In contrast, in the case of households taking part in the survey for the first time information was collected about all of the persons living in the household and their relationships to one another.

The household interview of the 2nd wave consists of the following seven subject blocks:

Table 7: Subject blocks of the household interviews in wave 2

Subject block
Household size and composition, family relationships, languages spoken
Household's ownership of consumer goods
Housing and housing costs
Income, assets, debts
Receipt of Unemployment Benefit II
Child-care
Living conditions

When collecting data on the household structure[16], information was collected not only about the size of the household and the age and sex of the household members but also about the relationships between the various household members[17]. Besides, ques-

[16] Questionnaire HHneu: HH1, HH25-HH38; HHalt: HH1-HH60; variable codes: HA and HD

[17] The information regarding the relationships between the household members is not contained in the scientific use files for data protection reasons. However, various generated variables are prepared on the basis of this data, for example indicators of partners and parents living in the household, synthetic benefit communities (‘Bedarfsgemeinschaften’), marital status or the household classification.

tions about the languages spoken[18] in the household were asked for the first time in the 2nd panel wave. The material situation of the households was recorded in two ways: in addition to asking the respondents about the net household income[19], various income components[20] as well as savings, debts and loans[21], a deprivation index[22] is used to gain information about ownership of consumer goods. Information about the housing situation[23], in particular the size and costs of housing, is also surveyed or updated. For re-interviewed households which had moved house since the previous wave, their reasons for moving were also recorded in the 2nd wave. Receipt of Unemployment Benefit II and Unemployment Benefit II cuts were recorded in spell form retrospectively from a particular reference date in the past until the interview date. In re-interviewed households, benefit receipt was recorded for the period since the last interview. Here spells which were ongoing as of the interview date in the previous wave were first updated. In households of the refreshment sample, benefit receipt in the period between the last change in the household composition and the interview date was surveyed if the household had experienced changes in composition after January 2006. If the last change in the household composition was before this date then cases of Unemployment Benefit II receipt since January 2006 were surveyed instead. In split-off households, in contrast, the date on which the split-off part of the household moved out of the original household was relevant as long as the person with whom the household interview was conducted in the original household in the 1st wave was not living in the split-off part of the household in the 2nd wave. If this was the case, however, data was gathered about receipt of Unemployment Benefit II in the split-off household since the date of the household interview of the original household in wave 1. The care of the children living in the household, the costs associated with child-care and any employment restrictions were also recorded in the 2nd wave at the end of the household questionnaires[24], before questions regarding the assessment of the current circumstances and expectations about future living conditions[25] concluded the questionnaire.

[18] Questionnaire HHneu: HH42; HHalt: HH85-HH89; variable code: HD
[19] Questionnaire HHneu: HH77-HH82; HHalt: HH121-HH126; variable code: HEK
[20] Questionnaire HHneu: HH20-HH23, HH63-HH76; HHalt: HH81-HH84, HH105-HH120; variable code: HEK
[21] Questionnaire HHneu: HH83-HH85; HHalt: HH127-HH129; variable code: HEK
[22] Questionnaire HHneu: HH2-HH4; HHalt: HH61-HH63; variable code: HLS
[23] Questionnaire HHneu:HH5-HH19; HHalt: HH64-HH80; variable code: HW
[24] Questionnaire HHneu: HH86-HH92; HHalt: HH130-HH136; variable code: HKI
[25] Questionnaire HHneu: HH93-HH94; HHalt: HH137-HH138; variable code: HA

4.3 Individual level

In the personal interview information is gathered about the personal situation of the particular household member. The aim is to interview all household members who are aged 15 or over. Older respondents from the age of 65 onwards receive a shorter version of the personal questionnaire, the senior citizens' questionnaire, which excludes questions that are not relevant for this group (regarding employment biographies, participation in employment and training measures etc.).

The personal interview consists of the following subject blocks:

Table 8: Subject blocks of the personal interview in wave 2

Subject block
Date of birth
Attitudes regarding standard of living
Education/training
Employment (employment history since January 2005; first/ last job; pooled measures on the entire employment biography) and current earned income
Unemployment and receipt of Unemployment Benefit I (history since January 2005; pooled measures on the entire unemployment history
Other activities since January 2005 (e.g. activity as housewife, househusband, retirement, parental leave)
Assumptions regarding self-efficacy, attitudes towards family, work and dealing with money
Contact to institutions providing basic social security
Participation in employment and training measures
Job-search
Social integration
Health
Care
Partnership
Children
Pensions
Social origin and migration
Concluding question

Asking for the date of birth again in the 2nd wave served among other things to check whether the personal interview was being conducted with the correct person. In addition to the questions asked at the beginning of the personal questionnaire about the living circumstances, questions were also asked during the course of the interview

about attitudes towards other subject areas. For instance, data was gathered on perceived self-efficacy, on attitudes towards work, family and occupation as well as on the allocation of resources and on decisions in partnerships. Respondents with children were also asked about their aspirations as parents with regard to their children's education and their attitudes towards their children's education and training.

People for whom no information was available from the previous wave about general-education and vocational qualifications were asked about these for the first time. Respondents who had already provided information about their qualifications in the previous wave were asked about when they had gained the particular qualifications and about qualifications gained since the previous wave.

Data on periods of employment, unemployment and economic inactivity were gathered on the basis of a new concept in the 2nd wave (see Chapter 4.4 for more details). Whereas in the 1st wave information was only collected about the employment status at the time of the interview, in the 2nd wave the employment history was surveyed in spell form. Here respondents were asked for retrospective information about periods of employment for which they had earned more than € 400 per month, and about spells of registered unemployment, for the period between January 2005 and the date of the interview in wave 2. If a comparison of the spells of employment and unemployment revealed gaps of more than three months, datings could be corrected or new spells could be included in the gap module. The gap module was also used when details about the current employment status were unclear. In this way it was also possible to record periods of economic inactivity which did not concern unemployment, for example military service, community service (as an alternative to military service) or periods covered by maternity protection. In addition to this, information was collected about the first and last job and pooled measures were recorded about periods of employment and unemployment referring to the entire working life. Receipt of Unemployment Benefit I was no longer recorded as a separate spell type in the 2nd wave, but in the context of unemployment spells with reference to a definite period of registered unemployment. Owing to the changes in the survey logic, the spell data on Unemployment Benefit I from the 1st wave were not updated. Net and gross earned income were only surveyed if the respondent was in employment on the interview date. Information about so-called 'mini-jobs' and income earned from such jobs was also restricted to the current status.

In the senior citizens' interview no employment-history information was collected – there the concept used in the 1st wave was retained.

The question complex on contact with the institutions providing basic social security (employment agencies working in cooperation with local authorities (ARGE) etc.) was revised only slightly. In the 2nd wave, too, questions were asked for example about personal contacts, the number and content of talks about occupational and personal circumstances, placement offers, integration assistance and about how the respondent assesses the contact. The module for recording participation in employment and training measures, on the other hand, was thoroughly re-worked (see Chapter 4.5), which is why a new dataset was created for spells in such measures.

Job-search activities was the special focus topic of the 1st wave. Accordingly, the set of questions had to be reduced slightly in the 2nd wave: questions were no longer asked about the reasons why the respondent was seeking work or why he/she was unable to take up or seek work.

As in the 1st wave, information continued to be collected about health (e.g. visits to the doctor, disabilities and serious health problems as well as subjective opinions) and about social integration (e.g. about close friends/relatives and activity in organisations).

The question block on the subject of pension payments was expanded considerably in the 2nd wave. Whereas only general information as to whether the respondent draws a pension and about the amount received is available for the survey period of the 1st wave, with the 2nd wave it is possible to distinguish between different types of pension and the respective payments.

The subject areas of social origin and migration were also expanded in the 2nd wave. In addition to collecting information about the respondent's parents' general-education and vocational qualifications, questions were also asked about the parents' occupational status and their jobs at the time when the respondent him/herself was 15 years old. Detailed questions about the migration background (up until the third generation) was retained in the 2nd wave, too. Furthermore, the module was extended to include questions about residence permits and about the languages spoken in the respondent's circle of friends.

4.4 Employment biographies

In the 2nd wave the concept for gathering information about unemployment, employment and economic inactivity as well as receipt of Unemployment Benefit I was thoroughly reworked. An integrated biography module was introduced which surveys periods of Unemployment Benefit I receipt, registered unemployment, employment and economic inactivity retrospectively in spell form. The interval surveyed here covers the period from January 2005 until the date of the interview for the 2nd wave. It is planned to continue the respondents' biographies in future waves of the survey. In wave 1, in contrast, only periods of Unemployment Benefit I receipt were recorded in spell form, whilst information about employment and economic inactivity was recorded as cross-sectional information at the time of the interview. By integrating the biography module, in which developments in the periods between the individual survey waves can now be recorded, a considerably more detailed picture of respondents' employment histories is obtained than was the case in the 1st wave of PASS and than is usual in "classic" panel studies. The decision to completely rework the survey concept had an impact both on the survey instrument and on the cleaning, editing, and structure of the data. These two factors are outlined in the following.

The employment-biography data are collected in modular form and consist of an employment module, an unemployment module and a "gap module". These three modules are run through one after the other in the interview. First, spells of employment and then spells of unemployment that fall in the time frame described above are recorded. Then a gap module is used to test whether there are still any unexplained time periods left. The relevant status and any other information for these gaps is then (re)surveyed or corrected. In order not to burden the respondents excessively with reconstructing and reporting the individual spells, gaps of very short duration (i.e. three months or less) were disregarded when checking the gaps in the data. An attempt was made to ensure that the sets of questions within the modules corresponded as far as possible with the contents of the first wave. In the employment module (P38–P63) all occupational activities are recorded which were performed in January 2005 or later and for which the respondents earned more than € 400 per month. For each activity reported the same characteristics are surveyed which were surveyed for the current employment as of the interview date in the 1st wave. In addition, the possibility to report parallel jobs

was explicitly intended in order to be able to record cases in which a person had several part-time jobs at the same time. Marginal employment (with an income of no more than € 400) are recorded outside the biography module and not in the form of employment-history data but, as in wave 1, as cross-sectional information about the current status. In the unemployment module (P108–P117), for the same period as was surveyed for the employment spells, information is requested about spells during which the individuals were registered as unemployed or were participating in an employment or training measure run by an employment agency. For each spell of unemployment reported, questions are then asked about whether Unemployment Benefit I was drawn during the spell of registered unemployment, and if so, for what period during the registered unemployment. If, as was described above, there are still unexplained periods lasting longer than three months after the employment and unemployment spells have been reported, the gap module begins (P126–P132). In order to fill gaps, this module can be used first to correct the dates of spells which were reported previously and second, to create new spells which make it possible to provide information not only regarding other spells of employment or unemployment but also, for example, about the status of economic inactivity for remaining periods of time. The gap module and the recording of employment biographies are completed when no more gaps can be found and information about the current status as of the interview date is available.

4.5 Participation in employment and training measures

The concept for surveying participation in employment and training measures was also thoroughly reworked. The experiences made in wave 1 revealed that in many cases it was not possible to identify clearly the exact type of measure on the basis of questions P101-P103 and the follow-up questions P104-P107 (incl. the open-ended question about the name of the measure. The only exception was so-called one-Euro jobs, which could be recorded directly as such in P102. For this reason, in wave 2 the type of measure in which a person had participated was first recorded directly using multiple choice questions (P159, P160). Then further information was collected in the form of looped sequences of questions about the reported measure types, covering such items as the starting date and duration/end of the measure and other details about the particular measure.

5. Dataset structure

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

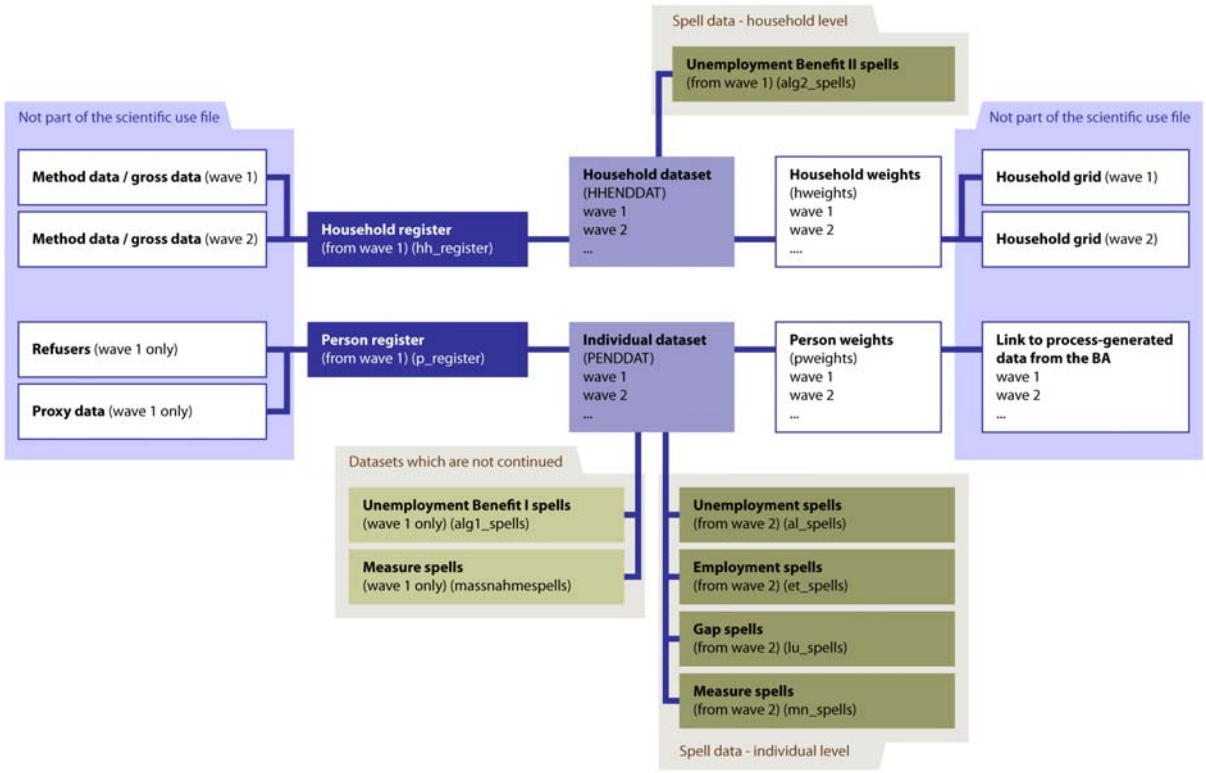
This data structure makes it possible to store the information using relatively little storage space. Which variables were surveyed in which year can be recognised immediately when looking into the datasets. The merging with additional information – via key variables such as household or personal identification numbers – is also comparatively simple. However, this structure which is usual for panel data also has disadvantages which make it quite difficult to work with these datasets. If analyses are to be conducted not only in the cross-section but also in the longitudinal section, then first all of the relevant variables from the individual datasets of the respective waves have to be integrated into a common dataset, whereby care must be taken to ensure that the constructs selected really are the same with regard to contents. For typical longitudinal analyses the cross-wave dataset created in this way then has to be reshaped into so-called long format. In contrast to wide format, in which the data matrix contains precisely one row for each observation unit (e.g. a household or an individual) and then several datasets exist for each survey wave, in long format all of the waves allocated to one observation unit are arranged below one another. Instead of arranging the 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 of the observation units.

Preparing the data in long format has both advantages and disadvantages. The decisive advantage of this variant is that the data are already available in the structure required for many longitudinal analyses (such as event history analyses). It is no longer necessary to invest additional time and effort in creating a cross-wave file. The switch from long format to wide format is also comparatively easy to perform. STATA for example provides a possibility to switch between the two formats with little effort using the “reshape” command. Until a few years ago, the central argument against using this type of dataset structure was the significantly larger memory space required, which

mainly results from the fact that even variables recorded in only one or a small number of survey waves always require a complete column across all waves in the dataset. In addition, the long files become relatively large with increasing duration of the panel, simply as a result of all annual waves being appended to one another, which significantly increases the storage space required and the time to perform individual operations using the data. The wide availability of fast processors and large storage capacities in even simple desktop PCs, makes this argument seem insignificant in the meantime. Another disadvantage is the merging with further information. Unlike the datasets prepared in wide format, an additional key variable is now required in order to be able to identify an observation clearly. This may be a wave identifier in the household or individual datasets, or alternatively the spell number in the spell datasets, which are also available in long format. Furthermore, it is not visible at first sight which variables were surveyed for which waves, as all of the variables ever surveyed are present in the dataset. These variables were marked accordingly in the variable labels and are given a special code (-9) for waves in which they were not surveyed.

When the advantages and disadvantages of long format for the user are weighed up, the advantages clearly outweigh the disadvantages in our opinion. Accordingly the household and individual datasets of PASS (*HHENDDAT*; *PENDDAT*) and the corresponding weighting data (*hweights*; *pweights*) were prepared in long format. At the household level the scientific use file contains the data on the household's receipt of Unemployment Benefit II processed in spell form (*alg2_spells*). At the individual level there are four spell datasets. These are (1) data about employment spells (*et_spells*), (2) periods of unemployment (*al_spells*) and (3) periods of economic inactivity (*lu_spells*), since January 2005 in each case, and (4) spell data on participation in employment and training measures (*mn_spells*) since January 2006. The dataset on participation in measures (*massnahmespells*) of the 1st wave and the spell dataset on receipt of Unemployment Benefit I (*alg1_spells*), which was also surveyed in the 1st wave, are not continued in the 2nd wave (see Chapter 4.5 on this subject).[26] The household and the person registers (*hh_register; p_register*) are available in wide format.

Figure 2: Dataset structure of PASS in wave 2



[26] A new spell dataset was created for participation in employment and training measures (mn_spells). Periods during which the respondent drew Unemployment Benefit I are no longer surveyed directly from the 2nd wave onwards but are surveyed in the context of the spells during which the respondent was registered as unemployed. The information about the periods of Unemployment Benefit I receipt is thus contained in the unemployment spells (al_spells).

6. Variable types and their names

General issues

For naming the variables of the data set we considered two main alternatives from which we had to choose one. The first option is naming the variables in accordance with their respective order in the questionnaire, as is done in the GSOEP, for example. The advantage of this type of naming convention is that the items corresponding to the variables are easy to find in the questionnaire, which significantly enhances the value of the questionnaire as a documentation instrument. The central disadvantage of this approach is that identical items are given different names due to changes in the order of questions in the questionnaire, resulting in considerable preparation being required for compiling and, if necessary, renaming the required variables even for simple trend analyses, as more and more panel waves become available.

The second main alternative is allocating independent variable names, which are kept constant across waves (apart from a wave indicator if necessary). The advantages and disadvantages of this strategy are the opposite to those of the first alternative: identifying the variables corresponding to an item across waves is unproblematic, whereas using the questionnaire as a documentation instrument becomes more difficult, as it is no longer possible to derive the position of an item in the questionnaire from the variable name.

In our opinion, the advantages of fixed variable names clearly outweigh the disadvantages in a long-term panel study. Moreover, the decision in favour of organising the data in long format as described above requires the use of uniform variable names.

Variable types

The codebook distinguishes between three different types of variables:

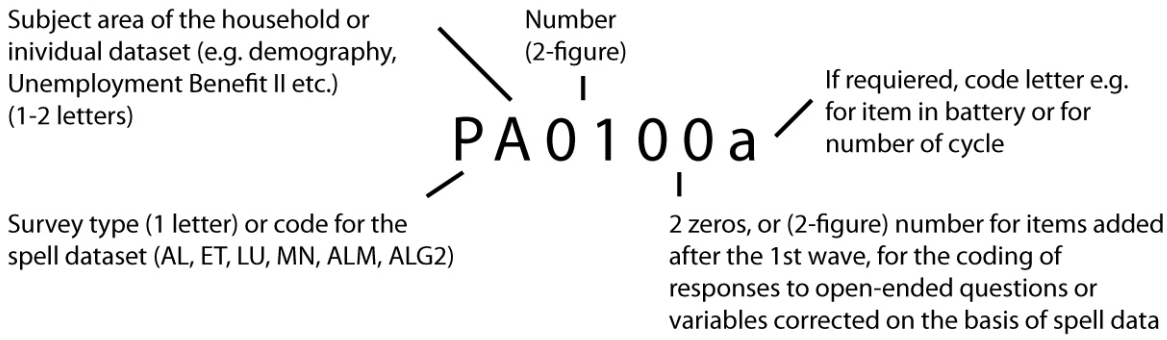
System variables:

1. System variables: system variables are variables created in the course of the survey process. They can be used, firstly, to comprehend the filters documented in the questionnaire. At least some of the system variables can also be of interest from a content-related or methodological point of view, for example the interview mode or

the number of children of a certain age group living in the household. System variables are allocated individual names, for which lower-case letters and numbers can be combined. The system variables also include the weights.

2. Surveyed variables: surveyed variables are variables that were collected in this form directly in the questionnaire. These variables are given entirely new, abstract variable names. The concept behind this naming process is illustrated in Figure 3 using an example.

Figure 3: The variable naming scheme



- The 1st letter of the variable name indicates the questionnaire level, i.e. household or individual dataset, by means of the letter H or P (upper-case) respectively.

- This is followed by one or two upper-case letters which indicate the subject area to which the variable belongs (see Appendix I for a complete list).

- In the datasets which are processed in spell form, there is no introductory P or H. Instead, the variables in these datasets are given a uniform subject-based name consisting of two or three letters or two letters and one number.

- The introductory letter combination is then followed by two consecutively allocated numbers, which indicate the number of the question within the subject area.

- These two numbers are followed by two zeros, which are intended to permit the addition of further variables in later waves. So far this option has only been used in cases where a second variant including coded information from an

open-ended survey question or response category has been made available in addition to the original version of the variable. The final zero is changed to “1” for these variables (e.g. *PA0101a* instead of *PA1100a*).

Table 9: List of subject-related indicators used in the variable names

Individual level		Household level	
Code	**Subject area**	**Code**	**Subject area**
PA	General	HA	General
PAS	Job-search	HD	Demography
PB	Education	HEK	Income
PD	Demography	HKI	Child-care
PEO	Attitudes and orientations	HLS	Standard of living
PEK	Income	HW	Housing
PET	Employment		
PG	Health		
PLS	Standard of living		
PMI	Migration		
PP	Care		
PSH	Social origin		
PSK	Social relations		
PTK	Contact to social security institutions		
AL	Spells of registered unemployment and receipt of Unemployment Benefit I since January 2005 (spell data, individual level, from 2nd wave onwards)	AL2	Receipt of Unemployment Benefit II (spell data, household level)
ET	Employment with earnings of more than € 400 per month since January 2005 (spell data, individual-level data, from wave 2 onwards)		
LU	Other activities since January 2005 (spell data, individual-level data, from wave 2 onwards)		
MN	Employment and training measures (spell data, individual level, from wave 2 onwards)		
AL	Receipt of Unemployment Benefit I (spell data, individual level, wave 1 only)		
ALM	Employment and training measures (spell data, individual level, wave 1 only)		

- In the case of variables for items from multi-item batteries or in a looped sequence of questions, a further lower-case letter may be added to identify the item or the current cycle within the loop.

3. Generated variables: the group of generated variables is divided again into three sub-groups. The generated variables in the strict sense are aggregated from various other variables, e.g. from open-ended and categorical income measures, or they are even more complex constructs such as equivalized household income or classifications for education (such as ISCED or Casmin) or status (e.g. EGP, ESEC). Generated variables in this strict sense are allocated individual names that are as clear and memorable as possible, in lower-case letters. For an overview of the generated variables, see Chapter 7.

   The generated variables also cover variables which are harmonised across the waves. It is always necessary to harmonise a variable if the way in which it is surveyed changed across the different survey waves, e.g. by a category being omitted or added. Although such a harmonisation could also be performed later by the data user, for key variables it is already done during the editing process for the scientific use file. The harmonised variables are also given clear names (see Chapter 7.2).

   The third group of generated variables includes those in which information from open-ended survey questions or response categories was added to another (closed) variable. Although these variables are, strictly speaking, also generated variables and are classified as such in the frequency tables of the codebook, they are not given clear names. Instead their names are based on those of the original variable, but with a ‘1’ as the final number rather than a ‘0’.

7. Generated variables

7.1 Coding of 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[27] which differed from the original variable only insofar as the information from the open-ended responses was coded to the corresponding categories where possible. Moreover, in some cases new categories were created on the basis of the information from open-ended questions. The naming of these additional variables differs from that of the original variable in the last digit only, where the '0' was replaced by a '1'. The items on country of birth, nationality, and the parents'/grandparents' country of residence before migration were also anonymised and given eloquent variable names[28]. Tables 10 and 11 give an overview of the open-ended survey questions which were coded in the 2nd wave[29].

[27] Other information from open-ended questions was not coded, for example the name of the institution providing basic social security (P138).
[28] *ogebland* (country of birth); *ostaatan* (nationality); *ozulanda* to *ozulandf* (parents'/grandparents' country of residence before migration)
[29] Variables for which information was gathered via open-ended questions and coded in the 1st wave but not in the 2nd wave are not listed (with the exception of the spell dataset for Unemployment Benefit II). For the observations of the 2nd wave these variables are allocated the code -9 (item not surveyed in wave) and are documented in the Datenreport of the 1st wave.

Table 10: Coding of responses to open-ended survey questions at the household level in wave 2

Questionnaire number Re-interviewed HH	New sample HH	Coded to variable	Dataset	Description
HH64	n. in Q. vers.	*HW0881a-j*	*HHENDDAT*	Other reason for moving out, not listed
HH85	HH42	*HD0601*	*HHENDDAT*	Language spoken in h'hold: other lang., not listed
HH87	HH44	*HD0801*	*HHENDDAT*	Language spoken in h'hold after follow-up question about other languages: other lang., not listed
HH88	HH45	*HD0901*	*HHENDDAT*	Language spoken in h'hold, equal use of two languages: first language is another lang. which is not listed
HH89	HH46	*HD1001*	*HHENDDAT*	Language spoken in h'hold, equal use of two languages: second language is another lang. which is not listed
HH99	HH56	*AL21301a-e* *AL21401a-e* *AL21501a-e* *AL21601a-e* *AL21701a-e* *AL21801a-e* *AL21851a-e* *AL22001a-e* *AL22101a-e* *AL22102a-e* *AL22103a-e*	*alg2_spells*	Other reason for benefit cut, not listed

Table 11: Coding of responses to open-ended survey questions at the individual level in wave 2

Questionnaire number Individuals	Senior cit's	Coded to variable	Dataset	Description
P8_6	n. in Q. vers	*PB0231*	*PENDDAT*	Other German school qualification, not listed (update)
P8_7	n. in Q. vers	*PB0231*	*PENDDAT*	Other foreign school qualification, not listed (update)
P10_9	P5_9	*PB0401*	*PENDDAT*	Other German school qualification not listed (first survey or contradiction to previous wave)
P10_10	P5_10	*PB0401*	*PENDDAT*	Other foreign school qualification not listed (first survey or contradiction to previous wave)
P11	n. in Q. vers	*PB1001*	*PENNDAT*	Other foreign school qualification, not listed (first survey or contradiction to previous wave)
P21_9	n. in Q. vers	*PB1181a-j*	*PENDDAT*	Other German vocational qualification, not listed (contradiction to previous wave + further qualification)

Questionnaire number Individuals	Senior cit's	Coded to variable	Dataset	Description
P21_10	n. in Q. vers	*PB1181a-j*	*PENDDAT*	Other foreign vocational qualification, not listed (contradiction to previous wave + further qualification)
P26_9	P7_9	*PB1301a-j*	*PENDDAT*	Other German vocational qualification, not listed (update or first survey)
P26_10	P7_10	*PB1301a-j*	*PENDDAT*	Other foreign vocational qualification, not listed (update or first survey)
P28	n. in Q. vers	*PB1601*	*PENDDAT*	Other qualification, not listed, to which the foreign qualification corresponds
P111	n. in Q. vers	*AL0601*	*al_spells*	Other reason, not listed, for no longer being registered as unemployed
P129	n. in Q. vers	*LU0101*	*lu_spells*	Other gap status, not listed
P143	n. in Q. vers	*PTK0321a-g*	*PENDDAT*	Other reason, not listed, for not having to seek employment
P162	n. in Q. vers	*MN0201a-h MN0202h*	*mn_spells*	Other component of measure, not listed
P167	n. in Q. vers	*MN1001a-e*	*mn_spells*	Other reason, not listed, for why the measure was ended prematurely
P184	n. in Q. vers	*PAS0901a-i*	*PENDDAT*	Other places, not listed, where target pers. obtained information about job vacancies
P219	P51	*PG0901a-g*	*PENDDAT*	Other health problems, not listed
P223	P54	*PG1301*	*PENDDAT*	Other health insurance, not listed
P264	P73	*ogebland*	*PENDDAT*	Other country of birth, not listed
P267	P76	*ostaatan*	*PENDDAT*	Other nationality, not listed
P274	P80	*ozulanda-f*	*PENDDAT*	Other country, not listed, from which parent/grandparent migrated
P275	P81	*PMI1111*	*PENDDAT*	Language spoken in circle of friends: other language, not listed
P276	P82	*PMI1121*	*PENDDAT*	Language spoken in circle of friends, equal use of two languages: first language is another language, not listed
P277	P83	*PMI1131*	*PENDDAT*	Language spoken in circle of friends, equal use of two languages: second language is another language, not listed
P278_9	n. in Q. vers	*PSH0201*	*PENDDAT*	Other German school qualification of mother, not listed
P278_10	n. in Q. vers	*PSH0201*	*PENDDAT*	Other foreign school qualification of mother, not listed
P279_7	n. in Q. vers	*PSH0301a-i*	*PENDDAT*	Other German vocational qualification of mother, not listed
P279_8	n. in Q. vers	*PSH0301a-i*	*PENDDAT*	Other foreign vocational qualification of mother, not listed
P278_9	n. in Q. vers	*PSH0501*	*PENDDAT*	Other German school qualification of father, not listed
P278_10	n. in Q. vers	*PSH0501*	*PENDDAT*	Other foreign school qualification of father, not listed
P279_7	n. in Q. vers	*PSH0601a-i*	*PENDDAT*	Other German vocational qualification of father, not listed
P279_8	n. in Q. vers	*PSH0601a-i*	*PENDDAT*	Other foreign vocational qualification of father, not listed

7.2 Harmonisation

For some variables there were changes in the survey instruments across the waves. This particularly concerns the employment status variable, for which the instrument was changed entirely between waves 1 and 2. In order to simplify cross-wave analyses in such cases, for important indicators variables are generated which are harmonised across the waves. This currently concerns only two variables. However, this number can be expected to increase with the duration of the panel.

Individual level:

PENDDAT	*erwerb2; sibkz*

7.3 Dependent interviewing

In various places in both the household interviews and the personal interviews information was gathered depending on responses given in the previous wave. Some variables were updated based on information from the previous wave. However, the information from the previous wave generally has to be combined with the changes observed since that time (update) in order to conduct analyses. Tables 12, 13 and 14 below provide a brief overview of all of the relevant places in the questionnaires and show in which variables the information from the previous wave and the update from wave 2 were combined. The cases where generated variables were updated or continued are additionally listed in Chapter 7.4 of this Datenreport.

Table 12: Updated information from the previous wave in wave 2, re-interviewed households

Household questionnaire for re-interviewed households (HHalt)			
Construct	**Q. No's.**	**Note**	**Updated in variable**
Household structure	HH1-HH60	Household size updated by TNS Infratest during the interview	*HA0100*
		Gender of the individuals in the household corrected if necessary by TNS Infratest during the interview	*HD0100a* to *HD0100o*
		Age of the individuals in the household updated by TNS Infratest during the interview	*HD0200a* to *HD0200o*
		Family relationships updated by TNS Infratest during the interview	not provided in the SUF
Year of move into current dwelling	HH66	Updated in generated variable	*einzugj*
Size of dwelling in sqm	HH65/HH69	Updated in generated variable	*wohnfl*
Receipt of Unemployment Benefit II	HH91-HH104	Updated in spell dataset for Unemployment Benefit II	Variables of the spell dataset for Unemployment Benefit II

Table 13: Updated information from the previous wave in wave 2, new sample households

Household questionnaire for new sample households (HHneu)			
Construct	**Q. No's.**	**Note**	**Update in variable**
Receipt of Unemployment Benefit II	HH48-HH60	Updated in spell dataset for Unemployment Benefit II	Variables of the spell dataset for Unemployment Benefit II

Table 14: Updated information from the previous wave in wave 2, personal questionnaire

Personal questionnaire			
Construct	**Q. No's.**	**Note**	**Update in variable**
Highest general school qualification	P7-P12	Updated in generated variable	*schul1* (without responses to open-ended questions) *schul2* (with responses to open-ended questions)
Year in which highest general school qual. gained	P6, P13	Updated in generated variable	*schulabj*
Vocational qualification	P20, P21, P23-P29	Highest vocational qualification updated in generated variable	*beruf1* (without responses to open-ended questions) *beruf2* (with responses to open-ended questions)
Year in which vocational qual. gained	P19, P22, P30	Updated in generated variable	*berabj*

7.4 Simple generated variables

This type of variable covers, for example, variables for which different items of one construct that were surveyed separately for technical reasons were aggregated or for which information from the current wave was combined with information from the previous wave (see Chapter 7.3) (such as the highest educational qualification) or for which important information was merged from other partial datasets (e.g. indicators for current receipt of Unemployment Benefit I or Unemployment Benefit II). The corresponding variables are shown in the dataset-specific Tables 15 to 22.

Table 15: Simple generated variables for wave 2 in the household dataset (*HHENDDAT*) (in alphabetical order)

Variable	Variable label and description	Source variables for generated variable in wave 2
alg2abez	*Current receipt of Unemployment Benefit II of the H'hold, 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 (HHalt: HH104; HHneu: HH61); *int_jahr*
bik	*BIK region size classes (GKBIK10), generated*	Supplied by survey institute
blneualt	*Western German States or Eastern German States, generated* Aggregation of German states into the Western German States of the former FRG (without Berlin) and the Eastern German States of the former GDR (with Berlin).	Information generated and supplied by the survey institute on the Federal State in which the household is resident as of survey date
einzugj	*Year of move into current dwelling, generated* Information as to the year in which the household moved into the current dwelling. In the case of re-interviewed households the year of the move into the current dwelling was only asked in the 2nd wave if the household had been living in a hostel or if it had moved house since the previous wave.	Information from wave 1: *HW0900* Information from wave 2: *HW0900, HW0200; umzug* (*HHENDAT*)
hhinckat	*Categorised household income per month (in euros), generated* Categorised information on the household's income aggregated from several survey items into one variable.	*HEK0700; HEK0800; HEK0900; HEK1000; HEK1100* (*HHENDDAT*)
hhincome	*Household income per month (in euros) incl. categorised information, generated* Generation of a variable integrating information from categorised and open-ended survey questions on the net household income.	*HEK0600; HEK0700; HEK0800; HEK0900; HEK1000; HEK1100* (*HHENDDAT*)
hintdat	*Date of household interview* Generated variable with the date on which the household interview was conducted in the form YMMDD.	*hintjahr, hintmon, hinttag* (*HHENDDAT*)

Variable	Variable label and description	Source variables for generated variable in wave 2
kindu4	*Control variable: child under age of 4 in the h'hold* The variable indicates that at least one person in the household is below 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 person aged four is also actually the child of another person living in the household.	*HD0200a - HD0200o* (*HHENDDAT*)
kindu13	*Control variable: child under age of 13 in the h'hold* The variable indicates that at least one person in the household is below 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 person aged 13 is also actually the child of another person living in the household.	*HD0200a - HD0200o* (*HHENDDAT*)
kindu15	*Control variable: child under age of 15 in the h'hold* The variable indicates that at least one person in the household is below 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 person aged 13 is also actually the child of another person 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.	*HD0200a - HD0200o*; categorical follow-up question about age group (in cases of no response in *HD0200*) (*HHENDDAT*)
wohnfl	*Living space in sqm, generated* 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 in the 2nd wave if the household had moved house or if the house/apartment had changed since the previous wave.	Information from wave 1: *HW1000* Information from wave 2: *HW100; HW0910; HW0920* (*HHENDDAT*)

Table 16: Simple generated variables for wave 2 in the individual dataset (*PENDDAT*) (in alphabetical order)

Variable	Variable label and description	Source variables for generated variable in wave 2
aktmassn	*Current participation in a measure funded/promoted by the employment agency, generated* Indicator: respondent is currently participating in a measure of active labour market policy as of interview date.	*MN0500; censored* (*mn_spells*) *PA0711b-f; PA0721a-f* (*PENDDAT*) Information from follow-up validation question P 178_X_Prüf (personal questionnaire)
alg1abez	*Current receipt of Unemp. Benefit I, generated* Indicator: respondent is in receipt of Unemployment Benefit I as of the interview date. In the 2nd wave the periods since January 2005 during which the respondent was registered as unemployed were surveyed. For each spell additional questions were asked as to whether the respondent received UB I and if so, during which period. This information was combined with a follow-up question for respondents who were aged 58 or over and were therefore entitled to Unemp. Benefit I without being registered as unemployed.	*AL0700; AL1000; AL1100; AL1200; alg1bj; alg1ej* (*al_spells*); *PA0405* (*PENDDAT*) ; information as to whether there is a further spell of unemployment (P117)
alg1s05	*Indicator: Receipt of Unemployment Benefit I since Jan. 2005? Generated (all waves)* Indicator: Respondent has received Unemployment Benefit I at some time since January 2005. In the 2nd wave the periods since January 2005 during which the respondent was registered as unemployed were surveyed. For each spell additional questions were asked as to whether the respondent received UB I and if so, during which period. This information was combined with a follow-up question for respondents who were aged 58 or over and were therefore entitled to Unemployment Benefit I without being registered as unemployed.	*AL0700; AL1200; alg1bj; alg1ej* (*al_spells*); *PET0911; PA0405* (*PENDDAT*); information as to whether there is a further spell of unemployment (P117)
apartner	*Control variable: cohabitee in the h'hold* Indicator: respondent has a cohabitee or a partner whose status is not specified in the household.	Information on relationships between household members (household grid); *PD0500 - PD0900* (*PENDDAT*)
arbzeit	*Weekly hours of work incl. details in the case of irregular working hours, generated.* Weekly hours of work in the job held by the respondent on the interview date, generated from responses to open-ended questions on working hours and categorical follow-up question in the case of irregular working hours. Integration of the variable *arbzeit,* already generated in wave 1, with wave 2 variables *ET2100* and *ET2200*.	*ET2100; ET2200* (*et_spells*); *PET0500; PET0700* (*PENDDAT*)
befrist	*Current job: fixed-term contract? Gen. (all waves)* Indicator: The job held by the respondent on interview date is on a fixed-term contract.	*PET2510a; PET2510b* (*PENDDAT*)

Variable	Variable label and description	Source variables for generated variable in wave 2
begjeewt	*Year when first job taken up, generated* Year in which the respondent first worked in a regular job. To generate the variable, information about the first regular job was combined with information from the employment spells if the respondent had already reported his/her first regular job during the questions on employment spells since January 2005.	*bjahr* (*et_spells*); *PET0150; PET0151; PET3200b* (*PENDDAT*)
begmeewt	*Month in which first job taken up, generated* Month in which the respondent first worked in a regular job (generation: see *begjeewt*).	*bmonat* (*et_spells*); *PET0150; PET0151; PET3200a* (*PENDDAT*)
berabj	*Year of highest vocational qualification* Year in which the respondent gained his/her highest vocational qualification as of the interview date. Note: The years in which the vocational qualifications reported in the 1st wave were gained were surveyed in the 2nd wave.	Information from wave 1: *PB1300a-j; PB1400* (*PENDDAT*) Information from wave 2: *PB1160a-j; PB1180a-j; PB1190a-j; PB1300a-j; PB1310am-jm; PB1310km; PB1310aj-jj; PB1310kj* (*PENDDAT*)
beruf1	*Highest vocational qualification, excl. foreign qualifications and information from open-ended questions, generated* Identification of the highest vocational qualification as of the interview date by hierarchising the vocational qualifications cited by the respondents, excl. information from open-ended questions.	Information from wave 1: *PB1200* Information from wave 2: *PB0100; PB0200; PB0300; PB1200a; PB1200b; PB1200c* (*PENDDAT*)
beruf2	*Highest vocational qualification, incl. foreign qualifications and information from open-ended questions, generated* As for *beruf1* with the following differences: 1. Inclusion of information from open-ended questions; 2. Inclusion of information on foreign qualifications; 3. Degrees are not distinguished by type of institution (e.g. university or other institution of higher education) but by the qualification level (Bachelor's degree; Master's degree; Ph.D.).	Information from wave 1: *PB1301a-j; PB1400; PB1500a; Pb1500b; PB1500c* (*PENDDAT*) Information from wave 2: *PB1160a-j;PB1180a-j; PB1300a-j; PB1601;PB0200; PB1500a; Pb1500b; PB1500c* (*PENDDAT*)
brutto	*Gross income incl. categorised information, generated* Generation of a variable integrating information from categorised and open-ended survey questions on gross income. Note: The concept for generating this variable was modified in some points in order to improve the imputation of the categorical follow-up question on gross income. The variable was also redeveloped for the 1st wave.	*PEK0100b; PEK0200; PEK0300; PEK0400; PEK0500; PEK0600;* (*PENDDAT*)
bruttokat	*Categorised gross income, generated* Aggregation of the categorised information on gross income, combined from several items on income categories.	*PEK0200; PEK0300; PEK0400; PEK0500; PEK0600* (*PENDDAT*)

Variable	Variable label and description	Source variables for generated variable in wave 2
ejhrlewt	*Time when last job ended (year)* Year in which the respondent was last in employment. To generate the variable, information from the employment spells was combined with the information on the last job if the respondent had not been in employment since January 2005.	*PET1200b* (*PENDDAT*); *ejahr; emonat2* (*et_spells*)
ekin1517	*Control variable: own child aged between 15 and 17 in the household* This variable indicates that 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.	Information on relationships between household members (household grid)
ekind	*Control variable: own child in the household* This variables indicates that the respondent has a natural child, a stepchild/adopted child or a child of non-specified status of any age in the household..	Information on relationships between household members (household grid)
ekin614	*Control variable: own child aged between 6 and 14 in the household* This variable indicates that 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.	Information on relationships between household members (household grid)
ekinu15	*Control variable: own child under age of 15 in household* This variable indicates that 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.	Information on relationships between household members (household grid)
ekinu18	*Control variable: own child under age of 18 in household* This variable indicates that 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.	Information on relationships between household members (household grid)
emonlewt	*Time when last job ended (month)* Month in which the respondent was last in employment. (Generation: see *ejhrlewt*).	*PET1200a* (*PENDDAT*); *emonat2* (*et_spells*)
epartner	*Control variable: spouse or registered partner in the household* This variable indicates that the respondent has a spouse or a same-sex registered partner in the household.	Information on relationships between household members (household grid)

Variable	Variable label and description	Source variables for generated variable in wave 2
erwerb2	*Employment status, generated (all waves)* Integrated employment status variable, harmonised for the 1st wave. The variable *erwerb* created in the 1st wave could not be continued due to the change-over to employment biographies in the 2nd wave. A new status variable was therefore created which, for the 1st wave, is based on the previous employment status variable *erwerb* and, for the 2nd wave, was generated on the basis of the economic inactivity status (including responses to open-ended questions), the status of school pupil/student/trainee, the current working hours and the spell-related information on currently held jobs. The basis for generating the variable is the information from the relevant spell dataset of wave 2 as to whether a certain type of spell is currently ongoing. In the case of a currently ongoing spell of economic inactivity from the gap dataset, the type of inactivity is identified via the variable *LU0101* (i.e. incl. information from open-ended survey questions). In the context of harmonisation, categories 2 ("unemployed") and 3 ("job-creation measure, one-Euro-job") of the 1st wave are combined to a joint category 2 ("unemployed"). The previous categories 8 ("apprenticeship/ training/further training/ retraining) and 10 ("student") were also merged into one category for the purpose of integration with wave 2.	For wave 1: *erwerb* (*PENDDAT*) For wave 2: *nichtew2; PB0100; arbzei* (*PENDDAT*) *ET0600* (*et_spells*)
famstand	*Marital status, generated* Generation of a marital status variable integrating information from the personal questionnaire and the control variable *epartner* generated from the household dataset.	*epartner; PD0500; PD0700* (*PENDDAT*)
gebhalbj	*Half-year of birth, generated* This variable indicates whether the date of birth is in the first or second half of the year of birth.	Information on month of birth
hhalg2	*Control variable: current receipt of UB II* This variable indicates that the household is receiving Unemployment Benefit II at time of h'hold interview.	*HA250b* (*HHENDDAT*) *AL20400; AL20500* (*alg2_spells*)
halg2s06	*Control variable: receipt of UB II since 2006* Indicator that shows whether the household in which the respondent lives received Unemployment Benefit II at least once between January 2006 and the interview date in the 2nd wave.	Information on receipt of Unemployment Benefit II since January 2006 generated on the basis of the information from the previous wave. *HA0250b; HA0300* (*HHENDDAT*)
kindzges	*Total number of own children (living in and outside the household), generated* Total number of respondent's children including the children living in his/her household and the children living outside the household.	Information on relationships between household members (household grid); *PD0900; PD1000; PD1100* (*PENDDAT*)

Variable	Variable label and description	Source variables for generated variable in wave 2
kindzihh	*Number of own children in the household, generated* Variable generated on the basis of the responses in the household questionnaire concerning the number of children that a person in the household has (total number of persons in the household (half) matrix who count as children of the respondent plus the number of persons in the household (half) matrix for whom the respondent is classified as being a parent). 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 household, excl. information from open-ended survey questions, generated* In the first wave the question on the mother's vocational qualification was only asked if the mother was not living in the survey household. If she was living in the household the information on her vocational qualification was taken from her personal interview. In the 2nd wave the question on the mother's vocational qualification was asked of all newly interviewed individuals, irrespective of whether the mother was living in the household or not. For people taking part in a repeat interview in the 2nd wave the values were taken over from the generated variable *mberuf1* from the 1st wave.	Information from wave 1: *mberuf1* (*PENDDAT*) Information from wave 2: *PSH0300a-i* (*PENDDAT*)
mberuf2	*Highest vocational qualification attained by the mother, incl. mother in the household, incl. information from open-ended survey questions, generated* Like *mberuf1* apart from the fact that responses to open-ended questions were also taken into account for the generation of *mberuf2*.	Information from wave 1: *mberuf2* Information from wave 2: *PSH0301a-i* (*PENDDAT*)
mhh	*Control variable: mother living in household* Variable indicating that the respondent's natural mother, stepmother, adoptive mother or mother of non-specified status is living in the household.	Information on relationships between household members (household grid)

Variable	Variable label and description	Source variables for generated variable in wave 2
migration	*Respondent's migration background, generated* Generated variable for four categories of migration background: no migration background; personal migration (1st generation); migration of at least one parent but no personal migration of the respondent (2nd generation); migration of at least one grandparent but no personal migration of respondent or of either parent (3rd generation). Note: The concept for generating this variable was revised. To generate the variable in the 1st wave, only the information on whether the respondent was born in Germany and on which generation/members of the family moved to Germany was used; now the information on whether a parent/grandparent was born outside Germany and, if applicable, which parent/grandparent, is also used. In order to guarantee a consistent logic across the waves, the variable for the 1st wave was also re-generated.	*PMI0100; PMI0700; PMI0800a-f; PMI0900a-f* (*PENDDAT*)
mschul1	*Highest general school qualification attained by the mother, incl. mother in household, excl. information from open-ended questions, generated* In the first wave the question on the mother's highest school qualification was only asked if the mother was not living in the survey household. If she was living in the household the information on her highest school qualification was taken from her personal interview. In the 2nd wave the question on the mother's highest school qualification was asked of all newly interviewed individuals, irrespective of whether their mother was living in the survey household or not. For people taking part in a repeat interview in the second wave the values were taken over from the generated variable *mschul1* from the 1st wave.	Information from wave 1: mschul1 Information from wave 2: *PSH0200* (*PENDDAT*)
mschul2	*Highest general school qualification attained by the mother, incl. mother in household, incl. information from open-ended questions, generated* Like *mschul1* apart from the fact that responses to open-ended survey questions were also taken into account for the generation of *mschul2*.	Information from wave 1: mschul2 (*PENDDAT*) Information from wave 2: *PSH0201* (*PENDDAT*)
mstib	*Mother's occupational status, code number, generated* Detailed occupational status of mother, generated from the individual variables.	*PSH0320; PSH0330; PSH0340; PSH0360; PSH0370; PSH0380* (*PENDDAT*)

Variable	Variable label and description	Source variables for generated variable in wave 2
netto	*Net income incl. categorised information, generated* Generation of a variable integrating information from categorised and open-ended survey questions on the respondent's net income Note: The concept for generating this variable was modified in some points in order to improve the imputation of the categorical follow-up question on net income. The variable was also redeveloped for the 1st wave.	*PEK0700b; PEK0800; PEK0900; PEK1000; PEK1100; PEK1200* (*PENDDAT*)
nettokat	*Categorised net income, generated* Aggregation of the categorised information on the respondent's net income, combined from several items on income categories.	*PEK0800; PEK0900; PEK1000; PEK1100; PEK1200* (*PENDDAT*)
nichterw	*Status: economic inactivity, generated (all waves)* Integrated variable for the respondent's status of economic inactivity. Generated from the variable *PET0800* in wave 1. In wave 2 *nichterw* is based on information regarding the type of the current economic inactivity from the gap module (*LU0100*, i.e. not taking into account the responses to open-ended survey questions) and information from the unemployment module regarding ongoing unemployment.	*LU0100* (*lu_spells*) *zensiert* (al_*spells*) *PET0151; PET0911* (*PENDDAT*) Indicator for cases where mistakenly no gap status was surveyed
nichtew2	*Status: economic inactivity, generated, incl. information from open-ended survey questions (all waves)* Integrated variable for the respondent's status of economic inactivity The responses to open-ended questions were also taken into account when generating *nichtew2*. Generated from the variable *PET0801* in wave 1. In wave 2 *nichtew2* is based on information regarding the type of the current economic inactivity from the gap module (*LU0101*, i.e. including the responses to open-ended questions) and information from the unemployment module regarding ongoing unemployment.	*LU0101* (*lu_spells*) *zensiert*(al_*spells*) *PET0151; PET0911 (PENDDAT)* Indicator for cases for which mistakenly no gap status was surveyed
palter	*Age (from p1), generated* Respondent's age, generated on the basis of the date of birth and the date of the personal interview in the current wave.	*p1; pintjahr, pintmon, pinttag* (*PENDDAT*)
panel	*Willingness to participate in panel*	Information supplied by the survey institute regarding the households' willingness to participate in the panel
pintdat	*Date of personal interview* Generated variable with the date on which the personal interview was conducted in the form YMMDD.	*pintjahr, pintmon, pinttag* (*PENDDAT*)

Variable	Variable label and description	Source variables for generated variable in wave 2
schul1	*Highest general school qualification, excl. foreign qualifications and information from open-ended survey questions* Variable for the highest general school qualification; equivalent eastern and western German qualifications were combined (e.g. EOS and Abitur); excl. information from open-ended questions.	Information from wave 1: *schul1* (*PENDDAT*) Information from wave 2: *PB0200; PB0210; PB0220; PB0230; PB0300; PB0400* (*PENDDAT*)
schul2	*Highest general school qualification, incl. foreign qualifications and information from open-ended survey questions* As for *schul1* with the following differences: 1. Inclusion of responses to open-ended questions; 2. Inclusion of information on foreign qualifications	Information from wave 1: *schul2* (*PENDDAT*) Information from wave 2: *PB0200; PB0210; PB0220; PB0231; PB0300; PB0401* (*PENDDAT*) + correction of individual cases from wave 1 (*PB0300; PB0400*)
schulabj	*Year of highest general school qualification* Year in which respondent gained his/her highest general school qualification. Note: The year in which the school qualification reported in the 1st wave was gained was surveyed in the 2nd wave. If no information was available from the 1st wave on the highest school qualification or if the individual was participating in the survey for the first time in wave 2, then the year in which the highest qual. was gained was also asked for directly. Re-interviewed respondents for whom information on the highest school qual. was already available from the 1st wave were not asked in the 2nd wave about the year when this qualification was gained if they had gained a new qualification since the previous wave. In this case the year in which the qualification was gained was estimated depending on the month and year of the interview. If the 2nd wave interview was conducted before May 2008 it was assumed that the qualification was gained in 2007, if the interview was conducted later than May, the qualification was assumed to have been gained in 2008.	Information from wave 1:*PB0400* (*PENDDAT*) Information from wave 2: *PB0210; PB0220; PB0230; PB0400; PB0410;pintjahr; pintmon* (*PENDDAT*)
stib	*Occupational status, code number, generated* Generation of the detailed code number for occupational status on the basis of the individual variables. Integration of the variable *stib,* which was already generated in wave 1 for the current job, with wave-2 information from the employment module (*ET0600-ET1200*). If there was more than one ongoing employment spell in wave 2, the one with the most hours of work was selected. If there was more than one ongoing spell with exactly the same number of hours, the one that began first was selected.	*ET0600 ET0700; ET0800; ET0900; ET1000; ET1100; ET1200; ET500b* (*et_spells*)

Variable	Variable label and description	Source variables for generated variable in wave 2
stibeewt	*Occupational status, first job, code number, generated* Detailed code number of the occupational status in the respondent's first regular job. To generate this variable, information on the first job was combined with information from the employment spells if the respondent reported having already given information about his/her first job during the questions about the employment spells since January 2005.	*PET3300b; PET3000; PET3400; PET3500; PET3600; PET3700; PET3800; PET3900 (PENDDAT) ET0600; ET0700; ET0800; ET0900; ET1000; ET1100; ET1200* (*et_spells*)
stibkz	*Current occupational status, simple classification, harmonised (anonymised)* Generation of the simple code number of occupational status on the basis of various items gathering information on occupational status.	Information from wave 1: *PET1500* (*PENDDAT*) Information from wave 2: *PET1510* (*PENDDAT*)
stiblewt	*Occupational status, last job, code number, generated* Detailed code number of the occupational status in the respondent's last job. To generate this variable, information from the employment spells was combined with information on the last job if the respondent had been out of work since Jan. 2005.	*PET1210b; PET1210; PET1220; PET1230; PET1240; PET1250; PET1260; PET1270 (PENDDAT) ET0600; ET0700; ET0800; ET0900; ET1000; ET1100; ET1200* (*et_spells*)
vberuf1	*Highest vocational qualification attained by the father, incl. father in the household, incl. information from open-ended survey questions, generated* Generation of variable for father's highest vocational qualification analogous to *mberuf1*.	Information from wave 1:*vberuf1* (*PENDDAT*) Information from wave 2: *PSH0600a-i* (*PENDDAT*)
vberuf2	*Highest vocational qualification attained by the father, incl. father in the household, excl. information from open-ended survey questions, generated* Generation of variable for father's highest vocational qualification (incl. information from open-ended survey questions) analogous to *mberuf2*.	Information from wave 1:*vberuf2* (*PENDDAT*) Information from wave 2: *PSH0601a-i* (*PENDDAT*)
vhh	*Control variable: father living in household* Variable indicating that the respondent's natural father, stepfather, adoptive father or father of non-specified status is living in the household.	Information on relationships between household members (household grid)
vschul1	*Highest general school qualification attained by father, incl. father in household, excl. information from open-ended survey questions, generated* Generation of variable for father's highest general school qualification analogous to *mschul1*.	Information from wave 1:*vschul1* (*PENDDAT*) Information from wave 2: *PSH0500* (*PENDDAT*)
vschul2	*Highest general school qualification attained by the father, incl. father in household, incl. information from open-ended survey questions, generated* Generation of variable for father's highest general school qualification (incl. information from open-ended survey questions) analogous to *mschul2*.	Information from wave 1:*vschul2* (*PENDDAT*) Information from wave 2: *PSH0501* (*PENDDAT*)
vstib	*Father's occupational status, code number, generated* Detailed occupational status of father, generated from the individual variables.	*PSH0620; PSH0630; PSH0640; PSH0660; PSH0670; PSH0680* (*PENDDAT*)

Table 17: Simple generated variables for wave 2 in the spell dataset for Unemployment Benefit II (*alg2_spells*) (in the same order as in the dataset)

Variable	Variable label and description	Source variables for generated variable in wave 2
bmonat	*Spell of UB II: starting month, generated* Month in which the spell of Unemployment Benefit II began. To generate the variable, if information was only available on the season when a spell started, it was converted into a definite 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 from the 2nd wave onwards. Details regarding the season in which a 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 the year → December	*AL20100* (*alg2_spells*)
bjahr	*Spell of UB II: starting year, generated* Year in which the spell of Unemployment Benefit II began. Note: see *bmonat*	*AL20200* (*alg2_spells*)
emonat	*Spell of UB II: ending month, generated* Month in which the spell of Unemployment Benefit II ended. To generate the variable, if information was only available on the season when a spell started, it was converted into a definite month and for right-censored spells (i.e. spells that were still ongoing when the household was interviewed) the interview month was entered. Note: see *bmonat*	*AL20300* (*alg2_spells*) *hintmon* (*HHENDDAT*)
ejahr	*Spell of UB II: ending year, generated* Year in which the spell of Unemployment Benefit II ended. In the case of right-censored spells (i.e. spells that were still ongoing when the household was interviewed) the interview year was entered. Note: see *bmonat*	*AL20400* (*alg2_spells*) *hintjahr* (*HHENDDAT*)

Variable	Variable label and description	Source variables for generated variable in wave 2
alg2kbma - *alg2kbme*	*UB II: 1st benefit cut: starting month, generated* Month in which the reduction of Unemployment Benefit II began. To generate the variable, if information was only available on the season when a spell started this was converted into a definite month. Note: The UB II cuts are embedded in the spells of UB II receipt. The information on the individual benefit-cut spells can be distinguished via the indicator at the end of the respective variable (a - e). 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 from the 2nd wave onwards.	1st benefit cut: *AL21000a* (*alg2_spells*) to 5th benefit cut: *AL21000e* (*alg2_spells*)
alg2kbja - *alg2kbje*	*UB II: 1st benefit cut: starting year, generated* Year when Unemployment Benefit II cut began. Note: see *alg2kma* - *alg2kbme*	1st benefit cut: *AL21100a* (*alg2_spells*) to 5th benefit cut: *AL21100e* (*alg2_spells*)
alg2kema - *alg2keme*	*UB II: 1st benefit cut: ending month, generated* Month in which the Unemployment Benefit II cut began. To generate the variable, if information was only available on the season when a spell started, it was converted into a definite month. If the respondent reported a duration for the benefit cut, this was used to calculate the end date of the benefit cut on the basis of the generated start date. Note: see *alg2kma* - *alg2kbme*	1st benefit cut: *alg2kbma; alg2kbmja; AL21200a; AL21201a; AL21202a* (*alg2_spells*) to 5th benefit cut: *alg2kbme; alg2kbmje; AL21200e; AL21201e; AL21202e* (*alg2_spells*)
alg2keja - *alg2keje*	*UB II: 1st benefit cut: ending year, generated* Year in which the Unemployment Benefit II cut ended. If the respondent reported a duration for the benefit cut, this was used to calculate the end date of the benefit cut on the basis of the generated start date. Note: see *alg2kma* - *alg2kbme*	1st benefit cut: *alg2kbma; alg2kbmja; AL21200a; AL21201a; AL21202a* (*alg2_spells*) to 5th benefit cut: *alg2kbme; alg2kbmje; AL21200e; AL21201e; AL21202e* (*alg2_spells*)

Variable	Variable label and description	Source variables for generated variable in wave 2
AL22150a to AL22150e	*UB II: benefit cut: which household member's benefit was reduced, generated* This variable contains coded information about which h'hold members' Unemployment Benefit II was cut. It is a string variable with 15 positions. Starting from the left, each position of this variable stands for the position of one person in the household grid. The first position of the variable, for example, indicates whether the benefit was cut for the first person in the h'hold in the particular benefit cut spell, the second position shows whether the second person's benefit was cut and so on. As the source information for the generation was only collected from the 2nd wave onwards, all 15 positions of the question are given the code "I" (item not surveyed in wave) for all benefit cuts reported in the first wave (see below). Each of the 15 positions of the variable which stands for one of the maximum of 15 individuals in the household structure is given one of the following codes indicating that person's benefit-cut 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 – Refused T – Not applicable (filter) F – Question mistakenly not asked U – Implausible value I – Item not administered in wave	Information about which household member's benefit was reduced in the particular benefit cut spell (HH102 in the household questionnaire for re-interviewed households; HH53 in the household questionnaire for split-off households and new sample households).
zensiert	*UB II spell: spell ongoing at time of last h'hold interview (right-censored.), generated* The censoring indicator shows whether a spell was still ongoing at the time of the last h'hold 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 in wave 1 which was not re-interviewed in wave 2. (b) A household surveyed in wave 2 reports in H91/H93 (HHalt) / H48/H50 (HHneu) that a spell of UB II is also still ongoing on the interview date of wave 2. Or in H91/H93 (HHalt) / H48/H50 (HHneu) an end date is reported which is identical to the interview date of wave 2 and it is confirmed in the follow-up question in H94 (HHalt) / HH51 (HHneu) 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 h'hold in wave 2 and was not interviewed in wave 2.	*AL20100; AL20500* (*alg2_spells*)

Table 18: Simple generated variables for wave 2 in the employment spell dataset (*et_spells*) (in the same order as in the dataset)

Variable	Variable label and description	Source variables for generated variable in wave 2
bmonat	*Employment: starting month, generated* Month in which the employment spell began. To generate the variable, if information was only available on the season when a spell started, it was converted into a definite month. Note: The generated date variables were checked for plausibility and corrected if necessary. The dates originally given 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 the year → December	*ET0100* (*et_spells*)
bjahr	*Employment: starting year, generated* Year in which the employment spell began. Note: see *bmonat*	*ET0200* (*et_spells*)
emonat	*Employment: ending month, generated* Month in which the employment spell ended. To generate the variable, if information was only available on the season when a spell started, it was converted into a definite month and for right-censored spells (i.e. spells that were still ongoing when the person was interviewed) the interview month was entered. Note: see *bmonat*	*ET0300; ET0500* (*et_spells*) *pintmon* (*PENDDAT*)
ejahr	*Employment: ending year, generated* Year in which the employment spell ended. For right-censored spells (i.e. spells that were still ongoing when the h'hold was interviewed) the interview year was entered. Note: see *bmonat*	*ET0400; ET0500* (*et_spells*) *pintjahr* (*PENDDAT*)

Variable	Variable label and description	Source variables for generated variable in wave 2
zensiert	*Employment: spell still ongoing (right-censoring)* The censoring indicator shows whether a spell was still ongoing at the time of the personal interview in the 2nd wave, i.e. whether it is a right-censored spell. Note: A spell is regarded as censored if one of the two following conditions is met: The person reports in question P42 concerning the end date of the employment spell that the employment is still ongoing on the interview date (P42 end = 0). Or in P42 an end date is reported which is identical to the interview date and it is confirmed in the follow-up question P43 that the employment spell is still currently ongoing. Additional employment spells reported in the gap module and corrected dates were taken into account before generating the variable.	*ET0300; ET0400; ET0500* (*et_spells*)
stib	*Occupational status, code number, generated* Generation of the detailed code number for occupational status on the basis of various items gathering information on occupational status.	*ET0600; ET0700; ET0800; ET0900; ET1000; ET1100; ET1200* (*et_spells*)
arbzeit	*Weekly hours of work incl. details in the case of irregular working hours, generated.* Integrated variable on weekly hours of work in the job held by the respondent, combining responses to open-ended questions on working hours and the categorical follow-up question. For the closed categories of the follow-up question the mean values for the categories were used, for the open-ended category (40 or more hours) the median of the weekly working hours reported in the open-ended questions was used.	*ET2100; ET2200* (*et_spells*)

Table 19: Simple generated variables for wave 2 in the unemployment spell dataset (*al_spells*) (in the same order as in the dataset)

Variable	Variable label and description	Source variables for generated variable in wave 2
bmonat	*Registered unemployment: starting month, generated* Month in which the spell of registered unemployment began. To generate the variable, if information was only available on the season when a spell started this was converted into a definite 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 the year → December	*AL0100* (*al_spells*)
bjahr	*Registered unemployment: starting year, generated* Year in which the spell of registered unemployment began. Note: see *bmonat*	*AL0200* (*al_spells*)
emonat	*Registered unemployment: ending month, generated* Month in which the spell of registered unemployment ended. To generate the variable, if information was only available on the season when a spell started this was converted into a definite month and for right-censored spells (i.e. spells that were still ongoing when the person was interviewed) the interview month was entered. Note: see *bmonat*	*AL0300; AL0500* (*al_spells*)
ejahr	*Registered unemployment: ending year, generated* Year in which the spell of registered unemployment ended. For right-censored spells (i.e. spells that were still ongoing when the household was interviewed) the interview year was entered. Note: see *bmonat*	*AL0400; AL0500* (*al_spells*)

Variable	Variable label and description	Source variables for generated variable in wave 2
alg1bm	*Receipt of UB I: starting month, generated* Month in which the spell of Unemployment Benefit I receipt began. To generate the variable, if information was only available on the season when a spell started this was converted into a definite month. Note: Periods of receipt of Unemployment Benefit I are embedded in the spells of registered unemployment. A maximum of one period of UB I receipt is provided for 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. Conversion of the month details, see *bmonat*.	*AL0800* (*al_spells*)
alg1bj	*Receipt of UB I: starting year, generated* Year in which the spell of Unemployment Benefit I receipt began. Note: see *alg1bm*	*AL0900* (*al_spells*)
alg1em	*Receipt of UB I: ending month, generated* Month in which the spell of Unemployment Benefit I receipt ended. To generate the variable, if information was only available on the season when a spell started this was converted into a definite month and for right-censored spells (i.e. spells that were still ongoing when the person was interviewed) the interview month was entered Note: see *alg2kma - alg2kbme*	*AL1000; AL1200* (*al_spells*) *pintmon* (*PENDDAT*)
alg1ej	*Receipt of UB I: ending year, generated* Year in which the spell of Unemployment Benefit I receipt ended. For right-censored spells (i.e. spells that were still ongoing when the person was interviewed) the interview year was entered. Note: see *alg2kma - alg2kbme*	*AL1100; AL1200* (*al_spells*) *pintjahr* (*PENDDAT*)

Variable	Variable label and description	Source variables for generated variable in wave 2
alg1akt	*Receipt of UB I: spell still ongoing (right-censoring)* The censoring indicator shows whether the spell of Unemployment Benefit I receipt was still ongoing at the time of the personal interview in the 2nd wave, i.e. whether it is a right-censored spell. Note: A spell is regarded as censored if one of the two following conditions is met: The person reports in question P114 concerning the end date of the spell of Unemployment Benefit I receipt that the benefit receipt is still ongoing on the interview date (P114 end = 0). Or in P114 an end date is reported which is identical to the interview date and it is confirmed in the follow-up question P115 that benefit receipt is still currently ongoing. The variable is generated on the basis of the generated date variables, which are checked for plausibility.	*emonat, ejahr, AL1000; AL1100; AL1200* (*al_spells*)
zensiert	*Registered unemployment: spell still ongoing (right-censoring)* The censoring indicator shows whether a spell was still ongoing at the time of the personal interview in the 2nd wave, i.e. whether it is a right-censored spell. Note: A spell is regarded as censored if one of the two following conditions is met: The person reports in question P109 concerning the end date of the spell of registered unemployment that he/she is still registered as unemployed on the interview date (P109 end = 0). Or in P109 an end date is reported which is identical to the interview date and it is confirmed in the follow-up question P110 that the spell of registered unemployment is still ongoing. Additional employment spells reported in the gap module and corrected dates were taken into account before the variable was generated.	*AL0300; AL0400; AL0500* (*al_spells*)

Table 20: Simple generated variables for wave 2 in the gap spell dataset (*lu_spells*) (in the same order as in the dataset)

Variable	Variable label and description	Source variables for generated variable in wave 2
bmonat	*Spell: starting month, generated* Month in which the spell of economic inactivity began. To generate the variable, if information was only available on the season when a spell started this was converted into a definite 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 the year → December	*LU0200* (*lu_spells*)
bjahr	*Spell: starting year, generated* Year in which the spell of economic inactivity began. Note: see *bmonat*	*LU0300* (*lu_spells*)
emonat	*Spell: ending month, generated* Month in which the spell of economic inactivity ended. To generate the variable, if information was only available on the season when a spell started this was converted into a definite month and for right-censored spells (i.e. spells that were still ongoing when the person was interviewed) the interview month was entered. Note: see *bmonat*	*LU0400; LU0600* (*lu_spells*) *pintjahr* (*PENDDAT*)
ejahr	*Spell: ending year, generated* Year in which the spell of economic inactivity ended. To generate the variable, if information was only available on the season when a spell started this was converted into a definite month and for right-censored spells (i.e. spells that were still ongoing when the person was interviewed) the interview year was entered. Note: see *bmonat*	*LU0500; LU0600* (*lu_spells*) *pintjahr* (*PENDDAT*)

Variable	Variable label and description	Source variables for generated variable in wave 2
zensiert	*Spell: spell still ongoing (right censoring)* The censoring indicator shows whether a spell was still ongoing at the time of the personal interview in the 2nd wave, i.e. whether it is a right-censored spell. Note: A spell is regarded as censored if one of the two following conditions is met: The person reports in question P130 concerning the end date that he/she is still economically inactive at the date of the interview (P130 end = 0). Or in P130 an end date is reported which is identical to the interview date and it is confirmed in the follow-up question P131 that the status of economic inactivity is still ongoing.	*LU0400; LU0500; LU0600* (*lu_spells*)

Table 21: Simple generated variables for wave 2 in the employment and training measure spell dataset (*mn_spells*) (in the same order as in the dataset)

Variable	Variable label and description	Source variables for generated variable in wave 2
bmonat	*Measure: starting month, generated* Month in which the measure of active labour market policy spell began. To generate the variable, if information was only available on the season when a spell started this was converted into a definite month. Note: The generated date variables were checked for plausibility and corrected if necessary. The dates originally reported by the respondent (apart from values identified as implausible when the range of values was checked) 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 the year → December	*MN0300* (*mn_spells*)
bjahr	*Measure: starting year, generated* Year in which the measure of active labour market policy spell began. Note: see *bmonat*	*MN0400* (*mn_spells*)
emonat	*Measure: ending month, generated* Month in which the measure of active labour market policy ended. To generate the variable, if information was only available on the season when a spell started this was converted into a definite month and for right-censored spells (i.e. spells that were still ongoing when the person was interviewed) the interview month was entered. If the duration of the measure was reported instead of an end date, then the end date was calculated from the start date and the duration. Note: see *bmonat*	*MN0300; MN0400; MN0500; MN0600; MN0700; MN1100; MN1200* (*mn_spells*) *pintjahr* (*PENDDAT*)
ejahr	*Measure: ending year, generated* Year in which the measure of active labour market policy spell ended. For right-censored spells (i.e. spells that were still ongoing when the person was interviewed) the interview year was entered. If the duration of the measure was reported instead of an end date, then the end date was calculated from the start date and the duration. Note: see *bmonat*	*MN0300; MN0400; MN0500; MN0600; MN0800; MN1100; MN1300* (*mn_spells*) *pintjahr* (*PENDDAT*)

Variable	Variable label and description	Source variables for generated variable in wave 2
zensiert	*Measure: spell still ongoing (right censoring)* The censoring indicator shows whether a spell was still ongoing at the time of the personal interview in the 2nd wave, i.e. whether it is a right-censored spell. Note: A spell is regarded as censored if the person reports in question P164 that he/she is currently still taking part in a measure. (P164=1)	*MN0500* (*mn_spells*)

Table 22: Simple generated variables for wave 2 in the person register dataset (p_register) (in alphabetical order)

Variable	Variable label and description	Source variables for generated variable in wave 2
alter1	*Person's age in wave 1 (2006/2007)* Variable contains the information provided by the household reference person in the household interview of wave 1 regarding the age of the respective household member.	*HD0200a* to *HD0200o* (*HHENDDAT*)
alter2	*Person's age in wave 2 (2007/2008)* Variable contains the "best" available information regarding a person's age. This is either (a) the age calculated from the date of birth reported in wave 2 or (b) if no date of birth is available from wave 2, then the age reported in the household interview. The information from *alter2* was also taken over into the household dataset and corresponds to the information in *HD0200a* to *HD0200o*. This procedure is consistent with that followed by TNS Infratest. Already during the fieldwork the age variable was filled with the respective "best" information. There, a variable in the database is first filled with the age according to the household interview. If a personal interview is conducted, this variable in the database is overwritten with the age calculated on the basis of the details given in the personal interview (date of birth, date of personal interview). Both the age details for wave 2 provided in the household dataset and those in the individual dataset are based on this variable of the database and are taken into account in the plausibility check and for the generation of the types of benefit communities (Bedarfsgemeinschaft) and households.	*p1, pintjahr, pintmon, pinttag* (*PENDDAT*) *HD0200a* to *HD0200o* (*HHENDDAT*)

Variable	Variable label and description	Source variables for generated variable in wave 2
korrsex	*Information on gender was corrected between the survey waves* For individuals who belonged to a sample h'hold in more than one wave this variable indicates whether the gender was corrected in the household interview.	*HD0100a* to *HD0100o* of waves 1 and 2 (*HHENDDAT*)
lastint	*Survey wave of the last interview at the personal level* This variable indicates the wave in which the last interview at the individual level was conducted with the person (personal interview or senior citizen's interview).	Personal interviews of waves 1 and 2 (*PENDDAT*)
neuj2	*Year when the person joined the current household, reported in wave 2 (2007/2008)* This variable indicates the year when the person joined the household of which he/she is a member in the 2nd wave. Note: Information on the date comes from the wave 2 interview with the re-interviewed household into which the person has moved or been born since the 1st wave.	Information on the date when a person joined the household. Asked in the household questionnaire for re-interviewed households (HH18, HH37)
neum2	*Month when the person joined the current household, reported in wave 2 (2007/2008)* This variable indicates the month when the person joined the household of which he/she is a member in the 2nd wave. Note: see *neuj2*	Information on the date when a person joined the household. Asked in the household questionnaire for re-interviewed households (HH18, HH37)
wegj2	*Year when the person left the previous household, reported in wave 2 (2007/20008)* This variable indicates the year when the person ceased to be a member of the household of the previous wave. Note: Information on the date comes from the wave 2 interview with the household in which the person was living in the 1st wave.	Information on the date when a person ceased to be in the household. Asked in the household questionnaire for re-interviewed households (HH8, HH28)
wegm2	*Month when the person left the previous household, reported in wave 2 (2007/20008)* This variable indicates the month when the person ceased to be a member of the household of the previous wave. Note: see *wegj2*	Information on the date when a person ceased to be in the household. Asked in the household questionnaire for re-interviewed households (HH8, HH28)
zmhh2	*Pointer: Personal identification number of target person's mother in the household in wave 2 (2007/2008)* Contains the personal identification number of the mother if she is living in the household. Natural mothers, stepmothers, adoptive or foster mothers, or mothers whose status is not specified are counted as the mother.	Information on relationships between household members of the 2nd wave (household grid);

Variable	Variable label and description	Source variables for generated variable in wave 2
zparthh2	*Pointer: Personal identification number of target person's partner in the household in wave 2 (2007/2008)* Contains the personal identification number of a partner living in the household. Spouses, same-sex registered partners, cohabitees and partners whose status is not specified are counted as a partner.	Information on relationships between household members of the 2nd wave (household grid);
zupanel	*Survey wave in which the person joined the panel* This variable indicates the wave in which the person was first a member of a sample household.	Information on people living in the household from waves 1 and 2 (Household grid)
zvhh2	*Pointer: Personal identification number of target person's father in the household in wave 2 (2007/2008)* Contains the personal identification number of the father if he is living in the household. Natural fathers, stepfathers, adoptive or foster fathers, or fathers whose status is not specified are counted as the father.	Information on relationships between household members of the 2nd wave (household grid);

The datasets at the individual level contain a multitude of generated variables and constructed variables. These also include variables (e.g. for occupational status) that can be found in more than one dataset. Figures 4 and 5 provide an overview of the simple and complex generated variables at the individual level.

Figure 4: Overview of generated variables at the individual level in wave 2

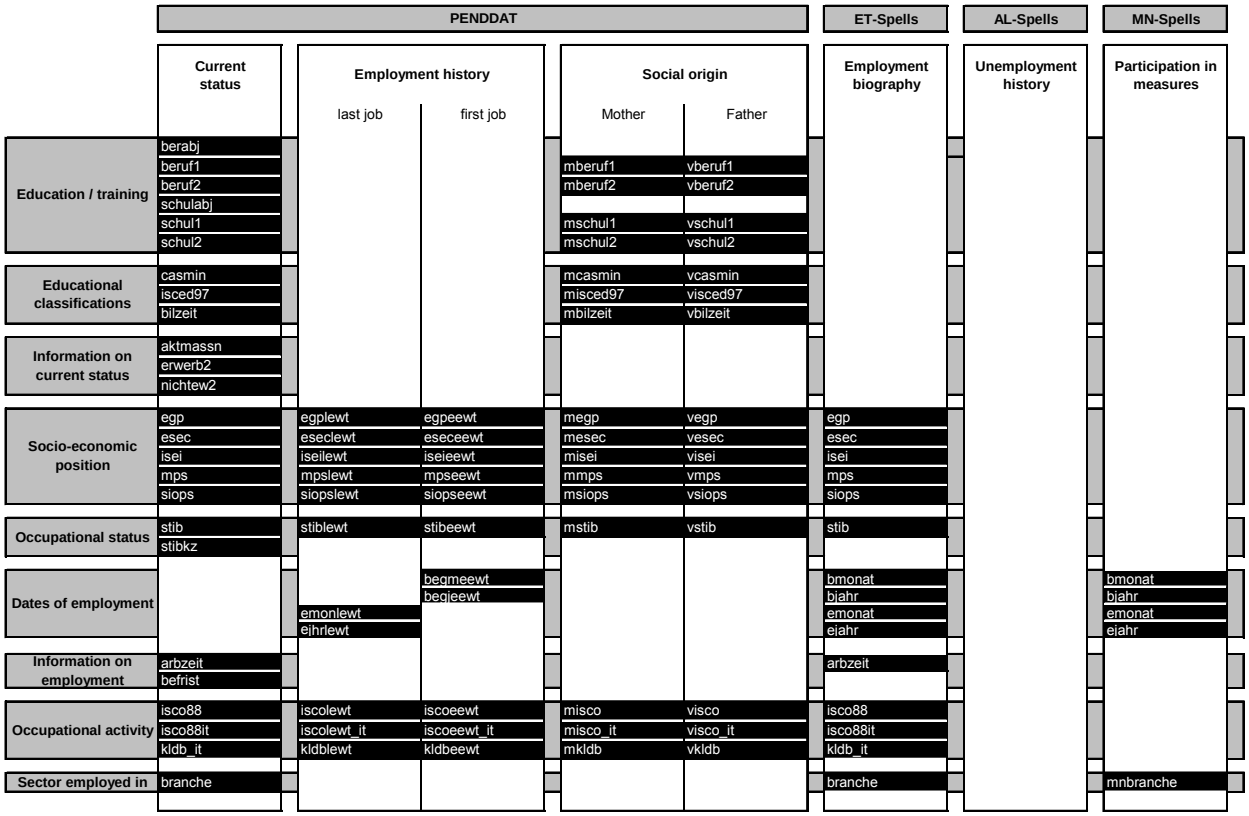
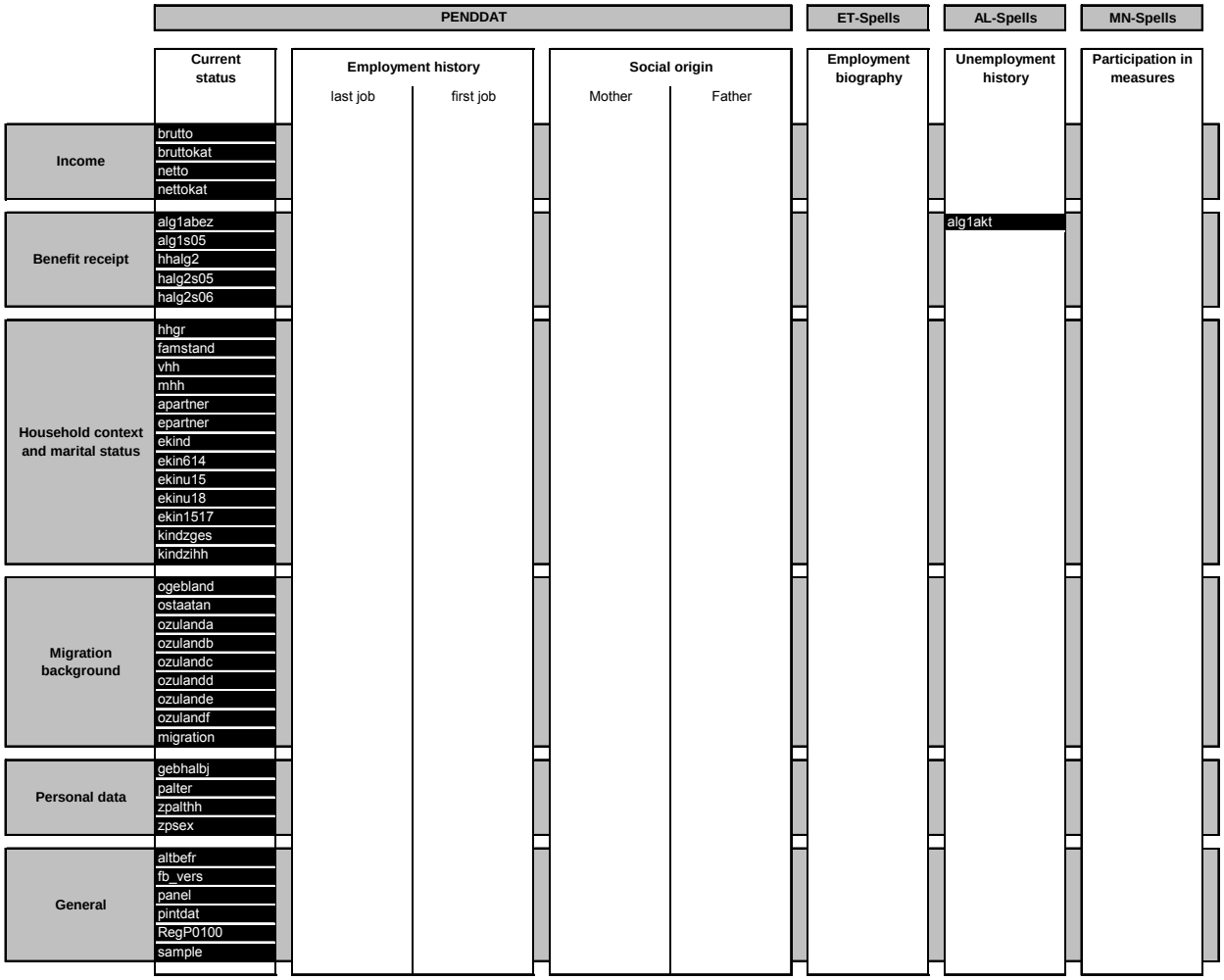


Figure 5: Overview of generated variables at the individual level in wave 2



7.5 Theory-based constructed variables

Theory-based constructed variables are variables whose generation requires more extensive re-coding and/or coding. In most cases these variables have been empirically tested elsewhere and have a foundation in theoretical concepts. Moreover, some of them are standardised instruments used in social sciences or economics. Examples of such standardised instruments are the European Socio-economic Classification (ESeC), the International Standard Classification of Education (ISCED) or the equivalised household income. This chapter provides a detailed description of the theory-based constructed variables made available in the PASS data as well as a short overview of their theoretical background and the most important references.

7.5.1 Individual level

Education in years

Variable name	*bilzeit*
Variable label	Duration of school education and vocational training in years, generated
Source variables	*schul2; beruf2*
Category / dataset	Education / individual-level data
Prepared by	Bernhard Christoph
Description	For many statistical models using a linear variable for education is more appropriate than using a categorical one. For schooling levels, it is fairly easy to convert the categorical information into linear information. The linear value simply corresponds to the time spent at school until attainment of the final school leaving qualification. Care must be taken here, however, to ensure that equivalent qualifications are always allocated identical durations. An upper secondary school leaving certificate, for example, should always be labelled with the same duration, irrespective of whether it was attained after twelve or thirteen years of education. Secondary school qualifications were allocated the following education durations for this variable:

Lower secondary school leaving certificate; lower secondary school leaving certificate from the former GDR (POS) after completion of grade 8; other lower secondary school leaving certificate:	9 years
Intermediate secondary school leaving certificate; intermediate secondary school leaving certificate from the former GDR (POS) after completion of grade 10:	10 years
Entrance qualification for University of Applied Sciences:	12 years
General qualification for university entrance or subject-specific higher education entrance qualification (incl. EOS – comparable qualification in the former GDR)	13 years

The situation is different for vocational qualifications, however. Due to the numerous different ways to gain a vocational qualification and the related potentially large differences in income even for qualifications with comparable training durations, the training duration may not be subjected to a simple one-to-one conversion process. This problem can be avoided by attempting to estimate the growth in human capital related to a certain vocational qualification. (see e.g. Helberger 1988).

This study uses such an approach. For the conversion process, 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 school education.

Training as a semi-skilled worker:	+1 year
Apprenticeship, vocational school, school for health-care occupations:	+1.5 years
Master craftsman's certificate	+3 years
College of advanced vocational studies:	+3 years
University of Applied Sciences/Bachelor:	+3 years
University/Masters degree:	+5 years
PhD.:	+8 years
Other German qualification:	+1.5 years
Other foreign qualification:	+1.5 years

Reference	Helberger (1988)

Education in years, mother

Variable name	*mbilzeit*
Variable label	Duration of school education and vocational training in years, generated
Source variables	*mschul2; mberuf2*
Category / dataset	Education / individual-level data
Prepared by	Bernhard Christoph
Description	General description: see 'Education in years'
	When generating the variable for the parents' years of education and training, the values added for vocational qualifications differ from those used when constructing the corresponding variable for the respondents, since information on vocational education/training was collected in less detail for the parents (especially as far as tertiary education is concerned). The values corresponding to particular courses of education/training are as follows:
	Training as a semi-skilled worker: +1 year Apprenticeship, vocational school, school for health care occupations: +1.5 years Master craftsman's certificate: +3 years College of advanced vocational studies: +3 years University of Applied Sciences: +3 years University: +5 years Other German qualification: +1.5 years Other foreign qualification: +1.5 years
Reference	Helberger (1988)

Education in years, father

Variable name	*vbilzeit*
Variable label	Duration of school education and vocational training in years, generated
Source variables	*vschul2; vberuf2*
Category / dataset	Education / individual-level data
Category	Education
Prepared by	Bernhard Christoph
Description	General description: see 'Education in years'
	When generating the variable for the parents' years of education and training, the values added for vocational qualifications differ from those used when constructing the corresponding variable for the respondents, since information on vocational education/training was collected in less detail for the parents (especially as far as tertiary education is concerned). The values corresponding to particular courses of education/training are as follows:
	Training as a semi-skilled worker: +1 year Apprenticeship, vocational school, school for health care occupations: +1.5 years Master craftsman's certificate: +3 years College of advanced vocational studies: +3 years University of Applied Sciences: +3 years University: +5 years Other German qualification: +1.5 years Other foreign qualification: +1.5 years
Reference	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
Description	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 school and vocational qualifications on an international scale (König et al. 1987). An updated version is now available (Brauns & Steinmann 1999). The procedures for re-coding qualifications acc. to CASMIN applied in the panel, especially for problematic cases, follow the procedures described in Lechert et al. (2006) and Granato (2000). For this, the slightly differing category values of the education variable in this dataset are of course taken into account. Details can be found in the table below. Cells containing valid combinations according to CASMIN are highlighted in light grey, those containing defined missing values are dark grey.

school / occup.	not surv.	pupil	not asked	not applic.	no details	don't know	no qual.	special needs school	lower sec. school	interm. sec. school	entrance qual. for univ. 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	-	-	-	-	-	-	-	-	-	-	-
not applic.	-	-	-	-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 qualif.	-	-	-	-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
apprent-iceship	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
f-t voc. school	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
health occ. sch.	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
master craftsm.	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
BA	-	-	-	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a
UAS/ bachelor	-	-	-	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a
univ./ masters	-	-	-	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b
PhD	-	-	-	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b
oth. Ger. qual.	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
oth. for.	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c

References	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
Description	General description: see CASMIN

Since the education variable has different category values for respondents and their parents, the coding pattern of *mcasmin* and *vcasmin* differs slightly from the pattern used in *casmin*. The following table shows the differences in detail.

school / occup.	not surv.	pers. int. missing	parent un-known	not asked	not applic.	no details	don't know	no qual.	special needs schook	lower sec. school	interm. sec. school	entrance qual. for univ. of app. sci.	upper sec. leaving cert.	other Ger. qual.	other for. qual.
not surv.	-10	-	-	-	-	-	-	-	-	-	-	-	-	-	-
implaus. value	-	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
pers. int. missing	-	-6	-	-	-	-	-	-	-	-	-	-	-	-	-
parent un-known	-	-	-5	-	-	-	-	-	-	-	-	-	-	-	-
not asked.	-	-	-	-4	-	-	-	-	-	-	-	-	-	-	-
not applic.	-	-	-	-	-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
apprent-iceship	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
master craftsm.	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
BA	-	-	-	-	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a
univ. of appl. sci.	-	-	-	-	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a
univ.	-	-	-	-	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b
oth. Ger. qual.	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
oth. for. qual.	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c

References

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
Description	General description: see CASMIN

Since the education variable has different category values for respondents and their parents, the coding pattern of *mcasmin* and *vcasmin* differs slightly from the pattern used in *casmin*. The following table shows the differences in detail.

school / occup.	not surv.	pers. int. missing	parent un-known	not asked	not applic.	no details	don't know	no qual.	special needs schook	lower sec. school	interm. sec. school	entrance qual. for univ. of app. sci.	upper sec. leaving cert.	other Ger. qual.	other for. qual.
not surv.	-10	-	-	-	-	-	-	-	-	-	-	-	-	-	-
implaus. value	-	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
pers. int. missing	-	-6	-	-	-	-	-	-	-	-	-	-	-	-	-
parent un-known	-	-	-5	-	-	-	-	-	-	-	-	-	-	-	-
not asked.	-	-	-	-4	-	-	-	-	-	-	-	-	-	-	-
not applic.	-	-	-	-	-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
apprent-iceship	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
master craftsm.	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
BA	-	-	-	-	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a
univ. of appl. sci.	-	-	-	-	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a
univ.	-	-	-	-	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b
oth. Ger. qual.	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
oth. for. qual.	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c

References

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

ISCED 97

Variable name	*isced*
Variable label	Education classified acc. to isced97, updated version, generated
Source variables	*schul2; beruf2*
Category / dataset	Education / individual-level data
Prepared by	Bernhard Christoph
Description	ISCED-97 (International Standard Classification of Education) developed by the OECD (OECD 1999, for an outline, see also BMBF 2003) is an education classification which can be used as an alternative to CASMIN.

What must be taken into account regarding the coding of the ISCED-97 classification is that it includes categories which cannot reasonably be assigned to the present data. The ISCED values '0' (pre-primary education/ kindergarten) and '1' (primary education) do not apply, because the respondents are at least 15 years of age. Instead a separate group was generated for individuals with an education below ISCED level 2 (ISCED 2 = lower or intermediate secondary school leaving certificate). Therefore, only ISCED levels 2 to 6 are considered in the coding procedure applied 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 surveyed	pupil	not asked	not applic.	no details	don't know	no qual.	special needs school	lower sec. school	interm. sec. school	entrance qual. for univ. of app. sci.	upper sec. leaving cert.	other German qual.	other foreign qual.
not surveyed	-10	-	-	-	-	-	-	-	-	-	-	-	-	-
implaus. value	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
pupil	-	-5	-	-	-	-	-	-	-	-	-	-	-	-
not asked	-	-	-4	-	-	-	-	-	-	-	-	-	-	-
not applic.	-	-	-	-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
apprent-iceship	-	-	-	-3	-2	-1	3b	3b	3b	3b	4a	4a	3b	3b
full-time voc. sch.	-	-	-	-3	-2	-1	3b	3b	3b	3b	4a	4a	3b	3b
health occ. sch.	-	-	-	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
master craftsm.	-	-	-	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
BA	-	-	-	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
UAS/ bachelor	-	-	-	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a
univ./ masters	-	-	-	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a
Ph.D.	-	-	-	6	6	6	6	6	6	6	6	6	6	6
oth. Ger. qual.	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2
other foreign	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2

References

BMBF (2003); OECD (1999)

MISCED 97

Variable name	*misced*
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
Description	For the theoretical background and generation details, see ISCED-97. In contrast to the ISCED-97 coding applied to data on the respondents' education, it is not possible generate ISCED level 6 for data on their parents. This is so, since data on the corresponding qualifications (i.e. PhD or equivalent) were not collected for the parents. Therefore only ISCED levels 2 to 5 are covered in the coding applied in this dataset. The following table shows the coding details.

school occup.	not surv.	pers. int. missing	parent un-known	not asked	not applic.	no details	don't know	no qual.	special needs school	lower sec. school	interm. sec. school	entrance qual. for univ. of app. sci.	upper sec. leaving cert.	other German qual.	other foreign qual.
not surv.	-10	-	-	-	-	-	-	-	-	-	-	-	-	-	-
implaus. value	-	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
pers. int. missing	-	-6	-	-	-	-	-	-	-	-	-	-	-	-	-
parent un-known	-	-	-5	-	-	-	-	-	-	-	-	-	-	-	-
not asked	-	-	-	-4	-	-	-	-	-	-	-	-	-	-	-
not applic.	-	-	-	-	-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 qualif.	-	-	-	-	-3	-2	-1	1	1	2	2	3a	3a	2	2
semi-skilled	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2
apprent-iceship	-	-	-	-	-3	-2	-1	3b	3b	3b	3b	4a	4a	3b	3b
master craftsm.	-	-	-	-	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
BA	-	-	-	-	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
univ. of appl. sci.	-	-	-	-	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a
univ.	-	-	-	-	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a
oth. Ger. qual.	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2
oth. for. qual.	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2

References

BMBF (2003); OECD (1999)

VISCED 97

Variable name	*visced*
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
Description	For the theoretical background and generation details, see ISCED-97. In contrast to the ISCED-97 coding applied to data on the respondents' education, it is not possible generate ISCED level 6 for data on their parents. This is so, since data on the corresponding qualifications (i.e. PhD or equivalent) were not collected for the parents. Therefore only ISCED levels 2 to 5 are covered in the coding applied in this dataset. The following table shows the coding details.

school occup.	not surv.	pers. int. missing	parent un-known	not asked	not applic.	no details	don't know	no qual.	special needs school	lower sec. school	interm. sec. school	entrance qual. for univ. of app. sci.	upper sec. leaving cert.	other German qual.	other foreign qual.
not surv.	-10	-	-	-	-	-	-	-	-	-	-	-	-	-	-
implaus. value	-	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
pers. int. missing	-	-6	-	-	-	-	-	-	-	-	-	-	-	-	-
parent un-known	-	-	-5	-	-	-	-	-	-	-	-	-	-	-	-
not asked	-	-	-	-4	-	-	-	-	-	-	-	-	-	-	-
not applic.	-	-	-	-	-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 qualif.	-	-	-	-	-3	-2	-1	1	1	2	2	3a	3a	2	2
semi-skilled	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2
apprent-iceship	-	-	-	-	-3	-2	-1	3b	3b	3b	3b	4a	4a	3b	3b
master craftsm.	-	-	-	-	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
BA	-	-	-	-	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
univ. of appl. sci.	-	-	-	-	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a
univ.	-	-	-	-	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a
oth. Ger. qual.	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2
oth. for. qual.	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2

References

BMBF (2003); OECD (1999)

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

Generated	Occupation	Variable name	Source variables
	Current	*isco88*	W1: P46; W2: P40_X
	Spell data (*et_spells*)	*isco88*	W2: P40_X
	First	*iscoeewt*	W2: P40_X, P91, P100
	Last	*iscolewt*	W2: P40_X, P91
	Of father	*visco*	W2: P299
	Of mother	*misco*	W2: P288
Variable labels	Current occupation: ISCO-88 (ZUMA coding), generated		
	Spell data (*et_spells*): ISCO-88 (ZUMA coding), generated		
	First occupation: ISCO-88 (ZUMA coding), first occupation, generated		
	Last occupation: ISCO-88 (ZUMA coding), last occupation, 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		
Description	The International Standard Classification of Occupations (ISCO) was developed by the International Labour Organization (ILO), as an internationally comparative classification. The special feature of the ISCO-88 is that in addition to the job performed, the qualification level generally necessary to perform the job is taken into account when assigning an occupation to a particular occupational code. This constitutes a major difference to the Classification of Occupations provided by the German Federal Statistical Office (KldB), which is also provided in this dataset.		
	The actual coding was carried out by the Centre for Survey Research and Methodology (Gesis, formerly ZUMA). In contrast to the coding variant used by TNS Infratest, ZUMA's ISCO-88 is generated using a coding procedure based on the original information from the open-ended job descriptions and is not derived from the KldB codes. The ISCO-based measures of occupational status and prestige provided in this dataset are generated on the basis of the ZUMA coding.		
	A list of ISCO-88 occupations and the respective values of occupational status and prestige scales are provided in Table 51 in Appendix I.		
Reference	ILO (1990)		

International Standard Classification of Occupations 1988 (ISCO88); TNS Infratest coding

Generated	Occupation	Variable name	Source variables
	Current	*isco88it*	W1: P46; W2: P40_X
	Spell data (*et_spells*)	*isco88it*	W2: P40_X
	First	*Iscoeewt_it*	W2: P40_X, P91, P100
	Last	*iscolewt_it*	W2: P40_X, P91
	Of father	*visco_it*	W2: P299
	Of mother	*misco_it*	W2: P288
Variable labels	Current occupation: ISCO-88 (TNS Infratest coding), generated		
	Spell data (*et_spells*): ISCO-88 (TNS Infratest coding), generated		
	First occupation: ISCO-88 (TNS Infratest coding), first occupation, generated		
	Last occupation: ISCO-88 (TNS Infratest coding), last occupation, generated		
	Father: ISCO-88 (TNS Infratest coding) of the father, generated		
	Mother: ISCO-88 (TNS Infratest coding) of the mother, generated		
Category / dataset	Occupation / individual-level data		
Contact person	Bernhard Christoph		
Description	The International Standard Classification of Occupations (ISCO) was developed by the International Labour Organization (ILO), as an internationally comparative classification. The special feature of the ISCO-88 is that in addition to the job performed, the qualification level generally necessary to perform the job is taken into account when assigning an occupation to a particular occupational code. This constitutes a major difference to the Classification of Occupations provided by the German Federal Statistical Office (KldB), which is also provided in this dataset.		
	The coding was carried out by TNS Infratest, the field institute of PASS for waves 1-3, using a procedure for deriving ISCO-88 codes from the German Federal Statistical Office's Classification of Occupations.		
	A detailed list of ISCO-88 occupations and the respective values of occupational status and prestige scales are provided in Table 51 of Appendix I.		
Reference	ILO (1990)		

Classification of Occupations 1992 (KldB92); TNS Infratest Coding

Generated	Occupation	Variable name	Source variables
	Current	*kldb_it*	W1: P46; W2: P40_X
	Spell data (*et_spells*)	*kldb_it*	W2: P40_X
	First	*kldbeewt*	W2: P40_X, P91, P100
	Last	*kldblewt*	W2: P40_X, P91
	Of father	*vkldb*	W2: P299
	Of mother	*mkldb*	W2: P288
Variable labels	Current occupation: KldB 92 (TNS Infratest coding), generated		
	Spell data (*et_spells*): KldB 92 (TNS Infratest coding), generated		
	First occup.: KldB 92 (TNS Infratest coding), first occupation, generated		
	Last occup.: KldB 92 (TNS Infratest coding), last occupation, generated		
	Father: KldB 92 (TNS Infratest coding) of the father, generated		
	Mother: KldB 92 (TNS Infratest coding) of the mother, generated		
Category / dataset	Occupation / individual-level data		
Contact person	Bernhard Christoph		
Description	The KldB92 is the current version of the Classification of Occupations published by the German Federal Statistical Office. It is a classification system that was specifically constructed to match the particularities of the German occupational structure. It is based solely on job descriptions.		
	The coding was carried out by TNS Infratest, the field institute of PASS for waves 1-3.		
Reference	StBA (1992).		

Class scheme according to Erikson, Goldthorpe and Portocarrero (EGP)

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

European Socio-economic Classification (ESeC)

Generated	Occupation	Variable name	Source variables
	Current	*esec*	*isco88, stib, PET2000, PET2700*
	Spell data (*et_spells*)	*esec*	*isco88, stib, ET1100, ET1300*
	First	*eseceewt*	*iscoeewt, stibeewt, PET1261,*
	Last	*eseclewt*	*iscolewt, stiblewt, PET3801*
	Of father	*vesec*	*visco, vstib, PSH0670*
	Of mother	*mesec*	*misco, mstib, PSH0370*
Variable labels	Current occup.: European Socio-economic Classification (ESeC), current occupation, generated		
	Spell data (*et_spells*): European Socio-economic Classification (ESeC), generated		
	First occup.: European Socio-economic Classification (ESeC), first occupation, generated		
	Last occup.: European Socio-economic Classification (ESeC), last occupation, generated		
	Father: European Socio-economic Classification (ESeC), occupation of father, generated		
	Mother: European Socio-economic Classification (ESeC), occupation of mother, generated		
Category / dataset	Socio-economic position / individual-level data		
Prepared by	Bernhard Christoph		
Description	With regard to its theoretical conception, the European Socio-economic Classification is largely based on the EGP class scheme. In contrast to the latter, however, great importance was attached to international comparability of operationalisation procedures and comprehensive validation of the classification scheme (for a general description, see: Rose & Harrison 2007, and Müller et al. 2006, 2007 for Germany).		
	The Stata do-file required to generate the ESeC was kindly provided by Heike Wirth from Gesis ZUMA (Fischer & Wirth 2007). We simply adjusted it to the requirements of this study. This do-file, originally written in standard SPSS syntax by Harrison & Rose (2006) as a standard program for the generation of the ESeC, was converted into Stata.		
References	Fischer & Wirth (2007); Harrison & Rose (2006); Müller et al. (2006, 2007); Rose & Harrison (2007)		

Magnitude-Prestige-Scale (MPS)

Generated	Occupation	Variable name	Source variables
	Current	*mps*	*isco88*
	Spell data (*et_spells*)	*mps*	*isco88*
	First	*mpseewt*	*iscoeewt*
	Last	*mpslewt*	*iscolewt*
	Of father	*vmps*	*visco*
	Of mother	*mmps*	*misco*
Variable labels	Current occup.: Magnitude-Prestige-Scale, current occupation, generated		
	Spell data (*et_spells*): Magnitude-Prestige-Scale, generated		
	First occup.: Magnitude-Prestige-Scale, first occupation, generated		
	Last occup.: Magnitude-Prestige-Scale, last occupation, generated		
	Father: Magnitude-Prestige-Scale, occupation of father, generated		
	Mother: Magnitude-Prestige-Scale, occupation of mother, generated		
Category / dataset	Socio-economic position / individual-level data		
Contact person	Bernhard Christoph		
Description	The Magnitude-Prestige-Scale [MPS] (Wegener 1985, 1988) is the only specifically German instrument available so far to operationalise social prestige on the basis of detailed occupation information. It was originally developed for the older 1968 version of the International Standard Classification of Occupations (ISCO-68). Since occupation coding in the PASS was conducted on the basis of the more recent ISCO-88 classification and the Classification of Occupations (KldB) developed by the Federal Statistical Office, a variant of the scale transferred to ISCO-88 was used (Christoph 2005). The data was merged by the Centre for Survey Research and Methodology (Gesis-ZUMA) as part of the occupational coding procedure.		
References	Christoph (2005); Wegener (1985, 1988)		

Standard International Occupational Prestige Scale (SIOPS/Treiman-Scale)

Generated	Occupation	Variable name	Source variables
	Current	*siops*	*isco88*
	Spell data (*et_spells*)	*siops*	*isco88*
	First	*siopseewt*	*iscoeewt*
	Last	*siopslewt*	*iscolewt*
	Of father	*vsiops*	*visco*
	Of mother	*msiops*	*misco*
Variable labels	Current occup.: Standard International Occupational Prestige Scale, current occupation, generated		
	Spell data (*et_spells*): Standard International Occupational Prestige Scale, generated		
	First occup.: Standard International Occupational Prestige Scale, first occupation, generated		
	Last occup.: Standard International Occupational Prestige Scale, last occupation, generated		
	Father: Standard International Occupational Prestige Scale, occupation of father, generated		
	Mother: Standard International Occupational Prestige Scale, occupation of mother, generated		
Category / dataset	Socio-economic position / individual-level data		
Contact person	Bernhard Christoph		
Description	The Treiman Prestige Scale, which was originally constructed by Treiman (1977) for the ISCO-68, is the first and only prestige scale available so far which can be used for internationally comparative research into occupations. Since its adaptation to the ISCO-88 (Ganzeboom & Treiman 1996, 2003) the scale has commonly been used under the name 'Standard International Occupational Prestige Scale'. The data were merged by the Centre for Survey Research and Methodology (Gesis-ZUMA) as part of the occupational coding procedure.		
References	Ganzeboom & Treiman (1996, 2003); Treiman (1977)		

International Socio-Economic Index (ISEI)

Generated	Occupation	Variable name	Source variables
	Current	*isei*	*isco88*
	Spell data (*et_spells*)	*isei*	*isco88*
	First	*iseieewt*	*iscoeewt*
	Last	*iseilewt*	*iscolewt*
	Of father	*visei*	*visco*
	Of mother	*misei*	*misco*
Variable labels	Current occup.: International Socio-Economic Index, current occupation, generated		
	Spell data (*et_spells*): International Socio-Economic Index, generated		
	First occup.: International Socio-Economic Index, first occupation, generated		
	Last occup.: International Socio-Economic Index, last occupation, generated		
	Father: International Socio-Economic Index, occupation of father, generated		
	Mother: International Socio-Economic Index, occupation of mother, generated		
Category / dataset	Socio-economic position / individual-level data		
Contact person	Bernhard Christoph		
Description	The International Socio-Economic Index is certainly one of the most common indices for operationalising social position. This is due not least to the fact that, in contrast to most other SEIs, the ISEI is based on an original theoretical concept which sees the occupation and its socio-economic status as an 'intervening variable' between education and income.		
	Initially, the ISEI was developed for the ISCO-68 (Ganzeboom et al. 1992) and was later adapted to the ISCO-88 (Ganzeboom & Treiman 1996, 2003).		
	The data were merged by the Centre for Survey Research and Methodology (Gesis-ZUMA) as part of the occupational coding procedure.		
References	Ganzeboom et al. (1992); Ganzeboom & Treiman (1996, 2003)		

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

Generated	Occupation	Variable name	Source variables
	Current	*branche*	P61_X
	Spell data (*et_spells*)	*branche*	P61_X
	Spell data (*mn_spells*)	*mnbranche*	P176_X
Variable labels	Current occup.: current activity: economic sector/industry (WZ2003)		
	Spell data (*et_spells*): economic sector/industry (WZ2003), generated		
	Spell data (*mn_spells*): measure: economic sector/industry (WZ2003)		
Category / dataset	Socio-economic position / individual-level data		
Contact person	Bernhard Christoph		
Description	The information from the open-ended survey question about the sector/ industry in which the respondent works was coded on the basis of the 2-digit code in the Classification of Economic Activities of the Federal Statistical Office (WZ2003). 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.		
	The coding was performed by TNS Infratest, the field institute of PASS for waves 1-3.		
References	StaBA (2002); EG (2002)		

7.5.2 Variables at the level of the household or benefit community (Bedarfsgemeinschaft)

Equivalised household income, old OECD scale

Variable name	*oecdinca*
Variable label	Equivalised household income, old OECD scale (rounded)
Source variables	*HD0200a-HD0200o; HA0100; hhincome*
Category / dataset	Socio-economic position / household-level data
Prepared by	Bernhard Christoph
Description	With what is called the 'equivalized household income', statisticians try to take into account the savings achievable by means of joint housekeeping in multi-person households compared to single households. To do this, the per-capita income in multi-person households is not calculated on the basis of the actual number of individuals living in the household, but by using a divisor which is usually below this figure and is calculated on the basis of the assumed needs of the household members (equivalised household size). According to the old OECD scale, only the first household member (aged 15 or over) is assigned a weighting factor of 1.0. Further household members aged 15 or over are assigned a weighting factor of 0.7; children up to the age of 14 are counted with a weighting factor of 0.5 to calculate the equivalised household size. For more information on the old OECD scale, see OECD (1982); an overview on the topic is provided by Hauser (1996).
References	Hauser (1996); OECD (1982)

Equivalised household income, modified OECD scale

Variable name	*oecdincn*
Variable label	Equivalised household income, modified OECD scale (rounded)
Source variables	*HD0200a-HD0200o; HA0100; hhincome*
Category / dataset	Socio-economic position / household-level data
Prepared by	Bernhard Christoph
Description	General description: see 'Equivalised household income, old OECD scale'. The modified OECD equivalence scale assumes a weighting factor of 1.0 only for the first household member (aged 15 or over). Any further household members aged 15 or over are assigned a weighting factor of 0.5; children up to the age of 14 are counted with a weighting factor of 0.3 to calculate the equivalised household size. For more information on the modified OECD scale, see Hagenaars et al. (1994).
Reference	Hagenaars et al. (1994)

Deprivation Index, unweighted

Variable name	*depindug*
Variable label	Deprivation index, unweighted (items missing for financial reasons; total of unweighted items: 26)
Source variables	*HLS0100a-HLS2600a; HLS0100b-HLS2600b*
Category / dataset	Material situation / household-level data
Prepared by	Bernhard Christoph
Description	Following a proposal by Ringen (1988), a distinction is usually made in poverty research between a direct and an indirect measurement of poverty. Indirect measurement focuses on the resources available to attain a certain standard of living, in particular the (equivalised household) income. For this reason this is also referred to as the resource-based approach to measuring poverty. In contrast, direct measurement attempts to record the households' actual ownership of goods and tries to determine the extent to which the households cannot afford certain goods or activities which are considered to be relevant, for financial reasons. This is also referred to as the deprivation approach (see e.g. Halleröd 1995). According to the general tenor of previous scientific research, the population classified as poor by the resource-based approach is not always identical to that defined by the deprivation approach. In order to define exactly who is to be considered poor in the narrow sense, it has therefore often been suggested to combine the measures of income-related poverty and deprivation and to count only those who are classified as poor by both approaches as belonging to the population living in poverty in the narrow sense (see Halleröd 1995; Nolan & Whelan 1996; Andreß and Lipsmeier 2001). The index is based on a list of 26 goods or activities. The households surveyed are asked to indicate whether they possess these goods or participate in the activities mentioned. The unweighted index calculated on this basis simply adds up the number of items which the respondents indicated that they do not possess or do not participate in. However, only items which are missing for financial reasons are counted, in order to avoid certain consumer preferences (e.g. a household deliberately doing without a car or a television) being misinterpreted as a reduction in the standard of living. Additionally, an item was only accepted as missing for financial reasons if the answers to both questions explicitly confirmed this. 'Don't know' or 'details refused' answers were evaluated either as if the particular good was available in the household or as if it was missing for a reason other than financial reasons. This assumption is certainly not applicable to all cases. Alternatively, it would have been possible not to calculate an index value for households that failed to answer a question for (at least) one particular good ('listwise deletion'). With respect to the total of 26 goods and activities surveyed, however, this method could quickly have led to a large number of missing index values. For this reason the first method described was selected. Nevertheless, compared to the listwise deletion procedure, there is a risk of the number of goods missing being underestimated with this method.
References	Andreß & Lipsmeier (2001); Halleröd (1995); Nolan & Whelan (1996); Ringen (1988)

Deprivation Index, weighted

Variable name	*depindg*
Variable label	Deprivation index, weighted (items missing for financial reasons; total of weighted items: 12.8)
Source variables	*HLS0100a-HLS2600a; HLS0100b-HLS2600b; PLS0100-PLS2600*
Category / dataset	Material situation / household-level data (weighted at the individual level)
Prepared by	Bernhard Christoph
Description	For a general description, see deprivation index, unweighted. With respect to unweighted indices such as the one described above, there is often criticism that all of the items included are given identical weightings. When comparing two items, for example the question as to whether the dwelling has an indoor toilet or the one as to whether there is a VCR/ DVD player in the household, it immediately becomes clear that there is a vast difference in the extent to which a household's standard of living would be restrained by the lack of one of these items. It therefore seems reasonable to weight the individual items, even if empirical research has proven that in most cases weighted and unweighted index variants do not deliver significantly different results (see Lipsmeier 1999). For the present survey, we decided to weight items according to the proportion of respondents who regarded a particular item as necessary. We chose this procedure not only because it is convincing in conceptual terms and is a commonly used procedure (applied by Halleröd 1995, for example), but also because it could be implemented without unreasonable costs. As the deprivation weightings to be determined for the individual questionnaire items can be assumed to be highly stable over time, these items need only be administered once or at comparably long intervals. Moreover, thanks to the large population of the PASS sample, we were able to split the population into several randomly selected subsamples, each of which was presented with only some of the items. Alternative weighting methods, such as restricting the index to those items which are considered necessary by a certain minimum proportion of the respondents (e.g. Andreß & Lipsmeier 1995, Andreß et al. 1996) or a theoretical restriction to a few fundamental items (e.g. Nolan & Whelan 1996), were not applied in this survey, but can be generated, if necessary, on the basis of the data provided. A discussion summarising the different methods of index weighting can be found in Andreß & Lipsmeier (2001, esp. pp. 28 ff.).
References	Andreß & Lipsmeier (1995, 2001); Andreß et al. (1996); Halleröd (1995); Lipsmeier (1999); Nolan & Whelan (1996)

Household typology

Variable name	*hhtyp*
Variable label	Household type, generated
Source variables	Household information regarding age and relationships between household members
Category / dataset	Household structure / household data
Prepared by	Daniel Gebhardt
Description	A number of variants and suggestions exist regarding the definition of household types (see e.g. Lengerer et al. 2005 for the Mikrozensus household typology, Porst (1984) and Beckmann & Trometer 1991 for the ALLBUS typology and Frick et al. (n.d.) for the SOEP). The household typology used in PASS follows the SOEP version. The decisive criteria of differentiation are existing partnerships, the number and age of children and existing family relationships. Whereas the SOEP typology is merely based on the relationship of the household members to the head of the household, PASS uses information on interrelationships between all household members. In addition, the PASS typology includes the age of the household members as indicated in the household interview and the household size. **Definition of relationships for generating the household type:** • Couples: married couples; registered partnerships; non-married partnerships and partnerships whose status is not further specified (missing value for the follow-up question about the type of partnership). • Child of a person: natural child; stepchild; adopted or foster child; child whose status is not further specified (missing value for the follow-up question about type of relationship to the child). • Parent of a person: natural parent: step-parent; adoptive or foster parent: parent whose status is not further specified (missing value in follow-up question about type of parentship). **Definition of household types:** • One-person household: household consisting of only one person. • Couple without children: household consists of two individuals living together as a couple. • One-parent household: household consists solely of one parent and his/her children. No restrictions are made with respect to the children's ages. • Couple with children under the age of 16: household consists solely of two individuals living as a couple and their respective and/or mutual children. All of the children are under the age of 16. • Couple with children aged 16 or over: household consists solely 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 under the age of 16 and children aged 16 or over: household consists solely of two individuals living as a couple and their respective and/or mutual children. There are both children under the age of 16 and children aged 16 or over living in the household. • Multi-generation household: household consists of members of at least three generations in linear succession. The core of the household is multi-generational, i.e. at least one person in the household is both a child and a parent of another member of the household. The other people living in the household are parents, children, siblings, partners of the central member(s) and partners' siblings. • Other household type: household which could not be assigned to one of the other defined household types. • Type generation not possible (missing values): basically all households with at least one missing value (-1,-2,-4) or implausible value (-8) in the main category of a relationship variable or the age variable. (Exception: for households with three or less members in unambiguous relationship constellations, the household type was also generated even if age details were missing.)
References	Beckmann & Trometer (1991); Frick et al. (n.d.); Lengerer et al. (2005); Porst (1984)

Benefit community ID, wave 2

Variable name	*bgnr2*
Variable label	Benefit community ID in wave 2
Source variables	Household information on age and relationships between household members
Category / dataset	Benefit community (Bedarfsgemeinschaft) / person register
Prepared by	Gerrit Müller
Description	The *bgnr2* variable is created at the individual level. It assigns an identification number to each household member indicating the person's affiliation to a particular benefit community ('Bedarfsgemeinschaft'). Consequently, household members with the same ID constitute together a benefit community. The *bgnr2* variable is composed of the known household number and a two-digit indicator to identify the benefit community within the household. The identification of a household member's affiliation to a benefit community in joint receipt of benefits is based solely on the information on the relationships between the different household members from the household grid table as well as on the members' ages according to the household interview. The benefit communities identified in this way are, therefore, to be regarded as 'synthetic' benefit communities. The identification process does not take into consideration information on actual benefit receipt or on the individual members' ability to work and qualification status. It is more a case of identifying groups of persons in the same household who are or would be regarded as household communities in joint receipt of benefits according to the provisions of the Social Code Book II in the event that they require benefits. This artificial allocation procedure is necessary, since information on the existence of a benefit community and the identification of individuals affiliated to this community cannot be collected directly in the context of an interview. With regard to implementation, the allocation of a person to a benefit community is based on the latest version of the German Social Code Book II, Section 7, sub-section 3 (last amended on 26/03/2007). According to this, each individual aged between 25 and 64 constitutes a separate benefit community, unless this person is living in a partnership and/or has a child/ children aged under 25 who has/have no own partner/children. In the latter case the benefit community comprises the person, his/her partner and the child(ren). If two individuals live in the same household with a joint child, but do not indicate in the household grid table that they are living in a partnership, a partnership is nevertheless assumed to exist in terms of Section 7, sub-section (3a), and the corresponding individuals and their child(ren) are assigned to the same benefit community. Individuals aged between 15 and 25 are in principle assigned to their parents unless they are already living together with a partner (or a child of their own) in a joint household. Individuals aged between 15 and 25 who live without their parents (or partner/ children) constitute a separate benefit community. Persons aged 65 and over are not covered by the Social Code Book II and are therefore not counted as members of a benefit community (code 0) unless they live together with a partner who is aged under 65 (or a child aged under 25) in the same household. Likewise, children under the age of 15 who live in a household without their parents are not counted as members of a benefit community (code 0). They are covered by the provisions of the Social Code Book XII. Allocations to benefit communities were not made for households with missing information on relationships and/or the age of certain household members; instead, all members of these households were assigned code 99. By approximation, such households may be interpreted as households consisting of one benefit community only.
Reference	German Social Code Book II (Sozialgesetzbuch, Zweites Buch - Grundsicherung für Arbeitssuchende (SGB II))

Benefit community typology, wave 2

Variable name	*bgtyp2*
Variable label	Type of benefit community in wave 2
Source variables	Household information on age and relationships between household members
Category / dataset	Benefit community (Bedarfsgemeinschaft) / person register
Prepared by	Gerrit Müller
Description	The benefit community typology is based on the same concept of the synthetic benefit community as was used for variable *bgnr2*. Up to the age of 25, children are counted as members of the benefit community of their parents unless they themselves have a partner or children of their own. This is handled differently from the BA statistics, where typologies are often still established on the basis of majority (18th birthday). As an example: households in which the youngest child is aged between 18 and 24 and which are classified as one-parent benefit communities according to our typology, are counted as single households in the BA statistics. This difference must be borne in mind when comparing PASS data with figures from the official statistics. Code 0 (no benefit community) was assigned to households in which one or more members are not covered by the Social Code Book II (see also code 0 for variable *bgnr2*). Code -5 (generation impossible due to missing values) was allocated to households with missing information on relationships and/or the age of individual household members (see code 99 for *bgnr2*).
Reference	–

Benefit community in receipt of unemployment benefit II as of the sampling date, wave 2

Variable name	*bgbezs2*
Variable label	Benefit community in receipt of UB II as of the sampling date in wave 2 (2006/2007)
Source variables	New sample households: *HH49, HH50, HH52, HH53, HH62, sample, hnr, bgnr2, hhgr* Re-interviewed households: *HH91, HH92, HH93, HH95, sample, hnr, bgnr2, hhgr*
Category / dataset	Benefit community (Bedarfsgemeinschaft) / person register
Prepared by	Mark Trappmann
Description	For each benefit community that was identified in accordance with the procedure described for variable *bgnr2* this variable indicates whether the benefit community was in fact receiving Unemployment Benefit II as of the sampling date of wave 2 or not.
Reference	–

Benefit Community in Receipt of Unemployment Benefit II as of the Survey Date, wave 2

Variable name	*bgbezb2*
Variable label	Benefit community in receipt of UB II as of the survey date in wave 2 (2006/2007)
Source variables	*AL20601, AL20701, zensiert* (*alg2_spells*), *sample, hhgr, bgnr2*
Category / dataset	Benefit community (Bedarfsgemeinschaft) / person register
Prepared by	Daniel Gebhardt
Description	For each benefit community that was identified in accordance with the procedure described for variable *bgnr2* this variable indicates whether the benefit community was in fact receiving Unemployment Benefit II as of the survey date of wave 2 or not.
References	–

Number of benefit communities within the household

Variable name	*anzbg*
Variable label	Number of synthetic benefit communities in the household, generated
Source variables	*bgnr2, hnr*
Category / dataset	Benefit community (Bedarfsgemeinschaft) / household dataset
Prepared by	Daniel Gebhardt
Description	This variable indicates the number of communities in joint receipt of benefits existing in the household. The benefit communities were identified in accordance with the procedure described for the generation of variable *bgnr2*.
References	–

Number of benefit communities in the household actually receiving benefits as of the sampling date

Variable name	*nbgbezug*
Variable label	Number of benefit communities in the household which were receiving benefits as of sampling date
Source variables	*bgbezs2, bgnr2, hnr*
Category / dataset	Benefit community (Bedarfsgemeinschaft) / household dataset
Prepared by	Daniel Gebhardt
Description	This variable indicates the number of benefit communities within the household which were in receipt of benefits in accordance with the Social Code Book II as of the sampling date. The value was calculated by aggregating via the household number the benefit communities within each household which were actually receiving benefits according to the variable *bgbezs2* from the person register.
References	–

8. Data editing

The information gathered in the interviews of the 2nd wave are first available to TNS Infratest in the form of ASCII data. In a first step, in which the IAB was not involved, TNS Infratest created the following standardised datasets from these raw data (see Büngeler et al 2009: 48ff.):

- Household dataset for re-interviewed households
- Household dataset for new sample households and split-off households
- Individual dataset (respondents aged from 15 up to and including 64)
- Gap dataset (information on gaps in the employment biographies of more than three months duration since January 2005)
- Senior citizens' dataset (respondents from the age of 65 onwards)

TNS Infratest conducted a basic check of the operation of the filter questions in these datasets, and questions that were not asked although they should have been were marked with a code. After the datasets had been prepared in this way they were delivered to the IAB. At the IAB the datasets were subjected to the second step of editing comprising further more detailed, formal and content-related checks, and were then prepared as the scientific use file.

In addition to this, TNS Infratest supplied datasets with information from open-ended survey questions (e.g. on the type of occupational activity), a gross dataset and other special datasets which are not obtained directly from the actual survey instruments (e.g. a dataset about the interviewer follow-up).

The data checks subsequently conducted at the IAB can be divided into three steps, which are described in more detail in the following sections. First the household structure of the re-interviewed households was checked and corrected if necessary. If serious problems were found in the structure, the corresponding interviews were removed (see Chapter 8.1 on this issue). This was followed by a detailed check of the filter questions (applying corrections if necessary). First, filter errors were marked and second, specific codes were set for missing values (see Chapter 8.2 on this issue). Af-

ter this, selected items were checked as regards plausibility of the information provided by the respondents. Clearly implausible or contradictory responses were marked as such by a specific missing code. However, such corrections of the data were carried out on a very restrictive basis.

The following table provides an overview of all of the steps conducted in the context of the data editing and their sequence:

Table 23: Overview of the steps involved in editing the data of the 2nd wave of PASS

No.	Step of the procedure
1	Conversion of the datasets supplied by TNS Infratest to STATA format
2	Check of the household structure of re-interviewed households (see Chapter 8.1)
3	Removal of problematic interviews (household and/or individual level) (see Chapter 8.1)
4	Integration of individual dataset and senior citizens' dataset
5	Correction of the household structure of re-interviewed households (see Chapter 8.1)
6	Filter checks at the household level (see Chapter 8.2)
7	Construction of a household grid dataset and plausibility checks on this (see Chapter 8.3)
8	Generation of the synthetic benefit communities (see description of variables, Chapter 7.5)
9	Generation of new control variables on the basis of the household data following filter checks and the household grid dataset after plausibility checks
10	Filter checks at the individual level (see Chapter 8.2)
11	Coding of information from open-ended survey questions (see Chapter 7.1)
12	Plausibility checks of the household and individual-level data (excluding spell data) (see Chapter 8.3)
13	Preparation, plausibility checks and construction of the spell datasets (see Chapter 8.6 as well as Chapters 4.4 and 4.5)
14	Simple variable generations (see Chapter 7.4)
15	Complex variable generations (see Chapter 7.5)
16	Generation of the data structure for the scientific use file (household dataset, individual dataset, register dataset)
17	Anonymisation (see Chapter 8.4)

8.1 Structure checks and interviews removed from the dataset

Before the filter checks were carried out in the 2nd wave a structure check was conducted. Here interviews with and within households which are regarded as not successfully surveyed in the sense of PASS were to be identified and were removed from the datasets for this reason if necessary.[30] In addition, the structure of the re-interviewed households was compared with the structure reported in the previous wave

in order to identify and, if necessary, correct implausible or problematic changes in the household composition and errors in the allocation of the personal interviews to their respective position in the household. For observing the households in the longitudinal section it is essential that the individuals are assigned consistently to their position in the household and that the respondents can be identified clearly across the waves. A definite personal identification number must not be allocated to different individuals in different waves. If the correct household composition was unclear, all of the interviews conducted with this household in the 2nd wave were removed from the dataset. If one of the personal interviews was conducted with the wrong person but without any further problems emerging in the household composition, then just the personal interview was removed.

Different checks were carried out to identify problematic cases:

- By comparing the first names reported in the 1st and 2nd waves, cases were identified in which changes in the household composition had not been recorded correctly. Instead of recording moves into and out of the household in the relevant places in the household interview, it sometimes happened that interviewers re-named household members or changed their age or sex. All cases where a first name had been changed and this could not be put down to the correction of umlauts and where the year of birth reported in the 1st wave differed by more than one year from that reported in the 2nd wave were subjected to a individual case review. Here a decision was made as to whether the change in the data was simply a matter of correcting the first name or whether the interviewer had made an inadmissible change to the household structure. The cases concerned were passed on to TNS Infratest for clarification. A final decision was then made at the IAB as to how to proceed with these cases.

- Furthermore it was checked whether more than one person with the same date of birth was living in the household. In the household context of the two waves it was decided whether these were plausible or implausible cases. The remaining cases then underwent a further check. For this households were identified in which a date of birth was reported in wave 1 and wave 2 by individuals in different posi-

[30] A household is regarded as successfully surveyed if both a full household interview and at least one full personal interview were conducted.

tions in the household structure. Here it seemed reasonable to suspect that the particular personal interview in the 2nd wave was conducted by a different person from that in the 1st wave. In the context of the household and individual-level data of the 1st and 2nd waves, individual case decisions were made regarding the respective household and personal interviews.

- In order to identify households which are regarded as not successfully surveyed in the sense of PASS, the datasets at the household and the individual level were merged. Personal interviews without a full household interview were marked, as were household interviews for which no interview at the individual level was available[31].

Individual case decisions were also made to deal with the cases which proved to be problematic during the structure checks. What was of significance here was how serious the particular problem was considered to be. In cases where the correct household composition in the 2nd wave was unclear, all of the interviews for the 2nd wave were removed. In the 3rd wave these households will be treated as households that did not take part in the 2nd wave. If there was merely a problem in assigning individuals to their respective position in the household, i.e. if it was suspected that a personal interview had been conducted with the wrong person in the 2nd wave, then only the personal or senior citizens' interview concerned was removed from the dataset. If the problem was a structural problem that had no serious consequences and could be solved, for example, by removing a personal interview from the dataset, additional corrections of the first name, age and sex were made at the household level. The incorrect information concerned was then put back to the last valid value from the 1st wave, or in the case of the age, to the value from the previous year + 1 year.

In addition, all interviews with individuals for whose household no full household interview was available 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. The households from the refreshment sample which were regarded as not successfully surveyed were removed

[31] In the case of new sample households households for which a household interview was available but no valid personal interview, the household interviews were removed from the dataset following the procedure used in the 1st wave. In contrast, the household interviews of re-interviewed households and split-off households were retained (N=140).

from the dataset following the procedure used in the 1st wave. In the case of re-interviewed households without interviews at the individual level, however, the household interview was not deleted.

Furthermore, TNS Infratest reported three households whose interviews were not conducted correctly. One of them was a test interview and two were confirmed as falsifications in the CAPI field. In these cases all of the interviews were removed.

The net variables (*hnettok2*, *hnettod2*, *pnettok2*, *pnettod2*) in the household register and person register datasets provide an indication that interviews have been removed. Via the corresponding variables in the household register it is possible to trace the re-interviewed households whose household interviews were removed later. By means of the net variables in the person register it is possible to trace the cases where only single individual-level interviews or all of the interviews of the household were deleted. In the case of households from the refreshment sample of the 2nd wave without a valid household interview but at least one personal interview it is not possible to trace deleted interviews in the register datasets as these households were not included in the datasets.

8.2 Filter checks

During the filter checks, the correct operation of the filter questions in the instruments was checked using a statistics program. If certain questions had been asked although the value of the relevant filter variable would have required something else (for example, if detailed information was requested on vocational training although the respondent had stated that he or she did not have a vocational qualification), these variables were set to the missing code '-3' (not applicable), which they would also have received through correct use of the filters.[32] Moreover, there were a few items not being surveyed although they ought to have been according to the relevant filter variable (e.g. if no further information was recorded on vocational training although the respondent had stated that he or she had undergone such training). In these cases, the specific missing code '-4' (question mistakenly not asked) was allocated.

[32] As is usual in such cases, the filter checks were conducted in the same order as the items in the questionnaire, beginning with the items which were asked first and then moving on to those asked later.

In an additional step of the filter checks, the missing codes allocated by the field institute and the system missings were replaced by standard values for all variables. Table 24 provides an overview of the allocated values. '-1' and '-2' are the standard recoding for the values 'Don't know' and 'Details refused' recorded during the survey. '-3' is the general 'Not applicable' code for questions not asked due to filters. As described above, the code '-4' was allocated if a question was not asked as a result of a filter error. Codes '-5' to '-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 above € 99,999 in the open question on income). These codes were only allocated as required.

Table 24: Overview of the missing codes used

Code	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 No. 1, only allocated as required
-6	Question-specific code No. 2, only allocated as required
-7	Question-specific code No. 3, only allocated as required
-8	"Implausible value"
-9	"Item not administered in wave"
-10	"Item not administered in questionnaire version"

The value '-8' is a specific missing code allocated during the plausibility checks (see Chapter 8.3 on the plausibility checks). The missing code '-9' is required for the first time in the second wave. It is always allocated if a certain item was not surveyed in a specific wave. Due to the dataset being prepared in long format, as was described above, variables that were no longer surveyed in the 2nd wave are given the value '-9' for the observations in that wave. The same is done for observations from the 1st wave. Variables that were surveyed for the first time in the 2nd wave are retroactively coded '-9' for observations of wave 1. The code '-10' can be used to take account of differences between the questionnaire versions, in other words between the standard questionnaire and the senior citizens' questionnaire or between the two versions of the household questionnaire.

8.3 Plausibility checks

For the plausibility checks an extensive list of theoretically possible contradictions in the respondents' statements was checked. For this the list of checks conducted in the 1st wave was adapted and extended. New checks became necessary for example as a result of changes in the survey concept, e.g. in the recording of employment biographies. In addition the household structure was checked for plausibility. Here the test conditions of the 1st wave were revised for the 2nd wave. Furthermore the spell data were also checked for plausibility – in particular with regard to inadmissible overlaps within the individual spell types. Here in principle only the data gathered in the cross-section of the 2nd wave were checked. No checks were carried out on the longitudinal section, in other words comparing the information provided in the 1st wave with that given in the 2nd wave.

In detail, the following steps were carried out:

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 regarded as comparatively minor. The latter applied, for example, if only a small number of cases were affected or if one missing code (e.g. "-3") was simply replaced by another one (e.g. "-8"). Two strategies were used for cleansing implausible statements: either the implausible responses were corrected directly or they were allocated a specific missing code.

   - Implausible responses were only corrected when it was highly probable that the interviewer had entered information incorrectly. An example of this is a statement of a monthly total rent of € 9,998. Here it was assumed in the plausibility check that the five-digit missing code "99998" ("don't know") had been entered incorrectly. This and other similar responses were recoded to the corresponding missing categories.

   - However, it was seldom the case that a value could be recognised as an incorrect entry with sufficient certainty. In most cases it was only possible to establish a contradiction between two statements but not to identify specific incorrect entries or such like that had led to the implausible statement. Therefore,

in these cases no corrections were made and the specific missing value code "-8" was allocated instead. It was decided on an individual basis whether the code was to be allocated to one of the two variables involved in the contradiction or to both of them.

2. Plausibility check of the household structure: this check was carried out on the basis of the information collected in the household interview on the family relationships between the household members, and the information on age, gender and first names. Prior to this check the information on relationships in the household was supplemented by the information on partnerships reported in the personal interview.

   - In order to identify implausible household structures, first the information on relationships was combined with the demographic information about the individual household members. For the households that were identified as implausible during these checks, individual case decisions were made which took into account the 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 were corrected on the basis of additional information on the household context, if it was highly probable that an error had occurred. One example: in the case of two people of the same sex who were both natural parents of a third member of the household the gender was corrected on the basis of the first name. If the first names also indicated that the two people were of the same gender, and if there was no other relevant information available, then the relationship was marked as implausible on the basis of the household structure.

   - In a second step checks were carried out comparing sets of three family relationships with one another for plausibility. An example of a relationship structure that would be classed as implausible in this check is: person A is person B's spouse. Person A is the natural parent of person C. Person C is a sibling of person B. If such a combination or another similarly implausible combination of relationships was identified during the plausibility checks, then here too an attempt was made to make the relationship plausible on the basis of the

household context. In the case described, the relationship data was corrected by person C being coded as a child of person B whose status was not further specified. In the 2nd wave it was possible to correct the content of all of the implausible relationships identified in this way.

3. The spell datasets were also subjected to a number of plausibility checks.

   - At the household level it was the spell dataset of Unemployment Benefit II receipt that underwent plausibility checks. Inadmissible overlaps and datings of spells of Unemployment Benefit II receipt or spells of benefit cuts embedded in these were corrected if necessary or were marked as implausible statements by setting to the value "-8". Changes were only made in the generated date variables (*bmonat; bjahr; emonat; ejahr*) of the spell of Unemployment Benefit II receipt, of the spells of benefit cuts (*alg2kbm; alg2kbj; alg2kem; alg2kej*) and in the censoring indicator for the spell of Unemployment Benefit II receipt (*zensiert*). Information on the season was also only recoded into definite months in the generated date variables[33]. If it was not possible to eliminate implausibilities by correcting the date variable then, on a small scale, spells of Unemployment Benefit II receipt were merged or spells of Unemployment Benefit II receipt or benefit cuts were deleted entirely.

   - At the individual level the spell datasets on employment and unemployment spells, the gap dataset and the spell dataset on participation in employment and training measures were checked for plausibility and corrected if necessary. As with the spell data on Unemployment Benefit II receipt, corrections and recodings were only carried out in the generated date variables. Here, too, seasons were recoded into months, "-8" values were allocated for implausible responses and date information was replaced or rendered plausible. As only the generated date variables were edited, the original information gathered in the survey is available to the user in the date variables *ET0100-ET0400*, *AL0100-AL0400* and *AL0800-AL1100*, *LU0200-LU0500*, thus permitting the user to conduct his/her own checks and corrections.
     In addition it seemed necessary to delete entire spells in some cases. Most of

these deletions can be attributed to faults in the gap module. For example in the gap module further spells were recorded for a person although the entire retrospective period was already covered. Or, as a result of mistakes in operating the gap module, interviewers recorded virtually identical spells more than once instead of using the available correction function. Spells that are completely outside the period surveyed but for which data were nonetheless collected were also deleted.

- In all of the spell datasets the spells were put into the correct chronological order if they had mistakenly been recorded in the incorrect order during the survey.

8.4 Retrospective alterations to the 1st wave

During the data editing process for the scientific use file of the 2nd wave, some changes were also made to the 1st wave of PASS, which had already been delivered. These alterations included corrections of errors that were detected after the completion of the scientific use file of the 1st wave. In addition, owing to a change in the concept in the 2nd wave, a new variable was generated for the 1st wave, too (*migration*). Furthermore, a variable that was recorded for the 1st wave was removed from the individual dataset (*kindu15*) and was incorporated into the household dataset together with two other variables (*kindu4*, *kindu13*). The revision of the weighting is dealt with in more detail in Chapter 9.2. Besides changes to the contents, the weighting datasets (*hweights; pweights*), which were originally in wide format and in which the individual weights were given a wave code, were converted into long format. This makes it easier to merge them with the household and individual datasets, which are also edited in long format. Tables 25 to 27 provide an overview of the alterations made retrospectively to the 1st wave of PASS[34].

[33] The seasons reported were coded into months as follows: Beginning of year/winter → January; spring/Easter → April; middle of year/summer → July; autumn → October; end of year → December

[34] Adjustments to value labels or variable labels are not taken into account here.

Table 25: Overview of retrospective alterations to the 1st wave at the household level

Altered variable	Dataset concerned	Altered wave	Type of alteration	Description of the alteration
hhincome	*HHENDDAT*	1	Correction	This variable integrates the response to the open-ended survey question about the net household income (*HEK0600*) and the categorical follow-up questions (*HEK0700* to *HEK1100*). For the individual categories in the follow-up questions mean values or the median are added to the generated variable. These values to be added were previously determined on the basis of the variable to which mean values or medians had already been added in preceding steps. The means and medians are now determined on the basis of the unchanged response to the open question on income and rounded. In addition the top-coding of the net household income is now taken into account when generating the variable.
hintmod (hintmod1)	*HHENDDAT (hh_register)*	1	Concept	In the 1st wave it was not possible to implement the Russian version with the software used for other languages, so it was reprogrammed. The cases concerned were previously labelled as “CAWI” in the interview mode. In actual fact these were CATI interviews conducted in Russian that were simply implemented using different software. The cases concerned were recoded to “CATI” and can be identified via the Russian interview language.
kindu4	*HHENDDAT*	1	Added	This variable is relevant for controlling the household dataset but was previously not included in the household dataset. It was added retrospectively for the 1st wave, too.
kindu13	*HHENDDAT*	1	Added	See *kindu4*.
kindu15	*HHENDDAT*	1	Added	See *kindu4*.
oecdinca	*HHENDDAT*	1	Correction	Variable is based on *hhincome* and is indirectly affected by the correction of *hhincome*.
oecdincn	*HHENDDAT*	1	Correction	Variable is based on *hhincome* and is indirectly affected by the correction of *hhincome*.

Table 26: Overview of retrospective alterations to the 1st wave at the individual level

Altered variable	Dataset concerned	Altered wave	Type of alteration	Description of the alteration
brutto	*PENDDAT*	1	Correction	This variable integrates the response to the open question about the gross income (*1st wave: PEK0100; 2nd wave: PEK0100b*) and the categorical follow-up questions (*PEK0200* to *PEK0600*). For the individual categories in the follow-up questions mean values or the median are added to the generated variable. These values to be added were previously determined on the basis of the variable to which mean values or medians had already been added in preceding steps. The means and medians are now determined on the basis of the unchanged response to the open question on income and rounded. In addition the top-coding of the gross income is now taken into account when generating the variable. A second alteration concerns the way in which missing values in the categorical follow-up questions are dealt with. If there are missing values in the follow-up questions, *brutto* was set to "-5". In contrast, missing values from the open-ended survey questions are still taken into the generated variable directly.
pintmod	*PENDDAT*	1	Concept	In the 1st wave it was not possible to implement the Russian version with the software used for other languages, so it was reprogrammed. The cases concerned were previously labelled as "CAWI" in the interview mode. In actual fact these were CATI interviews conducted in Russian that were simply implemented using different software. The cases concerned were recoded to "CATI" and can be identified via the Russian interview language.
kindu15	*PENDDAT*	1	Removed	The control variables *kindu4*, *kindu13* and *kindu15* were not included in the household dataset. One of these variables (*kindu15*) was contained in the individual dataset. The variable *kindu15,* which was not relevant for controlling the individual dataset, was removed as it can cause problems when merged with the household dataset into which control variables that had previously been missing were incorporated.

Altered variable	Dataset concerned	Altered wave	Type of alteration	Description of the alteration
kindzges	*PENDDAT*	1	Correction	Cases with the response “details refused” to the filter question on whether there are children living outside the household are filtered via the question about the number and are given a “not applicable” code in the editing process. These “not applicable” values were passed on into the generated variable *kindzges* although these are actually cases where a variable can not be generated owing to missing values (-5). These cases are now coded to “-5”.
migration	*PENDDAT*	1	Concept	For the 2nd wave the concept for generating this variable was revised. Missing codes are now taken into account with greater differentiation when creating the variable. In addition, when generating the variable information is now also included about whether the parents and grandparents were born in Germany or abroad (*PMI0800a* to *PMI0800f*).
netto	*PENDDAT*	1	Correction	This variable integrates the response to the open-ended survey question about the net income (*1st wave: PEK0700; 2nd wave: PEK0700b*) and the categorical follow-up questions (*PEK0800* to *PEK1200*). For the individual categories in the follow-up questions mean values or the median are added to the generated variable. These values to be added were previously determined on the basis of the variable to which mean values or medians had already been added in preceding steps. The means and medians are now determined on the basis of the unchanged response to the open question on income and rounded. In addition the top-coding of the net income is now taken into account when generating the variable. A second alteration concerns the way in which missing values in the categorical follow-up questions are dealt with. If there are missing values in the follow-up questions, *netto* was set to “-5”. In contrast, missing values from the open-ended survey questions are still taken into the generated variable directly.

Altered variable	Dataset concerned	Altered wave	Type of alteration	Description of the alteration
nettokat	*PENDDAT*	1	Correction	In wave 1 an error in the generation of *nettokat* occurred if the response to the categorical follow-up question P62 on net income was “don’t know” or “details refused”. So far no value was allocated for these cases in *nettokat* – the variable had the value -21, which can not be interpreted as regards contents. These cases are now set correctly to *nettokat*=11 (1000 or more, no further details) For the categories in *nettokat* values are imputated in *netto*. The correction of *nettokat* also results in changes to *netto*.

Table 27: Overview of retrospective alterations to the weighting datasets (*hweights*; *pweights*)

Altered variable	Dataset concerned	Altered wave	Type of alteration	Description of the alteration
wqmihh	*hweights*	1	Correction	See Chapter 9.2 on the revision of the weighting of the 1st wave
wqhh	*hweights*	1	Correction	See Chapter 9.2 on the revision of the weighting of the 1st wave
wqmip	*pweights*	1	Correction	See Chapter 9.2 on the revision of the weighting of the 1st wave
wqp	*pweights*	1	Correction	See Chapter 9.2 on the revision of the weighting of the 1st wave

8.5 Anonymisation

All data gathered by the IAB as a 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 a simplified form. These variables are generally identified as “anonymisiert” in the variable label. For the same reason it was also necessary to exclude available regional information, with the exception of the German federal states and information on East/West Germany derived from these. For data protection reasons neither the data on 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 on the household type and benefit community type (*hhtyp*[35], *bgtyp*[36]), indicator variables on partners in the household (*apartner; epartner*[37]), pointer variables for parents and partners in the household (*zmhh; zvhh; zparthh*[38]) and various indicator variables which show whether parents (*mhh; vhh*[39]) or children of the target person (e. g. *ekind*[40]) are living in the household.

[35] Contained in the household dataset (HHENDDAT), see Chapter 7.5
[36] Wave-specific variables contained in the person register (p_register), see Chapter 7.5
[37] Contained in the individual dataset (PENDDAT), see Chapter 7.4
[38] Wave-specific variables contained in the person register (p_register), see Chapter 7.5
[39] Contained in the individual dataset (PENDDAT), see Chapter 7.4
[40] Contained in the individual dataset (PENDDAT), see Chapter 7.4

Table 28 provides an overview of the variables concerned and the steps[41] taken to render them anonymous in the individual dataset. Table 29 shows the anonymised variables of the dataset on employment spells.

[41] If the use of the non-anonymised versions of one or more of these variables is indispensible for your research, please contact the Research Data Centre (Forschungsdatenzentrum) in order to find a suitable data access option. What form this may take will depend on the research project and the variables required.

Table 28: Overview of the anonymised variables in the individual dataset (*PENDDAT*)

Variable name	Question number		Variable label	Procedure
	Standard quest.	Sen. cit's quest.		
PD0100	P1	P1	Year of birth (date of birth, anonymised)	The precise date of birth was shortened to year of birth.
gebhalbj	generated	generated	1st/2nd half-year of birth, generated	The precise date of birth was shortened to an indicator for the 1st or 2nd half of the year.
PET1210	P84	not in q. version	Last occup. status, simple classification (before January 2005)(anon.)	For technical reasons, professional and regular soldiers were recorded separately in the survey. Due to the small case numbers and as this group is not usually asked about occupational status anyway, this group was merged with that of civil servants and judges.
PET1250	P87, P88	not in q. version	Last occup. status civil servant: detailed information, incl. soldiers (before January 2005)(anon.)	This variable contains additional cases. The professional and regular soldiers from P87 were added to the corresponding civil servant category. The variable for professional and regular soldiers (P87) is not supplied.
PET1211	generated	not in q. version	Last occup. status, simple classification (incl. spell info.) (anon.), gen.	Procedure as for *PET1210*.
PET1251	generated	not in q. version	Last occup. status civil servant: detailed information, incl. soldiers (incl. spell info.)(anon.), gen.	Procedure as for *PET1250*. The variable for professional and regular soldiers (P87) is not supplied.
stiblewt	generated	not in q. version	Occup. status, last job, code number, generated	When generating the occupational status variable, professional and regular soldiers are assigned to the corresponding civil servant category.
PET1510	generated	P12	Current occup. status, simple classification, surveyed from wave 2 onwards (anon.)	Procedure as for *PET1210*.
PET1900	generated	P15, P16	Current occup. status civil servant: detailed information, incl. soldiers (anon.)	Procedure as for *PET1250*. The variable for professional and regular soldiers surveyed in the senior citizens' interviews (P15) is not supplied. As regards the personal interviews no generated variable for prof. and regular soldiers is incorporated into the individual dataset from the employment spells (P47).

Variable name	Question number		Variable label	Procedure
	Standard quest.	Sen. cit's quest.		
stibkz	generated	generated	Current occup. status, simple classification, harmonised, (anon.)	When generating the occupational status variable, professional and regular soldiers are assigned to the corresponding civil servant category.
stib	generated	generated	Occup. status, code number, generated	Procedure as for *stiblewt*.
PET3300	P93	not in q. version	First occup. status, simple classification (anon.)	Procedure as for *PET1210*.
PET3700	P96, P97	not in q. version	First occup. status civil servant: detailed info., incl. soldiers	Procedure as for *PET1250*. The variable for professional and regular soldiers (P96) is not supplied.
PET3301	generated	not in q. version	First occup. status, simple classification, (merged, incl. spell info) (anon.), generated	Procedure as for *PET1210*.
PET3701	generated	not in q. version	First occup. status civil servant: detailed info., incl. soldiers, (merged, incl. spell info.) (anon.), generated.	Procedure as for *PET1250*. The variable for professional and regular soldiers (P96) is not supplied.
stibeewt	generated	not in q. version	Occup. status, first job, code number, generated	Procedure as for *stiblewt*.
PSH0320	P281	not in q. version	Mother's occup. status when respondent was 15, simple classification (anon.)	Procedure as for *PET1210*.
PSH0360	P284, P285	not in q. version	Mother's occup. status when respondent was 15: civil servant, incl. soldiers: detailed info (anon.)	Procedure as for *PET1250*. The variable for professional and regular soldiers (P284) is not supplied.
mstib	generated	not in q. version	Mother's occup. status, code number, generated	Procedure as for *stiblewt*.
PSH0620	P292	not in q. version	Father's occup. status when respondent was 15, simple classification (anon.)	Procedure as for *PET1210.*
PSH0660	P295, P296	not in q. version	Father's occup. status when respondent was 15: civil servant, incl. soldiers: detailed info (anon.)	Procedure as for *PET1250*. The variable for professional and regular soldiers (P284) is not supplied.
vstib	generated	not in q. version	Father's occup. status, code number, generated	Procedure as for *stiblewt*.

Variable name	Question number		Variable label	Procedure
	Standard quest.	Sen. cit's quest.		
PMI0200	P264	P73	Not born in Germany: country of birth	Countries with low case numbers were grouped into larger categories.
ogebland	generated	generated	Country of birth, incl. info. from open-ended questions, categories (anonymised)	Procedure as for *PMI0200*.
PMI0500	P267	P76	Nationality not German: which nationality? (anonymised)	Nationalities of countries with low case numbers were grouped into larger categories.
ostaatan	generated	generated	Nationality, incl. info. from open-ended questions, categories (anon.)	Procedure as for *PMI0500*.
PMI1000a	P274a	P80a	Father: country of residence before migration (anonymised)	Countries of residence before migration with low case numbers were grouped into larger categories.
PMI1000b	P274b	P80b	Mother: country of residence before migration (anonymised)	Procedure as for *PMI1000a*.
PMI1000c	P274c	P80c	Father's father: country of residence before migration (anonymised)	Procedure as for *PMI1000a*
PMI1000d	P274d	P80d	Father's mother: country of residence before migration (anonymised)	Procedure as for *PMI1000a*
PMI1000e	P274e	P80e	Mother's father: country of residence before migration (anonymised)	Procedure as for *PMI1000a*
PMI1000f	P274f	P80f	Mother's mother: country of residence before migration (anonymised)	Procedure as for *PMI1000a*
ozulanda	generated	generated	Father: country of residence before migration, incl. info. from open-ended questions, categories (anonymised)	Procedure as for *PMI1000a*
ozulandb	generated	generated	Mother: country of residence before migration, incl. info. from open-ended questions, categories (anonymised)	Procedure as for *PMI1000a*
ozulandc	generated	generated	Father's father: country of residence before migration, incl. info. from open-ended questions, categories (anonymised)	Procedure as for *PMI1000a*

Variable name	Question number		Variable label	Procedure
	Standard quest.	Sen. cit's quest.		
ozulandd	generated	generated	Father's mother: country of residence before migration, incl. info. from open-ended questions, categories (anonymised)	Procedure as for *PMI1000a*
ozulande	generated	generated	Mother's father: country of residence before migration, incl. info. from open-ended questions, categories (anonymised)	Procedure as for *PMI1000a*
ozulandf	generated	generated	Mother's mother: country of residence before migration, incl. info. from open-ended questions, categories (anonymised)	Procedure as for *PMI1000a*

Table 29: Overview of the anonymised variables in the employment spell dataset (*et_spells*)

Variable name	Question number		Variable label	Procedure
	Standard quest.	Sen. cit's quest.		
ET0600	P44		Occup. status, simple classification (anon.)	Procedure as for *PET1210.*
ET1000	P47, P48		Occ. status civil servant: detailed info. (anon.)	Procedure as for *PET1250*. The variable for professional and regular soldiers (P47) is not supplied.
stib	generated		Occ. status, code number, generated	Procedure as for *stiblewt*.

8.6 Receipt of Unemployment Benefit II

Receipt of Unemployment Benefit II at the household level was already recorded in spell form in the 1st wave. This concept was continued in wave 2 but with a slightly revised set of questions. In addition to the adjustments necessary for updating the spells that were still ongoing at the time of the interview for the 1st wave, the contents were expanded and minor corrections were made in response to the experiences made in the 1st wave.

Concept for updating the spells of Unemployment Benefit II receipt that were still ongoing in wave 1

In order to update the spells of Unemployment Benefit II receipt which were still ongoing in wave 1 and were therefore right-censored in the spell dataset, dependent interviewing questions were added to both versions of the household questionnaire (HH91 in the household questionnaire for re-interviewed households; HH48 in the household questionnaire for split-off households and new sample households). In cases where the household interviewed in the previous wave had split up, the censored spells of Unemployment Benefit II receipt were updated via the part of the household in which the person with whom the household interview was conducted in the previous wave is living (termed hereafter as "HRP"[42] for short).

If the HRP is a member of the household which is first reached at the old address / under the old telephone number, the spell is updated via the responses in the household questionnaire for re-interviewed households ("HHalt" for short). The procedure is different, however, if the part of the household reached at the old address / under the old telephone number gives the information that the HRP has moved out / has not been present for a year or longer / or remained at the place of residence of the previous wave. In these cases the part of the household that split off from the original household is regarded as a separate survey household and is interviewed using the questionnaire for new sample households ("HHneu" for short). If the HRP of the previous wave belongs to this split-off household, the spell of Unemployment Benefit II receipt of the original household that was still ongoing in wave 1 is updated using the details provided by the HRP in the split-off part of the household. Due to a control problem in wave 2, however, no split-off households were interviewed where a person who had moved out / had not been present for a year or longer / had remained in the place of residence of the household of the previous wave had been the HRP of a household in the 1st wave. Updates of censored spells of Unemployment Benefit II receipt were therefore only recorded in household interviews for re-interviewed households.

There are also differences between re-interviewed and split-off households with regard to the period for which information is collected about receipt of Unemployment Benefit II in the 2nd wave. Here, too, it is of importance whether the HRP of the previous wave is living in the household. If the HRP of wave 1 is living in the household, then spells of Unemployment Benefit II receipt since the interview date of the previous wave are

[42] HRP stands for "household reference person".

recorded. If the HRP is not living in the household, then only spells of Unemployment Benefit II receipt since the date when the HRP moved out or the date when the respondent moved out of the joint household with the HRP are recorded.

The households of the refreshment sample which were interviewed for the first time in wave 2 were asked about their receipt of Unemployment Benefit II during the period since the last change in the household composition. If this was before January 2006 or if no information was provided about changes in the household, then the household's receipt of Unemployment Benefit II from January 2006 onwards was recorded.

Structure of the spell dataset on Unemployment Benefit II

The structure and the contents of the spell dataset on Unemployment Benefit II change due to the integration of the spells of Unemployment Benefit II receipt reported in wave 2. Here it is necessary to distinguish between (1) new variables that refer to a certain wave, (2) new variables that do not refer to a particular wave, (3) variables that are no longer surveyed in wave 2 and (4) variables that are no longer included in the dataset.

1. The variables *AL20600*, *AL20700a* to *AL20700o*, *AL20800* and *AL20900* are seen as wave-specific cross-sectional information referring to wave 1. For the cross-sectional information gathered in wave 2, the new variables *AL20601*, *AL20701*a to *AL20701o*, *AL20801* and *AL20901* were added to the spell dataset of Unemployment Benefit II in the same way.

2. In wave 2 additional information was collected within the spells of benefit cuts. With variables *AL21850a* to *AL21850e* a new reason for benefit cuts was surveyed in wave 2 to which responses to open-ended questions are also coded in *AL21851a* to *AL21851e.* In addition, for all of the spells of benefit cuts reported in wave 2, information was gathered in *AL22050a* to *AL22050e* about which household member's Unemployment Benefit II was cut and the corresponding follow-up validation question in *AL22070a* to *AL22070e* was integrated into the dataset. Furthermore, for all of the benefit cuts contained in the Unemployment Benefit II spell dataset, generated variables were created for the start date (*alg2kbma* to *alg2kbme* and *alg2kbja* to *alg2kbje*) and for the end date (*alg2kema* to *alg2keme* and *alg2keja* to *alg2keje*).

3. The reason for benefit cuts *AL21900a* to *AL21900e* was no longer surveyed in wave 2. Accordingly no responses to open-ended questions were coded to the variables *AL21901a* to *AL21901e* any longer.

4. The variables *AL22200a* to *AL22200e* and *AL22300* are no longer included in the spell dataset from wave 2 onwards. These variables concern the follow-up question about whether there is another benefit cut or another spell of Unemployment Benefit II receipt. As both the spells of benefit cuts and the spells of Unemployment Benefit II receipt are sorted into chronological order during the editing process, it was decided not to supply these variables any longer.

Plausibility checks and corrections in the Unemployment Benefit II spell dataset

As was done in wave 1, the information on receipt of Unemployment Benefit II was also subjected to a number of plausibility checks in wave 2. Inadmissible overlaps and datings of spells of Unemployment Benefit II receipt or of benefit cuts were corrected if necessary. In principle changes were only made to the generated date variables (*bmonat; bjahr; emonat; ejahr*) of the spell of Unemployment Benefit II receipt, the spells of benefit cuts (*alg2kbm; alg2kbj; alg2kem; alg2kej*) and in the censoring indicator of the spell of Unemployment Benefit II receipt (*zensiert*). If it was not possible to remove implausibilities by correcting the date variables, then in a small number of cases spells of Unemployment Benefit II receipt were merged or spells of Unemployment Benefit II receipt or benefit cuts were deleted entirely.

Updating the spell dataset on Unemployment Benefit II receipt

After the spells of Unemployment Benefit II receipt that were reported in wave 2 had been converted into spell format and following the plausibility checks and corrections, where inadmissible overlaps and spells with implausible dates were corrected, the spells of Unemployment Benefit II receipt which were still ongoing at the time of the interview in the 1st wave were updated using the information gathered in wave 2. Four variants are to be distinguished here. In the first two, (1) and (2), only the censoring indicator *zensiert* is changed. The third variant (3) is an update of the spell which was censored in wave 1 using information gathered in wave 2 in the narrow sense. Here the censoring indicator is integrated into the spell of Unemployment Benefit II receipt which

was still ongoing in wave 1, as are the generated and surveyed end dates, the wave-specific cross-sectional information (see above) and information about new spells of benefit cuts. In addition to updating spells which were censored in wave 1, new spells that were reported in wave 2 are added to the spell dataset (4). These four variants are outlined briefly below:

1. *Cases in which the HRP of the previous wave no longer lives in the household.* As already mentioned, due to a control problem in wave 2 no split-off households were interviewed in which a person was living who had been the HRP of the original household in the previous wave. As a consequence in these cases no information was gathered about the development of a spell of Unemployment Benefit II receipt which was still ongoing in the original household in wave 1. In order to prevent the censored Unemployment Benefit II spells of the original household that were recorded in the first wave continuing to be evaluated as current benefit receipt of this household, the censoring indicator *zensiert* was set to "-5" (HRP of the previous wave not in the household and not interviewed) in these cases. The indicator *zensiert* was also set to "-5" in cases where the HRP of the 1st wave had died.

2. *Cases in which the household in wave 2 contradicts an ongoing spell of Unemployment Benefit II receipt as of the interview date in wave 1.*
   If the household has contradicted the information that there was an ongoing spell of Unemployment Benefit II receipt at the time of the previous wave, either explicitly or implicitly (by reporting an end date that preceded the interview date in wave 1) in the update question (HH91 in HHalt; HH48 in HHneu), then *zensiert* was set to "2" (no). The information provided in the interview of the previous wave is presumed to have been correct. As it is not possible to make any reliable statements about the continued duration of the benefit receipt beyond the date of the interview in wave 1, it is assumed that the benefit receipt ended in the month of the interview in wave 1. The generated end date of the spell of Unemployment Benefit II receipt (*emonat; ejahr*) was already set to the interview date of the 1st wave in the previous wave.

3. *Cases in which the household reports the end date of a spell of benefit receipt that was still ongoing in the previous wave.*
   If information about the end date of a spell of Unemployment Benefit II receipt that was censored in wave 1 is available in wave 2, then the spell which was censored

in wave 1 was updated using the current information. First the surveyed end date (*AL20300*; *AL20400*), the generated end date (*emonat; ejahr*), the follow-up question as to whether the receipt of Unemployment Benefit II is still ongoing (*AL20500*) and the censoring indicator (*zensiert*) are overwritten with the information gathered in wave 2. Furthermore the spells of benefit cuts reported in the 2nd wave and the cross-sectional data referring to wave 2 (*AL20601; AL20701a to AL20701o, AL20801, AL20901*) were included.

4. *Spells of Unemployment Benefit II receipt reported for the first time in wave 2 which do not update any spells that were censored in wave 1.*
   Spells reported for the first time in wave 2 were added to the Unemployment Benefit II spell dataset. Then the spell counter was generated again in order to create a variable without gaps *spellnr*.

9. Weighting

9.1 The weighting process of the 1st wave

The weighting process used in the first wave was based on a three-stage weighting concept:

- In the first stage, design weights were produced for the gross sample used.
- Subsequently, non-response was modelled in the second stage.
- Finally, in the third stage the weights were calibrated.

9.1.1 Stage 1: design weighting

The design weights are reciprocal selection probabilities for the gross sample used. The procedure used to generate the weights is described in detail in Rudolph and Trappmann (2007). The design weights are contained in the dataset *hweights*. The individual design weights supplied are:

dw_ba Design weight of a household in the BA sample (population: households in which there was at least one benefit community in joint receipt of benefits in accordance with Social Code Book II in July 2006)

dw_mi Design weight of a household in the Microm sample (population: households in the Federal Republic of Germany)

dw Design weight of a household in the total sample (population: households in the Federal Republic of Germany)

9.1.2 Stage 2: modelling of non-response

With the aid of two logit models, the participation probability was estimated for all households in the gross sample used. The first logit model explains the probability of a contact. The second logit model explains the participation (at least the household interview and one complete personal interview) in the case of successful contact. These logit models were calculated separately for the two subsamples, as the contact proce-

dures were based on two different processes. Only the micro-geographical variables supplied by Microm were used for modelling the Microm sample. In the case of the models for the BA sample, additional characteristics from the sampling frame (A2LL or XSozial) could be used. The models applied contain only variables with significant effects (likelihood ratio test, two-sided, 10% level). The tables below show the variables used and the coefficients of the models. A detailed description of the non-response modelling is also contained in Chapter 5.2 of the TNS Infratest method and field report (Hartmann et al. 2008).

The dataset *hweights* contains the variable *prop_t0*. This contains the product of the predicted probabilities of the two models.

Dividing the design weights by the estimated participation probabilities yields the modified design weights, which formed the starting point for the third stage – calibration.

9.1.3 Stage 3: calibration

A detailed documentation of the calibration process can be found in Kiesl (2009). We therefore merely outline the basic procedure here.

Household level

In an initial step, the two subsamples and the total sample were calibrated to official statistics at the household level.

The total and BA weights for benefit recipients in the two samples were calibrated to benchmark statistics from the Federal Employment Agency (reporting month July 2006). The total and Microm weights were additionally calibrated to benchmark statistics on private households in Germany for 2007 from the Federal Statistical Office. The benchmark figures used are detailed in Kiesl (2008).

All weights are household weights. The BA statistics, however, are based on values at the level of benefit communities. The link is created using the synthetic benefit communities, generated as described in Chapter 7.5 (variable *bgnr1* in the *p_register* dataset). Households are initially broken down into synthetic benefit communities. The characteristics used for the calibration process are then generated at the benefit community level. This also includes the characteristic of whether the benefit community was receiv-

ing Unemployment Benefit II as of the sampling date. After calibration, multiplying the characteristics of all benefit communities in receipt of benefits as of the sampling date by the projection factors for households yields the benchmark figures. Separate benefit communities in receipt of benefits within one household are therefore always given the same projection factors.

It is not always possible to determine accurately the benefit receipt of a household or even of a benefit community. As much data as possible is therefore provided in order to enable users to make independent decisions. Thus, for instance, the variable *alg2samp* at the household level is supplied in the *hh_register* dataset. This variable contains the benefit receipt as of the sampling date for all households in the categories: *0 no receipt, 1 receipt, 2 no receipt according to survey (but included in BA sample and thus receipt according to register data), 3 receipt unclear from survey (but included in BA sample and thus receipt according to register data), 4 receipt unclear from survey (Microm sample)*. In addition, every user can generate this variable him/herself using the unemployment benefit spell data (*alg2_spells* dataset). Other useful variables are *AL20600* and *AL20700a-o* (for which members does the household receive benefits?) and the variable *HA0400* from *HHENDDAT*, which, for households founded after July 2006, records whether at least one household member received benefits in July 2006. The variable *sample* in *hweights* indicates the sample from which each household originates.

To generate the weights, however, a clear decision is necessary on which benefit communities should be regarded as being in receipt of Unemployment Benefit II on the sampling date. The decisions upon which the weighting is based can be explained as follows:

At the household level it was decided that:

1. All households from the BA sample (*sample=1*) were in receipt of benefits as of the sampling date even if they denied this, provided at least one person aged between 15 and 64 lives in the household.

2. Households from the Microm sample for which benefit receipt can not be clearly established on the basis of the survey data are regarded as households receiving Unemployment Benefit II for the purpose of weighting if they report ever having re-

ceived Unemployment Benefit II (*HA0300=1*) and if the start or end date of at least one observation lies in 2006 (in cases of an undetermined end or start).

Transferring from the household to the benefit community level is wrought with even greater uncertainty. The reason for this is that it is not possible to obtain reliable retrospective information on which parts of the household received benefits in July 2006. In most cases, the entire household consists of only one benefit unit, making the question redundant, as the benefit unit receives benefits precisely when the household does so. In cases where the household consists of more than one benefit unit, the following approach was selected:

The information as to which individuals the household is currently receiving benefits for (*AL20600* and *AL20700a-o*) was used. A benefit unit was regarded as receiving benefits if at least one of its members was reported as a benefit recipient. In a household with more than one benefit unit and with no information as to which individuals the household is receiving benefits for (e.g. because the questionnaire responses state that no benefits are being claimed), all of the synthetic benefit communities were regarded as being in receipt of benefits. The result of this generation is contained in the variable *bgbezs1* in the *p_register* dataset.

The weights following calibration at the household level are also contained in the *hweights* dataset.

wqbahh	calibrated household weight of the BA sample
wqmihh	calibrated household weight of the Microm sample
wqhh	calibrated household weight of the total sample

Individual level

Following the calibration at the household level, the individuals who gave a personal or senior citizen's interview were calibrated to benchmark statistics at the individual level. The calibrated household weights were the starting point for this step.

The total and BA weights for benefit recipients in both subsamples were calibrated to benchmark statistics from the Federal Employment Agency (reporting month July 2006). The total and Microm weights were additionally calibrated to benchmark statis-

tics from the Federal Statistical Office on private households in Germany for 2007. The benchmark figures used are detailed in Kiesl (2008).

Senior citizens' interviews were calibrated to population statistics in the same way as the standard personal interviews. The BA statistics, however, do not contain figures on the number of senior citizens in households receiving benefits. Nor do they identify individuals living in households receiving benefits who are not part of a benefit unit. It was therefore impossible to obtain the BA person weights for these individuals by means of calibration. The participation probability of these individuals, given that their household takes part in the survey, was estimated using a logit model with the following covariates: number of individuals aged 15 and over in the household; interview mode; age and gender. The modified design weight was subsequently divided by this value.

The calibrated person weights are contained in the *pweights* dataset.

wqbap	calibrated person weight of the BA sample
wqmip	calibrated person weight of the Microm sample
wqgesp	calibrated person weight of the total sample

9.2 Revision of the weighting used in wave 1

Between the first and the second wave it became clear that the benchmark statistics on nationalities supplied by the Federal Statistical Office in wave 1 were incorrect. A new delivery was requested. The old and the new benchmark figures can be found in Kiesl (2009). The calibration procedure described above was repeated for the Microm sample and the total sample at the household and individual levels using the new benchmark statistics. This procedure results in changes in the weighting variables *wqhh*, *wqmihh*, *wqp* and *wqmip* of the first wave. This has systematic consequences in particular on the projected figures for foreigners and people with a migration background. However, for other analyses, too, the results will no longer correspond exactly with those of the previous year, as a change in a marginal distribution (in this case nationality) in the GREG has effects on all of the projection factors.

9.3 Datasets and variables

Like the individual and household datasets, the weighting datasets *hweights* (household weights) and *pweights* (person weights) are organised as long files from the 2nd wave onwards. This makes it easier to merge the datasets. For this reason some of the variables were given different names from those used in the 1st wave: the wave indicator at the end of the variable was dropped from all of the variables. In addition the two participation propensities for households in the sampling year (in wave 1: *pr_ba_1* and *pr_mi_1*) were combined into one variable *prop_t0*, which indicates the household's participation propensity in the respective sampling year. Researchers who do not wish to use the calibrated cross-sectional weights can still work with the quotient of the design weight and this variable, which is now defined for all households (apart from split-off households) in the year in which they were sampled.

A new feature in the datasets is the reciprocal re-participation probability *hpbleib* and *ppbleib*. They are used for the longitudinal weighting of the households and individuals respectively. The longitudinal weight for a household (or an individual) for the longitudinal section from wave 1 to wave 2 is obtained by multiplying the cross-sectional weight of the household (or the individual) for wave 1 by the reciprocal re-participation probability. The latter is provided in the dataset for all households (or individuals) that took part in waves 1 and 2. All cross-sectional weights continue to have the format of projection factors. i.e. within a wave the weights add up to the population size (not the sample size).

The file *hweights* therefore now contains the following variables:

Table 30: Overview of the variables in the household weights data file (*hweights*)

Name	Label	Remarks
hnr	Household number (current)	Used together with *welle* for linking the datasets
welle	Indicator for survey wave	Used together with *hnr* for linking the datasets
sample	Subsample	Indicates whether BA or Microm weights are used
dw_mi	Design weight - Microm sample	Is the selection probability (during sampling) in the respective subsample (gross)
dw_ba	Design weight - BA sample	Is the selection probability (during sampling) in the respective subsample (gross)
dw	Design weight - total sample	Is the selection probability (during sampling) in the total sample (gross)
prop_t0	Participation probability in the sampling year of the subsample	Is the probability of the household taking part in the year when the subsample was drawn, as predicted by means of a logit model
wqhh	Projection factor - household (total)	Projection factor for the cross-section of the respective wave (total)
wqmihh	Projection factor - household (Microm)	Projection factor for the cross-section of the respective wave (Microm)
wqbahh	Projection factor - household (BA)	Projection factor for the cross-section of the respective wave (BA)
hpbleib	Reciprocal re-participation probability - household (w->w+1)	Reciprocal value of the probability of the household participating in the survey again in the following wave, as predicted by means of a logit model

The file *pweights* contains the following variables

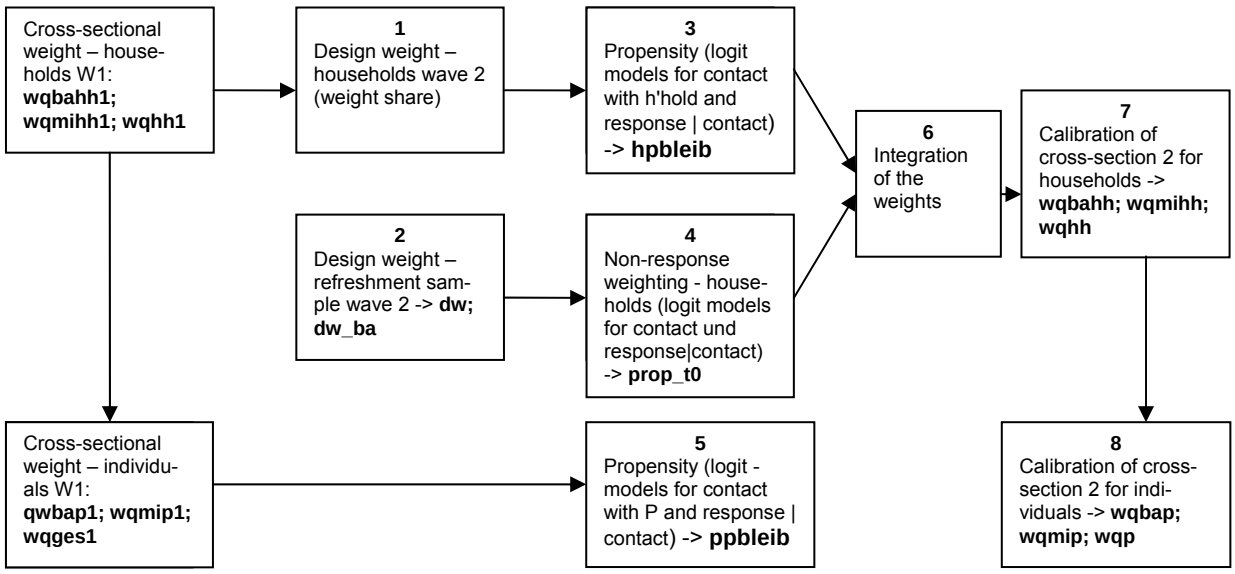
Table 31: Overview of the variables in the person weights data file (*pweights*)

Name	Label	Remarks
pnr	Unchanging personal ID number	Used together with *welle* for linking the datasets
welle	Indicator for survey wave	Used together with *pnr* for linking the datasets
sample	Subsample	Indicates whether BA or Microm weights are used
wqp	Projection factor - person (total)	Projection factor for the cross-section of the respective wave (total)
wqmihh	Projection factor - person (Microm)	Projection factor for the cross-section of the respective wave (Microm)
wqbahh	Projection factor - person (BA)	Projection factor for the cross-section of the respective wave (BA)
ppbleib	Reciprocal re-participation probability - person (w->w+1))	Reciprocal value of the probability of the individual participating in the survey again in the following wave, as predicted by means of a logit model

9.4 Construction of the weights from wave 2 onwards

The starting point for the weighting procedure for the second wave and for the longitudinal section from wave 1 to wave 2 is the cross-sectional weights from wave 1 for households and individuals. In wave 1 each household had two weights, *wqhh1* (calibrated total weight) and – depending on the sample – *wqbahh1* (calibrated BA weight) or *wqmihh1* (calibrated Microm weight), and each individual also had two weights, *wqgesp1* and – depending on the sample - *wqbap1* (calibrated BA weight) or *wqmip1* (calibrated Microm weight). All four weights are updated. Figure 6 shows the steps of the weighting procedure, which are explained below.

Figure 6: Generation of the weights for the 2nd wave



9.4.1 Design weights for the wave 1 households in the 2nd wave

New "household design weights" were generated for the 2nd wave from the cross-sectional weights for households of the first wave, taking into account people moving into households from within Germany. This is done using a weight share procedure. Births, deaths or moves out of households have no influence on the weight; moves into households from within Germany, on the other hand, increase the inclusion probability of a household as the individuals who have moved into the household also had the chance of being included in the sample in wave 1. For the preliminary weighting, if individuals had moved into the household from within Germany, the previous inclusion

probability was increased by the mean inclusion probability in the respective subsample (as it is not possible to reconstruct precisely what inclusion probability the new household members' households had in wave 1). The new design weight for subsample i dw_ihh_2 is therefore calculated from the old cross-sectional weight wq_ihh_1:

$$1/dw_ihh_2=1 / wq_ihh_1 + (n_{sample\ i} / n_{population\ i})$$

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

9.4.2 Design weights for the wave-2 refreshment sample

In the second wave the panel is only refreshed by sampling new households from the new inflows to benefit recipiency. All households that were in receipt of benefit in July 2007 but had had no probability of being selected for the register data sample in the same month of the previous year have a chance of being drawn. This refreshment of the sample can be done by selecting only benefit communities (Bedarfsgemeinschaften) in which no member was receiving benefits in July of the previous year. The refreshment sample is drawn in the 300 points of the first wave. Analogous with the special pps procedure used to draw the first register data sample, which is described in Rudolph and Trappmann (2007), the sample size is 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. However, in wave 2 the number of benefit communities in a household was no longer taken into account. For cases with *sample*=3 the design weight of the refreshment sample is included in the variable *dw_ba*.

9.4.3 Propensity to participate again - households

In this step, the probability of re-participating is estimated for each household on the basis of logit models for willingness to participate in a panel, loss of contact and refusal. In addition to variables from the household interview and the personal interview with the head of the household in the previous wave, other variables are included which are associated with the fieldwork, e.g. interview mode, number of contact attempts, relocation, household splits, household size, number of individuals willing to participate. The predicted propensities of all three models are multiplied. The reciprocal value of this

product can be found in the variable *hpbleib.* The longitudinal weight of a household for the period $[t_1;t_2]$ between two waves can then be calculated as the product of the cross-sectional weight for t_1 and *hpbleib*. The full lists of variables in the models and coefficients are described in the field and method report of TNS Infratest (Büngeler et al. 2009).

9.4.4 Non-response weighting for households from the wave-2 refreshment sample

For the households in the refreshment sample, non-response was modelled in a two-step procedure as was done for the first wave. The full lists of variables in the models and coefficients are described in the field and method report of TNS Infratest (Büngeler et al. 2009) The participation probability derived from this can be found in variable *prop_t0*.

9.4.5 Propensity to participate again – individuals

The decisive longitudinal weight is not the one at the household level but the one at the individual level, as the units here are stable over time. Participation propensities for individuals are modelled in the same way as the model for households shown in step 3. As the participation of the household is a precondition for the participation of the individual, the models contain similar variables. In addition, characteristics of the respective individual (e.g. age, item missings in the previous wave) are taken into account. The predicted propensities of the models are again multiplied. The reciprocal value of this product can be found in variable *ppbleib*. The longitudinal weight of an individual for the period $[t_1;t_2]$ between two waves can then be calculated as the product of the cross-sectional weight for t_1 and *ppbleib*. The full lists of variables in the models and coefficients are described in the field and method report of TNS Infratest.

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

This step involves combining the household weights of the refreshment sample and the panel households, which have been modified by the non-response modelling (from steps 3 and 4). The double selection probability of a newly sampled benefit recipient who was living in the same household as benefit recipients in the previous year but without being a member of the benefit unit him/herself is ignored. This is likely to be a

rare population as four conditions have to be fulfilled simultaneously: (i) benefit recipiency in 7/2007, (ii) no benefit recipiency in 7/2006, (iii) living in the same household as benefit recipients in 7/2006, (iv) not being a joint member of a benefit unit in 7/2007 together with a person who belonged to a benefit unit in 7/2006. As the frames are disjunctive under this assumption, the Projection factors of the register data sample alone remain unaffected by the integration of the refreshment sample (those of the Microm sample on its own of course, too). The new design weights of the benefit recipient sample project in the cross-section to all individuals who were living in a household containing at least one benefit unit in either 7/2006 or 7/2007. It is only when calculating new weights for the total sample that it becomes necessary to adjust the weights for all households in receipt of benefits in 7/2007. For this adjustment the inclusion probability in the respective other sample is estimated for cases from the Microm sample (wave 1) and the refreshment sample (wave 2). For cases from the refreshment sample, the mean wave-1 selection probability in the Microm sample in the respective postcode sector and the average participation probability (for W1 and W2) in that sample are assumed. For cases from the Microm sample, if they are (according to survey data) new recipients of Unemployment Benefit II who first received the benefit between the two sampling dates, the mean selection probability of a household in the refreshment sample in the respective postcode sector and the average participation probability in that sample are assumed. The two weights from 3 and 4 are then integrated to form a new total weight.

9.4.7 Calibration to the household weight, 2nd wave, cross-section

The steps described above are followed by another calibration - of the weights from step 6. At the household level, GREG is used to calibrate the weights to the benchmark statistics of the Federal Statistical Office for 2007 and for households in receipt of benefits the weights are adjusted to the statistics of the Federal Employment Agency for July 2007. There is a change in the procedure compared with that used in the previous year. For the BA sample a new benchmark statistic has been added: the number of new recipients of Unemployment Benefit II since the previous year at the level of benefit communities (734,927). Here those cases in the two samples of wave 1 which, according to wave 2 of the survey, were receiving benefit in July 2007 will be calibrated

to the benchmark statistics of the Federal Employment Agency on receipt of Unemployment Benefit II. The calibration process is described in detail in Kiesl (2009).

9.4.8 Calibration to the person weight, 2nd wave, cross-section

As in wave 1, the person weights were calibrated under the restriction that they differ as little as possible from the calibrated household weights. The calibration is therefore not based directly on the person weights of the first wave. Two changes were introduced compared with the weighting used in the first wave. For the BA sample a new benchmark statistic was added: the number of new recipients of Unemployment Benefit II since the previous year at the level of individuals aged between 15 and 64 (1,053,784). The second change has to do with the calibration to the unemployment figures of the Federal Employment Agency. With the second wave of PASS it is not possible to make a clear distinction between participants in labour market policy measures in accordance with SGB III and unemployed people. Therefore the weights of all individuals who were registered as unemployed or participating in a labour market policy measure run by the Federal Employment Agency at the time of the survey are calibrated to the benchmark figure for the survey months which had been weighted by the number of households who were interviewed in that month. Here those cases in the two samples from wave 1 which, according to wave 2 of the survey, were receiving benefits in July 2007 will be calibrated to the benchmark statistics of the Federal Employment Agency on receipt of Unemployment Benefit II. The calibration process is described in detail in Kiesl (2009).

9.4.9 Estimating the BA cross-sectional weights for households and individuals not in receipt of Unemployment Benefit II

Finally, some households and individuals remain that can not be assigned a BA cross-sectional household weight or a BA cross-sectional person weight by means of calibration. They belong to one of the following three groups which did not receive benefits in July 2007 but which belong to the population of the BA sample (households with receipt of Unemployment Benefit II in 7/2006 or 7/2007 and individuals in households with receipt of Unemployment Benefit II in 7/2006 or 7/2007).

1. From the refreshment sample: individuals in the household who are not members of a benefit unit: here the person weight is obtained from the BA household weight in wave 2 after calibration (*wqbahh*) by dividing it by the proportion of these individuals who gave a personal or senior citizens' interview – provided that their household was participating.

2. Wave 1 households in which nobody was in receipt of Unemployment Benefit II any longer in July 2007: the household retains the BA weight before calibration (from step 6). Individuals in these households with interviews in both waves are given a new BA person weight which is obtained by multiplying their old BA person weight from the previous wave by the reciprocal re-participation probability *ppbleib*. Individuals in these households who did not provide a personal interview in wave 1 are given a new BA person weight calculated by dividing the BA household weight of their household for wave 2 by the proportion of such individuals who participate provided that their household is taking part.

3. Individuals who are not members of a benefit unit in wave 1 households that were still in receipt of Unemployment Benefit II in July 2007: individuals in these households with interviews in both waves are given a new BA person weight which is obtained by multiplying their BA person weight of the previous wave by the reciprocal re-participation probability *ppbleib*.

10. Use of the datasets

10.1 Key variables

The survey datasets of PASS are processed as cross-wave long files in which all of the survey waves assigned to a person (individual dataset) or a household (household dataset) are arranged below one another. Individual households can be identified via the current household number (*hnr*) and can be related to their original household via the original household number (*uhnr*). Individuals have an constant personal ID number (*pnr*).

Processing in long format results in some unusual features in PASS with regard to the clear identification of individual observations and the linking of datasets. For instance, in both the household and the individual datasets the wave indicator (*welle*) is required in addition to the household number and/or the personal ID number in order to identify an observation clearly. In the spell datasets the spell number (*spellnr*) also has to be taken into account. In the following sections the key variables contained in PASS are presented and their use is demonstrated on the basis of typical examples. Table 32 provides an overview of the key variables contained in the PASS datasets.

Table 32: Overview of the key variables in the scientific use file of the 2nd wave

Key variable	Description
hnr	*Current household number* Eight-digit, constant ID number of a household, which is allocated when the household joins the panel. The first digit indicates the wave in which the household was first part of the gross sample of PASS e.g. : **1**0010008 – household in gross sample for first time in 1st wave **2**1011685 – household in gross sample for first time in 2nd wave
uhnr	*Original household number* Eight-digit, constant ID number of the original household. In the case of households that were drawn directly for one of the subsamples, the *uhnr* is the same as the respective *hnr*. In the case of households which have split off from panel households (split-off households) the *uhnr* corresponds to the *hnr* of the household from which the split-off household originated.

Key variable	Description
hnr1	*Household number in wave 1* Eight-digit, constant ID number of the household in the 1st wave of PASS. This variable is only contained in the register datasets processed in wide format.
hnr2	*Household number in wave 2* Eight-digit, constant ID number of the household in the 2nd wave of PASS. This variable is only contained in the register datasets processed in wide format.
pnr	*Constant personal ID number* Ten-digit, constant ID number of the individual. The *pnr* is allocated when a person first joins a PASS survey household. The first eight figures consist of the household number of the household to which the person belonged when he/she joined PASS and the last two figures are the serial number that this person had within this household. e. g.: 1001000801 – person joined the PASS in household 10010008 and had the serial number 01 in this household
zplfd1	*Serial number of the target person in the household in wave 1* Two-digit serial number within the household in the 1st wave, which indicates the person's position in the household structure.
zplfd2	*Serial number of the target person in the household in wave 2* Two-digit serial number within the household in the 2nd wave, which indicates the person's position in the household structure. Within a particular household the *zplfd* is constant in principle. If a person moves to a different household between the waves, then a new *zplfd* is allocated in the new household – in this case *zplfd1* and *zplfd2* differ. Serial numbers that were already used for a certain household in one of the previous waves are not allocated to anyone else. The numbering of new people in a household begins at N+1 (N = highest *zplfd* ever allocated in that household).
welle	*Indicator for survey wave* Both the household and individual datasets as well as the corresponding weighting files are processed in long format in PASS. For every interview that was conducted with a household or a person there is a row in the data matrix. By means of a wave indicator (*welle*) it is possible to assign these different observations for a household or a person to the respective survey wave.
spellnr	*Spell number* As the spell datasets are also processed in long format, here, too, another variable is necessary in addition to the household and personal ID numbers in order to identify observations clearly. In the different subject-related datasets the spells were put into chronological order and then each one was given a serial number, the spell number, within the household or the person. It is not possible to relate spell information clearly to a survey wave as the spells contain cross-wave information.

The following table shows which datasets contain which key variables:

Table 33: Overview of the key variables in the separate datasets

Dataset	Key variables contained								
	hnr	uhnr	hnr1	hnr2	pnr	zplfd1	zplfd2	welle	spellnr
Household level									
Household register *hh_register*	x	x	x	x					
Household dataset *HHENDDAT*	x	x						x	
Unemployment Benefit II spells *alg2_spells*	x								x
Household weights *hweights*	x							x	
Individual level									
Person register *p_register*		x	x	x	x	x	x		
Individual dataset *PENNDAT*	x	x			x			x	
Employment spells *et_spells*					x				x
Unemployment spells *al_spells*					x				x
Gap spells *lu_spells*					x				x
Measure spells *mn_spells*					x				x
Person weights *pweights*					x			x	

Unlike the survey and weighting datasets, the register datasets are not processed in long format. The wave indicator (*welle*) is therefore not contained in these datasets. The other key variables mentioned above *hnr* (household register), *uhnr* (household and person registers) and *pnr* (person register) can be found here, too, however. The person register also contains an additional identifier variable, *lasthnr,* which contains the household number of the household to which the person belonged when he/she was last part of a survey household.

EXAMPLE: MERGING HOUSEHOLD DATA WITH THE INDIVIDUAL DATASET

If household data are to be merged with the individual dataset (e. g. the information on household size which is contained in variable *HA0100*), then the individual dataset first has to be sorted according to the relevant key variables – the household number and the wave indicator. Then the household information can be merged stating the two key variables *(hnr* and *welle*).

```
use PENDDAT.dta
sort  hnr welle
merge hnr welle using HHENDDAT.dta, keep(HA0100)
tab _m welle
drop if _m == 2
```

The tabulation of the *_merge* variable shows that information from the household dataset was merged for 140 households for which no personal interviews were available. These 140 households were re-interviewed households in which no personal interview could be conducted in the 2nd wave. These cases are dropped here.

EXAMPLE: MERGING THE HOUSEHOLD WEIGHTS WITH THE HOUSEHOLD DATASET

The household dataset and the household weights are available in the same format and on the same level. Both of the datasets are already sorted according to the relevant key variables (*hnr* and *welle*). Accordingly the datasets can be merged directly. The same procedure is used for merging the individual dataset and the person weights.

```
use HHENDDAT.dta
merge hnr welle using hweights.dta
tab _m welle
```

The tabulation of the *_merge* variable shows that for all of the households in the 1st and 2nd waves it was possible to merge an observation from the household weight dataset. See Chapter 10.2 on the use of the weights.

EXAMPLE: MERGING INFORMATION FROM THE INDIVIDUAL DATASET WITH THE PERSON-SPECIFIC SPELL DATA

When merging spell data and the household or individual dataset it is always necessary to take into account the different logics of the datasets. Whilst the household and individual datasets contain wave-specific observations of the study units, the spells can not be assigned clearly to one particular wave. A spell of employment, for example, can span several survey dates. This spell is then visible in the data structure as a single observation with its respective start and end dates. If, for instance, individual-level information is to be merged with the person-specific spell data (spells of employment, unemployment, gaps, employment and training measures), then these different data structures have to be taken into consideration. As it is not possible to assign every spell clearly to a particular survey wave, only the personal ID number can be used as a key variable. The information from the individual dataset therefore first has to be converted to wide format and then merged with all of a person's spells. This is demonstrated below using the example of the date of the personal interview which is available in the individual dataset and is to be merged with the employment spells.

First the individual dataset, reduced to the relevant variables, is converted into wide format. For this the information on the interview date, which was so far stored in wave-specific observations, is restructured. Instead of there being one observation per survey wave, there is now only one single observation for each individual in the dataset. The information on the interview date is now stored in the wave-specific variables *pintdat1* and *pintdat2.* For many individuals the spell dataset contains more than one observation. By linking via the personal ID number, the respective interview dates of the 1st wave (*pintdat1*) and the 2nd wave (*pintdat2*) are added to each of a person's spells and are available for further calculations.

```
use PENDDAT.dta
keep pnr welle pintdat
reshape wide pintdat, i(pnr) j(welle)
la var pintdat1 "Datum des Personeninterviews in Welle 1"
la var pintdat2 "Datum des Personeninterviews in Welle 2"
sort pnr
save PINTDAT.dta
use et_spells.dta
sort pnr
merge pnr using PINTDAT.dta
```

```
tab _m
drop if _m == 2
```

The tabulation of the *_merge* variable shows that no employment spell is available for over 15,000 individuals. Some of these individuals were only interviewed in the 1st wave, some had not reported any employment spells in the 2nd wave and some were not asked about their employment owing to a filter. These cases are dropped.

10.2 Weights

10.2.1 Recommendations for the use of surveyset in Stata

All of the weights in PASS are so-called *probability weights*: the weight of a household or a person is equivalent to the reciprocal value of its/his/her inclusion probability (adjusted by non-response modelling and calibration). In Stata, starting with version 9, *probability weights* have to be set using the *surveyset* command (see StataCorp 2007). However, *surveyset* not only has the purpose of defining the weights to be used, but also of defining the aspects of the survey design that have an impact on the standard errors.

There are two different possibilities for doing this in Stata: by specifying the design or by using replication weights. In the first option, the aspects of the survey design that influence the standard error have to be entered in the command line. Besides the weights, these aspects are clusters, stratification characteristics and finite population corrections in sampling without replacement. The effect of calibration on the standard error and other factors such as pps-sampling cannot be taken into account. The second option, on the other hand, makes use of a set of replication weights, which are calculated for all units of the study using processes such as jackknifing, BRR or bootstrapping. This procedure also potentially permits the calibration to be taken into account.

There are no replication weights available for PASS to date, so researchers will have to use the first variant for the *surveyset* for PASS. However, the complex sample design of PASS cannot be used for variance estimation with the *surveyset* command in all details. We recommend the following approach:

```
svyset psu [pw=wqX], strata(strpsu)
```

Here *wqX* stands for the adequate weight for the intended analyses. An indicator for the *primary sampling units* (which are the same for both subsamples) is the variable *psu* in the household dataset *HHENDDAT*. The strata for the selection of the *primary sampling units* are represented by the variable *strpsu* in the same dataset. Strata with fewer than two units in the sample were collapsed. In the case of sampling with replacement, strata and clusters do not play a role in the variance estimation from the second level onwards (Särndal et al. 1992, 144ff.). If the sampling rate is very low, the variance estimation for sampling without replacement can be approximated very well using the formulae for sampling with replacement. This is the case for PASS (only approximately 3.6% of the postcodes in Germany were selected for the survey). There is therefore no need to indicate finite population corrections or further clusters (here: households). However, the recommended *surveyset* then takes neither calibration nor pps-sampling into account, nor the low finite population correction for sampling without replacement. The resulting standard errors are too large and thus should be considered conservative estimates.

In wave 2 there are cases where strata defined by the variable *strpsu* now only contain one single primary sampling unit because all of the respondents in the other PSU belonging to the stratum have dropped out. These cases are rare and only occur in the BA register data sample (*sample=1*) and its refreshment sample (*sample=3*) and affects the following strata:

BA register data sample: strpsu 15, 40

Refreshment sample of wave 2: strpsu 3, 44, 63

When a stratum consists of only one PSU, Stata can not calculate any standard errors. The easiest way to circumvent this problem is to retain the cases from both waves – even if only wave 2 is being analysed – and to declare the second wave to be a subpopulation using the *subpop* option of the survey commands (see StataCorp 2007, 53ff.).

If one works solely with the dataset of the 2nd wave instead (e.g. with the refreshment sample), Stata provides, from Version 10 onwards, various approximation procedures

for cases of strata with only one PSA (the *singleunit* option of the *svyset* command, see StataCorp 2007), but none of them solve the problem entirely satisfactorily.

Singleunit certainty assumes that the single PSU in the sample is also the only one in the population and that the variance between PSUs in this stratum is therefore zero. As there are several PSUs in every stratum in the population of PASS, the basic assumption is not correct. This setting thus results in the variance being underestimated.

In the case of *singleunit scaled* the stratum with missing variance is assumed to have a variance equal to the mean variance in the other strata. As these are rather small strata, however, the variance is likely to be larger in reality.

With *singleunit centered* a variance within the stratum with only one PSU is estimated by assuming that the (unknown) stratum mean is equal to the grand mean. The variance of the stratum is then estimated from the mean of the single PSU in the stratum and the grand mean.

In addition to using this command, it would also be possible to solve the problem by collapsing neighbouring strata. As the strata have been anonymised, however, it is not apparent from the number of a stratum which other stratum is its neighbour. From wave 2 onwards we therefore supply the variable *nextstra* in *HHENDDAT*, which indicates the number of the neighbouring stratum for all strata that consist of only one PSU.

Another remark is necessary on this subject: restrictions to subpopulations using "if" or "keep if" can also make it impossible to estimate standard errors if the restriction results in more strata with only one PSU. Here the recommendation is always to conduct restrictions using "subpop" and not with "if" or "keep if". The only exception is the restriction to one of the three subsamples. Here the restriction with "if" is appropriate. Examples are given in the next section.

10.2.2 Use of the cross-sectional weights

All of the cross-sectional weights are projection factors. Dividing these weights by their mean value results in weights that add up to the sample size. The design weights (*dw_mi, dw_ba, dw*) and the estimated participation propensities (*prop_t0*) are provided along with the panel study; however, we recommend using the calibrated weights. Researchers who wish to do without calibration should bear in mind that although division

of the household weights by the adequate participation propensities estimated for the respective subsample does yield modified household design weights, these do not take into account the fact that there were also cases of non-response within participating households. Use at the individual level thus initially requires an estimation of the person's participation propensity, given that the household takes part.

The following sections provide examples showing how to use the cross-sectional weights for various different research questions.

a) Analyses of benefit recipients in July 2006

The best way to obtain findings on the population of the BA sample in the 1st wave (households in which there was at least one benefit unit receiving benefits in accordance with SGB II as of July 2006; referred to below as "households receiving benefits in July 2006") is to use only the BA sample and the relevant weights. Proceeding in this way is more efficient than using the total sample, as the weights in the BA sample have less variance. Furthermore, the analyses have to be restricted to *sample==1,* as cases from the refreshment sample (benefit recipients in July 2007 who were not receiving benefits in July 2006) are also taken into account otherwise.

ANALYSES AT THE HOUSEHOLD LEVEL

To make analyses of households receiving benefits in July 2006, researchers should use *wqbahh*. The example below demonstrates its use in Stata 10. It is intended to calculate the number or percentage of households receiving benefits which are in possession of a car (variable *HLS0800a*). To start with, the household weights have to be merged with the household dataset, then the *surveyset* has to be carried out, and then the projected value can be calculated:

```
use HHENDDAT.dta, clear
merge hnr welle using hweights.dta
svyset psu [pw=wqbahh], strata(strpsu)
svy, subpop(if welle==1): tab HLS0800a if sample==1, ///
count cell format(%9.0g)
svy, subpop(if welle==1): tab HLS0800a if sample==1, cell ci ///
format(%9.0g)
```

Approximately 37.9% of the households receiving benefits in July 2006 had a car at the time of the survey in the 1st wave, 62.1% did not have a car, and the percentage with no valid response is extremely low. Whilst the first tabulation command shows the projected number and percentages of individuals with and without a car, the second tabulation gives the percentage and the corresponding 95% confidence intervals with the option "ci". The confidence interval is [36,0; 39,7]. It would also be possible to dispense with the restriction *if sample==1* as the weight *wqbahh* in wave 1 is only defined for the cases from *sample* 1 (BA register data sample as of the reference date in July 2006).

The values for the number and percentage of car owners in the same population at the time of the survey in the 2nd wave in the relevant population are obtained as follows:

```
svy, subpop(if welle==2): tab HLS0800a if sample==1, ///
count cell format(%9.0g)
```

Approximately 40.6% of the households receiving benefits in July 2006 had a car at the time of the survey in the 2nd wave. Here households that had split off from wave-1 households by moving out are also counted. The fact that the value increased com-

pared with that of the first wave could be associated with the fact that a not inconsiderable number of these households had probably managed to end benefit recipiency between the first and second waves. If researchers are solely interested in those households that are still in receipt of benefits at the time of the survey, then the command has to be restricted to this set. As it is not a separate sample, a restriction with "if" would result in an underestimation of the variances in this case. The restriction is to be carried out using *subpop* (see StataCorp 2007, 53ff.). The information as to whether a household is receiving benefits on the survey date is contained in the variable *alg2abez* in *HHENDDAT*. Here the value one means that the household was drawing benefits, the value two means that it was not in receipt of benefits and -5 means that it is not possible to establish clearly whether the household was receiving benefits from the information available. The command is therefore:

```
svy, subpop(if alg2abez==1 & welle==2): tab HLS0800a if ///
sample==1, count cell format(%9.0g)
```

Of the households which were receiving benefits in July 2006 and were also still in receipt of benefits on the survey date in the 2nd wave, only 33.1% have a car. This value has thus decreased compared with the first wave. The corresponding confidence intervals are requested using the option "ci".

```
svy, subpop(if alg2abez==1 & welle==2): tab HLS0800a if ///
sample==1, cell ci format(%9.0g)
```

[30.6%; 35.7%] is reported as the 95% confidence interval. This confidence interval lies entirely outside the corresponding interval for 2006. If the 2nd wave were selected using an *if* condition instead of the *subpop* option, in other words by entering the following command:

```
svy, subpop(if alg2abez==1): tab HLS0800a if welle==2 & ///
sample==1, cell ci format(%9.0g),
```

then the message "Note: missing standard errors because of stratum with single sampling unit" would appear. As was described in section a), the strata with only one PSU first have to be merged with neighbouring strata. In this case, however, another calculation problem would arise: the subpopulation with continuous benefit recipiency would no longer include anybody who had answered with "-2 = details refused" and Stata can

not issue any standard errors for this reason either. As we are not interested in the contents of this category, we recommend merging it with the other missing code "-1 don't know".

```
gen strpsu2=strpsu
replace strpsu2=nextstra if nextstra>0 & nextstra!=.
svyset psu [pw=wqbahh], strata(strpsu2)
gen HLS0800a_2= HLS0800a
replace HLS0800a_2=-1 if HLS0800a_2==-2
svy: tab HLS0800a_2 if welle==2 & sample==1, ///
subpop(if alg2abez==1) cell ci format(%9.0g)
```

The 95% confidence interval calculated in this way differs from the one calculated above in the second position after the decimal point. The intervals would be about 0.02 percentage points smaller. The differences are therefore very small.

ANALYSES AT THE BENEFIT UNIT LEVEL

Researchers working on recipiency of Unemployment Benefit II are often not interested in households but in benefit communities. If the above question on the percentage of households receiving benefits in July 2006 which are in possession of a car is to be transferred to benefit communities, the PASS data can be used to answer the question as to how many benefit communities live in a household that has a car (as the benefit communities were identified retrospectively, there are no questions in the questionnaire relating directly to benefit communities – it is therefore not possible to identify which benefit unit owns the car in a household consisting of several benefit communi-ties). This question is relatively easy to answer, using the variable *nbgbezug*, which states how many benefit communities in joint receipt of Unemployment Benefit II a household contains as of the sampling date[43]. The fastest way to do this is to multiply the household weights by this value.

```
use HHENDDAT.dta, clear
merge hnr using hweights.dta
```

[43] For this variable, the decisions required when the statements do not clearly identify how many benefit communities are receiving Unemployment Benefit II in the household were made in the same way as for the calibration process. Every user is of course free to make his or her own decisions on the basis of the Unemployment Benefit II spells.

```
gen bgweight=wqbahh*nbgbezug
svyset psu [pw=bgweight], strata(strpsu)
svy, subpop(if welle==1): tab HLS0800a if sample==1, ///
count cell format(%9.0g)
```

The percentages deviate slightly from those in the analysis presented above (37.9% of households receiving benefits, but 38.2% of the benefit communities receiving benefits had a car in their household in wave 1). Above all, however, the absolute numbers are different: the sum of all households receiving benefits was 3,882,013, whereas the sum of all benefit communities receiving benefits is 4,011,889, and matches the BA benchmark statistics due to the calibration.

In contrast, with PASS it is not possible to calculate the percentage of car owners as of the survey date of the 2nd wave for the benefit communities of the first wave. As the compositions of benefit communities are constantly changing due to deaths, births, moves into and out of the household, and also due to members reaching certain age limits (25 and 65 years of age), this kind of analysis across waves should be conducted at the level of more stable units.

ANALYSES AT THE INDIVIDUAL LEVEL

Analyses at the individual level are similarly simple. The weight *wqbap1* should be used in this case. An intermediate step becomes necessary, as the variables *psu*, *strpsu* and *nextstra* are only contained in the household dataset. The following example calculates the number of individuals aged 15 and above[44] in households receiving benefits who have a background of migration (variable *migration*).

```
use HHENDDAT.dta, clear
keep hnr welle psu strpsu
sort hnr welle
save psuinfo, replace
```

[44] As younger people are not interviewed in person, the PASS data can only be used to establish characteristics about them which are surveyed in the household questionnaires (e.g. age, gender). The household weights should be used in this case.

```
use PENDDAT.dta, clear
merge pnr welle using pweights.dta
drop _m

sort hnr welle
merge hnr welle using psuinfo
svyset psu [pw=wqbap], strata(strpsu)
svy, subpop(if welle==1): tab migration, ///
count cell format(%9.0g)
```

According to this calculation, just under 61.3% do not have a migration background, 24.4% migrated to Germany themselves, at least one parent migrated to Germany for a further 7.6%, and at least one grandparent for another 1.8%. The code "Item not surveyed in questionnaire"[45] applies to 3.6%. This is due to the fact that the data from the short questionnaire for people aged 65 and above are stored in the same dataset as data from the standard personal questionnaire. People aged 65 and above are assigned this code for questions that are not asked in the senior citizens' questionnaire. In order to run analyses excluding these individuals, researchers can limit the frequency count to data from the standard questionnaires (*fb_vers=1*)

```
svy, subpop(if welle==1 & fb_vers==1): tab migration, ///
count cell format(%9.0g)
```

In the same way as the procedure followed above for households, the analyses for individuals from households receiving benefits in July 2006 can also be run for the survey date of the 2nd wave (*welle==2*) and restricted to those people who were still living in a household in receipt of benefits on the survey date in the 2nd wave (*welle==2 & alg2abez==1*).

The person weights of the BA sample project to all individuals in households receiving benefits. Some households, however, consist of several synthetic benefit communities, not all of which receive benefits. Researchers wishing to project only to persons who are members of benefit communities under the provisions of Social Code Book II have to exclude individuals who did not belong to a benefit unit on the sampling date. The

[45] For a further 1.2%, the variable cannot be formed due to missing information.

variable *bgbezs1* from the dataset *p_register* provides information on a person's affiliation with a benefit unit in receipt of benefits as of the sampling date for wave 1:

```
drop _m
sort pnr
merge pnr using p_register.dta
keep if pnetto1==2 | pnetto1==3
svy, subpop(if bgbezs1==1 & fb_vers==1 & welle==1): ///
tab migration, count cell format(%9.0g)
```

The percentage of individuals who migrated to Germany themselves is therefore marginally higher among the people who are members of a benefit unit, at 25.5%, than among people living in a household receiving benefits (25.3%).

b) Analyses on the resident population of Germany

Analyses on the resident population of Germany can be carried out both with the total weights and with the Microm weights. In most cases the results will differ only slightly. The percentage of households with a car in the total population (in wave 1 in this case) is therefore calculated either with the following commands using the total weights:

```
use HHENDDAT.dta, clear
merge hnr welle using hweights.dta
svyset psu [pw=wqhh], strata(strpsu)
svy, subpop(if welle==1): tab HLS0800a, cell ci format(%10.0g)
```

or alternatively with the Microm weights

```
svyset psu [pw=wqmihh], strata(strpsu)
svy, subpop(if welle==1): tab HLS0800a, cell ci format(%10.0g)
```

In the first case, the percentage of households with a car is 75.9% (95% confidence interval of 73.9% to 77.7%), and in the second case 75.6% (95% confidence interval of 73.5% to 77.6%). The confidence interval is slightly narrower when the total weights are used, as in this case the part of the population receiving benefits under SGB II is represented much more precisely, which is why we prefer to use these weights. The same applies to the person weights.

c) Analyses on benefit recipients at different points in time

Section a) explained how the data can be projected onto the total population of the BA register data sample of the 1st wave (households with at least one benefit unit that was in receipt of benefits in accordance with Social Code Book II in July 2006). As a result of its design, however, PASS is more flexible and makes it possible in principle to make projections onto the benefit recipients at any point in time since the benefit was introduced in January 2005.

ANALYSES ON BENEFIT RECIPIENTS IN JULY 2007

PASS takes a first step in this direction with the annual refreshment samples of the register data sample. In wave 2 the refreshment sample consists of households in which there was at least one benefit unit receiving benefits in July 2007 but of which no member was living in a household with at least one benefit unit in receipt of benefits in July 2006 (*sample==3*). If the two samples are integrated, the result is a sample made up of households receiving benefits in July 2006 or July 2007 – admittedly an unusual population. However, if this combined population is restricted to households that were also still in receipt of benefits in accordance with Social Code Book II in July 2007, then these cases can be projected onto all households with Unemployment Benefit II recipiency in July 2007. The annual refreshment of the sample thus enables us to remain "representative" for the benefit recipients in July of the previous year using the integrated benefit recipient samples.

The indicator for benefit recipiency as of the sampling date (of the respective wave) at the household level is the variable *alg2abez* in *HHENDDAT*, which is available for each household in every wave. At the individual level it is the variable *bgbezs** in *p_register*. Here * is a placeholder for the respective wave (*bgbezs1* in wave 1, *bgbezs2* in wave 2 and so on).

We take up the examples from section a) again in the following when we calculate the percentage of households with a car and the percentage of individuals with a migration background as of the interview date of the 2nd wave; but restricted this time to all benefit recipients as of July 2007.

Households in receipt of benefits in July 2007

```
use HHENDDAT.dta, clear
merge hnr welle using hweights.dta
svyset psu [pw=wqbahh], strata(strpsu)
svy, subpop(if alg2abez==1 & welle==2): tab HLS0800a, cell ///
ci format(%9.0g)
```

34.1% of all households in receipt of benefits in July 2007 had a car on the interview date of the 2nd wave. A 95% confidence interval of 31.9% to 36.4% is obtained.

Individuals in receipt of benefits in July 2007

```
use PENDDAT.dta, clear
merge pnr welle using pweights.dta
drop _m
sort hnr welle
merge hnr welle using psuinfo
drop _m
sort pnr
merge pnr using p_register.dta
svyset psu [pw=wqbap], strata(strpsu)
svy, subpop(if bgbezs2==1 & welle==2 & fb_vers==1): ///
tab migration, count cell format(%9.0g)
```

Of all the individuals in receipt of benefits in accordance with Social Code Book II in July 2007, 25.9% migrated to Germany themselves, a further 5.2% have at least one parent who migrated to Germany and another 1.8% have at least one grandparent who migrated.

Analyses on benefit recipients using the latest available data

When working with the BA sample (*sample==1*) and the appropriate weights, the results refer to recipients in July 2006. For analyses of this population, this approach achieves the greatest statistical power, as the BA weights have a relatively low variance. However, researchers will wish to carry out many analyses – especially on fast-

changing characteristics – using the latest available data, to which many characteristics refer, such as employment status, income or employment volume. The survey date of the first wave is between 6 and 13 months after the sampling date, that of the second wave is even 18 to 25 months later. When working on the latest available data exclusively with the BA sample, researchers can only make statements about so-called 'stayers', those who continued to receive benefits from the sampling date until the survey date. In view of a rather high turnover (37% of people receiving benefits under SGB II in January 2005 were no longer doing so by December 2006 (Graf 2007)), this group may differ significantly from the current benefit recipients in its makeup. The refreshment of the benefit recipient sample can not solve this problem. It can be solved, however, by merging the benefit recipient sample with the population sample. The price for this is, however, a substantial loss of statistical power.

Analyses of benefit recipients using the latest available data at the household level

Representative results for current benefit recipients can therefore only be obtained using the total weights. The variable for whether the household is currently receiving benefits (*alg2abez*) is contained in the household dataset (*HHENDDAT*). Calculations are therefore relatively simple for analyses at the household level. The example below shows this, again using the question of car ownership. As in section a) we again calculate using collapsed strata and combined missing categories, so that standard errors can be calculated.

```
use HHENDDAT.dta, clear
merge hnr welle using hweights.dta
svyset psu [pw=wqhh], strata(strpsu)
svy, subpop(if alg2abez==1 & welle==2): tab HLS0800a, ///
cell ci format(%9.0g)
```

Of the households currently receiving benefits, 36.1% had a car on the survey date of the 2nd wave.

If this were estimated using the BA weights and the BA sample,

```
svyset psu [pw=wqba], strata(strpsu)
svy, subpop(if alg2abez==1 & welle==2): tab HLS0800a, ///
cell ci format(%9.0g)
```

a value of 34.1% would be calculated. However, as these data only include 'stayers', in other words households that were receiving benefits both on the sampling date in July 2006 (for *sample==1*) and/or July 2007 (for *sample==2*) and on the survey date, it is plausible that fewer of these households have cars than those that stopped or started receiving benefits during this period.

One consequence of using the total weights rather than the BA weights is the significant increase in the confidence intervals. The variance of the total weights is significantly larger due to the very different sampling rates in the two subsamples. The analyses on car ownership in households receiving Unemployment Benefit II in July 2007, for which we can only work with the BA register data sample, result in a 95% confidence interval of 31.9% to 36.4%. For the survey date, we obtain a substantially larger 95% confidence interval of 33.2% to 39.2%.

Analyses on benefit recipients using the latest data at the benefit unit level

In comparison to the analyses referring to the sampling date in the previous section, an additional step has to be taken as there is no variable equivalent to *nbgbezug* for recipiency of benefits as of the survey date. This variable first has to be generated using the variable *bgbezb2* in *p_register*, which indicates for each benefit unit whether this particular community was receiving Unemployment Benefit II on the survey date[46].

[46] In the sample code, 'recode bgbezb2 (-5=0)' is used to treat all benefit communities for which current recipiency of benefits is unclear on the basis of the survey data as non-recipients.

```
use p_register.dta, clear
collapse (mean) hnr2 bgbezb2, by(bgnr2)
recode bgbezb2 (-5=0)
by hnr2, sort: egen nbgbezak=sum(bgbezb2)
collapse nbgbezak, by(hnr2)
rename hnr2 hnr
sort hnr
save hnr_nbgbezak.dta, replace

use HHENDDAT.dta, clear
merge hnr welle using hweights.dta
drop _m
keep if welle==2
sort hnr
merge hnr using hnr_nbgbezak.dta
gen bgw_akt=wqhh*nbgbezak
svyset psu [pw=bgw_akt], strata(strpsu)
svy, subpop(if alg2abez==1): tab HLS0800a, cell ci format(%9.0g)
```

The estimated value of 36.2% does not differ from that obtained in the analysis at the household level. However, the value no longer refers to a sub-population of just under 3,310,000 households as in the section above, but to just below 3,348,000 benefit communities receiving benefits as of the survey date. During the survey period, the number of benefit communities varied between 3,577,000 (July 08) and 3,666,000 (March 08), according to the BA statistics. This benchmark value is thus not quite reached. The under-reporting arises from the fact that, unlike in the figures referring to the sampling date, information on benefit recipiency at the time of the survey is not available from the register data for all respondents. Thus the under-reporting of benefit recipiency[47] using the latest available data is not corrected by means of calibration.

[47] As recipiency of Unemployment Benefit II is a socially undesirable characteristic, a certain amount of under-reporting is not surprising.

ANALYSES ON BENEFIT RECIPIENTS USING THE LATEST AVAILABLE DATA AT THE INDIVIDUAL LEVEL

Analyses can be transferred to the individual level in much the same way as was done when using data referring to the sampling date. To start with, the person weights and the information for the *surveyset* should again be merged with the individual dataset. For analyses on individuals from households currently receiving benefits, the frequency counts should be limited to individuals with *alg2abez=1*. This variable has to be merged from the household dataset.

```
use HHENDDAT.dta, clear
keep hnr welle psu strpsu alg2abez
sort hnr welle
save psu_alg2_info, replace

use PENDDAT.dta, clear
merge pnr welle using pweights.dta
drop _m
sort hnr welle
merge hnr welle using psu_alg2_info
drop _m
svyset psu [pw=wqp], strata(strpsu)
svy, subpop(if alg2abez==1 & welle==2 & fb_vers==1): ///
tab migration, count cell format(%9.0g)
```

According to this, of the individuals in households currently receiving Unemployment Benefit II, 60.2% have no migration background, 30.1% migrated to Germany themselves, 6.0% have at least one parent who migrated and 2.2% one grandparent who migrated.

In most cases, however, analyses will not be limited to individuals in households receiving benefits, but to individuals in benefit communities receiving benefits. This characteristic is contained in the person register. The following series of commands produces the percentage of migrants among individuals in benefit communities aged between 15 and 64.

```
drop if welle==1
sort pnr
merge pnr using p_register.dta
svy, subpop(if bgbezb2==1 & fb_vers==1): tab migration, ///
count cell format(%9.0g)
```

Analyses on benefit recipients at other points in time

The biographical data on Unemployment Benefit II recipiency at the household level also make it possible in principle to perform analyses referring to other points in time which are between the sampling date and the date when the first wave of the survey was administered. However, variables such as *bgbezs1*, *bgbezb1* or *nbgbezug* are only provided for the two dates described above. Users who would like to run projections referring to other points in time will therefore have to generate analogous variables themselves. When doing this, both imprecision and the problem of benefit recipiency being under-reported will always have to be taken into account.

Comparison of benefit recipients with the general population

The large variety of options for studying benefit recipients and their households or benefit communities shown above results in an equally large variety of options for comparing benefit recipients with the general population. Table 34 provides an overview. The total weights are to be used in all cases.

Table 34: Variables and their possible uses for comparing SGB II benefit recipients with the general population

Variable	Dataset	Values	Suitable for comparing …
sample	*PENDDAT* *HHENDDAT*	1 BA sample 2 Microm sample 3 BA refreshment sample	a) Households in receipt of Unemployment Benefit II in July 2006 (*sample*=1) with households of the resident population (*sample*=2) and households new to the recipiency of Unemployment Benefit II in July 2007 (*sample*=3). b) Individuals in households receiving Unemployment Benefit II in July 2006 (*sample*=1) with individuals in households of the resident population (*sample*=2) and individuals in households new to the recipiency of Unemployment Benefit II in July 2007 (*sample*=3). Households in receipt of Unemployment Benefit II are defined via the subsample.
alg2samp	*hh_register*	0 no benefit recipiency 1 benefit recipiency 2 no benefit recipiency acc. to survey (BA-SP) 3 benefit recipiency unclear acc. to survey (BA-SP) 4 benefit recipiency unclear acc. to survey (Microm-SP)	a) Households in receipt of Unemployment Benefit II in July 2006 (*alg2samp*=1 & *sample*!=3) with households not receiving Unemployment Benefit II in July 2006 (*alg2samp*=0). b) Individuals in households receiving Unemployment Benefit II in July 2006 (*alg2samp*=1 & *sample*!=3) with individuals in households not receiving Unemployment Benefit II in July 2006 (*alg2samp*=0). The user may choose how to deal with cases that were receiving Unemployment Benefit II according to the sample but not according to the survey.
bgbezs1 *bgbezs2*	*p_register*	1 UB II recipiency as of sampling date 0 no UB II recipiency as of sampling date	Individuals in benefit communities receiving Unemployment Benefit II in July 2006 (bgbezs1=1) or July 2007 (bgbezs2=1) with individuals in benefit communities not receiving Unemployment Benefit II in July 2006 (bgbezs1=0) or July 2007 (bgbezs2=0) As this variable was used for the weighting process, a decision was made for every unclear case.
alg2abez	*HHENDDAT*	1 HH currently receiving UB II 2 HH currently not receiving UB II -5 No generation poss. (missing values)	a) Households receiving Unemployment Benefit II on the survey date (*alg2abez*=1) with households not receiving Unemployment Benefit II on the survey date (*alg2abez*=2) of the respective wave. b) Individuals in households receiving Unemployment Benefit II on the survey date (*alg2abez*=1) with individuals in households not receiving Unemployment Benefit II on the survey date (*alg2abez*=2) of the respective wave.

Variable	Dataset	Values	Suitable for comparing …
bgbezb1 *bgbezb2*	*p_register*	1 UB II recipiency in wave 1 0 No UB-II recipiency in wave 1 -5 No generation poss. (missing values)	Individuals in benefit communities receiving Unemployment Benefit II on the survey date of wave 1 (bgbezb1=1) or wave 2 (bgbezb2=1) with individuals in benefit communities not receiving Unemployment Benefit II on the survey date of wave 1 (bgbezb1=0) or wave 2 (bgbezb2=0)

10.2.3 Use of the longitudinal weights

The basic principle of the longitudinal weighting is simple: the reciprocal re-participation probabilities *hpbleib* and *ppbleib* are used for the longitudinal weighting of the households and individuals respectively. The longitudinal weight for a household (or an individual) for the longitudinal section from wave 1 to wave 2 is obtained by multiplying the cross-sectional weight of the household (or the individual) for wave 1 by the reciprocal re-participation probability. The reciprocal re-participation probability is provided in the dataset for all households (and individuals) that took part in both wave 1 and wave 2. Variety results from the fact that restrictions to certain subsamples or cases with certain characteristics or analyses at the different levels (household, benefit unit, individual) are possible. We would like to demonstrate the use of the longitudinal weights for some typical applications.

Individuals of the resident population

One possible research question involving the longitudinal section could be how many individuals (from the age of 15) of the resident population reported greater satisfaction with their standard of living in wave 2 than they did in wave 1 (variable *PA0300*). The population for such a question is all individuals who belonged to the resident population of Germany in wave 1 and wave 2.

Some preparations first have to be made, but they can also be used for the subsequent analyses. First wave 1 and the variables *psu* and *strpsu* are extracted from the household dataset

```
use HHENDDAT.dta, clear
keep hnr welle psu strpsu
keep if welle==1
drop welle
```

```
sort hnr
save psu_strpsu_w1.dta, replace
```

In a second step the weights from the first wave and the re-participation probabilities from wave 1 to wave 2 are stored.

```
use pweights.dta, clear
keep if welle==1
save pweights1.dta, replace
```

Now the individual dataset is retrieved. We have decided to run the analyses in wide format and therefore have to re-sort the dataset so that the variables *PA03001* (satisfaction with the standard of living in wave 1) and *PA03002* (satisfaction with the standard of living in wave 2) are retrieved. We only retain the variables that we require later.

```
use PENDDAT.dta, clear
keep pnr hnr welle PA0300
reshape wide PA0300 hnr, i(pnr) j(welle)
```

Now the three datasets are merged

```
rename hnr1 hnr
sort hnr
merge hnr using psu_strpsu_w1.dta
keep if _m==3
drop _m
sort pnr
merge pnr using pweights1.dta
drop _m
```

In order to make the tables clearer, a variable is created that indicates the relative level of satisfaction in wave 2 compared with wave 1.

```
gen rel_zufr=2 if PA03002>PA03001 & PA03001>=0 & PA03002>=0
replace rel_zufr=1 if PA03002==PA03001 & PA03001>=0 & PA03002>=0
replace rel_zufr=0 if PA03002<PA03001 & PA03001>=0 & PA03002>=0
replace rel_zufr=-1 if PA03001<0 | PA03002<0
```

```
label define rel_zufr_lb ///
2 "W2 zufriedener als W1" ///
1 "W1 und W2 gleich zufrieden" ///
0 "W2 weniger zufrieden als W1" ///
-1 "in mind. 1 Welle keine Angabe"

label values rel_zufr rel_zufr_lb
```

Finally the longitudinal weight is constructed and the weighted analysis follows

```
gen wp1_2=wqp*ppbleib
sort pnr
svyset psu [pw=wp1_2], strata(strpsu)
svy: tab rel_zufr, count cell format(%10.0g)
```

It refers to just under 64 million individuals who were at least 15 years old in wave 1 and were still resident in Germany on the survey date in wave 2. Of this group 34.6% were less satisfied in wave 2 than they were in wave 1. In contrast, 32.1% were more satisfied. For 33.1% the assessment had not changed.

Individuals in households receiving Unemployment Benefit II in July 2006

Now the same question can be also asked for the individuals in the benefit recipient sample of the first wave. How satisfied are these individuals in wave 2 compared with wave 1? The only difference to the previous analysis is that the BA weight has to be used instead of the total weight.

```
gen wbap1_2=wqbap*ppbleib
sort pnr
svyset psu [pw=wbap1_2], strata(strpsu)
svy: tab rel_zufr, count cell format(%10.0g)
```

Here 33.7% are less satisfied than in the previous wave, whereas 42.7% are more satisfied. The result refers to 5,709,000 individuals from the age of 15 who were living in a household which was receiving benefits in July 2006 and belonged to the resident population in wave 2. In this respect it is not surprising the majority is more satisfied than in wave 1, as some of them should have managed to leave benefit recipiency in

the meantime. Researchers will therefore perhaps be more interested in how the satisfaction levels changed for those people who were receiving benefits on both survey dates.

Individuals in receipt of Unemployment Benefit II on both survey dates

As was the case in the analyses described above, for the question as to changes in the satisfaction levels of people who are still in receipt of benefits, the variables that indicate benefit recipiency on the survey date are required again. These variables are contained in the person register, which is merged here.

```
merge pnr using p_register.dta
keep if _m==3
svyset psu [pw=wp1_2], strata(strpsu)
svy, subpop(if bgbezb2==1 & bgbezb1==1): tab rel_zufr, ///
count cell format(%10.0g)
```

As for the total population, a relatively consistent picture emerges here: 38.8% with increased satisfaction face 36.4% with a reduction in satisfaction. This preliminary work now also makes it possible to analyse rapidly the change in the satisfaction levels of people entering and leaving benefit recipiency

```
svy, subpop(if bgbezb2==0 & bgbezb1==1): tab rel_zufr, ///
count cell format(%10.0g)
svy, subpop(if bgbezb2==1 & bgbezb1==0): tab rel_zufr, ///
count cell format(%10.0g)
```

Of the individuals leaving benefit recipiency, 55.0% are more satisfied, but 27.7% are less satisfied; of the individuals entering benefit recipiency, 46.4% are less satisfied, but at least 26.8% are more satisfied. This of course leads to the question as to whether the relatively large proportions of people who are less satisfied than they were in the previous year despite leaving benefit recipiency or are more satisfied despite entering benefit recipiency are associated with the fact that their income has hardly changed. This would be going too far here, however.

Longitudinal weighting at the household level

First we present a simple example and then we address some of its problematic aspects. We answer the question as to how many households of the resident population acquired or gave up a car between wave 1 and wave 2. We use the same procedure as in the example for individuals described above: first the dataset is created.

```
use hweights.dta, clear
keep if welle==1
save hweights1.dta, replace

use HHENDDAT.dta, clear
keep hnr welle HLS0800a psu strpsu
reshape wide HLS0800a psu strpsu, i(hnr) j(welle)
* gen split=1 if hnr!=uhnr
* replace hnr=uhnr if uhnr!=hnr
* by hnr, sort: egen psu1x= mean(psu1)
* replace psu1=psu1x if psu1==.
* by hnr, sort: egen strpsu1x= mean(strpsu1)
* replace strpsu1=strpsu1x if strpsu1==.
* by hnr, sort: egen HLS0800a1x= mean(HLS0800a1)
* replace HLS0800a1=HLS0800a1x if HLS0800a1==.
sort hnr
merge hnr using hweights1.dta
keep if _m==3
drop _m
```

Then a variable is generated which expresses the change with regard to car ownership.

```
gen auto_neu=3 if HLS0800a1==2 & HLS0800a2==1
replace auto_neu=2 if HLS0800a1==1 & HLS0800a2==1
replace auto_neu=1 if HLS0800a1==2 & HLS0800a2==2
replace auto_neu=0 if HLS0800a1==1 & HLS0800a2==2
replace auto_neu=-1 if HLS0800a1<0 | HLS0800a2<0
label define auto_neu_lb ///
```

```
3 "Auto angeschafft" ///
2 "Auto behalten" ///
1 "weiterhin kein Auto" ///
0 "Auto abgeschafft" ///
-1 "in mind. 1 Welle keine Angabe"

label values auto_neu auto_neu_lb
```

Finally the weight is constructed and the table produced.

```
gen wh1_2=wqhh*hpbleib
svyset psu1 [pw=wh1_2], strata(strpsu1)
svy: tab auto_neu, count cell format(%10.0g)
```

1.7% of the households gave up a car, 2.2% acquired one, 76.2% kept one, 19.7% still do not have one. Instead of now once again distinguishing between households receiving benefits and those not receiving benefits, we wish to discuss something more fundamental here:

The result produced above applies for all households of the resident population at the end of 2006 and their successor households existing as of the survey date in 2007/2008. As households are not units that are stable over time, a longitudinal analysis of households always requires a definition of what is to be regarded as the successor of a household in cases where the household composition changes. If the calculation is done as in this example, then the rules applied by PASS when allocating household numbers are used:

a) If individuals move into a household, the household number does not change. The new, larger household is the successor of the smaller household from the previous wave.

b) If household members die or move abroad, the household number does not change. The new, smaller household is the successor of the larger household from the previous wave.

c) If parts of the old household form a new household within Germany, then the household that retains the household number (and is therefore defined as the successor household) is the one that is reached via the original contact information (depending on

the field this is either the telephone number or the address) or – if this does not apply for either of the new households – the one that is reached first. This household is defined as the successor here. This means that households which have split off from original households are not included in the analyses to date. This could be one explanation for the finding that there are more households which have acquired a car than households that no longer have one. Households that were merged were counted here too. In the case of households that have split up, only the half which remained at the old address or which was reached first was counted.

There is now the possibility to incorporate split-off households into the longitudinal analysis too. For this each split-off household has to be allocated the cross-sectional weight of the original household from wave 1 and a re-participation probability. The eight lines which were first commented out above have to be included for this. The split-off households are projected to about 200,000 additional households and increase the percentage of households that had a car in wave 1 but no longer had one in wave 2 to 1.9%.

The number of split-off households was presumably still clearly underestimated in this way, however, as the split-off households were allocated the re-participation probability of their original households. As was already explained in the chapter on response rates, split-off households are more difficult to interview. It would therefore presumably be more accurate to take the mean probability of interviewing a split-off household as the re-participation probability for these households. It was only possible to interview 46 of the 346 split-off households (13.3%) in wave 2. Another issue is that we are only able to diagnose the split if the original household was interviewed. This was the case for 7,342 of 12,774 households (58.9%) which were interviewed in wave 1 and still belonged to the population in wave 2[48]. If there is assumed to be the same proportion of split households among the households that were not interviewed, then it was probably only possible to interview about 13.3% of 58.9% = 7.8% of the households that had actually split off from original sample households. For split-off households a possible alternative would be to set 1/0.078 (or more precisely (12,774/7,342)*(346/46)) as the reciprocal re-participation probability. This would probably be closer to reality. Owing to

[48] These figures can be calculated in the household register: households interviewed in wave 1 which still belong to the population in wave 2: "`count if hnettod1==10 & hnettod2!=24`", households interviewed in waves 1 and 2: "`count if hnettod1==10 & hnettod2==10`".

the large weights of the split-off households, however, the confidence intervals may be very large.

```
replace hpbleib=(12774/7342)*(346/46) if split==1
replace wh1_2=wqh*hpbleib
svyset psu1 [pw=wh1_2], strata(strpsu1)
svy: tab auto_neu, count cell format(%10.0g)
```

The 46 split-off households now stand for 1,350,000 new households. We would now obtain the estimate that 3.6% of the households which had a car in wave 1 no longer had one in wave 2. Approximately half of this proportion now goes back to newly formed split-off households, which frequently do not have a car initially.

10.3 Constant characteristics

Variables which are assumed not to change over time are only surveyed once in PASS. However, despite the constant nature of the characteristics in reality, changes in these variables are sometimes possible since, for example, incorrect entries may be corrected in subsequent interviews (e.g. in the case of gender). The following sections provide a brief overview of the constant characteristics that are available in PASS. The intention here is to show the conditions under which the variable was surveyed for the first time and to indicate the dataset in which it can be found. The key variables are disregarded here.

a) *Gender*

Information as to a person's sex is gathered at the household level, either when the household in which the individual is living is first interviewed in the context of PASS or when the individual joins a sample household as a new member (e.g. when new individuals move into the household). In wave 2 the interviewers had the opportunity to correct details regarding gender which had been recorded incorrectly in the previous wave. During the plausibility checks of the household structure, too, changes were occasionally made to the gender variables in households that attracted attention as a result of implausible relationships between the household members. Here the gender data was checked on the basis of the first names. No retrospective changes of the data collected in wave 1 were made in either the household or the individual dataset.

Table 35: Information on constant characteristics – gender

Variable	Description	Dataset	Filled in for wave 1	Filled in for wave 2
HD0100a to o	Gender of individuals 1 to 15 in the h'hold	*HHENDDAT*	Yes, if person was in h'hold in wave 1	Yes, if person was in h'hold in wave 2
zpsex	Gender of target person	*PENDDAT*	Yes, if person interviewed in wave 1	Yes, if person interviewed in wave 2
sex	Gender of target person	*p_register*	Information not wave-specific but contains the respective last correction	

b) *Half-year of birth*

A person's half-year of birth was generated from the date of birth reported in the personal interview. Although it is an constant characteristic, the date of birth is asked for in every personal interview conducted. Among other things it serves to check whether the correct person is being interviewed. In wave 2 the interviewers had the opportunity to correct details which had been entered incorrectly in the previous wave. If the half-year of birth differs from that in the previous wave as a result of the date of birth being corrected in the personal interview, this was understood as the correction of an incorrect entry. No retrospective changes were made to the information collected in the previous wave.

Table 36: Information on constant characteristics – half-year of birth

Variable	Description	Dataset	Filled in for wave 1	Filled in for wave 2
gebhalbj	Target person's half-year of birth, generated	*PENDDAT*	Yes, if person interviewed in wave 1	Yes, if person interviewed in wave 2

c) *Migration background*

A person's migration background is also understood as an constant characteristic and is only surveyed in the personal questionnaire in the first interview conducted with a person. The information on nationality (*PMI0400*, *PMI0500*), on temporary residence permits (*PMI0600*) and the type of residence/settlement permit (*PMI0650*), on the other hand, is gathered in every wave as changeable characteristics. In the senior citizens' interviews of the 1st wave no information was collected about whether the respondent's parents and/or grandparents migrated to Germany, and if so from where. It was therefore not possible to establish the migration background for senior citizens in the same way as in the standard personal interviews because information was only available about whether the respondent

him/herself was born outside Germany. From the 2nd wave onwards the migration of parents and grandparents and their respective countries of origin are also surveyed in the senior citizens' interviews. In the 2nd wave all senior citizens were asked the questions. In subsequent waves this information will also be surveyed in the senior citizens' questionnaire in the first interview only.

Table 37: Information on constant characteristics – migration background

Variable	Description	Dataset	Filled in for wave 1	Filled in for wave 2
PMI0100	Target person born in Germany?	*PENDDAT*	Filled in for wave in which person was first interviewed	
PMI0200	Target person's country of birth, if not Germany?	*PENDDAT*	Filled in for wave in which person was first interviewed	
PMI0300a and *b*	Date of migration to Germany?	*PENDDAT*	Filled in for wave in which person was first interviewed	
PMI0700	Parents/grandparents born outside Germany?	*PENDDAT*	Filled in for wave in which person was first interviewed (Filled in for all senior citizens interviewed in 2nd wave)	
PMI0800a to *f*	Which parent/grandparent not born in Germany?	*PENDDAT*	Filled in for wave in which person was first interviewed (Filled in for all senior citizens interviewed in 2nd wave)	
PMI0900a to *f*	Which parent/grandparent migrated to Germany?	*PENDDAT*	Filled in for wave in which person was first interviewed (Filled in for all senior citizens interviewed in 2nd wave)	
PMI1000a to *f*	Country from which parent/ grandparent migrated to Germany?	*PENDDAT*	Filled in for wave in which person was first interviewed (Filled in for all senior citizens interviewed in 2nd wave)	

The respondent's country of birth and information about the countries from which the individual parents and grandparents migrated to Germany are also made available in generated variables in which the information that is collected once is also taken over into subsequent waves. These variables are shown in Table 38. Furthermore, for individuals who were not interviewed for the first time in wave 2, too, the generated variable *migration* (see Chapter 10.3) contains the information as to whether this person has a migration background and if so what this background is.

Table 38: Information on constant characteristics – generated variables on migration background

Variable	Description	Dataset	Filled in for wave 1	Filled in for wave 2
ogebland	Target person's country of birth, if not Germany? (incl. responses to open-ended questions, categorised)	*PENDDAT*	Yes, if person interviewed in wave 1	Yes, if person interviewed in wave 2
ozulanda to *f*[49]	Country from which parent/ grandparent migrated to Germany? (incl. responses to open-ended questions, categorised)	*PENDDAT*	Yes, if person interviewed in wave 1 (Not surveyed for senior citizens in wave 1)	Yes, if person interviewed in wave 2
migration[50]	Target person's migration background (generated)	*PENDDAT*	Yes, if person interviewed in wave 1 (Not generated for senior citizens in wave 1)	Yes, if person interviewed in wave 2

d) Parents' education and vocational training; parents' occupational status and occupational activity

In wave 1, individuals whose mother and/or father did not live in the same household were asked about their parents' respective highest school qualification and their vocational qualifications. If the mother or father was living in the household, the information they provided in their own personal interviews was assigned to the target person.

For individuals interviewed for the first time in wave 2, the parents' highest school qualifications and vocational qualifications were recorded as proxy information irrespective of whether the mother and/or father was living in the same household. Details about the qualifications which the parents may have given in their own personal interviews were thus no longer assigned to the children living in the household. People who had already been interviewed in the previous wave were not asked questions on this topic again. Furthermore, in wave 2 additional questions were incorporated about the mother's and father's occupational status and occupational activity at the time when the target person him/herself was 15 years old. This information is also collected only once.

Table 39: Information on constant characteristics – social origin

Variable	Description	Dataset	Filled in for wave 1	Filled in for wave 2
PSH0200	Target person's mother's highest general school qualification	*PENDDAT*	Filled in for wave in which person first interviewed	
PSH0300a to *i*	Target person's mother's vocational qualifications	*PENDDAT*	Filled in for wave in which person first interviewed	
PSH0310 - *PSH0380*	Mother's occup. status and type of occup. activity when target person was aged 15	*PENDDAT*	Not surveyed in wave 1	Filled in for all individuals interviewed in wave 2
PSH0500	Target person's father's highest general school qualification	*PENDDAT*	Filled in for wave in which person first interviewed	
PSH0600a to *i*	Target person's father's vocational qualifications	*PENDDAT*	Filled in for wave in which person first interviewed	
PSH0610 to *PSH0680*	Father's occup. status and type of occup. activity when target person was aged 15	*PENDDAT*	Not surveyed in wave 1	Filled in for all individuals interviewed in wave 2

In wave 2, however, information is also available on the mother's school and vocational qualifications for the people who were interviewed for the first time in wave 1. It is contained in the generated variables *mschul1*/ *mschul2* (mother's highest general school qualification without / with coding of responses to open-ended survey questions) and *mberuf1*/ *mberuf2* (mother's highest vocational qualification without / with coding of responses to open-ended survey questions). Corresponding information for the target person's father can be found in *vschul1*/ *vschul2* and *vberuf1*/ *vberuf2*. The information on the mother's and father's occupational status, which was gathered in wave 2, are available in *mstib* and *vstib* in the individual dataset, also as generated variables. The generated variables cited are described in the list of variables in this documentation. Moreover, the information about the parents' occupational activity at the time when the target person was 15 years old was coded both by TNS Infratest (*misco_it; visco_it*) and by Ge-

[49] The country from which the parents/grandparents migrated to Germany was surveyed for senior citizens for the first time in the 2nd wave.
[50] Not generated for senior citizens' interviews.

sis-ZUMA (*misco; visco*) according to the 1988 International Standard Classification of Occupations (ISCO-88) published by the International Labour Office (ILO). (See Chapter 7.5 on the differences between the two variants).

e) Sample indicator, sampling year and receipt of Unemployment Benefit II of the household on the sampling date

The sample indicator (*sample*), the sampling year (*jahrsamp*) and the receipt of Unemployment Benefit II of a household on the sampling date (*alg2samp*) are constant characteristics of the household which are defined once when the household joins the PASS sample. Individuals are assigned the sample indicator (*sample*) of the household to which they belong when they first become part of the PASS sample. Households which have split off from households already surveyed in the previous wave and are now surveyed as separate households in PASS take over the values of their original household in the variables *sample*, *jahrsamp* and *alg2samp*.

Table 40: Information on constant characteristics – sample information

Variable	Description	Dataset	Filled in for wave 1	Filled in for wave 2
sample	Sample indicator	*HHENDDAT* *PENDDAT* *hh_register* *p_register* *hweights* *pweights*	Information not wave-specific	
jahrsamp	Sampling year	*hh_register*	Information not wave-specific	
alg2samp	Receipt of Unemployment Benefit II of the household on sampling date	*hh_register*	Information not wave-specific	

10.4 Register data

In addition to the main datasets at the household and the individual levels (*HHENDDAT* and *PENDDAT* respectively), the various spell datasets (*alg2_spells*, *et_spells*, *al_spells*, *lu_spells*, *mn_spells*) and the weighting datasets (*hweights*, *pweights*), the scientific use file of PASS also contains a household register dataset and a person register dataset (*hh_register*, *p_register*). In contrast to the other datasets, these two files are processed in wide format, i.e. there is exactly one observation available per house-

hold or individual. Information referring to individual survey waves are stored in wave-specific variables. The wave to which a piece of information refers is indicated by a counter at the end of the respective variable – thus the variable *alter1* in the person register, for example, contains the person's age in the 1st wave, and *alter2* is accordingly the person's age in the 2nd wave. The register datasets are prepared in such a way that they can easily be converted from wide format to long format, for example using the "reshape" command in Stata. This makes it possible to merge also the register information rapidly with the survey datasets, which are available in long format. Households which are not interviewed in certain waves, individuals in households which are not interviewed and individuals who no longer belong to a sample household in a later wave can be identified via the respective net variables[51]. In addition, in these cases the wave-specific household number (*hnr1*, *hnr2*) is allocated the code "-6". In the following sections the structure and contents of the household register dataset and the person register dataset are presented and their use demonstrated using two examples.

All of the households which have been successfully surveyed at least once in the sense of PASS are contained in the household register. A household is regarded as successfully surveyed if the full household interview and at least one interview at the individual level with a household member aged over 15 are available. Accordingly, households from the gross samples of the individual waves which were not successfully surveyed and households that have split off from panel households and have not been interviewed are not contained in the household register. In addition to the identifiers, the register dataset contains in particular wave-specific information on the survey status of the households (*hnettok1*, *hnettod1*, *hnettok2*, *hnettod2*), on the sample (*sample*), the sampling year (*jahrsamp*), the Unemployment Benefit II receipt of the household on the sampling date (*alg2samp*) and on the number of benefit communities in the household. The household register therefore makes it possible to establish in which waves a household was interviewed in PASS and why no interview is available for certain waves. In this way a preliminary selection of households can be conducted – for example all of the households that were interviewed in all of the waves can be selected.

The person register dataset contains all individuals who were a member of a PASS survey household in at least one wave, irrespective of whether an interview at the

[51] These are described later in this chapter.

personal level has already been conducted with them or not. In addition to the constant personal ID number as the identifier and details regarding the person's gender (*sex*) and age (*alter1*, *alter2*), the person register dataset contains information about which household the person belonged to in the individual survey waves (*hnr1*, *hnr2*) and what position he/she occupied in the structure of these households (*zplfd1*, *zplfd2*). The person register thus makes it possible to allocate individuals to households in specific waves. Furthermore, the person register dataset contains information regarding the individuals' survey status in the individual survey waves (*pnettok1*, *pnettod1*, *pnettok2*, *pnettod2*), which can be used, for example, to identify fully surveyed households, to distinguish between reasons for non-response and to clarify people's whereabouts.

In addition to the person-related information, the person register dataset also contains information on the benefit community to which the individual was assigned. These benefit communities are so-called "synthetic" benefit communities created on the basis of the current legal situation at the particular time and based on information about the ages of the households members and relationships between them, irrespective of whether they are currently receiving Unemployment Benefit II (see Chapter 7.5). The information about the benefit communities is available as wave-specific information. It must be taken into account that this information was generated each time on the basis of the information available for the individual waves. Via the benefit community ID number (*bgnr1*, *bgnr2*) it is possible to identify the individuals who together constitute a benefit community. Here it must be taken into consideration that new numbers are allocated in each wave and that there is no continuation in the longitudinal section. Furthermore the dataset contains information on the type of benefit community (*bgtyp1*, *bgtyp2*) and on the benefit receipt of the benefit community on the sampling date and the survey date of the current wave (*bgbezs1*, *bgbezs2*, *bgbezb1*, *bgbezs2*).

The person register dataset also contains pointer variables referring to the mother living in the household (*zmhh1*, *zmhh2*), the father living in the household (*zvhh1*, *zvhh2*) and the partner living in the household (*zparthh1*, *zparthh2*). These pointers each contain the ten-digit personal ID number of the person who is the target person's mother/father/partner.

Examples to demonstrate the use of the register datasets are given below.

EXAMPLE: SELECTION OF THE HOUSEHOLDS THAT WERE SUCCESSFULLY SURVEYED IN THE 1ST AND 2ND WAVES AND WERE RECEIVING UNEMPLOYMENT BENEFIT II ON THE SAMPLING DATE

The net variables are available in two levels of detail – in a "short", single-digit variant (*hnettok1*, *hnettok2*) and a "detailed", two-digit variant (*hnettod1*, *hnettod2*). The two-digit net variables differentiate the single-digit codes further. The single-digit code "2" in *hnettok2* (household not successfully surveyed, only in gross sample) is further differentiated in *hnettod2* in the codes beginning with "2". This makes it possible to establish why the household could not be successfully surveyed in the 2nd wave, for example because the household could not be reached (*hnettod2=20*) or because it refused to participate (*hnettod2=21*). As only households that were successfully surveyed are to be selected here, the information in *hnettok1* and *hnettok2* is sufficient. After retaining only the cases that were successfully surveyed in both the 1st and the 2nd waves and were receiving Unemployment Benefit II on the sampling date, only the relevant variables (*hnr; alg2samp*) are retained, the dataset is sorted by household number, stored temporarily and merged with the household dataset, which has also been sorted according to *hnr*.

```
use hh_register.dta, clear
keep if hnettok1 == 1 & hnettok2 == 1 & alg2samp == 1
keep hnr alg2samp
sort hnr
save hh_register_vorb1.dta, replace
use HHENDDAT.dta, clear
sort hnr
merge hnr using hh_register_vorb1.dta
tab _merge alg2samp, m
```

An examination of the *_merge* variable indicates that 6210 observations (from 3105 households) from the household register dataset which were interviewed in both waves and were in receipt of Unemployment Benefit II on the sampling date were merged with the individual dataset.

EXAMPLE: IDENTIFICATION OF THE PERSONAL INTERVIEWS WITH THE HEADS OF HOUSEHOLDS

The household register dataset contains the wave-specific information about which person the household interview was conducted with. In order to mark the personal interviews of these heads of households, it is first necessary to prepare the household register and convert it into long format. First of all only the required variables are retained – the household number and the wave-specific pointer to the target persons of the household interview. Then the dataset is reshaped from wide format to long format. For this the household number serves as an ID variable that identifies an observation. In the course of the reshaping process a wave indicator (*welle*) is created which is needed for merging with the individual dataset.

However, before the register which has been converted into long format can be merged with the individual dataset, some observations have to be deleted. If a household was not interviewed in one wave, then the pointer variable referring to the head of the household was given the value "-6" (household not interviewed in wave or not in gross sample) for this wave. A household that was interviewed for the first time in the 2nd wave, for example in the context of the refreshment sample, has the value "-6" for the observation referring to the 1st wave. These observations can not be merged with the individual dataset and can therefore be deleted. After this, the pointer variable *pnrzp* is renamed *pnr*, as the data is to be merged via the constant personal ID number. After the register dataset has been prepared and sorted by *pnr* and *welle,* it is stored temporarily and merged with the individual dataset.

```
use hh_register.dta, clear
keep hnr pnrzp1 pnrzp2
reshape long pnrzp@, i(hnr) j(welle)
drop if pnrzp == -6
ren pnrzp pnr
drop hnr
sort pnr welle
save hh_register_vorb2.dta, replace
use PENDDAT.dta, clear
merge pnr welle using hh_register_vorb2.dta
```

```
tab _merge
drop if _merge == 2
gen hhvorst = _merge == 3
```

The tabulation of the *_merge* variable shows that in 414 cases there is no personal interview available with the person who completed the household interview in that wave. As there is no information about them from personal interviews, these observations that were merged from the person register can be deleted. All of the cases for which the merging was successful (*_merge* == *3*) were the head of the household in the particular wave and are flagged via the variable *hhvrst*.

10.5 Spell data

Already in the 1st wave spell datasets were supplied together with the scientific use file of PASS. Whereas the dataset on Unemployment Benefit II receipt of the household was continued in the 2nd wave, the survey concept for the other two spell datasets, Unemployment Benefit I receipt and participation in employment and training measures, was thoroughly revised. In the course of this revision process, it was decided not to continue the data structure of the spell datasets on employment and training measures and Unemployment Benefit I receipt used in the 1st wave but to create new datasets (see Chapters 4.5 and 4.4). Periods when the respondent received Unemployment Benefit I are surveyed from the 2nd wave onwards as part of the periods of registered unemployment in the biography module. For every period when the respondent was registered as unemployed, information is gathered as to whether he/she received Unemployment Benefit I and if so, from which start date and to which end date. Periods of Unemployment Benefit I receipt are therefore embedded in a period of registered unemployment and are no longer surveyed directly as they were in the 1st wave. The way in which participation in employment and training measures is surveyed was revised because it had emerged that in some cases it was not possible to identify the type of measure clearly with the concept used in the 1st wave. The type of measure is now identified right at the beginning of the employment and training measure module using a multiple choice question.

Another important innovation regarding the spell datasets results from the fact that the concept for surveying periods of employment, unemployment and economic inactivity

was altered in the 2nd wave. Instead of only asking about the status as of the interview date, as was done in the 1st wave, a biography module is used in the 2nd wave to record spells of employment and registered unemployment retrospectively for the period between January 2005 and the interview date. Gaps as of the interview date in the 2nd wave or periods of more than three months duration for which the respondent reported neither employment nor unemployment are caught by a gap module. If no period of employment or unemployment had been forgotten and if it was not a case of incorrect dating, the respondent was able to report the type of economic inactivity. These periods of economic inactivity are made available in the gap spells (*lu_spells*) in the scientific use file. The spell data on employment and unemployment (*et_spells*, *al_spells*) are also contained in the scientific use file for the first time.

The following table provides an overview of the spell datasets contained in the *scientific use file* and their contents and survey periods:

Table 41: Overview of the spell datasets in the scientific use file of the 2nd wave and the periods surveyed

Dataset	Contents	Period surveyed
Household level		
alg2_spells	Information on periods when the household received Unemployment Benefit II; periods of cuts in Unemployment Benefit II	1st wave: January 2005 (or, if later than that: date of last change in household composition) – interview date W1 2nd wave (re-interviewed households): interview date W1 – interview date W2 2nd wave (new sample households): January 2006 (or, if later than that: date of last change in household composition) – interview date W2
Individual level		
et_spells	Information on periods when the respondent was employed with a monthly income of more than € 400	2nd wave: January 2005 - interview date W2
al_spells	Information on periods when the respondent was registered as unemployed or was participating in a employment or training measure run by the Employment Agency	2nd wave: January 2005 - interview date W2
lu_spells	Information on periods when the respondent was not in employment (and not registered as unemployed)	2nd wave: January 2005 - interview date W2
mn_spells	Information on periods when the respondent was participating in an employment or training measure	2nd wave: January 2006 - interview date W2
Datasets of the 1st wave which are not continued		
massnahmespells	Information on periods when the respondent was participating in an employment or training measure	1st wave: January 2005 - interview date W1
alg1_spells	Information on periods when the respondent was receiving Unemployment Benefit I	1st wave: January 2005 - interview date W1

The spell datasets surveyed or updated in the 2nd wave have a comparable structure. In addition to an identifier (household or personal ID number) they also contain a spell number, which numbers the individual spells within a household (*alg2_spells*) or a person (*et_spells*, *al_spells*, *lu_spells*, *mn_spells*) consecutively in chronological order and makes it possible to identify them clearly together with the household or personal ID

number. Furthermore, generated date variables for the beginning (*bmonat*, *bjahr*) and the end (*emonat*, *ejahr*) of the respective spell can be found in the datasets. These variables were recoded (e.g. information on seasons was recoded into definite months) and cleansed (e.g. missing codes were set for implausible values). In addition, if these variables contained censored spells, the interview date was entered for the end of the spell. In contrast, the date variables as they were reported by the respondent (e.g. *ET0100*, *ET0200*, *ET0300*, *ET0400* in the *et_spells*), which are also included, were not altered[52]. Following content-related information on the various spell types, all of the spell datasets contain a censoring indicator (*zensiert*) for spells that were still ongoing on the respective last interview date, in other words, right-censored spells. Generated variables (e.g. ISCO-88 coding of occupational activities) can be found at the end of each list of variables in the spell datasets.

Finally, some important peculiarities of the spell data in PASS and some particularities of the spells of Unemployment Benefit II receipt should be pointed out. Due to the orientation towards actual spells here, it is generally not easy to relate the spells to specific waves, as spells may span more than one survey date. Furthermore, observations are not available for all households or individuals in the spell data. This may be the case if there were no relevant spells or if the corresponding questions were not asked due to the filters.

For identifying individual spells, the identifier variable (*hnr* or *pnr*) and the spell number are always required for a clear selection, as there are often several observations available per household or person. This also has to be taken into account when linking spell data and the household and individual datasets. As several spells are frequently available and there is also no wave indicator for the individual observations in the spell data, a wave-specific reference is not possible without further work.

Compared with the other spell datasets, the Unemployment Benefit II spells also have a special feature. They contain wave-specific cross-sectional information, embedded in the individual episodes. Cross-sectional variables for the 1st and 2nd waves contain the information gathered only for currently ongoing spells, details as to the household members for whom benefit is being received, the amount of benefit received and details about whether an Unemployment Benefit II cut has begun. If a spell refers only to

[52] Exceptions to this are the merging of two spells and the spells of Unemployment Benefit II receipt sur-

the 1st wave, then the cross-sectional variables[53] for the 2nd wave are given the special code "-9" (item not surveyed in wave) and vice versa. In spells that were currently ongoing in both survey waves, the variables referring to the interview date of the 1st wave and those referring to the interview date of the 2nd wave are filled in.

The following example demonstrates the generation of a variable containing the latest information about the amount of benefit received per month for each Unemployment Benefit II spell. Variables for the other cross-sectional information can be generated in the same way.

First a new variable is created, *hoehebez,* which is assigned code -3 (not applicable), as details about the amount of benefit received are only available for Unemployment Benefit II spells that were still ongoing on the interview date in at least one wave. Then in a loop the generated variable is filled with the information from *AL20800* (amount of benefit received per month in wave 1) or *AL20801* (amount of benefit received per month in wave 2). Information is only incorporated into *hoehebez,* however, when it does not involve the values "-3" (not applicable) or "-9" (item not surveyed in wave). A cross-sectional variable on the amount of benefit received is given the value "-3" if information about the spell was gathered in the respective wave (new details surveyed or previous details updated) but the spell was not ongoing on the interview date. The variable is assigned the code "-9" if no information was collected about this spell in the respective wave. First *hoehebez* is filled with the information on the amount of benefit received which is contained in the cross-sectional variable for wave 1 (*AL20800*) and then, in the second loop run, is replaced by the values of the cross-sectional variable referring to wave 2 (*AL20801*). In this way the latest available information for this spell is taken into *hoehebez*.

```
use al2_spells.dta, clear
gen hoehebez = -3
foreach var of varlist AL20800 AL20801 {
            replace hoehebez = `var' if `var' ~= -3 & `var' ~= -9
}
```

veyed in the first wave.

[53] AL20601, AL20701a to AL20701o, AL20801 and AL20901. The wave to which a cross-sectional variable refers is recorded in the variable label. See Chapter 8.6 on this subject.

11. References

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

Achatz, Juliane, Hirseland, Andreas und Promberger, Markus (2007), "IAB-Panelbefragung von Haushalten im Niedrigeinkommensbereich. Entwurf für ein Rahmenkonzept", S. 11-32 in Promberger, Markus (Hrsg.), Neue Daten für die Sozialstaatsforschung: Zur Konzeption der IAB-Panelerhebung 'Arbeitsmarkt und Soziale Sicherung', IAB Forschungsbericht 12/2007, Nürnberg.

Andreß Hans-Jürgen, Burkatzki, Eckhard, Lipsmeier, Gero, Salentin, Kurt, Schulte, Katja und Strengmann-Kuhn, Wolfgang (1996): Leben in Armut. Analysen der Verhaltensweisen armer Haushalte mit Umfragedaten. Endbericht des DFG-Projekts „Versorgungsstrategien privater haushalte im unteren Einkommensbereich (VuE)". Bielefeld.

Andreß, Hans-Jürgen und Lipsmeier, Gero (1995): Was gehört zum notwendigen Lebensstandard und wer kann ihn sich leisten? Ein neues Konzept zur Armutsmessung. Aus Politik und Zeitgeschichte, B 31-32/1995: 35-49.

Andreß, Hans-Jürgen und Lipsmeier, Gero (2001): Armut und Lebensstandard. Gutachten im Rahmen des Armuts- und Reichtumsberichts der Bundesregierung. BMAS. Bonn.

Beckmann, Petra und Trometer, Reiner (1991): Neue Dienstleistungen des ALLBUS: Haushalts- und Familientypologien, Klassenschema nach Goldthorpe. ZUMA-Nachrichten 28: 7-17.

Brauns, Hildegard and Steinmann, Susanne (1999): Educational Reform in France, West-Germany and the United Kingdom: Updating the Casmin Classification. ZUMA-Nachrichten 44: 7-45

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

Büngeler, Kathrin, Gensicke, Miriam, Hartmann, Josef, Jäckle, Robert, Tschersich, Nikolai (2009): IAB-Haushaltspanel m Niedrigeinkommensbereich Welle 2 (2007/2008). Methoden- und Feldbericht. FDZ-Methodenreport, 08/2009, Nürnberg.

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

Christoph, Bernhard, Müller, Gerrit, Gebhardt, Daniel, Wenzig, Claudia, Trappmann, Mark, Achatz, Juliane, Tisch, Anita und Gayer, Christine (2008): Codebuch und Dokumentation des Panel ‚Arbeitsmarkt und soziale Sicherung' Welle 2 (2006/2007). FDZ-Datenreport 05/2008, Nürnberg.

Christoph, Bernhard, Sachweh, Patrick und Müller, Tim (2005): Umfrage zum Forschungsprojekt „Die Akzeptanz des Wohlfahrtsstaats". Codebuch und Dokumentation. Mannheimer Zentrum für Europäische Sozialforschung (MZES), Mannheim.

Europäische Gemeinschaften (EG) (2202): Verordnung (EG) Nr. 29/2002 der Kommision 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, Brüssel.

Erikson, Robert and Goldthorpe, John H. (1992): The Constant Flux. A Study of Class Mobility in Industrial Society. Clarendon Press, Oxford.

Erikson, Robert, Goldthorpe, John H. and Portocarero, Lucienne (1979): Intergenerational Class Mobility in Three Western Societies: England, France and Sweden. British Journal of Sociology 30: 415-441.

Erikson, Robert, Goldthorpe, John H. and Portocarero, Lucienne (1982): Social Fluidity in Industrial Nations: England, France and Sweden. British Journal of Sociology 33: 1-34.

Fischer, Andreas and Wirth, Heike (2007): Constructing Version 4 of ESEC Classes from 3-digit ISCO (Stata-do file). Gesis-ZUMA, Mannheim

Frick, Joachim R.; Göbel, Jan and Krause, Peter (n.d.): $HGEN: Generated Household-Level Variables. Download: [http://www.diw.de/documents/dokumentenarchiv/17/60053/hgen.pdf (8.11.2007)].

Ganzeboom, Harry B. G. and Treiman, Donald (1996): Internationally Comparable Measures for Occupational Status for the 1988 International Standard Classification of Occupations. Social Science Research, 25: 201-239.

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

Ganzeboom, Harry B. G.; De Graaf, Paul M. and Treiman, Donald J. (1992). A Standard International Socio-Economic Index of Occupational Status. Social Science Research, 21: 1-56.

Graf, Tobias (2007). Bedarfsgemeinschaften 2005 und 2006: Die Hälfte war zwei Jahre lang durchgehend bedürftig. IAB-Kurzbericht 17/2007.

Granato, Nadia (2000): CASMIN-Bildungsklassifikation. Eine Umsetzung mit dem Mikrozensus 1996. ZUMA-Technischer Bericht 2000/12.

Hagenaars, Aldi, J. M. de Vos, Klaas and Zaidi, M. Ashgar (1994): Poverty Statistics in the Late 1980s: Research Based on Micro-data. Office for Official Publications of the European Communities, Luxembourg.

Halleröd, Björn (1995): The Truly Poor: Direct and Indirect Consensual Measurement of Poverty in Sweden. Journal of European Social Policy. Vol. 5, No. 2: 111-129.

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

Hartmann, Josef, Brink, Kathrin, Jäckle, Robert und Tschersich, Niko (2008). IAB-Haushaltspanel im Niedrigeinkommensbereich: Methoden- und Feldbericht. München.

Hauser, Richard (1996): Zur Messung individueller Wohlfahrt und Ihrer Verteilung. S. 13-38 in Statistisches Bundesamt (Hg.): Wohlfahrtsmessung. Aufgabe der Statistik im gesellschaftlichen Wandel. Metzler-Poeschel, Stuttgart.

Helberger, Christof (1988): Eine Überprüfung der Linearitätsannahme der Humankapitaltheorie. S. 151-170 in: H.-J. Bodenhöfer (Hrsg.): Bildung, Beruf, Arbeitsmarkt. Duncker & Humblot, Berlin.

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

Jäckle, Annette (2008): The Causes of Seam Effects in Panel Surveys. ISEP Working Paper Series, No. 2008-14, University of Essex.

Kiesl, Hans (2009): Kalibrierte Hochrechnung für das Panel "Arbeitsmarkt und soziale Sicherung". FDZ-Methodenreport x/2009. Forschungsdatenzentrum der BA im IAB, Nürnberg [im Erscheinen].

König, Wolfgang, Lüttinger, Paul und Müller, Walter (1987): Eine vergleichende Analyse der Entwicklung und Struktur von Bildungssystemen. Methodologische Grundlagen und Konstruktion einer vergleichbaren Bildungsskala. CASMIN-Projekt. Arbeitspapier Nr. 12.

Lechert, Yvonne; Schroedter, Julia; Lüttinger, Paul (2006): Die Umsetzung der Bildungsklassifikation CASMIN für die Volkszählung 1970, die Mikrozensus-Zusatzerhebung 1971 und die Mikrozensen 1976-2004. ZUMA-Methodenbericht 2006/12.

Lengerer, Andrea; Bohr, Jeanette und Janßen, Andrea (2005): Haushalte, Familien und Lebensformen im Mikrozensus – Konzepte und Typisierungen. ZUMA-Arbeitsbericht Nr. 2005/05; ZUMA; Mannheim.

Lipsmeier, Gero (1999): Die Bestimmung des notwendigen Lebensstandards – Einschätzungsunterschiede und Entscheidungsprobleme. Zeitschrift für Soziologie, Jg. 28, Heft 4: 281-300.

Müller, Walter, Wirth, Heike, Bauer, Gerrit, Pollak, Reinhard und Weiss, Felix (2006): ESeC – Kurzbericht zur Validierung und Operationalisierung einer europäischen sozioökonomischen Klassifikation. ZUMA-Nachrichten 59: 111–119.

Müller, Walter, Wirth, Heike, Bauer, Gerrit, Pollak, Reinhard und Weiss, Felix (2007): Entwicklung einer europäischen sozioökonomischen Klassifikation. Wirtschaft und Statistik, 5/2007: 527-530.

Nolan, Brian and Whelan, Christopher T. (1996). Measuring Poverty Using Income and Deprivation Indicators: Alternative Approaches. Journal of European Social Policy, Vol. 6, No. 3: 225-240

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

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

Promberger, Markus (2007) (Hrsg.): Neue Daten für die Sozialstaatsforschung: Zur Konzeption der IAB-Panelerhebung 'Arbeitsmarkt und Soziale Sicherung', IAB Forschungsbericht 12/2007, Nürnberg.

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

Rendtel, Ulrich, Harms, Torsten (2009): Weighting and calibration for household panels, pp. 265-286 in Lynn, Peter (ed.), Methodology of Longitudinal Surveys. Chichester, Wiley.

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

Rose, Richard and Harrison, Eric (2007): The European Socio-Economic Classification: A New Social Class Schema for Comparative European Research. European Societies 9, 3: 459-490.

Rudolph, Helmut und Trappmann, Mark (2007), "Design und Stichprobe des Panels „Arbeitsmarkt und Soziale Sicherung" (PASS)", S. 60-101 in Promberger, Markus. (Hrsg.), Neue Daten für die Sozialstaatsforschung: Zur Konzeption der IAB-Panelerhebung "Arbeitsmarkt und Soziale Sicherung", IAB Forschungsbericht 12/2007, Nürnberg.

Särndal, Carl-Erik, Swensson, Bengt and Wretman, Jan (1992): Model Assisted Survey Sampling. Springer, New York.

Schnell, Rainer (2007). Alternative Verfahren zur Stichprobengewinnung für ein Haushaltspanelsurvey mit Schwerpunkt im Niedrigeinkommens- und Transferleistungsbezug, S. 33-59, in Promberger, Markus. (Hrsg.), Neue Daten für die Sozialstaatsforschung: Zur Konzeption der IAB-Panelerhebung "Arbeitsmarkt und Soziale Sicherung", IAB Forschungsbericht 12/2007, Nürnberg.

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

Stata Corp (2007): Survey Data Reference Manual Release 10. Stata Press, College Station.

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

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

Trappmann, Mark, Christoph, Bernhard, Achatz, Juliane and Wenzig, Claudia (2007). "Labour Market and Social Security": A New Panel Study for Research on Long Term Unemployment. Paper Presented at the International Conference of the German Association of Political Economy, Nuremberg, October 11th, 2007. http://doku.iab.de/veranstaltungen/2007/cape_2007_trappmann.pdf

Trappmann, Mark, Christoph, Bernhard, Achatz, Juliane, Wenzig, Claudia, Müller, Gerrit, Gebhardt, Daniel (2009): Design and stratification of PASS. A New Panel Study for Research on Long Term Unemployment. IAB-Discussion Paper, 5/2009, Nuremberg.

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

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

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

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

Appendix I: Correspondence lists and ISCO-88 occupational classification

Table 42: List of correspondence between the household questionnaires and the household dataset (HHENDDAT) in wave 2

Variable name in SUF	Surveyed in		Description	No. in wave 2 questionn.	
	W1	W2		HHneu	HHalt
hnr	yes	yes	Household number (current)	system	system
uhnr	yes	yes	Original household number	system	system
welle	yes	yes	Indicator for survey wave	system	system
hintdat	yes	yes	Date of household interview	generated	generated
hintjahr	yes	yes	Date of household interview - year, generated	system	system
hintmon	yes	yes	Date of household interview - month, generated	system	system
hinttag	yes	yes	Date of household interview - day, generated	system	system
hintmod	yes	yes	System variable: interview mode of household interview	system	system
hsprache	yes	yes	System variable: language in which household interview conducted	system	system
sample	yes	yes	Sample indicator	system	system
psu	yes	yes	Primary sampling unit	system	system
strpsu	yes	yes	Stratification of primary sampling unit	system	system
nextstra	yes	yes	Stratification of primary sampling unit: code of neighbouring stratum	generated	generated
blneualt	yes	yes	Western or eastern Germany, generated	generated	generated
bundesld	yes	yes	Federal state, generated	generated	generated
bik	no	yes	BIK region size classes (GKBIK10), generated	generated	generated
HW0100	yes	no	Indicator: primary or secondary residence of TP?	not in wave	not in wave
HW0200	yes	yes	Indicator: form of accommodation?	Anf36	generated
HW0300	yes	yes	Indicator: type of tenancy	Anf37	generated
HW0400	yes	yes	Indicator: type of hostel/home/hall of residence etc.	Anf38	generated
HW0500	yes	no	Owner, principle tenant, rent-free: other h'holds living in dwelling?	not in wave	not in wave
HW0600	yes	no	Owner, principle tenant, rent-free: no. of other h'holds in dwelling?	not in wave	not in wave
HW0700	yes	no	Sub-tenant: are other h'holds living in the dwelling?	not in wave	not in wave
HW0800	yes	no	Sub-tenant: no. of other h'holds in dwelling?	not in wave	not in wave
HW0880a	no	yes	Reason for moving out: size and equipment/furnishings of dwelling	n. in Q vers.	HH64a
HW0880b	no	yes	Reason for moving out: residential area	n. in Q vers.	HH64b
HW0880c	no	yes	Reason for moving out: work-related reasons/distance from work	n. in Q vers.	HH64c
HW0880d	no	yes	Reason for moving out: family-related reasons	n. in Q vers.	HH64d
HW0880e	no	yes	Reason for moving out: accommodation costs too high	n. in Q vers.	HH64e
HW0880f	no	yes	Reason for moving out: landlord gave notice	n. in Q vers.	HH64f
HW0880g	no	yes	Reason for moving out: own house/apartment	n. in Q vers.	HH64g
HW0880h	no	yes	Reason for moving out: emp. agency/jobcentre wanted the move	n. in Q vers.	HH64h
HW0880i	no	yes	Reason for moving out: other reason	n. in Q vers.	HH64i
HW0881a	no	yes	Reason for moving out: size and equipment/furnishings of dwelling, incl. open-ended info., gen.	n. in Q vers.	generated
HW0881b	no	yes	Reason for moving out: residential area, incl. open-ended info., gen.	n. in Q vers.	generated
HW0881c	no	yes	Reason for moving out: work-related reasons/distance from work, incl. open-ended info., gen.	n. in Q vers.	generated
HW0881d	no	yes	Reason for moving out: family-related reasons, incl. o-e info., gen.	n. in Q vers.	generated
HW0881e	no	yes	Reason for moving out: accommodation costs too high, incl. open-ended info., gen.	n. in Q vers.	generated
HW0881f	no	yes	Reason for moving out: landlord gave notice, incl. o-e info., gen.	n. in Q vers.	generated
HW0881g	no	yes	Reason for moving out: own house/apartment, incl. open-ended info., gen.	n. in Q vers.	generated
HW0881h	no	yes	Reason for moving out: emp. agency/jobcentre wanted the move, incl. open-ended info., gen.	n. in Q vers.	generated
HW0881i	no	yes	Reason for moving out: dwelling in unacceptable condition, incl. open-ended info., gen.	n. in Q vers.	generated
HW0881j	no	yes	Reason for moving out: other reason, incl. open-ended info., gen.	n. in Q vers.	generated
HW0890	no	yes	Change in size of dwelling since last interview	n. in Q vers.	HH65
HW0900	yes	yes	Year of move into current dwelling	HH5	HH66
HW0910	no	yes	In hostel/home/hall of residence: own rooms?	HH6	HH67
HW0920	no	yes	In hostel/home/hall of residence: bedroom shared with other pers.	HH7	HH68
HW1000	yes	yes	Living space in m^2	HH8	HH69
HW1100	yes	yes	Hostel/home/hall of residence: rent incl. running costs and energy	HH9	HH70
HW1200	yes	yes	Tenant: rent paid to landlord (incl. additional costs, if applicable)	HH10	HH71

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn. HHneu	No. in wave 2 questionn. HHalt
HW1300	yes	yes	Tenant: other housing costs	HH11	HH72
HW1400	yes	yes	No details provided on rent: total rent incl. all additional costs	HH12	HH73
HW1500	yes	yes	Owner: regular costs for apartment/house	HH15	HH76
HW1600	yes	yes	Owner: mortgage/loan repayments for apartment/house?	HH13	HH74
HW1700	yes	yes	Owner: amount of mortgage/loan repayments	HH14	HH75
HW1800	yes	yes	Indicator: receipt of housing assistance?	HH16	HH77
HW1900	yes	yes	Housing assistance: monthly amount received	HH17	HH78
HW2000	yes	yes	Condition of house/building, subjective assessment?	HH18	HH79
HW2100	yes	yes	Quality of residential area, subjective assessment?	HH19	HH80
HA0100	yes	yes	Size of household	HH1	generated
HA0200	yes	no	Assessment: protection against poverty	not in wave	not in wave
HA0250a	no	yes	UB II: ending month of spell ongoing in previous wave	HH48	HH91
HA0250b	no	yes	UB II: ending year of spell ongoing in previous wave	HH48	HH91
HA0300	yes	yes	Indicator: UB II since previous wave/house move/January of the year before start of fieldwork in this wave?	HH49	HH92
HA0400	yes	yes	Indicator: UB II in July of 1st field year of wave?	HH62	n. in Q vers.
HA0500	yes	yes	Household's current standard of living: subjective assessment	HH93	HH137
HA0600	yes	yes	Household's future standard of living: subjective expectations	HH94	HH138
HLS0100a	yes	yes	Deprivation: dwelling with enough rooms?	HH2a_01	HH61a_01
HLS0100b	yes	yes	Deprivation: why no dwelling with enough rooms?	HH2a_02	HH61a_02
HLS0200a	yes	yes	Deprivation: dwelling without damp walls/floors?	HH2b_01	HH61b_01
HLS0200b	yes	yes	Deprivation: why no dwelling without damp walls/floors?	HH2b_02	HH61b_02
HLS0300a	yes	yes	Deprivation: bathroom inside dwelling?	HH2c_01	HH61c_01
HLS0300b	yes	yes	Deprivation: why no bathroom inside dwelling?	HH2c_02	HH61c_02
HLS0400a	yes	yes	Deprivation: indoor toilet?	HH2d_01	HH61d_01
HLS0400b	yes	yes	Deprivation: why no indoor toilet?	HH2d_02	HH61d_02
HLS0500a	yes	yes	Deprivation: central heating/heating system covering one floor/district heating?	HH2e_01	HH61e_01
HLS0500b	yes	yes	Deprivation: why no central heating/heating system covering one floor/district heating?	HH2e_02	HH61e_02
HLS0600a	yes	yes	Deprivation: garden/balcony/terrace?	HH2f_01	HH61f_01
HLS0600b	yes	yes	Deprivation: why no garden/balcony/terrace?	HH2f_02	HH61f_02
HLS0700a	yes	yes	Deprivation: sufficient winter clothing?	HH2g_01	HH61g_01
HLS0700b	yes	yes	Deprivation: why no sufficient winter clothing?	HH2g_02	HH61g_02
HLS0800a	yes	yes	Deprivation: car?	HH2h_01	HH61h_01
HLS0800b	yes	yes	Deprivation: why no car?	HH2h_02	HH61h_02
HLS0900a	yes	yes	Deprivation: television?	HH2i_01	HH61i_01
HLS0900b	yes	yes	Deprivation: why no television?	HH2i_02	HH61i_02
HLS1000a	yes	yes	Deprivation: video recorder / DVD player?	HH2j_01	HH61j_01
HLS1000b	yes	yes	Deprivation: why no video recorder / DVD player?	HH2j_02	HH61j_02
HLS1100a	yes	yes	Deprivation: computer with internet access?	HH2k_01	HH61k_01
HLS1100b	yes	yes	Deprivation: why no computer with internet access?	HH2k_02	HH61k_02
HLS1200a	yes	yes	Deprivation: washing machine?	HH2l_01	HH61l_01
HLS1200b	yes	yes	Deprivation: why no washing machine?	HH2l_02	HH61l_02
HLS1300a	yes	yes	Deprivation: freezer/refrigerator with freezer compartment?	HH2m_01	HH61m_01
HLS1300b	yes	yes	Deprivation: why no freezer/refrigerator with freezer compartment?	HH2m_02	HH61m_02
HLS1400a	yes	yes	Deprivation: buying new clothes occasionally?	HH3a_01	HH62a_01
HLS1400b	yes	yes	Deprivation: why no new clothes bought occasionally?	HH3a_02	HH62a_02
HLS1500a	yes	yes	Deprivation: one warm meal per day?	HH3b_01	HH62b_01
HLS1500b	yes	yes	Deprivation: why not one warm meal per day?	HH3b_02	HH62b_02
HLS1600a	yes	yes	Deprivation: a one-week holiday per year?	HH3c_01	HH62c_01
HLS1600b	yes	yes	Deprivation: why not a one-week holiday per year?	HH3c_02	HH62c_02
HLS1700a	yes	yes	Deprivation: inviting friends for dinner once per month?	HH3d_01	HH62d_01
HLS1700b	yes	yes	Deprivation: why not inviting friends for dinner once per month?	HH3d_02	HH62d_02
HLS1800a	yes	yes	Deprivation: eating out once per month?	HH3e_01	HH62e_01
HLS1800b	yes	yes	Deprivation: why not eating out once per month?	HH3e_02	HH62e_02
HLS1900a	yes	yes	Deprivation: trip to the theatre/cinema/a concert once per month?	HH3f_01	HH62f_01
HLS1900b	yes	yes	Deprivation: why no trip to theatre/cinema/a concert once per month?	HH3f_02	HH62f_02
HLS2000a	yes	yes	Deprivation: saving a regular amount per month?	HH3g_01	HH62g_01
HLS2000b	yes	yes	Deprivation: why not saving a regular amount per month?	HH3g_02	HH62g_02
HLS2100a	yes	yes	Deprivation: replace worn-out furniture?	HH3h_01	HH62h_01
HLS2100b	yes	yes	Deprivation: why not replacing worn-out furniture?	HH3h_02	HH62h_02

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn. HHneu	No. in wave 2 questionn. HHalt
HLS2200a	yes	yes	Deprivation: pay unexpected expenses oneself?	HH3i_01	HH62i_01
HLS2200b	yes	yes	Deprivation: why not paying unexpected expenses oneself?	HH3i_02	HH62i_02
HLS2300a	yes	yes	Deprivation: health treatment not reimbursed by health insurance?	HH3j_01	HH62j_01
HLS2300b	yes	yes	Deprivation: why no health treatment not reimbursed by health ins.?	HH3j_02	HH62j_02
HLS2400a	yes	yes	Deprivation: pay rent punctually?	HH3l_01	HH62l_01
HLS2400b	yes	yes	Deprivation: why not paying rent punctually?	HH3l_02	HH62l_02
HLS2500a	yes	yes	Deprivation: pay utilities bills punctually?	HH3m_01	HH62m_01
HLS2500b	yes	yes	Deprivation: why not paying utilities bills punctually?	HH3m_02	HH62m_02
HLS2600a	yes	yes	Deprivation: over-the-counter medicines?	HH3k_01	HH62k_01
HLS2600b	yes	yes	Deprivation: why no over-the-counter medicines?	HH3k_02	HH62k_02
HD0100a	yes	yes	Person 1: gender	HH27_01 / generated	generated
HD0100b	yes	yes	Person 2: gender	HH27_02 / generated	generated
HD0100c	yes	yes	Person 3: gender	HH27_03 / generated	generated
HD0100d	yes	yes	Person 4: gender	HH27_04 / generated	generated
HD0100e	yes	yes	Person 5: gender	HH27_05 / generated	generated
HD0100f	yes	yes	Person 6: gender	HH27_06 / generated	generated
HD0100g	yes	yes	Person 7: gender	HH27_07 / generated	generated
HD0100h	yes	yes	Person 8: gender	HH27_08 / generated	generated
HD0100i	yes	yes	Person 9: gender	HH27_09 / generated	generated
HD0100j	yes	yes	Person 10: gender	HH27_10 / generated	generated
HD0100k	yes	yes	Person 11: gender	HH27_11 / generated	generated
HD0100l	yes	yes	Person 12: gender	HH27_12 / generated	generated
HD0100m	1	1	Person 13: gender	HH27_13 / generated	generated
HD0100n	1	1	Person 14: gender	HH27_14 / generated	generated
HD0100o	1	1	Person 15: gender	HH27_15 / generated	generated
HD0200a	1	1	Person 1: age	HH28_01 / generated	generated
HD0200b	1	1	Person 2: age	HH28_02 / generated	generated
HD0200c	1	1	Person 3: age	HH28_03 / generated	generated
HD0200d	1	1	Person 4: age	HH28_04 / generated	generated
HD0200e	1	1	Person 5: age	HH28_05 / generated	generated
HD0200f	1	1	Person 6: age	HH28_06 / generated	generated
HD0200g	1	1	Person 7: age	HH28_07 / generated	generated
HD0200h	1	1	Person 8: age	HH28_08 / generated	generated
HD0200i	1	1	Person 9: age	HH28_09 / generated	generated
HD0200j	1	1	Person 10: age	HH28_10 / generated	generated
HD0200k	1	1	Person 11: age	HH28_11 / generated	generated

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn. HHneu	No. in wave 2 questionn. HHalt
HD0200l	1	1	Person 12: age	HH28_12 / generated	generated
HD0200m	1	1	Person 13: age	HH28_13 / generated	generated
HD0200n	1	1	Person 14: age	HH28_14 / generated	generated
HD0200o	1	1	Person 15: age	HH28_15 / generated	generated
HD0300	1	1	Ind.: change in HH since Jan. of year before 1st field year of wave?	HH39	n. in Q vers.
HD0400a	1	1	Date of last move into / out of household (month)	HH40	n. in Q vers.
HD0400b	1	1	Date of last move into / out of household (year)	HH40	n. in Q vers.
HD0500a	1	1	Date of household formation (month)	HH41	n. in Q vers.
HD0500b	1	1	Date of household formation (year)	HH41	n. in Q vers.
HD0600	no	1	Main language spoken in household	HH42	HH85
HD0601	no	1	Main language spoken in household, incl. open-ended info., gen.	generated	generated
HD0700	no	1	Indicator: other language spoken in household?	HH43	HH86
HD0800	no	1	Name of other household language	HH44	HH87
HD0801	no	1	Name of other household language, incl. open-ended info., gen.	generated	generated
HD0900	no	1	Two household languages: 1st language	HH45	HH88
HD0901	no	1	Two household languages: 1st lang., incl. open-ended info., gen.	generated	generated
HD1000	no	1	Two household languages: 2nd language	HH46	HH89
HD1001	no	1	Two household languages: 2nd lang. incl. open-ended info., gen.	generated	generated
HEK0100	1	1	Ind.: receipt of social assistance or social security for sen. citizens?	HH63	HH105
HEK0110	no	1	Soc. assistance or soc. security for sen. cit's: amount per month	HH64	HH106
HEK0120	no	1	Indicator: receipt of nursing allowance	HH65	HH107
HEK0130	no	1	Nursing allowance: amount received per month	HH66	HH108
HEK0200	1	1	Indicator: receipt of payments from other individuals	HH67	HH109
HEK0300	1	1	Payments from other individuals: amount per month	HH68	HH110
HEK0400	1	1	Indicator: HH pays money to other individuals	HH69	HH111
HEK0500	1	1	Payments to other individuals: amount per month	HH70	HH112
HEK0600	1	1	Household income, open-ended info. (in euros)	HH77	HH121
HEK0700	1	1	Household income, indicator under/over € 1,000	HH78	HH122
HEK0800	1	1	Household income, categories under € 1,000	HH79	HH123
HEK0900	1	1	Household income, indicator under/over € 3,000	HH80	HH124
HEK1000	1	1	Household income, categories € 1,000 to € 3,000	HH81	HH125
HEK1100	1	1	Household income, categories over € 3,000	HH82	HH126
HEK1200	1	1	Household savings (in euros)	HH83	HH127
HEK1300a	1	1	Debts and loans: consumer credit	HH84a	HH128a
HEK1300b	1	1	Debts and loans: business loans	HH84b	HH128b
HEK1300c	1	1	Debts and loans: personal loan	HH84c	HH128c
HEK1300d	1	1	Debts and loans: overdraft credit	HH84d	HH128d
HEK1300e	1	1	Debts and loans: other debts	HH84e	HH128e
HEK1400	1	1	Total amount of debts/loans (in euros)	HH85	HH129
HEK1500	1	1	Indicator: receipt of child allowance	HH71	HH113
HEK1600	1	1	Child allowance: no. of children for whom child allowance paid	HH72	HH114
HEK1610	no	1	Indicator.: receipt of parental allowance, parental benefit, maternity benefit	HH72a	HH115
HEK1620	no	1	Parent. allowance, parent. benefit, mat. benefit: amount / month	HH72b	HH116
HEK1700	1	1	Indicator: receipt of orphan's pension or similar	HH73	HH117
HEK1710	no	1	Orphan's pension or similar: amount received per month	HH74	HH118
HEK1800	1	1	Indicator: receipt of advance child maintenance payment	HH75	HH119
HEK1810	no	1	Advance child maintenance payment: amount received per month	HH76	HH120
HEK1900	1	no	Advance child maintenance payment: no. of children for whom payment received	not in wave	not in wave
HEK2000	no	1	Indicator: income from rent and leasing	HH20	HH81
HEK2100	no	1	Income from rent and leasing: amount per month	HH21	HH82
HEK2200	no	1	Indicator: other income from assets	HH22	HH83
HEK2300	no	1	Other income from assets: amount per year	HH23	HH84
HKI0100b	1	1	Child (Position 2): attends primary school?	HH87	HH131
HKI0200b	1	1	Child (Position 2): attends child-care facility?	HH88	HH132
HKI0300b	1	1	Child (Position 2): type of school attended	HH89	HH133
HKI0400b	1	1	Child (Position 2): child-care outside school	HH90	HH134
HKI0100c	1	1	Child (Position 3): attends primary school?	HH87	HH131

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn. HHneu	No. in wave 2 questionn. HHalt
HKI0200c	1	1	Child (Position 3): attends child-care facility?	HH88	HH132
HKI0300c	1	1	Child (Position 3): type of school attended	HH89	HH133
HKI0400c	1	1	Child (Position 3): child-care outside school	HH90	HH134
HKI0100d	1	1	Child (Position 4): attends primary school?	HH87	HH131
HKI0200d	1	1	Child (Position 4): attends child-care facility?	HH88	HH132
HKI0300d	1	1	Child (Position 4): type of school attended	HH89	HH133
HKI0400d	1	1	Child (Position 4): child-care outside school	HH90	HH134
HKI0100e	1	1	Child (Position 5): attends primary school?	HH87	HH131
HKI0200e	1	1	Child (Position 5): attends child-care facility?	HH88	HH132
HKI0300e	1	1	Child (Position 5): type of school attended	HH89	HH133
HKI0400e	1	1	Child (Position 5): child-care outside school	HH90	HH134
HKI0100f	1	1	Child (Position 6): attends primary school?	HH87	HH131
HKI0200f	1	1	Child (Position 6): attends child-care facility?	HH88	HH132
HKI0300f	1	1	Child (Position 6): type of school attended	HH89	HH133
HKI0400f	1	1	Child (Position 6): child-care outside school	HH90	HH134
HKI0100g	1	1	Child (Position 7): attends primary school?	HH87	HH131
HKI0200g	1	1	Child (Position 7): attends child-care facility?	HH88	HH132
HKI0300g	1	1	Child (Position 7): type of school attended	HH89	HH133
HKI0400g	1	1	Child (Position 7): child-care outside school	HH90	HH134
HKI0100h	1	1	Child (Position 8): attends primary school?	HH87	HH131
HKI0200h	1	1	Child (Position 8): attends child-care facility?	HH88	HH132
HKI0300h	1	1	Child (Position 8): type of school attended	HH89	HH133
HKI0400h	1	1	Child (Position 8): child-care outside school	HH90	HH134
HKI0100i	1	1	Child (Position 9): attends primary school?	HH87	HH131
HKI0200i	1	1	Child (Position 9): attends child-care facility?	HH88	HH132
HKI0300i	1	1	Child (Position 9): type of school attended	HH89	HH133
HKI0400i	1	1	Child (Position 9): child-care outside school	HH90	HH134
HKI0100j	1	1	Child (Position 10): attends primary school?	HH87	HH131
HKI0200j	1	1	Child (Position 10): attends child-care facility?	HH88	HH132
HKI0300j	1	1	Child (Position 10): type of school attended	HH89	HH133
HKI0400j	1	1	Child (Position 10): child-care outside school	HH90	HH134
HKI0100k	1	1	Child (Position 11): attends primary school?	HH87	HH131
HKI0200k	1	1	Child (Position 11): attends child-care facility?	HH88	HH132
HKI0300k	1	1	Child (Position 11): type of school attended	HH89	HH133
HKI0400k	1	1	Child (Position 11): child-care outside school	HH90	HH134
HKI0100l	1	1	Child (Position 12): attends primary school?	HH87	HH131
HKI0200l	1	1	Child (Position 12): attends child-care facility?	HH88	HH132
HKI0300l	1	1	Child (Position 12): type of school attended	HH89	HH133
HKI0400l	1	1	Child (Position 12): child-care outside school	HH90	HH134
HKI0100m	1	1	Child (Position 13): attends primary school?	HH87	HH131
HKI0200m	1	1	Child (Position 13): attends child-care facility?	HH88	HH132
HKI0300m	1	1	Child (Position 13): type of school attended	HH89	HH133
HKI0400m	1	1	Child (Position 13): child-care outside school	HH90	HH134
HKI0100n	1	1	Child (Position 14): attends primary school?	HH87	HH131
HKI0200n	1	1	Child (Position 14): attends child-care facility?	HH88	HH132
HKI0300n	1	1	Child (Position 14): type of school attended	HH89	HH133
HKI0400n	1	1	Child (Position 14): child-care outside school	HH90	HH134
HKI0100o	1	1	Child (Position 15): attends primary school?	HH87	HH131
HKI0200o	1	1	Child (Position 15): attends child-care facility?	HH88	HH132
HKI0300o	1	1	Child (Position 15): type of school attended	HH89	HH133
HKI0400o	1	1	Child (Position 15): child-care outside school	HH90	HH134
HKI0500	1	1	Child-care: monthly expenditure	HH91	HH135
HKI0600	1	1	Child-care: person responsible	HH92	HH136
HKI0700	1	no	Child-care: mother's employment restricted	not in wave	not in wave
HKI0800	1	no	Child-care: father's employment restricted	not in wave	not in wave
hhtyp	1	1	Household type, generated	generated	generated
kindu4	1	1	Control variable: child under age of 4 in HH	generated	generated
kindu13	1	1	Control variable: child under age of 13 in HH	generated	generated
kindu15	1	1	Control variable: child under age of 15 in HH	generated	generated
hhinckat	1	1	Categorised household income per month (in euros), generated	generated	generated
hhincome	1	1	Household income per month (in euros) incl. categorised info., generated	generated	generated
anzbg	1	1	Number of synthetic benefit communities in HH, generated	generated	generated

Variable name in SUF	Surveyed in W1	W2	Description	No. in wave 2 questionn. HHneu	HHalt
alg2abez	1	1	Current receipt of UB II of the HH, generated	generated	generated
nbgbezug	1	1	No. of benefit communities receiving benefits in HH as of sampling date	generated	generated
oecdinca	1	1	Equivalised household income, old OECD scale (rounded)	generated	generated
oecdincn	1	1	Equivalised household income, modified OECD scale (rounded)	generated	generated
depindug	1	1	Deprivation index, unweighted (items missing for financial reasons; total of unweighted items:26)	generated	generated
depindg	1	1	Deprivation index, weighted (items missing for financial reasons; total of weighted items:12.8)	generated	generated
wohnfl	1	1	Living space in m^2, generated	generated	generated
einzugj	1	1	Year of move into current dwelling, generated	generated	generated
umzug	no	1	Control variable: moved house since previous wave	n. in Q vers.	system

Table 43: List of correspondence between the standard personal and senior citizens' questionnaires and the individual dataset (*PENDDAT*) in wave 2

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn.: Personal	No. in wave 2 questionn.: Senior cit.
pnr	yes	yes	Constant personal ID number	system	system
hnr	yes	yes	Household number (current)	system	system
uhnr	yes	yes	Original household number	system	system
welle	yes	yes	Indicator for survey wave	system	system
pintdat	yes	yes	Date of personal interview	generated	generated
pintjahr	yes	yes	Date of personal interview - year, generated	system	system
pintmon	yes	yes	Date of personal interview - month, generated	system	system
pinttag	yes	yes	Date of personal interview - day, generated	system	system
pintmod	yes	yes	System variable: interview mode of personal interview	system	system
psprache	yes	yes	System variable: interview language of personal interview	system	system
fb_vers	yes	yes	System variable: questionnaire type	system	system
nachbef	yes	no	Follow-up questions: P76-P90, P93-P115, P116-P138	not in wave	not in wave
sample	yes	yes	Sample indicator	generated	generated
altbefr	no	yes	Control variable: re-interviewed person	system	system
zpsex	yes	yes	Control variable: gender of respondent (from HH grid)	system	system
zpalthh	yes	yes	Control variable: age of respondent (from HH grid)	system	system
palter	yes	yes	Age (from p1), generated	generated	generated
hhgr	yes	yes	Control variable: household size	system	system
PD0100	yes	yes	Year of birth (date of birth, anonymised)	P1	P1
PD0200	yes	yes	Indicator: member of a religious community	P208	P40
PD0300	yes	yes	Religious denomination	P209	P41
PD0400	yes	yes	Religiousness, self-assessment	P210	P42
PD0500	yes	yes	Marital status	P237	P66
PD0600	yes	yes	Indicator: steady partner outside the household?	P238	P67
PD0700	yes	yes	Correction loop: inconsistent details about spouse	P239	P68
PD0800	yes	yes	Correction loop: inconsistent details about partner	P241	P70
PD0900	yes	yes	Respondent with children in HH: children outside HH?	P243	n. in Q vers.
PD1000	yes	yes	Respondent without children in HH: children outside HH?	P244	n. in Q vers.
PD1100	yes	yes	Number of children outside the HH	P245	n. in Q vers.
PA0100	yes	yes	Satisfaction with health	P2a	P2a
PA0200	yes	yes	Satisfaction with dwelling	P2b	P2b
PA0300	yes	yes	Satisfaction with standard of living	P2c	P2c
PA0400	yes	no	Indicator: receipt of UB I since Jan. 2005? (wave 1 only)	not in wave	not in wave
PA0405	no	yes	Indicator: receipt of UB I in acc. with rule for people over 58?	P118	n. in Q vers.
PA0410	no	yes	Start of UB I in acc. with rule for people over 58 (month)	P119m	n. in Q vers.
PA0415	no	yes	Start of UB I in acc. with rule for people over 58 (year)	P119j	n. in Q vers.
PA0420	no	yes	Current receipt of UB I in acc. with rule for people over 58?	P120	n. in Q vers.
PA0425	no	yes	End of UB I in acc. with rule for people over 58 (month)	P121m	n. in Q vers.
PA0430	no	yes	End of UB I in acc. with rule for people over 58 (year)	P121j	n. in Q vers.
PA0435	no	yes	Ever registered as unemployed?	P122	n. in Q vers.
PA0440	no	yes	Number of times registered as unemployed	P123	n. in Q vers.
PA0445	no	yes	Duration of unemployment, total (months)	P124m	n. in Q vers.
PA0450	no	yes	Duration of unemployment, total (years)	P124j	n. in Q vers.
PA0455	no	yes	Duration of longest spell of unemployment (months)	P125m	n. in Q vers.
PA0460	no	yes	Duration of longest spell of unemployment (years)	P125j	n. in Q vers.
PA0500	yes	no	Currently no receipt of UB II:	not in wave	not in wave
PA0600	yes	no	Indicator: participation in youth training/employment measures	not in wave	not in wave
PA0700	yes	no	Indicator: participation in measure in acc. with SGBII	not in wave	not in wave
PA0710b	no	yes	Particip.: job-creation measure/work opportunity in the paid variant	P159b	n. in Q vers.
PA0710c	no	yes	Participation: job application course	P159c	n. in Q vers.
PA0710d	no	yes	Participation: course/ short-term training measure/ work experience	P159d	n. in Q vers.
PA0710e	no	yes	Participation: training/retraining with vocational qualification	P159e	n. in Q vers.
PA0710f	no	yes	Participation: other measure	P159f	n. in Q vers.
PA0711b	no	yes	Participation: job-creation measure/work opportunity in the paid variant (cleansed spell data)	generated	n. in Q vers.
PA0711c	no	yes	Participation: job application course (cleansed spell data)	generated	n. in Q vers.
PA0711d	no	yes	Participation: course/ short-term training measure/ work experience (cleansed spell data)	generated	n. in Q vers.
PA0711e	no	yes	Participation: training/retraining with vocational qualification (cleansed spell data)	generated	n. in Q vers.
PA0711f	no	yes	Participation: other measure (cleansed spell data)	generated	n. in Q vers.
PA0720a	no	yes	Participation: one-euro job	P160a	n. in Q vers.
PA0720b	no	yes	Particip.: job-creation measure/work opportunity in the paid variant	P160b	n. in Q vers.
PA0720c	no	yes	Participation: job application course	P160c	n. in Q vers.
PA0720d	no	yes	Participation: course/ short-term training measure/ work experience	P160d	n. in Q vers.

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn.: Personal	No. in wave 2 questionn.: Senior cit.
PA0720e	no	yes	Participation: training/retraining with vocational qualification	P160e	n. in Q vers.
PA0720f	no	yes	Participation: other measure	P160f	n. in Q vers.
PA0721a	no	yes	Participation: one-euro job (cleansed spell data)	generated	n. in Q vers.
PA0721b	no	yes	Participation: job-creation measure/work opportunity in the paid variant (cleansed spell data)	generated	n. in Q vers.
PA0721c	no	yes	Participation: job application course (cleansed spell data)	generated	n. in Q vers.
PA0721d	no	yes	Participation: course/ short-term training measure/ work experience (cleansed spell data)	generated	n. in Q vers.
PA0721e	no	yes	Participation: training/retraining with vocational qualification (cleansed spell data)	generated	n. in Q vers.
PA0721f	no	yes	Participation: other measure (cleansed spell data)	generated	n. in Q vers.
PA0800	yes	yes	Social participation	P199	P34
PA0900	yes	yes	Social position: top-bottom scale	P200	P35
PA0950a	no	yes	Leisure activities: television/DVD/video	P201a	n. in Q vers.
PA0950b	no	yes	Leisure activities: playing on computer	P201b	n. in Q vers.
PA0950c	no	yes	Leisure activities: surfing Internet / chatting online	P201c	n. in Q vers.
PA0950d	no	yes	Leisure activities: working on computer	P201d	n. in Q vers.
PA0950e	no	yes	Leisure activities: manual work/DIY	P201e	n. in Q vers.
PA0950f	no	yes	Leisure activities: listening to music	P201f	n. in Q vers.
PA0950g	no	yes	Leisure activities: playing music	P201g	n. in Q vers.
PA0950h	no	yes	Leisure activities: sports	P201h	n. in Q vers.
PA0950i	no	yes	Leisure activities: going to friends' parties	P201i	n. in Q vers.
PA0950j	no	yes	Leisure activities: discos/bars/pubs	P201j	n. in Q vers.
PA0950k	no	yes	Leisure activities: theatre/concert/cinema	P201k	n. in Q vers.
PA0950l	no	yes	Leisure activities: reading	P201l	n. in Q vers.
PA0950m	no	yes	Leisure activities: voluntary activities	P201m	n. in Q vers.
PA0950n	no	yes	Leisure activities: doing nothing	P201n	n. in Q vers.
PA0950o	no	yes	Leisure activities: spending time with partner	P201o	n. in Q vers.
PA0950p	no	yes	Leisure activities: meeting friends	P201P	n. in Q vers.
PA0950q	no	yes	Leisure activities: youth centre/ youth club / community centre	P201q	n. in Q vers.
PA0950r	no	yes	Leisure activities: family activities	P201r	n. in Q vers.
PA0950s	no	yes	Leisure activities: going to church / religious events	P201s	n. in Q vers.
PA0960	no	yes	Receipt of social assistance before Jan. 2005?	P261	n. in Q vers.
PA0970a	no	yes	Duration of receipt of social assistance, total (month)	P262m	n. in Q vers.
PA0970b	no	yes	Duration of receipt of social assistance, total (year)	P262j	n. in Q vers.
PA1000	yes	yes	General satisfaction with life	P300	P84
PB0100	yes	yes	Indicator: school pupil/student/trainee?	P4	n. in Q vers.
PB0200	yes	yes	School pupil: general-education or vocational school?	P5	n. in Q vers.
PB0210	no	yes	Year in which general school qualification reported in wave 1 attained (wave 2 only)	P6	n. in Q vers.
PB0220	no	yes	General school qualification gained since last interview?	P7	n. in Q vers.
PB0230	no	yes	Type of school qualification gained since last interview	P8	n. in Q vers.
PB0231	no	yes	Type of school qual. gained since last interview, incl. open info.	generated	n. in Q vers.
PB0300	yes	yes	Indicator: general-education school qualification?	P9	P4
PB0400	yes	yes	Highest general school qualification	P10	P5
PB0401	yes	yes	Highest general school qualification, incl. open-ended info.	generated	generated
PB0410	no	yes	Year in which highest general school qualification attained (first surveyed from wave 2 onwards)	P13	n. in Q vers.
PB0500	yes	yes	School pupil: highest school qualification aimed for	P14	n. in Q vers.
PB0600	yes	yes	School pupil: vocational qualification/studies after leaving school?	P15	n. in Q vers.
PB0700	yes	yes	School pupil: type of vocational qualification aimed for	P16	n. in Q vers.
PB0800a	yes	yes	School pupil: last grade in German	P17a	n. in Q vers.
PB0800b	yes	yes	School pupil: last grade in mathematics	P17b	n. in Q vers.
PB0800c	yes	yes	School pupil: last grade in English	P17c	n. in Q vers.
PB0900	yes	yes	School pupil: repeated a class?	P18	n. in Q vers.
PB1000	yes	yes	Foreign school qualification: equivalent German qualification	P11	n. in Q vers.
PB1001	yes	yes	Foreign school qual.: equivalent German qual., incl. open info.	generated	n. in Q vers.
PB1100	yes	yes	Foreign school qual.: equiv. German qual., duration of schooling	P12	n. in Q vers.
PB1160a	no	yes	Year qualification gained: training as semi-skilled worker, compact vocational training	P19a	n. in Q vers.
PB1160b	no	yes	Year qualification gained: apprenticeship, in-firm training	P19b	n. in Q vers.
PB1160c	no	yes	Year qualification gained: training at school for health-care occupations	P19c	n. in Q vers.
PB1160d	no	yes	Year qualification gained: training at full-time vocational school	P19d	n. in Q vers.
PB1160e	no	yes	Year qualification gained: master craftsman's cert., technician cert. or similar	P19e	n. in Q vers.
PB1160f	no	yes	Year qualification gained: cert. from college of adv. voc. studies	P19f	n. in Q vers.
PB1160g	no	yes	Year qualification gained: degree from univ. of applied sciences or college of education	P19g	n. in Q vers.
PB1160h	no	yes	Year qualification gained: university degree	P19h	n. in Q vers.

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn.: Personal	No. in wave 2 questionn.: Senior cit.
PB1160i	no	yes	Year qualification gained: other German qualification	P19i	n. in Q vers.
PB1160j	no	yes	Year qualification gained: foreign qualification	P19j	n. in Q vers.
PB1170	no	yes	Forgotten qualification	P20	n. in Q vers.
PB1180a	no	yes	Vocational qualification: training as semi-skilled worker, compact vocational training	P21a	n. in Q vers.
PB1180b	no	yes	Vocational qualification: apprenticeship, in-firm training	P21b	n. in Q vers.
PB1180c	no	yes	Vocational qual.: training at school for health-care occupations	P21c	n. in Q vers.
PB1180d	no	yes	Vocational qual.: training at full-time vocational school	P21d	n. in Q vers.
PB1180e	no	yes	Vocational qual.: master craftsman's cert., technician cert. or similar	P21e	n. in Q vers.
PB1180f	no	yes	Vocational qual.: cert. from college of adv. vocational studies	P21f	n. in Q vers.
PB1180g	no	yes	Vocational qual.: degree from univ. of applied sciences or college of education	P21g	n. in Q vers.
PB1180h	no	yes	Vocational qual.: university degree	P21h	n. in Q vers.
PB1180i	no	yes	Vocational qual.: other German qualification	P21i	n. in Q vers.
PB1180j	no	yes	Vocational qual.: foreign qualification	P21j	n. in Q vers.
PB1181a	no	yes	Vocational qual.: training as semi-skilled worker, compact vocational training, incl. open info.	generated	n. in Q vers.
PB1181b	no	yes	Vocational qual.: apprenticeship, in-firm training, incl. open info.	generated	n. in Q vers.
PB1181c	no	yes	Vocational qual.: school for health-care occup's, incl. open info.	generated	n. in Q vers.
PB1181d	no	yes	Vocational qual.: training at full-time voc. school, incl. open info.	generated	n. in Q vers.
PB1181e	no	yes	Vocational qual.: master craftsman's cert., technician cert. or similar, incl. open info.	generated	n. in Q vers.
PB1181f	no	yes	Vocational qual.: cert. from coll. of adv. voc. studies, incl. open info.	generated	n. in Q vers.
PB1181g	no	yes	Vocational qual.: degree from univ. of appl. sciences or college of education, incl. open info.	generated	n. in Q vers.
PB1181h	no	yes	Vocational qual.: university degree, incl. open-ended info.	generated	n. in Q vers.
PB1181i	no	yes	Vocational qual.: other German qualification, incl. open-ended info.	generated	n. in Q vers.
PB1181j	no	yes	Vocational qual.: foreign qualification, incl. open-ended info.	generated	n. in Q vers.
PB1190a	no	yes	Year qual. gained: training as semi-sk. worker, compact voc. training	P22a	n. in Q vers.
PB1190b	no	yes	Year qual. gained: apprenticeship, in-firm training	P22b	n. in Q vers.
PB1190c	no	yes	Year qual. gained: school for health-care occupations	P22c	n. in Q vers.
PB1190d	no	yes	Year qual. gained: training at full-time vocational school	P22d	n. in Q vers.
PB1190e	no	yes	Year qual. gained: master craftsman's cert., technician cert. or similar	P22e	n. in Q vers.
PB1190f	no	yes	Year qual. gained: cert. from college of adv. vocational studies	P22f	n. in Q vers.
PB1190g	no	yes	Year qual. gained: degree from univ. of appl. sciences or college of education	P22g	n. in Q vers.
PB1190h	no	yes	Year qual. gained: university degree	P22h	n. in Q vers.
PB1190i	no	yes	Year qual. gained: other German qualification	P22i	n. in Q vers.
PB1190j	no	yes	Year qual. gained: foreign qualification	P22j	n. in Q vers.
PB1200	yes	no	Indicator: vocational qual. or degree? (wave 1 only)	not in wave	not in wave
PB1200a	no	yes	Indicator: vocational qual. or degree since last interview? (from wave 2)	P23	n. in Q vers.
PB1200b	no	yes	Indicator: vocational qual. or degree before current training? (from wave 2)	P24	n. in Q vers.
PB1200c	no	yes	Indicator: vocational qual. or degree at any time? (from wave 2)	P25	P6
PB1300a	yes	yes	Vocational qual.: training as semi-skilled worker, compact vocational training	P26_1	P7a
PB1300b	yes	yes	Vocational qual.: apprenticeship, in-firm training	P26_2	P7b
PB1300c	yes	yes	Vocational qual.: school for health-care occupations	P26_3	P7c
PB1300d	yes	yes	Vocational qual.: training at full-time vocational school	P26_4	P7d
PB1300e	yes	yes	Vocational qual.: master craftsman's cert., technician cert. or similar	P26_5	P7e
PB1300f	yes	yes	Vocational qual.: cert. from college of adv. vocational studies	P26_6	P7f
PB1300g	yes	yes	Vocational qual.: degree from univ. of appl. sciences or college of education	P26_7	P7g
PB1300h	yes	yes	Vocational qual.: university degree	P26_8	P7h
PB1300i	yes	yes	Vocational qual.: other German qualification	P26_9	P7i
PB1300j	yes	yes	Vocational qual.: foreign qualification	P26_10	P7j
PB1301a	yes	yes	Vocational qual.: training as semi-skilled worker, compact vocational training, incl. open-ended info.	generated	generated
PB1301b	yes	yes	Vocational qual.: apprenticeship, in-firm training, incl. open info.	generated	generated
PB1301c	yes	yes	Vocational qual.: school for health-care occupations, incl. open info.	generated	generated
PB1301d	yes	yes	Vocational qual.: training at full-time voc. school, incl. open info.	generated	generated
PB1301e	yes	yes	Vocational qual.: master craftsman's cert., technician cert. or similar, incl. open-ended info.	generated	generated
PB1301f	yes	yes	Vocational qual.: cert. from college of adv. vocational studies, incl. open info.	generated	generated
PB1301g	yes	yes	Vocational qual.: degree from univ. of appl. sciences or college of education, incl. open-ended info.	generated	generated

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn.: Personal	No. in wave 2 questionn.: Senior cit.
PB1301h	yes	yes	Vocational qual.: university degree, incl. open-ended info.	generated	generated
PB1301i	yes	yes	Vocational qual.: other German qualification, incl. open-ended info.	generated	generated
PB1301j	yes	yes	Vocational qual.: foreign qualification, incl. open-ended info.	generated	generated
PB1310am	no	yes	End of training: training as semi-skilled worker, compact vocational training (month)	P30am	n. in Q vers.
PB1310aj	no	yes	End of training: training as semi-skilled worker, compact vocational training (year)	P30aj	n. in Q vers.
PB1310bm	no	yes	End of training: apprenticeship, in-firm training (month)	P30bm	n. in Q vers.
PB1310bj	no	yes	End of training: apprenticeship, in-firm training (year)	P30bj	n. in Q vers.
PB1310cm	no	yes	End of training: school for health-care occupations (month)	P30cm	n. in Q vers.
PB1310cj	no	yes	End of training: school for health-care occupations (year)	P30cj	n. in Q vers.
PB1310dm	no	yes	End of training: training at full-time vocational school (month)	P30dm	n. in Q vers.
PB1310dj	no	yes	End of training: training at full-time vocational school (year)	P30dj	n. in Q vers.
PB1310em	no	yes	End of training: master craftsman, technician or similar (month)	P30em	n. in Q vers.
PB1310ej	no	yes	End of training: master craftsman, technician or similar (year)	P30ej	n. in Q vers.
PB1310fm	no	yes	End of training: college of adv. vocational studies (month)	P30fm	n. in Q vers.
PB1310fj	no	yes	End of training: college of adv. vocational studies (year)	P30fj	n. in Q vers.
PB1310gm	no	yes	End of training: degree from univ. of appl. sciences or college of education (month)	P30gm	n. in Q vers.
PB1310gj	no	yes	End of training: degree from univ. of appl. sciences or college of education (year)	P30gj	n. in Q vers.
PB1310hm	no	yes	End of training: university degree (month)	P30hm	n. in Q vers.
PB1310hj	no	yes	End of training: university degree (year)	P30hj	n. in Q vers.
PB1310im	no	yes	End of training: other German qualification (month)	P30im	n. in Q vers.
PB1310ij	no	yes	End of training: other German qualification (year)	P30ij	n. in Q vers.
PB1310jm	no	yes	End of training: foreign qualification (month)	P30jm	n. in Q vers.
PB1310jj	no	yes	End of training: foreign qualification (year)	P30jj	n. in Q vers.
PB1310km	no	yes	End of training: current qualification (month)	P30km	n. in Q vers.
PB1310kj	no	yes	End of training: current qualification (year)	P30kj	n. in Q vers.
PB1400	yes	no	Indicator: training as semi-skilled worker, compact vocational training?	not in wave	not in wave
PB1500a	yes	yes	Degree: bachelor	P27_1	n. in Q vers.
PB1500b	yes	yes	Degree: Diplom, Master, Magister or Staatsexamen	P27_2	n. in Q vers.
PB1500c	yes	yes	Degree: PhD or habilitation	P27_3	n. in Q vers.
PB1600	yes	yes	Foreign vocational qual.: equivalent German qualification	P28	n. in Q vers.
PB1601	yes	yes	Foreign vocational qual.: equivalent German qual., incl. open info.	generated	n. in Q vers.
PB1700	yes	yes	Foreign vocational qual.: officially recognised in Germany?	P29	n. in Q vers.
PB1800	yes	no	Foreign vocational qual.: Duration of training/studies in years	not in wave	not in wave
PB1900	yes	yes	In training: qualification aimed for	P31	n. in Q vers.
PEO0100a	yes	yes	Skills: a solution to every problem	P3a	P3a
PEO0100b	yes	yes	Skills: also coping with unexpected occurrences	P3b	P3b
PEO0100c	yes	yes	Skills: no difficulties in putting aims into practice	P3c	P3c
PEO0100d	yes	yes	Skills: acting in unexpected situations	P3d	P3d
PEO0100e	yes	yes	Skills: always able to solve difficult problems	P3e	P3e
PEO0200a	yes	no	Work orientations: work is just a way to earn money	not in wave	not in wave
PEO0200b	yes	no	Work orientations: work is the most important thing in life	not in wave	not in wave
PEO0200c	yes	no	Work orientations: work gives me a feeling of belonging	not in wave	not in wave
PEO0200d	yes	no	Work orientations: work also without being dependent on wage	not in wave	not in wave
PEO0300a	yes	no	Important in job: earning a lot of money	not in wave	not in wave
PEO0300b	yes	no	Important in job: a job that is enjoyable	not in wave	not in wave
PEO0300c	yes	no	Important in job: good opportunities for advancement	not in wave	not in wave
PEO0300d	yes	no	Important in job: a secure job	not in wave	not in wave
PEO0300e	yes	no	Important in job: a job in which I can put my abilities to good use	not in wave	not in wave
PEO0400a	yes	yes	Family/work: a woman should be willing to reduce working hours for family	P133a	P29a
PEO0400b	yes	yes	Family/work: what women really want is a home and children	P133b	P29b
PEO0400c	yes	yes	Family/work: a working mother can have just as close a relationship with her children	P133c	P29c
PEO0400d	yes	yes	Family/work: husband's task: earning money, wife's task: household/family	P133d	P29d
PEO0410	no	yes	Dealing with money in partnership	P134	P30
PEO0420	no	yes	Free availability of money for oneself	P135	P31
PEO0430	no	yes	Amount of money available for oneself	P136	P32
PEO0440	no	yes	Decisions: larger purchases	P137a	P33a
PEO0450	no	yes	Decisions: organisation of leisure time	P137b	P33b
PEO0500	yes	yes	General aspirations for children's education	P246	n. in Q vers.
PEO0600a	yes	yes	Expected school qualification of child 1	P24701	n. in Q vers.

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn.: Personal	No. in wave 2 questionn.: Senior cit.
PEO0600b	yes	yes	Expected school qualification of child 2	P24702	n. in Q vers.
PEO0600c	yes	yes	Expected school qualification of child 3	P24703	n. in Q vers.
PEO0600d	yes	yes	Expected school qualification of child 4	P24704	n. in Q vers.
PEO0600e	yes	yes	Expected school qualification of child 5	P24705	n. in Q vers.
PEO0600f	yes	yes	Expected school qualification of child 6	P24706	n. in Q vers.
PEO0600g	yes	yes	Expected school qualification of child 7	P24707	n. in Q vers.
PEO0600h	yes	yes	Expected school qualification of child 8	P24708	n. in Q vers.
PEO0600i	yes	yes	Expected school qualification of child 9	P24709	n. in Q vers.
PEO0600j	yes	yes	Expected school qualification of child 10	P24710	n. in Q vers.
PEO0600k	yes	yes	Expected school qualification of child 11	P24711	n. in Q vers.
PEO0600l	yes	yes	Expected school qualification of child 12	P24712	n. in Q vers.
PEO0600m	yes	yes	Expected school qualification of child 13	P24713	n. in Q vers.
PEO0600n	yes	yes	Expected school qualification of child 14	P24714	n. in Q vers.
PEO0600o	yes	yes	Expected school qualification of child 15	P24715	n. in Q vers.
PEO0700a	yes	yes	Attitudes towards education and training: children should earn money quickly	P248a	n. in Q vers.
PEO0700b	yes	yes	Attitudes towards education and training: a good training is worth personal sacrifices	P248b	n. in Q vers.
PEO0700c	yes	yes	Attitudes towards education and training: education definitely important	P248c	n. in Q vers.
PEO0700d	yes	yes	Attitudes towards education and training: training costs a heavy burden	P248d	n. in Q vers.
PET0100	yes	yes	Indicator: employment for at least 1 hour per week	n. in Q vers.	P8
PET0150	no	yes	Indicator: employment since January 2005?	P38	n. in Q vers.
PET0151	no	yes	Indicator: employment since January 2005? (cleansed spell data)	generated	n. in Q vers.
PET0200	yes	no	In employment: one or more than one job?	not in wave	not in wave
PET0300	yes	no	In employment: apprenticeship/vocational training?	not in wave	not in wave
PET0400	yes	no	In employment: assisted employment?	not in wave	not in wave
PET0500	yes	yes	In employment: mini-job?	n. in Q vers.	P9
PET0510	no	yes	In employment: mini-job?	P80	n. in Q vers.
PET0600	yes	no	Mini-job: earnings in euros	not in wave	not in wave
PET0610	no	yes	Mini-job: earnings in euros	P81	P10
PET0700	yes	yes	Mini-job/1-euro job: weekly working hours	n. in Q vers.	P11
PET0800	yes	no	Type of economic inactivity (wave 1 only)	not in wave	not in wave
PET0801	yes	no	Type of economic inactivity, incl. open info. (wave 1 only)	not in wave	not in wave
PET0900	yes	no	Indicator: registered as unemployed? (surveyed in wave 1)	not in wave	not in wave
PET0910	no	yes	Indicator: registered as unempl./ particip. in measure since Jan. 2005? (first surveyed in wave 2)	P106	n. in Q vers.
PET0911	no	yes	Indicator: registered as unempl./ particip. in measure since Jan. 2005? (cleansed spell data from wave 2 onwards)	generated	n. in Q vers.
PET1000a	yes	yes	Pers. registered as unemployed: month of unemp. registration	generated	n. in Q vers.
PET1000b	yes	yes	Pers. registered as unemployed: year of unemployment registration	generated	n. in Q vers.
PET1100	yes	no	Pers. not in employment: in employment previously?	not in wave	not in wave
PET1110	no	yes	Pers. not in employment: in employment previously?	P82	n. in Q vers.
PET1200	yes	no	End of last job (year)	not in wave	not in wave
PET1200a	no	yes	Last job: ending month	P83m	n. in Q vers.
PET1200b	no	yes	Last job: ending year	P83j	n. in Q vers.
PET1210	no	yes	Last occupational status, simple classification (before January 2005) (anon.)	P84	n. in Q vers.
PET1220	no	yes	Last occup. status blue-collar worker: detailed info. (before Jan. 2005)	P85	n. in Q vers.
PET1230	no	yes	Last occup. status white-collar worker: detailed info. (before January 2005)	P86	n. in Q vers.
PET1250	no	yes	Last occup. status civil servant: detailed info., incl. soldiers (before January 2005) (anon.)	P87, P88	n. in Q vers.
PET1260	no	yes	Last occup. status self-employed: no. of employees (before Jan. 2005)	P89	n. in Q vers.
PET1270	no	yes	Last occup. status farmer: area farmed (before January 2005)	P90	n. in Q vers.
PET1211	no	yes	Last occup. status, simple class. (incl. spell info.) (anon.), gen.	generated	n. in Q vers.
PET1221	no	yes	Last occup. status blue-collar worker: detailed info. (incl. spell info.), gen.	generated	n. in Q vers.
PET1231	no	yes	Last occup. status while-collar worker: detailed info. (incl. spell info.), gen.	generated	n. in Q vers.
PET1251	no	yes	Last occup. status civil servant: detailed info., incl. soldiers (incl. spell info.) (anon.), gen.	generated	n. in Q vers.
PET1261	no	yes	Last occup. status self-employed: no. of employees (incl. spell info.), gen.	generated	n. in Q vers.
PET1271	no	yes	Last occup. status farmer: area farmed (incl. spell info.), gen.	generated	n. in Q vers.
PET1290	no	yes	Current job: weekly working hours (excl. overtime), gen.	generated	n. in Q vers.

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn.: Personal	No. in wave 2 questionn.: Senior cit.
PET1300	yes	yes	Current job: actual weekly working hours (incl. overtime), gen.	generated	P11
PET1400	yes	no	Irregular working hours: average working hours per week	not in wave	not in wave
PET1450	no	yes	Current job: average working hours per week with irregular hours of work	generated	n. in Q vers.
PET1500	yes	no	Current occup. status, simple classification, surv'd in W1 (anon.)	not in wave	not in wave
PET1510	no	yes	Current occup. status, simple classification, surv'd from W2 (anon.)	generated	P12
PET1600	yes	yes	Current occup. status blue-collar worker: detailed info.	generated	P13
PET1700	yes	yes	Current occup. status white-collar worker: detailed info.	generated	P14
PET1900	yes	yes	Current occup. status civil servant: detailed info., incl. soldiers (anon.)	generated	P15, P16
PET2000	yes	yes	Current occup. status self-employed: no. of employees	generated	P17
PET2100	yes	yes	Current occup. status farmer: area farmed	generated	P18
PET2200	yes	no	Current occup. status trainee: detailed info.	not in wave	not in wave
PET2300	yes	no	Current job: training required	not in wave	not in wave
PET2390	no	yes	Current job: since when (month)?	generated	n. in Q vers.
PET2400	yes	yes	Current job: since when (year)?	generated	n. in Q vers.
PET2500	yes	no	Current job: fixed-term contract? (wave 1 only)	not in wave	not in wave
PET2510a	no	yes	Current job: initially fixed-term contract?	generated	n. in Q vers.
PET2510b	no	yes	Current job: conversion into permanent employment relationship	generated	n. in Q vers.
PET2600	yes	yes	Current job: temporary employment?	generated	n. in Q vers.
PET2700	yes	yes	Current job: supervisory function?	generated	n. in Q vers.
PET2800	yes	yes	Current job: number of staff supervised	generated	n. in Q vers.
PET2900	yes	yes	Current job: public sector?	generated	n. in Q vers.
PET3000	yes	yes	Current job: no. of employees in establishment/local office	generated	n. in Q vers.
PET3100	yes	no	Other job: weekly working hours	not in wave	not in wave
PET3200a	no	yes	First regular job: starting month	P92m	n. in Q vers.
PET3200b	no	yes	First regular job: starting year	P92j	n. in Q vers.
PET3300	no	yes	First occup. status, simple classification (anon.)	P93	n. in Q vers.
PET3400	no	yes	First occup. status blue-collar worker, detailed info.	P94	n. in Q vers.
PET3500	no	yes	First occup. status white-collar worker: detailed info.	P95	n. in Q vers.
PET3700	no	yes	First occup. status civil servant: detailed info., incl. soldiers	P96, P97	n. in Q vers.
PET3800	no	yes	First occup. status self-employed: no. of employees	P98	n. in Q vers.
PET3900	no	yes	First occup. status farmer: area farmed	P99	n. in Q vers.
PET3301	no	yes	First occup. status, simple class. (merged, incl. spell info.) (anon.), gen.	generated	n. in Q vers.
PET3401	no	yes	First occup. status blue-collar worker: detailed info., (merged, incl. spell info.), gen.	generated	n. in Q vers.
PET3501	no	yes	First occup. status white-collar worker: detailed info., (merged, incl. spell info), gen.	generated	n. in Q vers.
PET3701	no	yes	First occup. status civil servant: detailed info., incl. soldiers, (merged, incl. spell info) (anon.), gen.	generated	n. in Q vers.
PET3801	no	yes	First occup. status self-employed: no. of employees, (merged, incl. spell info), gen.	generated	n. in Q vers.
PET3901	no	yes	First occup. status farmer: area farmed, (merged, incl. spell info), gen.	generated	n. in Q vers.
PET4000	no	yes	Frequency of employer changes	P102	n. in Q vers.
PET4100a	no	yes	Total employment duration (months)	P103m	n. in Q vers.
PET4100b	no	yes	Total employment duration (years)	P103j	n. in Q vers.
PET4200a	no	yes	Duration of longest period of employment (months)	P104m	n. in Q vers.
PET4200b	no	yes	Duration of longest period of employment (years)	P104j	n. in Q vers.
PEK0100	yes	no	Gross income, open-ended info.	not in wave	not in wave
PEK0100a	no	yes	Training allowance (gross), open-ended info.	P32	n. in Q vers.
PEK0100b	no	yes	Gross income (excl. trainees), open-ended info. (in euros)	P64	P20
PEK0200	yes	yes	Gross income, indicator under/over € 1,000	P65	n. in Q vers.
PEK0300	yes	yes	Gross income, categories under € 1,000	P66	n. in Q vers.
PEK0400	yes	yes	Gross income, indicator under/over € 3,000	P67	n. in Q vers.
PEK0500	yes	yes	Gross income, categories € 1,000 to € 3,000	P68	n. in Q vers.
PEK0600	yes	yes	Gross income, categories over € 3,000	P69	n. in Q vers.
PEK0700	yes	no	Net income, open-ended info.	not in wave	not in wave
PEK0700a	no	yes	Training allowance (net), open-ended info.	P33	n. in Q vers.
PEK0700b	no	yes	Net income (excl. trainees), open-ended info. (in euros)	P70	P21
PEK0800	yes	yes	Net income, indicator under/over € 1,000	P71	n. in Q vers.
PEK0900	yes	yes	Net income, categories under € 1,000	P72	n. in Q vers.
PEK1000	yes	yes	Net income, indicator under/over € 2,000	P73	n. in Q vers.
PEK1100	yes	yes	Net income, categories € 1,000 to € 2,000	P74	n. in Q vers.
PEK1200	yes	yes	Net income, categories over € 2,000	P75	n. in Q vers.
PEK1300	yes	no	Other job: net income, open-ended info.	not in wave	not in wave
PEK1350a	no	yes	Indicator: receipt of extra pay (trainees only)?	P34	n. in Q vers.
PEK1350b	no	yes	Indicator: receipt of extra pay	P76	n. in Q vers.
PEK1360a	no	yes	Amount of extra pay (trainees only)	P35	n. in Q vers.

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn.: Personal	No. in wave 2 questionn.: Senior cit.
PEK1360b	no	yes	Amount of extra pay	P77	n. in Q vers.
PEK1400	yes	yes	Indicator: receipt of payments from statutory pension insurance?	P249	P22
PEK1400a	yes	no	Statutory old-age pension (sen. citizen's questionnaire wave 1 only)	not in wave	not in wave
PEK1400b	yes	no	Statutory widow's pension (sen. citizen's questionnaire wave 1 only)	not in wave	not in wave
PEK1410	no	yes	Indicator: receipt of old-age pension from statutory pension insurance / civil servant's pension?	P251	P23
PEK1415	no	yes	Amount of old-age pension	P252	P24
PEK1420	no	yes	Indicator: receipt of widow's pension or child-raising benefits?	P253	P25
PEK1425	no	yes	Amount of widow's pension or child-raising benefits (net)	P254	P26
PEK1430	no	yes	Indicator: receipt of orphan's pension?	P255	n. in Q vers.
PEK1435	no	yes	Amount of orphan's pension (net)	P256	n. in Q vers.
PEK1440	no	yes	Indicator: receipt of pension for reduction in earning capacity?	P257	n. in Q vers.
PEK1445	no	yes	Amount of pension for reduction in earning capacity	P258	n. in Q vers.
PEK1450	no	yes	Indicator: receipt of private pension / company pension?	P259	P27
PEK1455	no	yes	Amount of private pension / company pension (net)	P260	P28
PEK1500	yes	no	Payments from statutory pension insurance fund: amount of monthly payment	not in wave	not in wave
PEK1600	yes	yes	Indicator: receipt of training allowance/student grant	P36	n. in Q vers.
PEK1700	yes	yes	Amount of training allowance/student grant	P37	n. in Q vers.
PEK1800	no	yes	Indicator: receipt of state benefits (e.g. short-time working allowance)	P78	n. in Q vers.
PEK1900	no	yes	Amount of state benefits	P79	n. in Q vers.
PTK0100	yes	yes	Local name of institution resp. for administration of SGB II (e.g. job centre)	P138	n. in Q vers.
PTK0200	yes	yes	Job centre: number of visits	P140	n. in Q vers.
PTK0300a	yes	yes	No visits to job centre: only contact by phone or in writing	P141_1	n. in Q vers.
PTK0300b	yes	yes	No visits to job centre: plan to visit soon	P141_2	n. in Q vers.
PTK0300c	yes	yes	No visits to job centre: contact via other household member	P141_3	n. in Q vers.
PTK0300d	yes	yes	No visits to job centre: other reason	P141_4	n. in Q vers.
PTK0310	no	yes	Obligation to seek work	P142	n. in Q vers.
PTK0320a	no	yes	Reason for no obligation to seek work: over age of 58	P143_1	n. in Q vers.
PTK0320b	no	yes	Reason for no obligation to seek work: health reasons	P143_2	n. in Q vers.
PTK0320c	no	yes	Reason for no obligation to seek work: care of child(ren)	P143_3	n. in Q vers.
PTK0320d	no	yes	Reason for no obligation to seek work: care of relative	P143_4	n. in Q vers.
PTK0320e	no	yes	Reason for no obligation to seek work: other reason	P143_5	n. in Q vers.
PTK0321a	no	yes	Reason for no oblig. to seek work: over age of 58, incl. open info.	generated	n. in Q vers.
PTK0321b	no	yes	Reason for no oblig. to seek work: health reasons, incl. open info.	generated	n. in Q vers.
PTK0321c	no	yes	Reason for no oblig. to seek work: care of child(ren), incl. open info.	generated	n. in Q vers.
PTK0321d	no	yes	Reason for no oblig. to seek work: care of relative, incl. open info.	generated	n. in Q vers.
PTK0321e	no	yes	Reason for no oblig. to seek work: other reason, excl. coded open info.	generated	n. in Q vers.
PTK0321f	no	yes	Reason for no oblig. to seek work: employment, new category from open-ended info.	generated	n. in Q vers.
PTK0321g	no	yes	Reason for no obligation to seek work: training, new category from open-ended info.	generated	n. in Q vers.
PTK0400	yes	yes	Contact to job centre: detailed counselling?	P144	n. in Q vers.
PTK0500	yes	yes	Contact to job centre: frequency of detailed counselling	P145	n. in Q vers.
PTK0600a	yes	yes	Counselling interview: strengths/weaknesses profile drawn up?	P146a	n. in Q vers.
PTK0600b	yes	yes	Counselling interview: discussion of personal employment prospects?	P146b	n. in Q vers.
PTK0600c	yes	yes	Counselling interview: advice on job interviews	P146c	n. in Q vers.
PTK0600d	yes	yes	Counselling interview: advice on definite further training measures	P146d	n. in Q vers.
PTK0600e	yes	yes	Counselling interview: advice on payments available for taking up employment / becoming self-employed	P146e	n. in Q vers.
PTK0700	yes	yes	Job centre: regular contact person?	P147	n. in Q vers.
PTK0800a	yes	yes	Job centre offer: mini/midi-job	P148a	n. in Q vers.
PTK0800b	yes	yes	Job centre offer: regular full-time job	P148b	n. in Q vers.
PTK0800c	yes	yes	Job centre offer: regular part-time job	P148c	n. in Q vers.
PTK0800d	yes	yes	Job centre offer: training place	P148d	n. in Q vers.
PTK0800e	yes	yes	Job centre offer: placement voucher	P148e	n. in Q vers.
PTK0800f	yes	yes	Job centre offer: training voucher	P148f	n. in Q vers.
PTK0900a	yes	yes	Action arranged by job centre: vocational aptitude test?	P149	n. in Q vers.
PTK0900b	yes	yes	Would vocational aptitude test be/have been important?	P150	n. in Q vers.
PTK1000a	yes	yes	Action arranged by job centre: advice for personal problems?	P151	n. in Q vers.
PTK1000b	yes	yes	Would advice for personal problems be/have been important?	P152	n. in Q vers.
PTK1100a	yes	yes	Action arranged by job centre: child-care?	P153	n. in Q vers.
PTK1100b	yes	yes	Would child-care be/have been important?	P154	n. in Q vers.
PTK1200	yes	yes	Indicator: integration agreement concluded?	P155	n. in Q vers.
PTK1300	yes	no	No. of integration agreements	not in wave	not in wave
PTK1400	yes	no	Integration agreement: currently still valid?	not in wave	not in wave

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn.: Personal	No. in wave 2 questionn.: Senior cit.
PTK1500	yes	yes	Integration agreement: why none concluded?	P156	n. in Q vers.
PTK1600a	yes	yes	Experience with job centre: staff make to many regulations	P157a	n. in Q vers.
PTK1600b	yes	yes	Experience with job centre: they really want to help me	P157b	n. in Q vers.
PTK1600c	yes	yes	Experience with job centre: counselling improves my situation	P157c	n. in Q vers.
PTK1600d	yes	yes	Experience with job centre: support me in finding work again	P157d	n. in Q vers.
PTK1600e	yes	yes	Experience with job centre: only demands, no support	P157e	n. in Q vers.
PTK1600f	yes	yes	Experience with job centre: trust the staff	P157f	n. in Q vers.
PTK1600g	yes	yes	Experience with job centre: my ideas are taken into account	P157g	n. in Q vers.
PTK1600h	yes	yes	Experience with job centre: staff friendly and helpful	P157h	n. in Q vers.
PAS0100	yes	yes	Employed person: job-search in last four weeks?	P179	n. in Q vers.
PAS0200a	yes	yes	Employed person: starting month of job-search	P180m	n. in Q vers.
PAS0200b	yes	yes	Employed person: starting year of job-search	P180j	n. in Q vers.
PAS0300a	yes	no	Reason for job-search: household income not sufficient	not in wave	not in wave
PAS0300b	yes	no	Reason for job-search: no longer want to commute	not in wave	not in wave
PAS0300c	yes	no	Reason for job-search: employment to end soon	not in wave	not in wave
PAS0300d	yes	no	Job-search: inconvenient working hours	not in wave	not in wave
PAS0300e	yes	no	Job-search: seeking secure job	not in wave	not in wave
PAS0300f	yes	no	Job-search: inadequate wage	not in wave	not in wave
PAS0300g	yes	no	Reason for job-search: working conditions	not in wave	not in wave
PAS0300h	yes	no	Reason for job-search: want job suited to qualification level	not in wave	not in wave
PAS0400	yes	yes	Economically inactive pers.: job-search in last four weeks?	P181	n. in Q vers.
PAS0500a	yes	yes	Economically inactive pers: starting month of job-search	P182m	n. in Q vers.
PAS0500b	yes	yes	Economically inactive pers: starting year of job-search	P182j	n. in Q vers.
PAS0600	yes	yes	Indicator: taking up employment possible in 2 weeks?	P183	n. in Q vers.
PAS0700a	yes	no	Taking up employment not poss. because: in training	not in wave	not in wave
PAS0700b	yes	no	Taking up employment not poss. because: care of relatives	not in wave	not in wave
PAS0700c	yes	no	Taking up employment not poss. because: health reasons	not in wave	not in wave
PAS0700d	yes	no	Taking up employment not poss. because: other reasons	not in wave	not in wave
PAS0701a	yes	no	Taking up employment not poss. because: in training, incl. open info.	not in wave	not in wave
PAS0701b	yes	no	Taking up employment not poss. because: care of relatives, incl. open info.	not in wave	not in wave
PAS0701c	yes	no	Taking up employment not poss. because: health reasons, incl. open info.	not in wave	not in wave
PAS0701d	yes	no	Taking up employment not poss. because: other reasons, incl. open info.	not in wave	not in wave
PAS0800a	yes	no	No job-search because: household income is sufficient	not in wave	not in wave
PAS0800b	yes	no	No job-search because: financial situation would not be better	not in wave	not in wave
PAS0800c	yes	no	No job-search because: already found new job	not in wave	not in wave
PAS0800d	yes	no	No job-search because: health reasons	not in wave	not in wave
PAS0800e	yes	no	No job-search because: care of child(ren)	not in wave	not in wave
PAS0800f	yes	no	No job-search because: not enough jobs	not in wave	not in wave
PAS0800g	yes	no	No job-search because: sought work previously without success	not in wave	not in wave
PAS0800h	yes	no	No job-search because: retired	not in wave	not in wave
PAS0900a	yes	yes	Job-search via: newspaper advertisements	P184a	n. in Q vers.
PAS0900b	yes	yes	Job-search via: online jobs portal of the Employment Agency	P184b	n. in Q vers.
PAS0900c	yes	yes	Job-search via: other Internet sources	P184c	n. in Q vers.
PAS0900d	yes	yes	Job-search via: friends or relatives	P184d	n. in Q vers.
PAS0900e	yes	yes	Job-search via: placement officers at Employment Agency	P184e	n. in Q vers.
PAS0900f	yes	yes	Job-search via: private placement services	P184f	n. in Q vers.
PAS0900g	yes	yes	Job-search via: other	P184g	n. in Q vers.
PAS0900h	yes	yes	Job-search via: none of these	P184h	n. in Q vers.
PAS0901a	no	yes	Job-search via: newspaper advertisements, incl. open info.	generated	n. in Q vers.
PAS0901b	no	yes	Job-search via: online jobs portal of the Employment Agency, incl. open info.	generated	n. in Q vers.
PAS0901c	no	yes	Job-search via: other Internet sources, incl. open info.	generated	n. in Q vers.
PAS0901d	no	yes	Job-search via: friends or relatives, incl. open info.	generated	n. in Q vers.
PAS0901e	no	yes	Job-search via: placement officers at Employment Agency, incl. open info.	generated	n. in Q vers.
PAS0901f	no	yes	Job-search via: private placement services, incl. open info.	generated	n. in Q vers.
PAS0901g	no	yes	Job-search via: other, incl. open info.	generated	n. in Q vers.
PAS0901h	no	yes	Job-search via: none of these, incl. open info.	generated	n. in Q vers.
PAS0901i	no	yes	Job-search via: speculative applications, generated	generated	n. in Q vers.
PAS1000a	yes	yes	Job-search activities: replied to job advertisement	P185a	n. in Q vers.
PAS1000b	yes	yes	Job-search activities: submitted own advertisement	P185b	n. in Q vers.
PAS1000c	yes	yes	Job-search activities: enquired at firms	P185c	n. in Q vers.
PAS1000d	yes	yes	Job-search activities: speculative applications	P185d	n. in Q vers.
PAS1100	yes	no	Last 4 weeks: no. of applications for distant jobs (relocation)	not in wave	not in wave
PAS1110	no	yes	Last 4 weeks: no. of applications for jobs at least 100 km away	P186	n. in Q vers.
PAS1200	yes	yes	Last 4 weeks: no. of job interviews	P187	n. in Q vers.

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn.: Personal	No. in wave 2 questionn.: Senior cit.
PAS1300	yes	yes	Desired volume of employment	P188	n. in Q vers.
PAS1400a	yes	yes	Acceptable difficulty: long journey to work	P189a	n. in Q vers.
PAS1400b	yes	yes	Acceptable difficulty: low income	P189b	n. in Q vers.
PAS1400c	yes	yes	Acceptable difficulty: inconvenient working hours	P189c	n. in Q vers.
PAS1400d	yes	yes	Acceptable difficulty: work below skills level	P189d	n. in Q vers.
PAS1400e	yes	yes	Acceptable difficulty: physical strain/exposure on the job	P189e	n. in Q vers.
PAS1400f	yes	yes	Acceptable difficulty: relocation	P189f	n. in Q vers.
PAS1500	yes	yes	Pers. not seeking work: ever sought work?	P190	n. in Q vers.
PAS1600	yes	yes	Pers. seeking work: type of job sought	P191	n. in Q vers.
PAS1700	yes	yes	Pers. not seeking work: type of job last sought	P192	n. in Q vers.
PAS1800	yes	yes	Expected net wages	P193	n. in Q vers.
PAS1900	yes	yes	Expected net wages: estimated working hours	P194	n. in Q vers.
PAS2000	yes	yes	Indicator: lower net wages acceptable?	P195	n. in Q vers.
PAS2100	yes	yes	Minimum net wages	P196	n. in Q vers.
PAS2200	yes	yes	Details refused expected net wages: minimum net wages	P197	n. in Q vers.
PAS2300	yes	yes	Minimum net wages: estimated working hours	P198	n. in Q vers.
PLS0100	yes	no	Necessary: dwelling with enough rooms	not in wave	not in wave
PLS0200	yes	no	Necessary: dwelling without damp walls/floors	not in wave	not in wave
PLS0300	yes	no	Necessary: bathroom inside the dwelling	not in wave	not in wave
PLS0400	yes	no	Necessary: indoor toilet	not in wave	not in wave
PLS0500	yes	no	Necessary: central heating/one-floor heating/district heating	not in wave	not in wave
PLS0600	yes	no	Necessary: garden/balcony/terrace	not in wave	not in wave
PLS0700	yes	no	Necessary: enough winter clothes	not in wave	not in wave
PLS0800	yes	no	Necessary: car	not in wave	not in wave
PLS0900	yes	no	Necessary: television	not in wave	not in wave
PLS1000	yes	no	Necessary: video recorder/DVD player	not in wave	not in wave
PLS1100	yes	no	Necessary: computer with Internet access	not in wave	not in wave
PLS1200	yes	no	Necessary: washing machine	not in wave	not in wave
PLS1300	yes	no	Necessary: freezer/refrigerator with freezer compartment	not in wave	not in wave
PLS1400	yes	no	Necessary: new clothes occasionally	not in wave	not in wave
PLS1500	yes	no	Necessary: a warm meal every day	not in wave	not in wave
PLS1600	yes	no	Necessary: a one-week holiday every year	not in wave	not in wave
PLS1700	yes	no	Necessary: inviting friends for dinner once a month	not in wave	not in wave
PLS1800	yes	no	Necessary: eating out once a month	not in wave	not in wave
PLS1900	yes	no	Necessary: going to cinema/theatre/concert once a month	not in wave	not in wave
PLS2000	yes	no	Necessary: saving a regular amount of money every month	not in wave	not in wave
PLS2100	yes	no	Necessary: replacing worn furniture	not in wave	not in wave
PLS2200	yes	no	Necessary: paying unexpected expenses oneself	not in wave	not in wave
PLS2300	yes	no	Necessary: health treatment not reimbursed by health insurance	not in wave	not in wave
PLS2400	yes	no	Necessary: paying rent punctually	not in wave	not in wave
PLS2500	yes	no	Necessary: paying utilities bills punctually	not in wave	not in wave
PLS2600	yes	no	Necessary: over-the-counter medicines	not in wave	not in wave
PLS2700a	yes	no	More or less able to afford: fashionable clothes	not in wave	not in wave
PLS2700b	yes	no	More or less able to afford: hobbies such as sports or music	not in wave	not in wave
PLS2700c	yes	no	More or less able to afford: visits to cinema/bars/discos	not in wave	not in wave
PLS2700d	yes	no	More or less able to afford: holidays/outings	not in wave	not in wave
PSK0050	no	yes	Relocation over 100 km distance?	P202	n. in Q vers.
PSK0060	no	yes	Frequency of relocations over 100 km distance	P203	n. in Q vers.
PSK0100	yes	yes	Indicator: close friends / relatives outside the HH?	P204	P36
PSK0200	yes	yes	Number of close friends / relatives outside the HH	P205	P37
PSK0300	yes	yes	Frequency of misunderstandings/tensions/conflicts?	P206	P38
PSK0400a	yes	yes	Active in: trade union	P207a	P39a
PSK0400b	yes	yes	Active in: political party	P207b	P39b
PSK0400c	yes	yes	Active in: church/parish	P207c	P39c
PSK0400d	yes	yes	Active in: club/association such as music club, sports club, cultural exchange association	P207d	P39d
PSK0400e	yes	yes	Active in: other organisation	P207e	P39e
PG0100	yes	yes	Number of visits to doctor, last 3 months	P211	P43
PG0200	yes	yes	Indicator: stay in hospital in the last 12 months?	P212	P44
PG0300	yes	yes	Number of nights in hospital, last 12 months?	P213	P45
PG0400	yes	yes	Number of stays in hospital, last 12 months?	P214	P46
PG0500	yes	yes	Indicator: officially diagnosed disability?	P215	P47
PG0600	yes	yes	Officially diagnosed disability: degree of disability	P216	P48
PG0700	yes	yes	Officially diagnosed disability: year of acknowledgement	P217	P49
PG0800	yes	yes	Indicator: other serious health problems?	P218	P50
PG0900a	yes	yes	Type of disability/health problem: physical	P219a	P51a
PG0900b	yes	yes	Type of disability/health problem: vision/hearing disability	P219b	P51b
PG0900c	yes	yes	Type of disability/health problem: seizures	P219c	P51c
PG0900d	yes	yes	Type of disability/health problem: internal disease/ organ damage	P219d	P51d
PG0900e	yes	yes	Type of disability/health problem: psychological disorders/disability	P219e	P51e

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn.: Personal	No. in wave 2 questionn.: Senior cit.
PG0900f	yes	yes	Type of disability/health problem: something else	P219f	P51f
PG0901a	yes	yes	Type of disability/health problem: physical, incl. open info.	generated	generated
PG0901b	yes	yes	Type of disability/health problem: vision/hearing disability, incl. open info.	generated	generated
PG0901c	yes	yes	Type of disability/health problem: seizures, incl. open info.	generated	generated
PG0901d	yes	yes	Type of disability/health problem: internal disease/ organ damage, incl. open info.	generated	generated
PG0901e	yes	yes	Type of disability/health problem: psychological disorders/disability, incl. open info.	generated	generated
PG0901f	yes	yes	Type of disability/health problem: something else, incl. open info.	generated	generated
PG0901g	yes	yes	Type of disability/health problem: allergy, generated	generated	generated
PG1000	yes	yes	Indicator: unable to work (official assessment)?	P220	n. in Q vers.
PG1100	yes	yes	Psychological problems	P221	P52
PG1200	yes	yes	Subjective assessment of state of health	P222	P53
PG1300	yes	yes	Type of health insurance	P223	P54
PG1301	yes	yes	Type of health insurance, incl. open info.	generated	generated
PP0100	yes	no	Indicator: regular care of relatives/friends? (wave1 only)	not in wave	not in wave
PP0110	no	yes	Indicator: regular care of relatives/friends? (from wave 2 onwards)	P224	P55
PP0200	yes	yes	Care: total number of individuals cared for	P225	P56
PP0300a	yes	no	1st person cared for: relationship to respondent	not in wave	not in wave
PP0300b	yes	no	2nd person cared for: relationship to respondent	not in wave	not in wave
PP0300c	yes	no	3rd person cared for: relationship to respondent	not in wave	not in wave
PP0300d	yes	no	4th person cared for: relationship to respondent	not in wave	not in wave
PP0300e	yes	no	5th person cared for: relationship to respondent	not in wave	not in wave
PP0310a	no	yes	1st person cared for: relationship to respondent	P2261	P57_01
PP0310b	no	yes	2nd person cared for: relationship to respondent	P2262	P57_02
PP0310c	no	yes	3rd person cared for: relationship to respondent	P2263	P57_03
PP0310d	no	yes	4th person cared for: relationship to respondent	P2264	P57_04
PP0310e	no	yes	5th person cared for: relationship to respondent	P2265	P57_05
PP0400a	yes	no	Help required for 1st person: errands/shopping	not in wave	not in wave
PP0400b	yes	no	Help required for 2nd person: errands/shopping	not in wave	not in wave
PP0400c	yes	no	Help required for 3rd person: errands/shopping	not in wave	not in wave
PP0400d	yes	no	Help required for 4th person: errands/shopping	not in wave	not in wave
PP0400e	yes	no	Help required for 5th person: errands/shopping	not in wave	not in wave
PP0500a	yes	no	Help required for 1st person: housekeeping/meals	not in wave	not in wave
PP0500b	yes	no	Help required for 2nd person: housekeeping/meals	not in wave	not in wave
PP0500c	yes	no	Help required for 3rd person: housekeeping/meals	not in wave	not in wave
PP0500d	yes	no	Help required for 4th person: housekeeping/meals	not in wave	not in wave
PP0500e	yes	no	Help required for 5th person: housekeeping/meals	not in wave	not in wave
PP0600a	yes	no	Help required for 1st person: simple nursing	not in wave	not in wave
PP0600b	yes	no	Help required for 2nd person: simple nursing	not in wave	not in wave
PP0600c	yes	no	Help required for 3rd person: simple nursing	not in wave	not in wave
PP0600d	yes	no	Help required for 4th person: simple nursing	not in wave	not in wave
PP0600e	yes	no	Help required for 5th person: simple nursing	not in wave	not in wave
PP0700a	yes	no	Help required for 1st person: difficult nursing	not in wave	not in wave
PP0700b	yes	no	Help required for 2ndt person: difficult nursing	not in wave	not in wave
PP0700c	yes	no	Help required for 3rd person: difficult nursing	not in wave	not in wave
PP0700d	yes	no	Help required for 4th person: difficult nursing	not in wave	not in wave
PP0700e	yes	no	Help required for 5th person: difficult nursing	not in wave	not in wave
PP0800a	yes	no	1st person cared for: payments from nursing care insurance?	not in wave	not in wave
PP0800b	yes	no	2nd person cared for: payments from nursing care insurance?	not in wave	not in wave
PP0800c	yes	no	3rd person cared for: payments from nursing care insurance?	not in wave	not in wave
PP0800d	yes	no	4th person cared for: payments from nursing care insurance?	not in wave	not in wave
PP0800e	yes	no	5th person cared for: payments from nursing care insurance?	not in wave	not in wave
PP0810a	no	yes	1st person cared for: payments from nursing care insurance?	P2271	P58_01
PP0810b	no	yes	2nd person cared for: payments from nursing care insurance?	P2272	P58_02
PP0810c	no	yes	3rd person cared for: payments from nursing care insurance?	P2273	P58_03
PP0810d	no	yes	4th person cared for: payments from nursing care insurance?	P2274	P58_04
PP0810e	no	yes	5th person cared for: payments from nursing care insurance?	P2275	P58_05
PP0820a	no	yes	1st person cared for: level of nursing care required	P2281	P59_01
PP0820b	no	yes	2nd person cared for: level of nursing care required	P2282	P59_02
PP0820c	no	yes	3rd person cared for: level of nursing care required	P2283	P59_03
PP0820d	no	yes	4th person cared for: level of nursing care required	P2284	P59_04
PP0820e	no	yes	5th person cared for: level of nursing care required	P2285	P59_05
PP0830a	no	yes	1st person cared for: accommodation	P2291	P60_01
PP0830b	no	yes	2nd person cared for: accommodation	P2292	P60_02
PP0830c	no	yes	3rd person cared for: accommodation	P2293	P60_03
PP0830d	no	yes	4th person cared for: accommodation	P2294	P60_04
PP0830e	no	yes	5th person cared for: accommodation	P2295	P60_05
PP0900a	yes	no	1st person cared for: assistance from nursing service?	not in wave	not in wave

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn.: Personal	No. in wave 2 questionn.: Senior cit.
PP0900b	yes	no	2nd person cared for: assistance from nursing service?	not in wave	not in wave
PP0900c	yes	no	3rd person cared for: assistance from nursing service?	not in wave	not in wave
PP0900d	yes	no	4th person cared for: assistance from nursing service?	not in wave	not in wave
PP0900e	yes	no	5th person cared for: assistance from nursing service?	not in wave	not in wave
PP0910a	no	yes	1st person cared for: assistance from nursing service?	P2301	P61_01
PP0910b	no	yes	2nd person cared for: assistance from nursing service?	P2302	P61_02
PP0910c	no	yes	3rd person cared for: assistance from nursing service?	P2303	P61_03
PP0910d	no	yes	4th person cared for: assistance from nursing service?	P2304	P61_04
PP0910e	no	yes	5th person cared for: assistance from nursing service?	P2305	P61_05
PP0920a	no	yes	1st person cared for: nursing service partly paid by long-term nursing care fund?	P2311	P62_01
PP0920b	no	yes	2nd person cared for: nursing service partly paid by long-term nursing care fund?	P2312	P62_02
PP0920c	no	yes	3rd person cared for: nursing service partly paid by long-term nursing care fund?	P2313	P62_03
PP0920d	no	yes	4th person cared for: nursing service partly paid by long-term nursing care fund?	P2314	P62_04
PP0920e	no	yes	5th person cared for: nursing service partly paid by long-term nursing care fund?	P2315	P62_05
PP1000a	yes	no	1st person cared for: amount of nursing care allowance	not in wave	not in wave
PP1000b	yes	no	2nd person cared for: amount of nursing care allowance	not in wave	not in wave
PP1000c	yes	no	3rd person cared for: amount of nursing care allowance	not in wave	not in wave
PP1000d	yes	no	4th person cared for: amount of nursing care allowance	not in wave	not in wave
PP1000e	yes	no	5th person cared for: amount of nursing care allowance	not in wave	not in wave
PP1100	yes	yes	Care: hours/week spent on care, total	P232	P63
PP1150	no	yes	Regular amount of money received for care?	P233	P64
PP1160	no	yes	Amount of money received	P234	P65
PP1200a	yes	yes	Action arranged by job centre: help with finding nursing staff?	P235	n. in Q vers.
PP1200b	yes	yes	Would support in finding nursing staff be important?	P236	n. in Q vers.
PMI0100	yes	yes	Indicator: born in Germany?	P263	P72
PMI0200	yes	yes	Not born in Germany: country of birth (anonymised)	P264	P73
PMI0300a	yes	yes	Not born in Germany: month of move to Germany	P265m	P74m
PMI0300b	yes	yes	Not born in Germany: year of move to Germany	P265j	P74j
PMI0400	yes	yes	Indicator: German nationality?	P266	P75
PMI0500	yes	yes	Not German nationality: which nationality? (anonymised)	P267	P76
PMI0600	yes	yes	Not German nationality: temporary or permanent residence permit?	P268	n. in Q vers.
PMI0650	no	yes	Type of right of residence	P269	n. in Q vers.
PMI0700	yes	yes	Indicator: parents or grandparents born outside Germany?	P271	P77
PMI0800a	yes	yes	Father: country of birth	P272a	P78a
PMI0800b	yes	yes	Mother: country of birth	P272b	P78b
PMI0800c	yes	yes	Father's father: country of birth	P272c	P78c
PMI0800d	yes	yes	Father's mother: country of birth	P272d	P78d
PMI0800e	yes	yes	Mother's father: country of birth	P272e	P78e
PMI0800f	yes	yes	Mother's mother: country of birth	P272f	P78f
PMI0900a	yes	yes	Father: migrated to Germany?	P273a	P79a
PMI0900b	yes	yes	Mother: migrated to Germany?	P273b	P79b
PMI0900c	yes	yes	Father's father: migrated to Germany?	P273c	P79c
PMI0900d	yes	yes	Father's mother: migrated to Germany?	P273d	P79d
PMI0900e	yes	yes	Mother's father: migrated to Germany?	P273e	P79e
PMI0900f	yes	yes	Mother's mother: migrated to Germany?	P273f	P79f
PMI1000a	yes	yes	Father: country of residence before migration (anonymised)	P274a	P80a
PMI1000b	yes	yes	Mother: country of residence before migration (anonymised)	P274b	P80b
PMI1000c	yes	yes	Father's father: country of residence before migration (anonymised)	P274c	P80c
PMI1000d	yes	yes	Father's mother: country of res. before migration (anonymised)	P274d	P80d
PMI1000e	yes	yes	Mother's father: country of residence before migration (anonymised)	P274e	P80e
PMI1000f	yes	yes	Mother's mother: country of res. before migration (anonymised)	P274f	P80f
PMI1100	yes	no	Language spoken in household	not in wave	not in wave
PMI1110	no	yes	1st language spoken in circle of friends	P275	P81
PMI1111	no	yes	1st language spoken in circle of friends, incl. open info.	generated	generated
PMI1120	no	yes	Equal use of two languages in circle of friends: 1st language	P276	P82
PMI1121	no	yes	Equal use of two lang's in circle of friends: 1st lang., incl. open info.	generated	generated
PMI1130	no	yes	Equal use of two languages in circle of friends: 2nd language	P277	P83
PMI1131	no	yes	Equal use of two languages in circle of friends: 2nd language, incl. open info.	generated	generated
PSH0100	yes	no	Internal filter: mother in HH?	not in wave	not in wave
PSH0200	yes	yes	Mother's highest general school qualification	P278	n. in Q vers.
PSH0201	yes	yes	Mother's highest general school qualification, incl. open info.	generated	n. in Q vers.
PSH0300a	yes	yes	Mother's vocational qualification: training as semi-skilled worker, compact vocational training	P279_1	n. in Q vers.
PSH0300b	yes	yes	Mother's vocational qual.: apprenticeship/training	P279_2	n. in Q vers.

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn.: Personal	No. in wave 2 questionn.: Senior cit.
PSH0300c	yes	yes	Mother's vocational qual.: master craftsman's/technician's cert. or similar	P279_3	n. in Q vers.
PSH0300d	yes	yes	Mother's vocational qual.: cert. from college of adv. vocational studies	P279_4	n. in Q vers.
PSH0300e	yes	yes	Mother's vocational qualification: degree from univ. of applied sciences or college of education	P279_5	n. in Q vers.
PSH0300f	yes	yes	Mother's vocational qual.: university degree	P279_6	n. in Q vers.
PSH0300g	yes	yes	Mother's vocational qual.: other German qualification	P279_7	n. in Q vers.
PSH0300h	yes	yes	Mother's vocational qual.: other foreign qualification	P279_8	n. in Q vers.
PSH0300i	yes	yes	Mother's vocational qual.: no qualification	P279_9	n. in Q vers.
PSH0301a	yes	yes	Mother's vocational qual.: training as semi-skilled worker, compact vocational training, incl. open info.	generated	n. in Q vers.
PSH0301b	yes	yes	Mother's vocational qual.: apprenticeship/training, incl. open info.	generated	n. in Q vers.
PSH0301c	yes	yes	Mother's vocational qual.: master craftsman's/technician's cert. or similar, incl. open info.	generated	n. in Q vers.
PSH0301d	yes	yes	Mother's vocational qual.: cert. from college of adv. vocational studies, incl. open info.	generated	n. in Q vers.
PSH0301e	yes	yes	Mother's vocational qual.: degree from univ. of applied sciences or college of education, incl. open info.	generated	n. in Q vers.
PSH0301f	yes	yes	Mother's vocational qual.: university degree, incl. open info.	generated	n. in Q vers.
PSH0301g	yes	yes	Mother's vocational qual.: other German qual., incl. open info.	generated	n. in Q vers.
PSH0301h	yes	yes	Mother's vocational qual.: other foreign qual., incl. open info.	generated	n. in Q vers.
PSH0301i	yes	yes	Mother's vocational qual.: no qualification, incl. open info.	generated	n. in Q vers.
PSH0310	no	yes	Was mother employed when respondent was 15 years old?	P280	n. in Q vers.
PSH0320	no	yes	Mother's occup. status at that time, simple classification (anon.)	P281	n. in Q vers.
PSH0330	no	yes	Mother's occup. status at time blue-collar worker: detailed info.	P282	n. in Q vers.
PSH0340	no	yes	Mother's occup. status at time white-collar worker, detailed info.	P283	n. in Q vers.
PSH0360	no	yes	Mother's occup. status at time civil servant, incl. soldiers: detailed info. (anon.)	P284, P285	n. in Q vers.
PSH0370	no	yes	Mother's occup. status at that time self-employed: no. of employees	P286	n. in Q vers.
PSH0380	no	yes	Mother's occup. status at that time farmer: area farmed	P287	n. in Q vers.
PSH0400	yes	no	Internal filter: father in HH?	not in wave	not in wave
PSH0500	yes	yes	Father's highest general school qualification	P289	n. in Q vers.
PSH0501	yes	yes	Father's highest general school qualification, incl. open info.	generated	n. in Q vers.
PSH0600a	yes	yes	Father's vocational qual.: training as semi-skilled worker, compact vocational training	P290_1	n. in Q vers.
PSH0600b	yes	yes	Father's vocational qual.: apprenticeship/training	P290_2	n. in Q vers.
PSH0600c	yes	yes	Father's voc. qual.: master craftsman's/technician's cert. or similar	P290_3	n. in Q vers.
PSH0600d	yes	yes	Father's vocational qual.: cert. from college of adv. voc. studies	P290_4	n. in Q vers.
PSH0600e	yes	yes	Father's vocational qual.: degree from univ. of applied sciences or college of education	P290_5	n. in Q vers.
PSH0600f	yes	yes	Father's vocational qual.: university degree	P290_6	n. in Q vers.
PSH0600g	yes	yes	Father's vocational qual.: other German qualification	P290_7	n. in Q vers.
PSH0600h	yes	yes	Father's vocational qual.: other foreign qualification	P290_8	n. in Q vers.
PSH0600i	yes	yes	Father's vocational qual.: no qualification	P290_9	n. in Q vers.
PSH0601a	yes	yes	Father's vocational qual.: training as semi-skilled worker, compact vocational training, incl. open info.	generated	n. in Q vers.
PSH0601b	yes	yes	Father's vocational qual.: apprenticeship/training, incl. open info.	generated	n. in Q vers.
PSH0601c	yes	yes	Father's vocational qual.: master craftsman's/technician's cert. or similar, incl. open info.	generated	n. in Q vers.
PSH0601d	yes	yes	Father's vocational qual.: cert. from college of adv. vocational studies, incl. open info.	generated	n. in Q vers.
PSH0601e	yes	yes	Father's vocational qual.: degree from univ. of applied sciences or college of education, incl. open info.	generated	n. in Q vers.
PSH0601f	yes	yes	Father's vocational qual.: university degree, incl. open info.	generated	n. in Q vers.
PSH0601g	yes	yes	Father's vocational qual.: other German qual., incl. open info.	generated	n. in Q vers.
PSH0601h	yes	yes	Father's vocational qual.: other foreign qual., incl. open info.	generated	n. in Q vers.
PSH0601i	yes	yes	Father's vocational qual.: no qualification, incl. open info.	generated	n. in Q vers.
PSH0610	no	yes	Was father employed when respondent was 15 years old?	P291	n. in Q vers.
PSH0620	no	yes	Father's occup. status at that time, simple classification (anon.)	P292	n. in Q vers.
PSH0630	no	yes	Father's occup. status at time blue-collar worker: detailed info.	P293	n. in Q vers.
PSH0640	no	yes	Father's occup. status at time white-collar worker, detailed info.	P294	n. in Q vers.
PSH0660	no	yes	Father's occup. status at time civil servant, incl. soldiers: detailed info. (anon.)	P295, P296	n. in Q vers.
PSH0670	no	yes	Father's occup. status at time self-employed: no. of employees	P297	n. in Q vers.
PSH0680	no	yes	Father's occup. status at time farmer: area farmed	P298	n. in Q vers.
RegP0100	yes	yes	Consent to merging of data	P105	n. in Q vers.
panel	yes	yes	Willingness to participate in panel	system	system
vhh	yes	yes	Control variable: father living in HH	system	n. in Q vers.
mhh	yes	yes	Control variable: mother living in HH	system	n. in Q vers.

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn.: Personal	No. in wave 2 questionn.: Senior cit.
apartner	yes	yes	Control variable: unmarried partner living in HH	system	n. in Q vers.
epartner	yes	yes	Control variable: spouse or registered partner in HH	system	system
famstand	yes	yes	Marital status, generated	generated	generated
hhalg2	yes	yes	Control variable: current receipt of UB II	system	n. in Q vers.
ekin1517	yes	yes	Control variable: own child aged between 15 and 17 in HH	system	n. in Q vers.
ekinu18	yes	yes	Control variable: own child under age of 18 in HH	system	n. in Q vers.
ekin614	yes	yes	Control variable: own child aged between 6 and 14 in HH	system	n. in Q vers.
ekinu15	yes	yes	Control variable: own child under age of 15 in HH	system	n. in Q vers.
ekind	yes	yes	Control variable: own child in HH	system	n. in Q vers.
kindzges	yes	yes	Total number of own children (living in and outside HH), generated	generated	n. in Q vers.
kindzihh	yes	yes	Own children living in the HH, generated	generated	generated
halg2s05	yes	no	Control variable: receipt of UB II since 2005	not in wave	not in wave
halg2s06	no	yes	Control variable: receipt of UB II since 2006	generated	generated
schul1	yes	yes	Highest general school qual., excl. foreign qual's and open info.	generated	generated
schul2	yes	yes	Highest general school qual., incl. foreign qual's and open info.	generated	generated
schulabj	no	yes	Year of highest general school qualification	generated	n. in Q vers.
beruf1	yes	yes	Highest vocational qual., excl. foreign qual's and open info. generated	generated	generated
beruf2	yes	yes	Highest vocational qual., incl. foreign qual's and open info. generated	generated	generated
berabj	no	yes	Year of highest vocational qualification	generated	n. in Q vers.
casmin	yes	yes	Education classified acc. to CASMIN, updated version, generated	generated	generated
isced97	yes	yes	Education classified acc. to ISCED97, generated	generated	generated
bilzeit	yes	yes	Duration of school education and vocational training in years, generated	generated	generated
mschul1	yes	yes	Mother's highest general school qual., incl. mother in HH, excl. open info. gen.	generated	generated
mschul2	yes	yes	Mother's highest general school qual., incl. mother in HH, incl. open info. gen.	generated	generated
mberuf1	yes	yes	Mother's highest vocational qual., incl. mother in HH, excl. open info. gen.	generated	generated
mberuf2	yes	yes	Mother's highest vocational qual., incl. mother in HH, incl. open info. gen.	generated	generated
mcasmin	yes	yes	Mother's education, classified acc. to CASMIN, updated version, generated	generated	generated
misced97	yes	yes	Mother's education, classified acc. to ISCED97, generated	generated	generated
mbilzeit	yes	yes	Duration of mother's school education and vocational training in years, generated	generated	generated
mstib	no	yes	Mother's occupational status, code number, generated	generated	n. in Q vers.
misco	no	yes	ISCO 88 (ZUMA coding) of the mother, generated	generated	n. in Q vers.
mmps	no	yes	Magnitude-Prestige-Scale, mother's occupation, gen.	generated	n. in Q vers.
msiops	no	yes	Standard International Occupational Prestige Scale, mother's occupation, gen.	generated	n. in Q vers.
misei	no	yes	International Socio-Economic Index, mother's occupation, gen.	generated	n. in Q vers.
megp	no	yes	Class scheme acc. to Goldthorpe & Portocarrero (EGP), mother's occupation, gen.	generated	n. in Q vers.
mesec	no	yes	European Socio-economic Classification (ESeC), mother's occupation, gen.	generated	n. in Q vers.
misco_it	no	yes	ISCO 88 (Infratest coding) of the mother, generated	generated	n. in Q vers.
mkldb	no	yes	Classification of Occupations 1992 (Infratest coding) of the mother, generated	generated	n. in Q vers.
vschul1	yes	yes	Father's highest general school qual., incl. father in HH, excl. open info. gen.	generated	generated
vschul2	yes	yes	Father's highest general school qual., incl. father in HH, incl. open info. gen.	generated	generated
vberuf1	yes	yes	Father's highest vocational qual., incl. father in HH, excl. open info. gen.	generated	generated
vberuf2	yes	yes	Father's highest vocational qual., incl. father in HH, incl. open info. gen.	generated	generated
vcasmin	yes	yes	Father's education, classified acc. to CASMIN, updated version, generated	generated	generated
visced97	yes	yes	Father's education, classified acc. to ISCED97, generated	generated	generated
vbilzeit	yes	yes	Duration of father's school education and vocational training in years, generated	generated	generated
vstib	no	yes	Father's occupational status, code number, generated	generated	n. in Q vers.
visco	no	yes	ISCO 88 (ZUMA coding) of the father, generated	generated	n. in Q vers.
vmps	no	yes	Magnitude-Prestige-Scale, father's occupation, gen.	generated	n. in Q vers.
vsiops	no	yes	Standard International Occupational Prestige Scale, father's occupation, gen.	generated	n. in Q vers.
visei	no	yes	International Socio-Economic Index, father's occupation, gen.	generated	n. in Q vers.

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn.: Personal	No. in wave 2 questionn.: Senior cit.
vegp	no	yes	Class scheme acc. to Erikson, Goldthorpe & Portocarrero (EGP), father's occupation, gen.	generated	n. in Q vers.
vesec	no	yes	European Socio-economic Classification (ESeC), father's occupation, gen.	generated	n. in Q vers.
visco_it	no	yes	ISCO 88 (Infratest coding) of the father, generated	generated	n. in Q vers.
vkldb	no	yes	Classification of Occupations 1992 (Infratest coding) of the father, generated	generated	n. in Q vers.
netto	yes	yes	Net income, incl. categorised info., generated	generated	generated
nettokat	yes	yes	Categorised net income, generated	generated	n. in Q vers.
brutto	yes	yes	Gross income, incl. categorised info., generated	generated	generated
bruttokat	yes	yes	Categorised gross income, generated	generated	n. in Q vers.
aktmassn	yes	yes	Current participation in a measure funded/promoted by the employment agency, generated	generated	n. in Q vers.
alg1abez	yes	yes	Current receipt of UB I, generated	generated	n. in Q vers.
stibkz	yes	yes	Current occupational status, simple classification, harmonised (anonymised)	generated	generated
stib	yes	yes	Occupational status, code number, generated	generated	generated
erwerb	yes	no	Employment status, generated (wave 1 only)	not in wave	not in wave
erwerb2	yes	yes	Employment status, generated (all waves)	generated	n. in Q vers.
nichterw	yes	yes	Status economic inactivity, generated (all waves)	generated	n. in Q vers.
nichtew2	yes	yes	Status economic inactivity, generated, incl. open info. (all waves)	generated	n. in Q vers.
befrist	yes	yes	Current job: fixed-term contract? generated (all waves)	generated	n. in Q vers.
arbzeit	yes	yes	Weekly hours of work incl. details in the case of irregular working hours, gen.	generated	generated
alg1s05	yes	yes	Indicator: receipt of UB I since Jan. 2005? generated (all waves)	generated	n. in Q vers.
isco88	yes	yes	ISCO 88 (ZUMA coding), generated	generated	generated
mps	yes	yes	Magnitude-Prestige-Scale, current occupation, gen.	generated	generated
siops	yes	yes	Standard International Occupational Prestige Scale, current occupation, gen.	generated	generated
isei	yes	yes	International Socio-Economic Index, current occupation, gen	generated	generated
egp	yes	yes	Class scheme acc. to Erikson, Goldthorpe & Portocarrero (EGP), current occupation, gen.	generated	generated
esec	yes	yes	European Socio-economic Classification (ESeC) current occupation, gen.	generated	generated
isco88it	yes	yes	ISCO 88 (Infratest coding),generated	generated	generated
kldb_it	yes	yes	Classification of Occupations 1992 (Infratest coding),generated	generated	generated
berpr_it	yes	no	Occupational coding, problem indicator (Infratest coding), generated	not in wave	not in wave
branche	no	yes	Current job: economic sector/industry (WZ2003)	generated	n. in Q vers.
emonlewt	no	yes	Time when last job ended (month)	generated	n. in Q vers.
ejhrlewt	no	yes	Time when last job ended (year)	generated	n. in Q vers.
stiblewt	no	yes	Occupational status, last job, code number, generated	generated	n. in Q vers.
iscolewt	no	yes	ISCO 88 (ZUMA coding), last job, generated	generated	n. in Q vers.
mpslewt	no	yes	Magnitude-Prestige-Scale, last job, gen.	generated	n. in Q vers.
siopslewt	no	yes	Standard International Occupational Prestige Scale, last job, generated	generated	n. in Q vers.
iseilewt	no	yes	International Socio-Economic Index, last job, gen.	generated	n. in Q vers.
egplewt	no	yes	Class scheme acc. to Erikson, Goldthorpe & Portocarrero (EGP), last job	generated	n. in Q vers.
eseclewt	no	yes	European Socio-economic Classification (ESeC), last job, generated	generated	n. in Q vers.
iscolewt_it	no	yes	ISCO 88 (Infratest coding), last job, generated	generated	n. in Q vers.
kldblewt	no	yes	Classification of Occupations 1992 (Infratest coding), last job,	generated	n. in Q vers.
begmeewt	no	yes	Month in which first job taken up, generated	generated	n. in Q vers.
begjeewt	no	yes	Year in which first job taken up, generated	generated	n. in Q vers.
stibeewt	no	yes	Occupational status, first job, code number, generated	generated	n. in Q vers.
iscoeewt	no	yes	ISCO 88 (ZUMA coding), first job, generated	generated	n. in Q vers.
mpseewt	no	yes	Magnitude-Prestige-Scale, first job, gen.	generated	n. in Q vers.
siopseewt	no	yes	Standard International Occupational Prestige Scale, first job, gen.	generated	n. in Q vers.
iseieewt	no	yes	International Socio-Economic Index, first job, gen.	generated	n. in Q vers.
egpeewt	no	yes	Class scheme acc. to Erikson, Goldthorpe & Portocarrero (EGP), first job	generated	n. in Q vers.
eseceewt	no	yes	European Socio-economic Classification (ESeC), first job, generated	generated	n. in Q vers.
iscoeewt_it	no	yes	ISCO 88 (Infratest coding), first job, generated	generated	n. in Q vers.
kldbeewt	no	yes	Classification of Occupations 1992 (Infratest coding), first job, gen.	generated	n. in Q vers.
gebhalbj	yes	yes	Half-year of birth, generated	generated	generated
ogebland	yes	yes	Country of birth, incl. open info., categories (anonymised)	generated	generated
ostaatan	yes	yes	Nationality, incl. open info., categories (anonymised)	generated	generated
ozulanda	yes	yes	Father: country of residence before migration, incl. open info., categories (anonymised)	generated	generated

Variable name in SUF	Surveyed in		Description	No. in wave 2 questionn.:	
	W1	W2		Personal	Senior cit.
ozulandb	yes	yes	Mother: country of residence before migration, incl. open info., categories (anonymised)	generated	generated
ozulandc	yes	yes	Father's father: country of residence before migration, incl. open info., categories (anonymised)	generated	generated
ozulandd	yes	yes	Father's mother: country of residence before migration, incl. open info., categories (anonymised)	generated	generated
ozulande	yes	yes	Mother's father: country of residence before migration, incl. open info., categories (anonymised)	generated	generated
ozulandf	yes	yes	Mother's mother: country of residence before migration, incl. open info., categories (anonymised)	generated	generated
migration	yes	yes	Respondent's migration background, generated	generated	n. in Q vers.

Table 44: List of correspondence between the household questionnaires and the Unemployment Benefit II spell dataset (alg2_spells) in wave 2

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn. HHneu	No. in wave 2 questionn. HHalt
hnr	yes	yes	Household number (current)	system	system
spellnr	yes	yes	Spell of UB II: spell indicator	generated	generated
bmonat	yes	yes	Spell of UB II: starting month, generated	generated	generated
bjahr	yes	yes	Spell of UB II: starting year, generated	generated	generated
emonat	yes	yes	Spell of UB II: ending month, generated	generated	generated
ejahr	yes	yes	Spell of UB II: ending year, generated	generated	generated
AL20100	yes	yes	Spell of UB II: starting month	HH50	HH93
AL20200	yes	yes	Spell of UB II: starting year	HH50	HH93
AL20300	yes	yes	Spell of UB II: ending month	HH50	HH93
AL20400	yes	yes	Spell of UB II: ending year	HH50	HH93
AL20500	yes	yes	Spell of UB II: still ongoing in interview month	HH51	HH94
AL20600	yes	no	UB II: receipt for all HH members in wave 1?	not in wave	not in wave
AL20700a	yes	no	UB II: HH receives UB II for person 1 in wave 1	not in wave	not in wave
AL20700b	yes	no	UB II: HH receives UB II for person 2 in wave 1	not in wave	not in wave
AL20700c	yes	no	UB II: HH receives UB II for person 3 in wave 1	not in wave	not in wave
AL20700d	yes	no	UB II: HH receives UB II for person 4 in wave 1	not in wave	not in wave
AL20700e	yes	no	UB II: HH receives UB II for person 5 in wave 1	not in wave	not in wave
AL20700f	yes	no	UB II: HH receives UB II for person 6 in wave 1	not in wave	not in wave
AL20700g	yes	no	UB II: HH receives UB II for person 7 in wave 1	not in wave	not in wave
AL20700h	yes	no	UB II: HH receives UB II for person 8 in wave 1	not in wave	not in wave
AL20700i	yes	no	UB II: HH receives UB II for person 9 in wave 1	not in wave	not in wave
AL20700j	yes	no	UB II: HH receives UB II for person 10 in wave 1	not in wave	not in wave
AL20700k	yes	no	UB II: HH receives UB II for person 11 in wave 1	not in wave	not in wave
AL20700l	yes	no	UB II: HH receives UB II for person 12 in wave 1	not in wave	not in wave
AL20700m	yes	no	UB II: HH receives UB II for person 13 in wave 1	not in wave	not in wave
AL20700n	yes	no	UB II: HH receives UB II for person 14 in wave 1	not in wave	not in wave
AL20700o	yes	no	UB II: HH receives UB II for person 15 in wave 1	not in wave	not in wave
AL20800	yes	no	UB II: Amount of UB II received per month in wave 1	not in wave	not in wave
AL20900	yes	no	UB II: UB II cut begun? (reported in wave 1)	not in wave	not in wave
AL20601	no	yes	UB II: receipt for all HH members in wave 2?	HH52	HH95
AL20701a	no	yes	UB II: HH receives UB II for person 1 in wave 2	HH53	HH96
AL20701b	no	yes	UB II: HH receives UB II for person 2 in wave 2	HH53	HH96
AL20701c	no	yes	UB II: HH receives UB II for person 3 in wave 2	HH53	HH96
AL20701d	no	yes	UB II: HH receives UB II for person 4 in wave 2	HH53	HH96
AL20701e	no	yes	UB II: HH receives UB II for person 5 in wave 2	HH53	HH96
AL20701f	no	yes	UB II: HH receives UB II for person 6 in wave 2	HH53	HH96
AL20701g	no	yes	UB II: HH receives UB II for person 7 in wave 2	HH53	HH96
AL20701h	no	yes	UB II: HH receives UB II for person 8 in wave 2	HH53	HH96
AL20701i	no	yes	UB II: HH receives UB II for person 9 in wave 2	HH53	HH96
AL20701j	no	yes	UB II: HH receives UB II for person 10 in wave 2	HH53	HH96
AL20701k	no	yes	UB II: HH receives UB II for person 11 in wave 2	HH53	HH96
AL20701l	no	yes	UB II: HH receives UB II for person 12 in wave 2	HH53	HH96
AL20701m	no	yes	UB II: HH receives UB II for person 13 in wave 2	HH53	HH96
AL20701n	no	yes	UB II: HH receives UB II for person 14 in wave 2	HH53	HH96
AL20701o	no	yes	UB II: HH receives UB II for person 15 in wave 2	HH53	HH96
AL20801	no	yes	UB II: Amount of UB II received per month in wave 2	HH54	HH97
AL20901	no	yes	UB II: UB II cut begun? (reported in wave 2)	HH55	HH98
alg2kbma	yes	yes	UB II: 1st benefit cut: starting month, generated	generated	generated
alg2kbja	yes	yes	UB II: 1st benefit cut: starting year, generated	generated	generated
alg2kema	yes	yes	UB II: 1st benefit cut: ending month, generated	generated	generated
alg2keja	yes	yes	UB II: 1st benefit cut: ending year, generated	generated	generated
AL21000a	yes	yes	UB II: 1st benefit cut: starting month	HH57	HH100
AL21100a	yes	yes	UB II: 1st benefit cut: starting year	HH57	HH100
AL21200a	yes	yes	UB II: 1st benefit cut: duration of benefit cut in days	HH58a	HH101a
AL21201a	no	yes	UB II: 1st benefit cut: ending month	HH58b	HH101b
AL21202a	no	yes	UB II: 1st benefit cut: ending year	HH58b	HH101b
AL21300a	yes	yes	Reason for 1st benefit cut: refusal to conclude integration agreement	HH56a	HH99a
AL21400a	yes	yes	Reason for 1st benefit cut: advisory interview appointment not kept	HH56b	HH99b

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn. HHneu	No. in wave 2 questionn. HHalt
AL21500a	yes	yes	Reason for 1st benefit cut: insufficient own job-search effort	HH56c	HH99c
AL21600a	yes	yes	Reason for 1st benefit cut: employment, training, work opportunity, measure not taken up	HH56d	HH99d
AL21700a	yes	yes	Reason for 1st benefit cut: training, work opportunity, measure ended prematurely	HH56e	HH99e
AL21800a	yes	yes	Reason for 1st benefit cut: income/assets not reported correctly	HH56f	HH99f
AL21850a	no	yes	Reason for 1st benefit cut: accommodation too large/too expensive	HH56g	HH99g
AL21900a	yes	no	Reason for 1st benefit cut: uneconomical behaviour	not in wave	not in wave
AL22000a	yes	yes	Reason for 1st benefit cut: increase in income/assets	HH56h	HH99h
AL22100a	yes	yes	Reason for 1st benefit cut: other reasons, yes/no	HH56i	HH99i
AL21301a	yes	yes	Reason for 1st benefit cut: refusal to conclude integration agreement, + open info.	generated	generated
AL21401a	yes	yes	Reason for 1st benefit cut: advisory interview appointment not kept, + open info.	generated	generated
AL21501a	yes	yes	Reason for 1st ben. cut: insufficient own job-search effort, + open info.	generated	generated
AL21601a	yes	yes	Reason for 1st benefit cut: employment, training, work opportunity, measure not taken up, + open info.	generated	generated
AL21701a	yes	yes	Reason for 1st benefit cut: training, work opportunity, measure ended prematurely, + open info.	generated	generated
AL21801a	yes	yes	Reason for 1st benefit cut: income/assets not reported correctly, + open info.	generated	generated
AL21851a	no	yes	Reason for 1st benefit cut: accomm. too large/too expensive, + open info.	generated	generated
AL21901a	yes	no	Reason for 1st benefit cut: uneconomical behaviour, + open info.	not in wave	not in wave
AL22001a	yes	yes	Reason for 1st benefit cut: increase in income/assets, + open info.	generated	generated
AL22101a	yes	yes	Reason for 1st ben. cut: other sanctions, yes/no; open response	generated	generated
AL22102a	yes	no	Reason for 1st benefit cut: size of accommodation, rent inappropriate; open response	not in wave	not in wave
AL22103a	yes	yes	Reason for 1st benefit cut: other reasons (not sanctions); open response	generated	generated
AL22150a	no	yes	UB II: 1st benefit cut: which HH member's benefit reduced, gen.	HH59	HH102
AL22170a	no	yes	UB II: 1st benefit cut: follow-up validation question, whether a HH member had their benefit cut	HH59a	HH102a
alg2kbmb	yes	yes	UB II: 2nd benefit cut: starting month, generated	generated	generated
alg2kbjb	yes	yes	UB II: 2nd benefit cut: starting year, generated	generated	generated
alg2kemb	yes	yes	UB II: 2nd benefit cut: ending month, generated	generated	generated
alg2kejb	yes	yes	UB II: 2nd benefit cut: ending year, generated	generated	generated
AL21000b	yes	yes	UB II: 2nd benefit cut: starting month	HH57	HH100
AL21100b	yes	yes	UB II: 2nd benefit cut: starting year	HH57	HH100
AL21200b	yes	yes	UB II: 2nd benefit cut: duration of benefit cut in days	HH58a	HH101a
AL21201b	no	yes	UB II: 2nd benefit cut: ending month	HH58b	HH101b
AL21202b	no	yes	UB II: 2nd benefit cut: ending year	HH58b	HH101b
AL21300b	yes	yes	Reason for 2nd benefit cut: refusal to conclude integration agreement	HH56a	HH99a
AL21400b	yes	yes	Reason for 2nd benefit cut: advisory interview appointment not kept	HH56b	HH99b
AL21500b	yes	yes	Reason for 2nd benefit cut: insufficient own job-search effort	HH56c	HH99c
AL21600b	yes	yes	Reason for 2nd benefit cut: employment, training, work opportunity, measure not taken up	HH56d	HH99d
AL21700b	yes	yes	Reason for 2nd benefit cut: training, work opportunity, measure ended prematurely	HH56e	HH99e
AL21800b	yes	yes	Reason for 2nd benefit cut: income/assets not reported correctly	HH56f	HH99f
AL21850b	no	yes	Reason for 2nd benefit cut: accommodation too large/too expensive	HH56g	HH99g
AL21900b	yes	no	Reason for 2nd benefit cut: uneconomical behaviour	not in wave	not in wave
AL22000b	yes	yes	Reason for 2nd benefit cut: increase in income/assets	HH56h	HH99h
AL22100b	yes	yes	Reason for 2nd benefit cut: other reasons, yes/no	HH56i	HH99i
AL21301b	yes	yes	Reason for 2nd benefit cut: refusal to conclude integration agreement, + open info.	generated	generated
AL21401b	yes	yes	Reason for 2nd benefit cut: advisory interview appointment not kept, + open info.	generated	generated
AL21501b	yes	yes	Reason for 2nd benefit cut: insufficient own job-search effort, + open info.	generated	generated

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn. HHneu	No. in wave 2 questionn. HHalt
AL21601b	yes	yes	Reason for 2nd benefit cut: employment, training, work opportunity, measure not taken up, + open info.	generated	generated
AL21701b	yes	yes	Reason for 2nd benefit cut: training, work opportunity, measure ended prematurely, + open info.	generated	generated
AL21801b	yes	yes	Reason for 2nd benefit cut: income/assets not reported correctly, + open info.	generated	generated
AL21851b	no	yes	Reason for 2nd benefit cut: accomm. too large/too expensive, + open info.	generated	generated
AL21901b	yes	no	Reason for 2nd benefit cut: uneconomical behaviour, + open info.	not in wave	not in wave
AL22001b	yes	yes	Reason for 2nd benefit cut: increase in income/assets, + open info.	generated	generated
AL22101b	yes	yes	Reason for 2nd benefit cut: other sanctions, yes/no; open response	generated	generated
AL22102b	yes	no	Reason for 2nd benefit cut: size of accommodation, rent inappropriate; open response	not in wave	not in wave
AL22103b	yes	yes	Reason for 2nd benefit cut: other reasons (not sanctions); open response	generated	generated
AL22150b	no	yes	UB II: 2nd benefit cut: which HH member's benefit reduced, gen.	HH59	HH102
AL22170b	no	yes	UB II: 2nd benefit cut: follow-up validation question, whether a HH member had their benefit cut	HH59a	HH102a
alg2kbmc	yes	yes	UB II: 3rd benefit cut: starting month, generated	generated	generated
alg2kbjc	yes	yes	UB II: 3rd benefit cut: starting year, generated	generated	generated
alg2kemc	yes	yes	UB II: 3rd benefit cut: ending month, generated	generated	generated
alg2kejc	yes	yes	UB II: 3rd benefit cut: ending year, generated	generated	generated
AL21000c	yes	yes	UB II: 3rd benefit cut: starting month	HH57	HH100
AL21100c	yes	yes	UB II: 3rd benefit cut: starting year	HH57	HH100
AL21200c	yes	yes	UB II: 3rd benefit cut: duration of benefit cut in days	HH58a	HH101a
AL21201c	no	yes	UB II: 3rd benefit cut: ending month	HH58b	HH101b
AL21202c	no	yes	UB II: 3rd benefit cut: ending year	HH58b	HH101b
AL21300c	yes	yes	Reason for 3rd benefit cut: refusal to conclude integration agreement	HH56a	HH99a
AL21400c	yes	yes	Reason for 3rd benefit cut: advisory interview appointment not kept	HH56b	HH99b
AL21500c	yes	yes	Reason for 3rd benefit cut: insufficient own job-search effort	HH56c	HH99c
AL21600c	yes	yes	Reason for 3rd benefit cut: employment, training, work opportunity, measure not taken up	HH56d	HH99d
AL21700c	yes	yes	Reason for 3rd benefit cut: training, work opportunity, measure ended prematurely	HH56e	HH99e
AL21800c	yes	yes	Reason for 3rd benefit cut: income/assets not reported correctly	HH56f	HH99f
AL21850c	no	yes	Reason for 3rd benefit cut: accomm. too large/too expensive	HH56g	HH99g
AL21900c	yes	no	Reason for 3rd benefit cut: uneconomical behaviour	not in wave	not in wave
AL22000c	yes	yes	Reason for 3rd benefit cut: increase in income/assets	HH56h	HH99h
AL22100c	yes	yes	Reason for 3rd benefit cut: other reasons, yes/no	HH56i	HH99i
AL21301c	yes	yes	Reason for 3rd benefit cut: refusal to conclude integration agreement, + open info.	generated	generated
AL21401c	yes	yes	Reason for 3rd benefit cut: advisory interview appointment not kept, + open info.	generated	generated
AL21501c	yes	yes	Reason for 3rd benefit cut: insufficient own job-search effort, + open info.	generated	generated
AL21601c	yes	yes	Reason for 3rd benefit cut: employment, training, work opportunity, measure not taken up, + open info.	generated	generated
AL21701c	yes	yes	Reason for 3rd benefit cut: training, work opportunity, measure ended prematurely, + open info.	generated	generated
AL21801c	yes	yes	Reason for 3rd benefit cut: income/assets not reported correctly, + open info.	generated	generated
AL21851c	no	yes	Reason for 3rd benefit cut: accomm. too large/too expensive, + open info.	generated	generated
AL21901c	yes	no	Reason for 3rd benefit cut: uneconomical behaviour, + open info.	not in wave	not in wave
AL22001c	yes	yes	Reason for 3rd benefit cut: increase in income/assets, + open info.	generated	generated
AL22101c	yes	yes	Reason for 3rd benefit cut: other sanctions, yes/no; open response	generated	generated
AL22102c	yes	no	Reason for 3rd benefit cut: size of accommodation, rent inappropriate; open response	not in wave	not in wave
AL22103c	yes	yes	Reason for 3rd benefit cut: other reasons (not sanctions); open response	generated	generated
AL22150c	no	yes	UB II: 3rd benefit cut: which HH member's benefit reduced, gen.	HH59	HH102

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn. HHneu	No. in wave 2 questionn. HHalt
AL22170c	no	yes	UB II: 3rd benefit cut: follow-up validation question, whether a HH member had their benefit cut	HH59a	HH102a
alg2kbmd	yes	yes	UB II: 4th benefit cut: starting month, generated	generated	generated
alg2kbjd	yes	yes	UB II: 4th benefit cut: starting year, generated	generated	generated
alg2kemd	yes	yes	UB II: 4th benefit cut: ending month, generated	generated	generated
alg2kejd	yes	yes	UB II: 4th benefit cut: ending year, generated	generated	generated
AL21000d	yes	yes	UB II: 4th benefit cut: starting month	HH57	HH100
AL21100d	yes	yes	UB II: 4th benefit cut: starting year	HH57	HH100
AL21200d	yes	yes	UB II: 4th benefit cut: duration of benefit cut in days	HH58a	HH101a
AL21201d	no	yes	UB II: 4th benefit cut: ending month	HH58b	HH101b
AL21202d	no	yes	UB II: 4th benefit cut: ending year	HH58b	HH101b
AL21300d	yes	yes	Reason for 4th benefit cut: refusal to conclude integration agreement	HH56a	HH99a
AL21400d	yes	yes	Reason for 4th benefit cut: advisory interview appointment not kept	HH56b	HH99b
AL21500d	yes	yes	Reason for 4th benefit cut: insufficient own job-search effort	HH56c	HH99c
AL21600d	yes	yes	Reason for 4th benefit cut: employment, training, work opportunity, measure not taken up	HH56d	HH99d
AL21700d	yes	yes	Reason for 4th benefit cut: training, work opportunity, measure ended prematurely	HH56e	HH99e
AL21800d	yes	yes	Reason for 4th benefit cut: income/assets not reported correctly	HH56f	HH99f
AL21850d	no	yes	Reason for 4th benefit cut: accomm. too large/too expensive	HH56g	HH99g
AL21900d	yes	no	Reason for 4th benefit cut: uneconomical behaviour	not in wave	not in wave
AL22000d	yes	yes	Reason for 4th benefit cut: increase in income/assets	HH56h	HH99h
AL22100d	yes	yes	Reason for 4th benefit cut: other reasons, yes/no	HH56i	HH99i
AL21301d	yes	yes	Reason for 4th benefit cut: refusal to conclude integration agreement, + open info.	generated	generated
AL21401d	yes	yes	Reason for 4th benefit cut: advisory interview appointment not kept, + open info.	generated	generated
AL21501d	yes	yes	Reason for 4th benefit cut: insufficient own job-search effort, + open info.	generated	generated
AL21601d	yes	yes	Reason for 4th benefit cut: employment, training, work opportunity, measure not taken up, + open info.	generated	generated
AL21701d	yes	yes	Reason for 4th benefit cut: training, work opportunity, measure ended prematurely, + open info.	generated	generated
AL21801d	yes	yes	Reason for 4th benefit cut: income/assets not reported correctly, + open info.	generated	generated
AL21851d	no	yes	Reason for 4th benefit cut: accomm. too large/too expensive, + open info.	generated	generated
AL21901d	yes	no	Reason for 4th benefit cut: uneconomical behaviour, + open info.	not in wave	not in wave
AL22001d	yes	yes	Reason for 4th benefit cut: increase in income/assets, + open info.	generated	generated
AL22101d	yes	yes	Reason for 4th benefit cut: other sanctions, yes/no; open response	generated	generated
AL22102d	yes	no	Reason for 4th benefit cut: size of accommodation, rent inappropriate; open response	not in wave	not in wave
AL22103d	yes	yes	Reason for 4th benefit cut: other reasons (not sanctions); open response	generated	generated
AL22150d	no	yes	UB II: 4th benefit cut: which HH member's benefit reduced, gen.	HH59	HH102
AL22170d	no	yes	UB II: 4th benefit cut: follow-up validation question, whether a HH member had their benefit cut	HH59a	HH102a
alg2kbme	yes	yes	UB II: 5th benefit cut: starting month, generated	generated	generated
alg2kbje	yes	yes	UB II: 5th benefit cut: starting year, generated	generated	generated
alg2keme	yes	yes	UB II: 5th benefit cut: ending month, generated	generated	generated
alg2keje	yes	yes	UB II: 5th benefit cut: ending year, generated	generated	generated
AL21000e	yes	yes	UB II: 5th benefit cut: starting month	HH57	HH100
AL21100e	yes	yes	UB II: 5th benefit cut: starting year	HH57	HH100
AL21200e	yes	yes	UB II: 5th benefit cut: duration of benefit cut in days	HH58a	HH101a
AL21201e	no	yes	UB II: 5th benefit cut: ending month	HH58b	HH101b
AL21202e	no	yes	UB II: 5th benefit cut: ending year	HH58b	HH101b
AL21300e	yes	yes	Reason for 5th benefit cut: refusal to conclude integration agreement	HH56a	HH99a
AL21400e	yes	yes	Reason for 5th benefit cut: advisory interview appointment not kept	HH56b	HH99b
AL21500e	yes	yes	Reason for 5th benefit cut: insufficient own job-search effort	HH56c	HH99c
AL21600e	yes	yes	Reason for 5th benefit cut: employment, training, work opportunity, measure not taken up	HH56d	HH99d

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn. HHneu	No. in wave 2 questionn. HHalt
AL21700e	yes	yes	Reason for 5th benefit cut: training, work opportunity, measure ended prematurely	HH56e	HH99e
AL21800e	yes	yes	Reason for 5th benefit cut: income/assets not reported correctly	HH56f	HH99f
AL21850e	no	yes	Reason for 5th benefit cut: accomm. too large/too expensive	HH56g	HH99g
AL21900e	yes	no	Reason for 5th benefit cut: uneconomical behaviour	not in wave	not in wave
AL22000e	yes	yes	Reason for 5th benefit cut: increase in income/assets	HH56h	HH99h
AL22100e	yes	yes	Reason for 5th benefit cut: other reasons, yes/no	HH56i	HH99i
AL21301e	yes	yes	Reason for 5th benefit cut: refusal to conclude integration agreement, + open info.	generated	generated
AL21401e	yes	yes	Reason for 5th benefit cut: advisory interview appointment not kept, + open info.	generated	generated
AL21501e	yes	yes	Reason for 5th benefit cut: insufficient own job-search effort, + open info.	generated	generated
AL21601e	yes	yes	Reason for 5th benefit cut: employment, training, work opportunity, measure not taken up, + open info.	generated	generated
AL21701e	yes	yes	Reason for 5th benefit cut: training, work opportunity, measure ended prematurely, + open info.	generated	generated
AL21801e	yes	yes	Reason for 5th benefit cut: income/assets not reported correctly, + open info.	generated	generated
AL21851e	no	yes	Reason for 5th benefit cut: accomm. too large/too expensive, + open info.	generated	generated
AL21901e	yes	no	Reason for 5th benefit cut: uneconomical behaviour, + open info.	not in wave	not in wave
AL22001e	yes	yes	Reason for 5th benefit cut: increase in income/assets, + open info.	generated	generated
AL22101e	yes	yes	Reason for 5th benefit cut: other sanctions, yes/no; open response	generated	generated
AL22102e	yes	no	Reason for 5th benefit cut: size of accommodation, rent inappropriate; open response	not in wave	not in wave
AL22103e	yes	yes	Reason for 5th benefit cut: other reasons (not sanctions); open response	generated	generated
AL22150e	no	yes	UB II: 5th benefit cut: which HH member's benefit reduced, gen.	HH59	HH102
AL22170e	no	yes	UB II: 5th benefit cut: follow-up validation question, whether a HH member had their benefit cut	HH59a	HH102a
zensiert	yes	yes	UB II spell: spell still ongoing at time of last HH interview (right-censored), gen.	generated	generated

Table 45: List of correspondence between the personal questionnaire and the unemployment spell dataset (al_spells) in wave 2

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn.: Personal	No. in wave 2 questionn.: Senior cit.
pnr	no	yes	Constant personal ID number	system	n. in Q vers.
spellnr	no	yes	Spell number	generated	n. in Q vers.
bmonat	no	yes	Registered unemployment: starting month, generated	generated	n. in Q vers.
bjahr	no	yes	Registered unemployment: starting year, generated	generated	n. in Q vers.
emonat	no	yes	Registered unemployment: ending month, generated	generated	n. in Q vers.
ejahr	no	yes	Registered unemployment: ending year, generated	generated	n. in Q vers.
AL0100	no	yes	Registered unemployment: starting month	P108	n. in Q vers.
AL0200	no	yes	Registered unemployment: starting year	P108	n. in Q vers.
AL0300	no	yes	Registered unemployment: ending month	P109	n. in Q vers.
AL0400	no	yes	Registered unemployment: ending year	P109	n. in Q vers.
AL0500	no	yes	Currently unemployed/participating in a measure run by the Fed. Empl. Agency	P110	n. in Q vers.
AL0600	no	yes	Reason why no longer registered as unemployed	P111	n. in Q vers.
AL0601	no	yes	Reason why no longer registered as unemployed, incl. open info.	generated	n. in Q vers.
AL0700	no	yes	Indicator: receipt of UB I since Jan. 2005?	P112	n. in Q vers.
AL0800	no	yes	Receipt of UBI: starting month	P113	n. in Q vers.
AL0900	no	yes	Receipt of UBI: starting year	P113	n. in Q vers.
AL1000	no	yes	Receipt of UBI: ending month	P114	n. in Q vers.
AL1100	no	yes	Receipt of UBI: ending year	P114	n. in Q vers.
AL1200	no	yes	Receipt of UBI: still ongoing in interview month	P115	n. in Q vers.
alg1bm	no	yes	Receipt of UBI: starting month, generated	generated	n. in Q vers.
alg1bj	no	yes	Receipt of UBI: starting year, generated	generated	n. in Q vers.
alg1em	no	yes	Receipt of UBI: ending month, generated	generated	n. in Q vers.
alg1ej	no	yes	Receipt of UBI: ending year, generated	generated	n. in Q vers.
alg1akt	no	yes	Receipt of UBI: spell currently still ongoing (right-censoring)	generated	n. in Q vers.
AL1300	no	yes	UB I: amount of benefit received per month	P116	n. in Q vers.
zensiert	no	yes	Registered unemployment: spell currently still ongoing (right-censoring)	generated	n. in Q vers.

Table 46: List of correspondence between the personal questionnaire and the employment spell dataset (et_spells) in wave 2

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn.: Personal	No. in wave 2 questionn.: Senior cit.
pnr	no	yes	Constant personal ID number	system	n. in Q vers.
spellnr	no	yes	Spell number	generated	n. in Q vers.
bmonat	no	yes	Employment: starting month, generated	generated	n. in Q vers.
bjahr	no	yes	Employment: starting year, generated	generated	n. in Q vers.
emonat	no	yes	Employment: ending month, generated	generated	n. in Q vers.
ejahr	no	yes	Employment: ending year, generated	generated	n. in Q vers.
ET0100	no	yes	Employment: starting month	P41	n. in Q vers.
ET0200	no	yes	Employment: starting year	P41	n. in Q vers.
ET0300	no	yes	Employment: ending month	P42	n. in Q vers.
ET0400	no	yes	Employment: ending year	P42	n. in Q vers.
ET0500	no	yes	Currently still exercising this occupation?	P43	n. in Q vers.
ET0600	no	yes	Occupational status, simple classification (anon.)	P44	n. in Q vers.
ET0700	no	yes	Occup. status blue-collar worker: detailed info.	P45	n. in Q vers.
ET0800	no	yes	Occup. status white-collar worker: detailed info.	P46	n. in Q vers.
ET1000	no	yes	Occup. status civil servant: detailed info. (anon.)	P48, P47	n. in Q vers.
ET1100	no	yes	Occup. status self-employed: number of employees	P49	n. in Q vers.
ET1200	no	yes	Occup. status farmer: area farmed	P50	n. in Q vers.
ET1300	no	yes	Employment: supervisory function?	P51	n. in Q vers.
ET1400	no	yes	Employment: number of staff supervised	P52	n. in Q vers.
ET1500	no	yes	Employment: temporary employment?	P53	n. in Q vers.
ET1600	no	yes	Employment: fixed-term contract?	P54	n. in Q vers.
ET1700	no	yes	Employment: conversion into permanent employment relationship later?	P55	n. in Q vers.
ET1800	no	yes	Employment: public sector?	P56	n. in Q vers.
ET1900	no	yes	Employment: no. of employees in establishment/local office	P57	n. in Q vers.
ET2000	no	yes	Employment: weekly working hours (excl. overtime)	P58	n. in Q vers.
ET2100	no	yes	Employment: actual weekly working hours (incl. overtime)	P59	n. in Q vers.
ET2200	no	yes	Employment: average working hours per week with irregular hours of work	P60	n. in Q vers.
ET2300	no	yes	Reason for termination of employment contract	P62	n. in Q vers.
zensiert	no	yes	Employment: spell currently still ongoing (right-censoring)	generated	n. in Q vers.

Table 47: List of correspondence between the personal questionnaire and the gap spell dataset (lu_spells) in wave 2

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn.: Personal	No. in wave 2 questionn.: Senior cit.
pnr	no	yes	Constant personal ID number	system	n. in Q vers.
spellnr	no	yes	Spell number	generated	n. in Q vers.
bmonat	no	yes	Spell: starting month, generated	generated	n. in Q vers.
bjahr	no	yes	Spell: starting year, generated	generated	n. in Q vers.
emonat	no	yes	Spell: ending month, generated	generated	n. in Q vers.
ejahr	no	yes	Spell: ending year, generated	generated	n. in Q vers.
LU0100	no	yes	Employment status / economic inactivity status gap	P128	n. in Q vers.
LU0101	no	yes	Employment status / economic inactivity status gap (+ open info.)	generated	n. in Q vers.
LU0200	no	yes	Spell: starting month	P130	n. in Q vers.
LU0300	no	yes	Spell: starting year	P130	n. in Q vers.
LU0400	no	yes	Spell: ending month	P130	n. in Q vers.
LU0500	no	yes	Spell: ending year	P130	n. in Q vers.
LU0600	no	yes	Gap status currently still ongoing?	P131	n. in Q vers.
LU0700	no	yes	Person doing compulsory military service/community service/gap year for voluntary work or similar: monthly pay/remuneration	P132	n. in Q vers.
zensiert	no	yes	Spell: spell currently still ongoing (right-censoring)	generated	n. in Q vers.

Table 48: List of correspondence between the personal questionnaire and the employment and training measure spell dataset (mn_spells) in wave 2

Variable name in SUF	Surveyed in		Description	No. in wave 2 questionn.:	
	W1	W2		Personal	Senior cit.
pnr	no	yes	Constant personal ID number	system	n. in Q vers.
spellnr	no	yes	Spell number	generated	n. in Q vers.
bmonat	no	yes	Measure: starting month, generated	generated	n. in Q vers.
bjahr	no	yes	Measure: starting year, generated	generated	n. in Q vers.
emonat	no	yes	Measure: ending month, generated	generated	n. in Q vers.
ejahr	no	yes	Measure: ending year, generated	generated	n. in Q vers.
MN0100	no	yes	Measure: type of scheme	P161	n. in Q vers.
MN0200a	no	yes	Part of measure: vocational aptitude test	P162	n. in Q vers.
MN0200b	no	yes	Part of measure: German language course	P162	n. in Q vers.
MN0200c	no	yes	Part of measure: course in another language	P162	n. in Q vers.
MN0200d	no	yes	Part of measure: course to learn new skills	P162	n. in Q vers.
MN0200e	no	yes	Part of measure: period of work experience	P162	n. in Q vers.
MN0200f	no	yes	Part of measure: employment in practice firm	P162	n. in Q vers.
MN0200g	no	yes	Part of measure: preparation for self-employment	P162	n. in Q vers.
MN0200h	no	yes	Part of measure: other	P162	n. in Q vers.
MN0201a	no	yes	Part of measure: vocational aptitude test, incl. open info.	generated	n. in Q vers.
MN0201b	no	yes	Part of measure: German language course, incl. open info.	generated	n. in Q vers.
MN0201c	no	yes	Part of measure: course in another language, incl. open info.	generated	n. in Q vers.
MN0201d	no	yes	Part of measure: course to learn new skills, incl. open info.	generated	n. in Q vers.
MN0201e	no	yes	Part of measure: period of work experience, incl. open info.	generated	n. in Q vers.
MN0201f	no	yes	Part of measure: employment in practice firm, incl. open info.	generated	n. in Q vers.
MN0201g	no	yes	Part of measure: preparation for self-employment, incl. open info.	generated	n. in Q vers.
MN0201h	no	yes	Part of measure: other	generated	n. in Q vers.
MN0202h	no	yes	Part of measure: job application training	generated	n. in Q vers.
MN0300	no	yes	Start of measure: starting month	P163	n. in Q vers.
MN0400	no	yes	Start of measure: starting year	P163	n. in Q vers.
MN0500	no	yes	Measure: current participation in measure?	P164	n. in Q vers.
MN0600	no	yes	Measure: duration of measure in days (completed measure)	P165_a	n. in Q vers.
MN0700	no	yes	Measure: ending month (completed measure)	P165_b	n. in Q vers.
MN0800	no	yes	Measure: ending year (completed measure)	P165_b	n. in Q vers.
MN0900	no	yes	Measure: premature termination of measure?	P166	n. in Q vers.
MN1000a	no	yes	Reason for ending measure prematurely: dismissal	P167	n. in Q vers.
MN1000b	no	yes	Reason for ending measure prematurely: no interest	P167	n. in Q vers.
MN1000c	no	yes	Reason for ending measure prematurely: health reasons	P167	n. in Q vers.
MN1000d	no	yes	Reason for ending measure prematurely: training, studies or job	P167	n. in Q vers.
MN1000e	no	yes	Reason for ending measure prematurely: other reasons	P167	n. in Q vers.
MN1001a	no	yes	Reason for ending measure prematurely: dismissal, incl. open info.	generated	n. in Q vers.
MN1001b	no	yes	Reason for ending measure prematurely: no interest, incl. open info.	generated	n. in Q vers.
MN1001c	no	yes	Reason for ending measure prematurely: health reasons, incl. open info.	generated	n. in Q vers.
MN1001d	no	yes	Reason for ending measure prematurely: training/studies/job, incl. open info.	generated	n. in Q vers.
MN1001e	no	yes	Reason for ending measure prematurely: other reasons, incl. open info.	generated	n. in Q vers.
MN1100	no	yes	Measure: duration in days (current measure)	P168_a	n. in Q vers.
MN1200	no	yes	Measure: ending month (current measure)	P168_b	n. in Q vers.
MN1300	no	yes	Measure: ending year (current measure)	P168_b	n. in Q vers.
MN1400	no	yes	Measure: measure specifically for young people	P169	n. in Q vers.
MN1500	no	yes	Measure: name of measure	P170	n. in Q vers.
MN1600	no	yes	Measure: initiative to participate	P171	n. in Q vers.
MN1700a	no	yes	Assessment of measure: improves employment prospects	P172	n. in Q vers.
MN1700b	no	yes	Assessment of measure: is in line with my personal abilities	P172	n. in Q vers.
MN1700c	no	yes	Assessment of measure: improvement of financial situation	P172	n. in Q vers.

Variable name in SUF	Surveyed in		Description	No. in wave 2 questionn.:	
	W1	W2		Personal	Senior cit.
MN1700d	no	yes	Assessment of measure: doing something useful	P172	n. in Q vers.
MN1700e	no	yes	Assessment of measure: degrading	P172	n. in Q vers.
MN1700f	no	yes	Assessment of measure: participation because of threatened benefit cut	P172	n. in Q vers.
MN1700g	no	yes	Assessment of measure: good to meet people again	P172	n. in Q vers.
MN1800	no	yes	Measure: average hours per week	P173	n. in Q vers.
MN1900	no	yes	Measure: training requirements?	P174	n. in Q vers.
MN2000	no	yes	Measure: work identical to that of permanent employees?	P175	n. in Q vers.
MN2100	no	yes	Measure: social education worker present?	P177	n. in Q vers.
mnbranche	no	yes	Measure: economic sector/industry (WZ 2003)	generated	n. in Q vers.
zensiert	no	yes	Measure: spell currently still ongoing (right-censoring)	generated	n. in Q vers.

Table 49: List of variables for the household register dataset (*hh_register*)

Variable name in SUF	Surveyed in		Description	No. in wave 2 questionn.:	
	W1	W2		HHneu	HHalt
hnr	1	1	Household number (current)	system	system
uhnr	1	1	Original household number	system	system
sample	1	1	Sample indicator	system	system
jahrsamp	1	1	Sampling year	system	system
alg2samp	1	1	Receipt of UB II of the household on the sampling date	generated	generated
hnr1	1	0	Household number in wave 1 (2006/2007)	generated	generated
hhnetto1	1	0	Survey status in wave 1 (2006/2007)	generated	generated
hnettok1	1	0	Survey status of the HH (1-digit) in wave 1 (2006/2007)	generated	generated
hnettod1	1	0	Survey status of the HH (2-digit) in wave 1 (2006/2007)	generated	generated
hhgr1	1	0	Household size in wave 1 (2006/2007)	generated	generated
anzbg1	1	0	Number of synthetic benefit communities in household in wave 1 (2006/2007)	generated	generated
hinmod1	1	0	Interview mode of household interview in wave 1 (2006/2007)	generated	generated
pnrzp1	1	0	Personal ID number of household reference person in wave 1 (2006/2007)	generated	generated
hnr2	0	1	Household number in wave 2 (2007/2008)	generated	generated
hnettok2	0	1	Survey status of the HH (1-digit) in wave 2 (2007/2008)	generated	generated
hnettod2	0	1	Survey status of the HH (2-digit) in wave 2 (2007/2008)	generated	generated
hhgr2	0	1	Household size in wave 2 (2007/2008)	generated	generated
nweg2	0	1	Individuals of last interview no longer living in HH in wave 2 (2007/2008)	generated	generated
nneu2	0	1	Individuals who have joined HH since last interview in wave 2 (2007/2008)	generated	generated
anzbg2	0	1	Number of synthetic benefit communities in household in wave 2 (2007/2008)	generated	generated
hinmod2	0	1	Interview mode of household interview in wave 2 (2007/2008)	generated	generated
pnrzp2	0	1	Personal ID number of household reference person in wave 2 (2007/2008)	generated	generated
hnettod2	0	1	Survey status of the HH (2-digit) in wave 2 (2007/2008)	generated	generated

Table 50: List of variables for the person register dataset (*p_register*)

Variable name in SUF	Surveyed in W1	Surveyed in W2	Description	No. in wave 2 questionn.: HHneu	No. in wave 2 questionn.: HHalt
pnr	1	1	Constant personal ID number	system	system
lasthnr	1	1	Last household number	generated	generated
uhnr	1	1	Original household number	system	system
sample	1	1	Sample indicator	system	system
sex	1	1	Gender	generated	generated
korrsex	1	1	Info. on gender was corrected between survey waves	generated	generated
zupanel	1	1	Survey wave in which person joined panel	generated	generated
lastint	1	1	Survey wave of last interview at individual level	generated	generated
hnr1	1	0	Household number in wave 1 (2006/2007)	generated	generated
zplfd1	1	0	Serial number of person in household in wave 1 (2006/2007)	generated	generated
alter1	1	0	Age of person in wave 1 (2006/2007)	generated	generated
bgnr1	1	0	Benefit community number in wave 1 (2006/2007)	generated	generated
bgtyp1	1	0	Type of benefit community in wave 1 (2006/2007)	generated	generated
bgbezs1	1	0	Receipt of UB II of benefit community on sampling date in wave 1 (2006/2007)	generated	generated
bgbezb1	1	0	Receipt of UB II of benefit community on interview date in wave 1 (2006/2007)	generated	generated
pnetto1	1	0	Survey status of person in wave 1 (2006/2007)	generated	generated
pnettok1	1	0	Survey status of person (1-digit) in wave 1 (2006/2007)	generated	generated
pnettod1	1	0	Survey status of person (2-digit) in wave 1 (2006/2007)	generated	generated
zmhh1	1	0	Pointer: personal ID number of target person's mother in HH in wave 1 (2006/2007)	generated	generated
zvhh1	1	0	Pointer: personal ID number of target person's father in HH in wave 1 (2006/2007)	generated	generated
zparthh1	1	0	Pointer: personal ID number of target person's partner in HH in wave 1 (2006/2007)	generated	generated
hnr2	0	1	Household number in wave 2 (2007/2008)	generated	generated
zplfd2	0	1	Serial number of person in household in wave 2 (2007/2008)	generated	generated
alter2	0	1	Age of person in wave 2 (2007/2008)	generated	generated
bgnr2	0	1	Benefit community number in wave 2 (2007/2008)	generated	generated
bgtyp2	0	1	Type of benefit community in wave 2 (2007/2008)	generated	generated
bgbezs2	0	1	Receipt of UB II of benefit community on sampling date in wave 2 (2007/2008)	generated	generated
bgbezb2	0	1	Receipt of UB II of benefit community on interview date in wave 2 (2007/2008)	generated	generated
pnettok2	0	1	Survey status of person (1-digit) in wave 2 (2007/2008)	generated	generated
pnettod2	0	1	Survey status of person (2-digit) in wave 2 (2007/2008)	generated	generated
wegm2	0	1	Month since which person no longer living in previous HH, reported in wave 2 (2007/2008)	generated	generated
wegj2	0	1	Year since which person no longer living in previous HH, reported in wave 2 (2007/20008)	generated	generated
neum2	0	1	Month in which person joined current HH, reported in wave 2 (2007/2008)	generated	generated
neuj2	0	1	Year in which person joined current HH, reported in wave 2 (2007/2008)	generated	generated
zmhh2	0	1	Pointer: personal ID number of target person's mother in HH in wave 2 (2007/2008)	generated	generated
zvhh2	0	1	Pointer: personal ID number of target person's father in HH in wave 2 (2007/2008)	generated	generated
zparthh2	0	1	Pointer: personal ID number of target person's partner in HH in wave 2 (2007/2008)	generated	generated

Table 51: Overview of the ISCO-88 codes and the corresponding values of SIOPS, ISEI, MPS

ISCO88	LABEL	SIOPS	ISEI	MPS	Comments
110	Armed forces			85.3	
111	Armed forces – soldier	39	40		SIOPS & ISEI for soldiers and officers only shown separately
112	Armed forces - officer	65	64		
1000	Legislators, senior officials and managers	51	55	119.3	
1100	Legislators and senior officials	67	70	133.6	
1110	Legislators	64	77	160.3	
1120	Senior government officials	71	77	160.3	
1130	Traditional chiefs and heads of villages	63	66	131.1	
1140	Senior officials of special-interest organisations	63	58	121.1	
1141	Senior officials of political-party organisations	63	58	119.2	
1142	Senior officials of employers', workers' and other economic-interest organisations	63	58	121.3	
1143	Senior officials of humanitarian and other special-interest organisations	63	58	119.2	
1200	Corporate managers	60	68	129.7	
1210	Directors and chief executives	70	70	150.8	
1220	Production and operations department managers	63	67	115.6	
1221	Production and operations department managers in agriculture, hunting, forestry and fishing	60	67	112.3	
1222	Production and operations department managers in manufacturing	60	67	93.9	
1223	Production and operations department managers in construction	60	67	112.3	
1224	Production and operations department managers in wholesale and retail trade	60	59	112.3	
1225	Production and operations department managers in restaurants and hotels	60	59	112.3	
1226	Production and operations department managers in transport, storage and communications	60	59	112.3	
1227	Production and operations department managers in business services	60	87	153.5	
1228	Production and operations department managers in personal care, cleaning and related services	60	59	112.3	
1229	Production and operations department managers not elsewhere classified	60	67	153.8	
1230	Other department managers	60	61	135.3	
1231	Finance and administration department managers	60	69	131.6	
1232	Personnel and industrial relations departments managers	60	69	145.2	
1233	Sales and marketing department managers	60	56	125.4	
1234	Advertising and public relations department managers	60	69	135.8	
1235	Supply and distribution department managers	60	69	160.5	
1236	Computing services department managers	60	69	149.6	
1237	Research and development department managers	60	69	135.8	
1239	Other department managers not elsewhere classified	60	69	134.8	
1300	General managers	50	51	102.7	
1310	General managers	50	51	102.7	
1311	General managers in agriculture, hunting, forestry and fishing	47	43	104.4	
1312	General managers in manufacturing	52	56	108.7	
1313	General managers in construction	52	51	101.1	
1314	General managers in wholesale and retail trade	46	49	105.6	
1315	General managers of restaurants and hotels	38	44	81.6	
1316	General managers in transport, storage and communications	52	51	81.3	
1317	General managers of business services	52	51	144.5	
1318	General managers in personal care, cleaning and related services	52	51	104.4	
1319	General managers not elsewhere classified	52	51	124.1	
2000	Professionals	62	70	135.1	
2100	Physical, mathematical and engineering science professionals	63	69	129.3	
2110	Physicists, chemists and related professionals	69	74	142.3	
2111	Physicists and astronomers	75	74	141.9	
2112	Meteorologists	72	74	141.9	
2113	Chemists	69	74	142.5	
2114	Geologists and geophysicists	67	74	141.9	
2120	Mathematicians, statisticians and related professionals	56	71	129.6	
2121	Mathematicians and related professionals	69	71	129.6	
2122	Statisticians	55	71	129.6	
2130	Computing professionals	51	71	121.0	
2131	Computer systems designers and analysts	51	71	113.8	
2132	Computer programmers	51	71	119.9	
2139	Computing professionals not elsewhere classified	51	71	128.0	
2140	Architects, engineers and related professionals	63	73	130.8	
2141	Architects, town and traffic planners	72	69	133.7	
2142	Civil engineers	70	69	129.1	
2143	Electrical engineers	65	68	126.2	
2144	Electronics and telecommunications engineers	65	68	129.5	

ISCO88	LABEL	SIOPS	ISEI	MPS	Comments
2145	Mechanical engineers	66	67	130.4	
2146	Chemical engineers	66	71	133.4	
2147	Mining engineers, metallurgists and related professionals	61	67	131,5	
2148	Cartographers and surveyors	58	56	120.2	
2149	Architects, engineers and related professionals not elsewhere classified	56	69	135.1	
2200	Life science and health professionals	70	80	161.4	
2210	Life science professionals	62	78	136.9	
2211	Biologists, botanists, zoologists and related professionals	69	77	138.6	
2212	Pharmacologists, pathologists and related professionals	68	77	134.6	
2213	Agronomists and related professionals	56	79	134.6	
2220	Health professionals (except nursing)	73	85	176.3	
2221	Medical doctors	78	88	179.6	
2222	Dentists	70	85	160.5	
2223	Veterinarians	61	83	160.5	
2224	Pharmacists	64	74	173.3	
2229	Health professionals (except nursing) not elsewhere classified	73	85	160.5	
2230	Nursing and midwifery professionals	54	43	70.9	
2290	Life science, health and related professionals not elsewhere classified	66	78	144.6	ZUMA special code
2300	Teaching professionals	61	69	149.4	
2310	College, university and higher education teaching professionals	78	77	159.8	
2320	Secondary education teaching professionals	60	69	149.2	
2330	Primary and pre-primary education teaching professionals	57	66	125.5	
2331	Primary education teaching professionals	57	66	130.3	
2332	Pre-primary education teaching professionals	49	43	67.8	
2340	Special education teaching professionals	62	66	149.2	
2350	Other teaching professionals	62	66	149.2	
2351	Education methods specialists	68	70	149.2	
2352	School inspectors	68	70	149.2	
2359	Other teaching professionals not elsewhere classified	62	65	149.2	
2400	Other professionals	60	68	129.1	
2410	Business professionals	57	69	123.9	
2411	Accountants	62	69	129.2	
2412	Personnel and careers professionals	56	69	96.5	
2419	Business professionals not elsewhere classified	57	69	126.7	
2420	Legal professionals	73	85	164.5	
2421	Lawyers	73	85	170.9	
2422	Judges	76	90	186.8	
2429	Legal professionals not elsewhere classified	71	82	140.9	
2430	Archivists, librarians and related information professionals	54	65	143.2	
2431	Archivists and curators	54	65	143.2	
2432	Librarians and related information professionals	54	65	143.2	
2440	Social science and related professionals	58	65	115.0	
2441	Economists	60	78	132.0	
2442	Sociologists, anthropologists and related professionals	67	71	117.2	
2443	Philosophers, historians and political scientists	67	71	117.2	
2444	Philologists, translators and interpreters	62	65	117.2	
2445	Psychologists	67	71	147.1	
2446	Social work professionals	52	51	91.5	
2450	Writers and creative or performing artists	57	61	109.6	
2451	Authors, journalists and other writers	58	65	113.0	
2452	Sculptors, painters and related artists	57	54	110.7	
2453	Composers, musicians and singers	45	64	101.4	
2454	Choreographers and dancers	40	64	110.7	
2455	Film, stage and related actors and directors	57	64	110.7	
2460	Religious professionals	60	53	142.2	
2470	Administrative professionals in the public service	62	70	130.4	ISCO88-com code
3000	Technicians and associate professionals	48	54	84.9	
3100	Physical and engineering science associate professionals	48	50	82.5	
3110	Physical and engineering science technicians	47	49	78.5	
3111	Chemical and physical science technicians	46	45	69.3	
3112	Civil engineering technicians	39	45	78.8	
3113	Electrical engineering technicians	46	46	83.9	
3114	Electronics and telecommunications engineering technicians	46	46	77.3	
3115	Mechanical engineering technicians	46	54	83.2	
3116	Chemical engineering technicians	46	54	79.1	
3117	Mining and metallurgical technicians	53	54	79.1	
3118	Draughtspersons	55	51	71.7	
3119	Physical and engineering science technicians not elsewhere classified	46	53	78.3	
3120	Computer associate professionals	53	52	108.1	
3121	Computer assistants	53	52	115.8	
3122	Computer equipment operators	53	52	88.9	
3123	Industrial robot controllers	53	52	106.0	
3130	Optical and electronic equipment operators	46	52	93.6	

ISCO88	LABEL	SIOPS	ISEI	MPS	Comments
3131	Photographers and image and sound recording equipment operators	46	48	93.6	
3132	Broadcasting and telecommunications equipment operators	49	57	93.6	
3133	Medical equipment operators	58	57	93.6	
3139	Optical and electronic equipment operators not elsewhere classified	44	52	93.6	
3140	Ship and aircraft controllers and technicians	57	57	127.8	
3141	Ship's engineers	60	52	127.8	
3142	Ship's deck officers and pilots	55	52	127.8	
3143	Aircraft pilots and related associate professionals	60	69	127.8	
3144	Air traffic controllers	50	69	127.8	
3145	Air traffic safety technicians	46	50	127.8	
3150	Safety and quality inspectors	54	50	66.1	
3151	Building and fire inspectors	54	50	67.6	
3152	Safety, health and quality inspectors	54	50	66.0	
3200	Life science and health associate professionals	51	48	77.1	
3210	Life science technicians and related associate professionals	52	50	82.9	
3211	Life science technicians	52	50	82.9	
3212	Agronomy and forestry technicians	47	50	82.9	
3213	Farming and forestry advisers	55	50	82.9	
3220	Modern health associate professionals (except nursing)	51	55	86.8	
3221	Medical assistants	53	51	86.1	
3222	Sanitarians	48	51	86.1	
3223	Dieticians and nutritionists	52	51	86.1	
3224	Optometrists and opticians	60	60	86.1	
3225	Dental assistants	44	51	86.1	
3226	Physiotherapists and related associate professionals	51	60	87.8	
3227	Veterinary assistants	48	51	86.1	
3228	Pharmaceutical assistants	44	51	86.1	
3229	Modern health associate professionals (except nursing) not elsewhere classified	45	51	86.1	
3230	Nursing and midwifery associate professionals	44	38	70.9	
3231	Nursing associate professionals	44	38	70.9	
3232	Midwifery associate professionals	44	38	70.2	
3240	Traditional medicine practitioners and faith healers	29	49	77.0	
3241	Traditional medicine practitioners	29	51	77.0	
3242	Faith healers	22	38	77.0	
3300	Teaching associate professionals	50	38	82.8	
3310	Primary education teaching associate professionals	50	38	82.9	
3320	Pre-primary education teaching associate professionals	50	38	67.8	
3330	Special education teaching associate professionals	50	38	82.9	
3340	Other teaching associate professionals not elsewhere classified	50	38	87.3	
3400	Other associate professionals (middle qualification level)	48	55	87.6	
3410	Finance and sales associate professionals	47	55	87.3	
3411	Securities and finance dealers and brokers	50	61	88.6	
3412	Insurance representatives	44	54	95.4	
3413	Estate agents	49	59	100.3	
3414	Travel consultants and organisers	43	56	88.6	
3415	Technical and commercial sales representatives	46	56	90.5	
3416	Buyers	49	50	76.9	
3417	Appraisers, valuers and auctioneers	46	56	88.6	
3419	Finance and sales associate professionals not elsewhere classified	46	55	78.6	
3420	Business services agents and trade brokers	42	55	93.9	
3421	Trade brokers	55	55	93.9	
3422	Clearing and forwarding agents	50	55	93.9	
3423	Employment agents and labour contractors	49	55	93.9	
3429	Business services agents and trade brokers not elsewhere classified	42	55	93.9	
3430	Administrative associate professionals	49	54	83.1	
3431	Administrative secretaries and related associate professionals	53	54	73.2	
3432	Legal and related business associate professionals	49	59	87.6	
3433	Bookkeepers	49	51	90.2	
3434	Statistical, mathematical and related associate professionals	51	61	82.9	
3439	Administrative associate professionals not elsewhere classified	53	54	69.9	
3440	Customs, tax and related government associate professionals	52	56	85.7	
3441	Customs and border inspectors	44	56	89.8	
3442	Government tax and excise officials	52	57	86.1	
3443	Government social benefits officials	55	56	79.7	
3444	Government licensing officials	54	46	89.8	
3449	Customs, tax and related government associate professionals not elsewhere classified	55	56	88.9	
3450	Police inspectors and detectives	45	56	107.4	
3460	Social work associate professionals	49	43	88.2	
3470	Artistic, entertainment and sports associate professionals	45	52	90.0	
3471	Decorators and commercial designers	49	53	91.9	
3472	Radio, television and other announcers	50	64	86.1	

ISCO88	LABEL	SIOPS	ISEI	MPS	Comments
3473	Street, night-club and related musicians, singers and dancers	32	50	86.1	
3474	Clowns, magicians, acrobats and related associate professionals	33	50	86.1	
3475	Athletes, sportspersons and related associate professionals	49	54	86.1	
3479	Artistic, entertainment and sports associate professionals not elsewhere classified	43	45	87.3	ZUMA special code
3480	Religious associate professionals	50	38	88.2	
4000	Clerks	37	45	72.3	
4100	Office clerks	37	45	73.0	
4110	Secretaries and keyboard-operating clerks	45	51	73.1	
4111	Stenographers and typists	42	51	73.1	
4112	Word processor and related operators	42	50	73.1	
4113	Data entry operators	45	50	73.1	
4114	Calculating machine operators	45	51	73.1	
4115	Secretaries	53	53	73.1	
4120	Numerical clerks	44	51	92.4	
4121	Accounting and book-keeping clerks	45	51	93.6	
4122	Statistical and finance clerks	36	51	92.1	
4129	Numerical clerks not elsewhere classified	41	51	92.9	ZUMA special code
4130	Material-recording and transport clerks	32	36	62.8	
4131	Stock clerks	30	32	46.7	
4132	Production clerks	44	43	80.1	
4133	Transport clerks	37	45	76.6	
4139	Material-recording and transport clerks not elsewhere classified	37	40	67.8	
4140	Library, mail and related clerks	37	39	45.4	
4141	Library and filing clerks	36	39	47.9	
4142	Mail carriers and sorting clerks	33	39	45.1	
4143	Coding, proof-reading and related clerks	41	39	47.9	
4144	Scribes and related workers	37	39	47.9	
4190	Other office clerks	37	39	73.1	
4200	Customer services clerks	39	49	65.1	
4210	Cashiers, tellers and related clerks	37	48	67.1	
4211	Cashiers and ticket clerks	34	53	67.4	
4212	Tellers and other counter clerks	42	46	67.1	
4213	Bookmakers and croupiers	34	40	67.4	
4214	Pawnbrokers and money lenders	15	40	67.4	
4215	Debt-collectors and related workers	27	40	67.4	
4220	Client information clerks	38	52	60.2	
4221	Travel agency and related clerks	38	52	60.2	
4222	Receptionists and information clerks	38	52	60.2	
4223	Telephone switchboard operators	38	52	60.2	
4290	Client information clerks not elsewhere classified	34	48	64.4	ZUMA special code
5000	Service workers and shop and market sales workers	32	40	63.0	
5100	Personal and protective services workers	32	38	67.5	
5110	Travel attendants and related workers	32	34	53.4	
5111	Travel attendants and travel stewards	50	34	57.3	
5112	Transport conductors	32	34	53.1	
5113	Travel guides	29	34	57.3	
5120	Housekeeping and restaurant services workers	26	32	51.1	
5121	Housekeepers and related workers	37	30	55.4	
5122	Cooks	31	30	49.8	
5123	Waiters, waitresses and bartenders	21	34	55.4	
5130	Personal care and related workers	27	25	57.3	
5131	Child-care workers	23	25	56.9	
5132	Institution-based personal care workers	42	25	57.3	
5133	Home-based personal care workers	17	25	56.9	
5139	Personal care and related workers not elsewhere classified	29	25	56.9	
5140	Other personal services workers	29	30	77.5	
5141	Hairdressers, barbers, beauticians and related workers	32	29	77.9	
5142	Companions and valets	17	19	75.7	
5143	Undertakers and embalmers	34	54	75.7	
5149	Other personal services workers not elsewhere classified	29	19	75.7	
5150	Astrologers, fortune tellers and related workers	37	43	68.3	
5151	Astrologers and related workers	37	43	68.3	
5152	Fortune-tellers, palmists and related workers	37	43	68.3	
5160	Protective services workers	37	47	74.3	
5161	Fire-fighters	35	42	73.1	
5162	Police officers	40	50	80.0	
5163	Prison guards	39	40	70.1	
5169	Protective services workers not elsewhere classified	30	40	56.1	
5200	Models, salespersons and demonstrators	31	43	53.8	
5210	Fashion and other models	28	43	53.8	
5220	Shop salespersons and demonstrators	32	43	53.8	
5230	Stall and market salespersons	24	37	53.8	

ISCO88	LABEL	SIOPS	ISEI	MPS	Comments
5900	Service workers not elsewhere classified	31	35	61.9	ZUMA special code
6000	Skilled agricultural and fishery workers	37	23	45.3	
6100	Market-oriented skilled agricultural and fishery workers	38	23	45.3	
6110	Market gardeners and crop growers	40	23	42.1	
6111	Field crop and vegetable growers	40	23	41.9	
6112	Tree and shrub crop growers	40	23	36.6	
6113	Gardeners, horticultural and nursery growers	40	23	45.8	
6114	Mixed-crop growers	40	23	41.9	
6120	Market-oriented animal producers and related workers	40	23	38.5	
6121	Dairy and livestock producers	40	23	38.4	
6122	Poultry producers	40	23	39.2	
6123	Apiarists and sericulturists	40	23	39.2	
6124	Mixed-animal producers	40	23	39.2	
6129	Market-oriented animal producers and related workers not elsewhere classified	40	23	39.2	
6130	Market-oriented crop and animal producers	38	23	46.3	
6140	Forestry and related workers	24	22	60.0	
6141	Forestry workers and loggers	24	22	60.0	
6142	Charcoal burners and related workers	16	22	60.0	
6150	Fishery workers, hunters and trappers	28	28	44.2	
6151	Aquatic-life cultivation workers	23	28	44.2	
6152	Inland and coastal waters fishery workers	23	28	44.2	
6153	Deep-sea fishery workers	28	28	44.2	
6154	Hunters and trappers	6	28	44.2	
6210	Subsistence agricultural and fishery workers	38	16	44.0	
7000	Craft and related trades workers	38	34	50.6	
7100	Extraction and building trades workers	34	31	50.4	
7110	Miners, shotfirers, stone cutters and carvers	34	30	45.1	
7111	Miners and quarry workers	34	30	45.9	
7112	Shotfirers and blasters	36	30	42.9	
7113	Stone splitters, cutters and carvers	34	27	42.9	
7120	Building frame and related trades workers	34	30	47.1	
7121	Builders, traditional materials	36	29	47.1	
7122	Bricklayers and stonemasons	34	29	45.3	
7123	Concrete placers, concrete finishers and related workers	34	26	40.3	
7124	Carpenters and joiners	37	29	48.7	
7129	Building frame and related trades workers not elsewhere classified	28	30	53.4	
7130	Building finishers and related trades workers	37	34	52.6	
7131	Roofers	31	19	47.2	
7132	Floor layers and tile setters	31	30	56.8	
7133	Plasterers	31	31	51.5	
7134	Insulation workers	28	34	45.6	
7135	Glaziers	26	26	52.4	
7136	Plumbers and pipe fitters	34	33	51.0	
7137	Building and related electricians	44	37	56.0	
7139	Building finishers and related trades workers not elsewhere classified			52.4	ISCO88-com code
7140	Painters, building structure cleaners and related trades workers	31	29	50.0	
7141	Painters and related workers	31	29	52.5	
7142	Varnishers and related painters	29	32	40.3	
7143	Building structure cleaners	20	29	41.2	
7200	Metal, machinery and related trades workers	40	34	49.7	
7210	Metal moulders, welders, sheet-metal workers, structural metal preparers, and related trades workers	38	31	43.9	
7211	Metal moulders and coremakers	38	29	44.5	
7212	Welders and flamecutters	39	30	38.3	
7213	Sheet metal workers	34	33	47.1	
7214	Structural-metal preparers and erectors	44	30	45.4	
7215	Riggers and cable splicers	32	30	44.5	
7216	Underwater workers	26	30	44.5	
7220	Blacksmiths, tool-makers and related trades workers	37	35	50.1	
7221	Blacksmiths, hammer-smiths and forging-press workers	35	33	49.6	
7222	Tool-makers and related workers	40	40	52.6	
7223	Machine-tool setters and setter operators	38	34	48.5	
7224	Metal wheel-grinders, polishers and tool sharpeners	27	24	49.6	
7230	Machinery mechanics and fitters	43	34	49.9	
7231	Motor vehicle mechanics and fitters	43	34	52.9	
7232	Aircraft engine mechanics and fitters	50	42	58.2	
7233	Agricultural- or industrial-machinery mechanics and fitters	42	33	47.4	
7240	Electrical and electronic equipment mechanics and fitters	38	40	55.5	
7241	Electrical mechanics and fitters	38	40	49.9	
7242	Electronics fitters	48	39	62.3	
7243	Electronics mechanics and servicers	42	41	65.4	
7244	Telegraph and telephone installers and servicers	35	40	60.5	

ISCO88	LABEL	SIOPS	ISEI	MPS	Comments
7245	Electrical line installers, repairers and cable jointers	36	38	55.0	
7249	Electrical and electronic equipment mechanics and fitters not elsewhere classified	40	48	58.6	ZUMA special code
7300	Precision, handicraft, printing and related trades workers	39	34	58.3	
7310	Precision workers in metal and related materials	45	38	62.5	
7311	Precision-instrument makers and repairers	47	38	62.5	
7312	Musical instrument makers and tuners	33	38	63.0	
7313	Jewellery and precious-metal workers	43	38	63.0	
7320	Potters, glass-makers and related trades workers	28	28	37.1	
7321	Abrasive wheel formers, potters and related workers	25	27	36.1	
7322	Glass makers, cutters, grinders and finishers	37	29	38.1	
7323	Glass engravers and etchers	31	29	36.1	
7324	Glass, ceramics and related decorative painters	31	29	36.1	
7330	Handicraft workers in wood, textile, leather and related materials	31	29	58.8	
7331	Handicraft workers in wood and related materials	31	29	58.8	
7332	Handicraft workers in textile, leather and related materials	21	29	58.8	
7340	Printing and related trades workers	42	40	62.8	
7341	Compositors, typesetters and related workers	42	40	63.1	
7342	Stereotypers and electrotypers	41	40	64.2	
7343	Printing engravers and etchers	41	42	60.7	
7344	Photographic and related workers	42	40	64.2	
7345	Bookbinders and related workers	32	37	64.2	
7346	Silk-screen, block and textile printers	52	38	64.2	
7400	Other craft and related trades workers	33	33	51.2	
7410	Food processing and related trades workers	28	30	53.0	
7411	Butchers, fishmongers and related food preparers	24	30	49.9	
7412	Bakers, pastry-cooks and confectionery makers	33	31	55.0	
7413	Dairy-products makers	34	30	53.1	
7414	Fruit, vegetable and related preservers	35	30	53.1	
7415	Food and beverage tasters and graders	34	30	53.1	
7416	Tobacco preparers and tobacco products makers	34	30	53.1	
7420	Wood treaters, cabinet-makers and related trades workers	29	33	52.2	
7421	Wood treaters	29	33	52.1	
7422	Cabinet markers and related workers	40	33	53.1	
7423	Woodworking machine setters and setter-operators	36	33	42.1	
7424	Basketry weavers, brush makers and related workers	21	33	52.1	
7430	Textile, garment and related trades worker	34	36	37.3	
7431	Fibre preparers	29	29	41.5	
7432	Weavers, knitters and related workers	32	29	41.5	
7433	Tailors, dressmakers and hatters	40	45	41.5	
7434	Furriers and related workers	35	36	41.5	
7435	Textile, leather and related pattern-makers and cutters	40	36	41.5	
7436	Sewers, embroiderers and related workers	26	33	41.5	
7437	Upholsterers and related workers	31	28	35.6	
7440	Pelt, leather and shoemaking trades workers	27	31	51.1	
7441	Pelt dressers, tanners and fellmongers	22	31	51.1	
7442	Shoemakers and related workers	27	31	51.1	
7490	Other craft and related trades workers not elsewhere classified	31	32	47.9	ZUMA special code
7510	Fitters, metalworkers	38	33	49.1	ZUMA special code
7900	Master craftsmen, foremen and supervisors in production	52	45	65.4	ZUMA special code
8000	Plant and machine operators and assemblers	34	31	40.6	
8100	Stationary-plant and related operators	36	30	43.1	
8110	Mining and mineral-processing-plant operators	31	35	45.6	
8111	Mining-plant operators	34	35	45.6	
8112	Mineral-ore- and stone-processing-plant operators	32	35	45.6	
8113	Well drillers and borers and related workers	31	35	45.6	
8120	Metal-processing-plant operators	40	30	37.5	
8121	Ore and metal furnace operators	45	31	43.7	
8122	Metal melters, casters and rolling-mill operators	36	30	33.9	
8123	Metal-heat-treating-plant operators	38	28	43.7	
8124	Metal drawers and extruders	28	30	43.7	
8130	Glass, ceramics and related plant operators	31	22	45.6	
8131	Glass and ceramics kiln and related machine operators	31	22	45.6	
8139	Glass, ceramics and related plant operators not elsewhere classified	31	22	45.6	
8140	Wood-processing- and papermaking-plant operators	28	27	31.6	
8141	Wood-processing-plant operators	29	27	31.6	
8142	Paper-pulp plant operators	28	27	31.6	
8143	Papermaking-plant operators	28	27	31.6	
8150	Chemical-processing-plant operators	42	35	45.9	
8151	Crushing-, grinding- and chemical-mixing-machinery operators	43	35	44.8	
8152	Chemical-heat-treating-plant operators	43	35	44.8	
8153	Chemical-filtering- and separating-equipment operators	43	35	44.8	
8154	Chemical-still and reactor operators (except petroleum and natural gas)	43	35	44.8	

ISCO88	LABEL	SIOPS	ISEI	MPS	Comments
8155	Petroleum- and natural-gas-refining-plant operators	37	35	44.8	
8159	Chemical-processing-plant operators not elsewhere classified	43	35	46.0	
8160	Power-production and related plant operators	38	32	48.6	
8161	Power-production plant operators	42	33	51.6	
8162	Steam-engine and boiler operators	35	27	45.8	
8163	Incinerator, water-treatment and related plant operators	34	33	49.0	
8170	Automated-assembly-line and industrial-robot operators	30	26	45.6	
8171	Automated-assembly-line operators	30	26	45.6	
8172	Industrial-robot operators	30	26	45.6	
8200	Machine operators and assemblers	34	32	40.7	
8210	Metal- and mineral-products machine operators	37	36	42.8	
8211	Machine-tool operators	38	36	42.7	
8212	Cement and other mineral products machine operators	30	30	43.3	
8220	Chemical-products machine operators	43	30	40.5	
8221	Pharmaceutical- and toiletry-products machine operators	43	30	46.8	
8222	Ammunition- and explosive-products machine operators	43	30	46.8	
8223	Meta finishing-, plating-, and coating-machine operators	28	30	37.6	
8224	Photographic-products machine operators	43	30	46.8	
8229	Chemical-products machine operators not elsewhere classified	43	30	46.8	
8230	Rubber- and plastic-products machine operators	30	30	40.5	
8231	Rubber-products machine operators	30	30	41.4	
8232	Plastic-products machine operators	30	30	39.9	
8240	Wood-products machine operators	31	29	29.3	
8250	Printing-, binding- and paper-products machine operators	41	38	44.0	
8251	Printing-machine operators	41	38	55.0	
8252	Bookbinding-machine operators	32	38	44.2	
8253	Paper-products machine operators	28	38	36.1	
8260	Textile-, fur- and leather-products machine operators	28	30	48.6	
8261	Fibre-preparing-, spinning- and winding-machine operators	29	29	44.2	
8262	Weaving- and knitting-machine operators	29	29	44.2	
8263	Sewing-machine operators	25	32	44.2	
8264	Bleaching-, dyeing- and cleaning-machine operators	25	24	50.9	
8265	Fur- and leather-preparing-machine operators	26	32	44.2	
8266	Shoemaking- and related machine operators	28	32	44.2	
8269	Textile-, fur- and leather-products machine operators not elsewhere classified	26	32	44.2	
8270	Food and related products machine operators	33	29	48.3	
8271	Meat- and fish-processing-machine operators	31	29	48.3	
8272	Dairy-products machine operators	34	29	48.3	
8273	Grain- and spice-milling-machine operators	33	29	48.3	
8274	Baked-goods, cereal and chocolate-products machine operators	33	29	48.3	
8275	Fruit-, vegetable- and nut-processing-machine operators	35	29	48.3	
8276	Sugar production machine operators	45	29	48.3	
8277	Tea-, coffee- and cocoa-processing-machine operators	34	29	48.3	
8278	Brewers, wine and other beverage machine operators	34	29	48.3	
8279	Tobacco production machine operators	39	29	48.3	
8280	Assemblers	33	31	31.9	
8281	Mechanical-machinery assemblers	30	30	31.9	
8282	Electrical equipment assemblers	48	34	31.9	
8283	Electronic equipment assemblers	48	34	31.9	
8284	Metal-, rubber- and plastic-products assemblers	30	30	31.9	
8285	Wood and related products assemblers	31	30	31.9	
8286	Paperboard, textile and related products assemblers	28	30	31.9	
8287	Assemblers of products of varied materials			31.9	ISCO88-com code
8290	Other machine operators and assemblers	33	26	31.8	
8300	Drivers and mobile-plant operators	33	32	39.9	
8310	Locomotive-engine drivers and related workers	36	36	56.2	
8311	Locomotive-engine drivers	43	41	55.7	
8312	Railway brakers, signallers and shunters	29	32	57.0	
8320	Motor-vehicle drivers	32	34	39.6	
8321	Motor-cycle drivers	31	30	39.5	
8322	Car, taxi and van drivers	31	30	38.3	
8323	Bus and tram drivers	32	30	40.5	
8324	Heavy-truck and lorry drivers	33	34	40.7	
8330	Agricultural and other mobile-plant operators	32	26	34.8	
8331	Motorised farm and forestry plant operators	31	26	34.7	
8332	Earth-moving and related plant operators	32	26	36.8	
8333	Crane, hoist and related plant operators	33	28	41.5	
8334	Lifting-truck operators	28	28	26.7	
8340	Ships' deck crews and related workers	29	32	40.0	
9000	Elementary occupations	21	20	34.7	
9100	Sales and services elementary occupations	23	25	40.3	
9110	Street vendors and related workers	25	29	38.3	

ISCO88	LABEL	SIOPS	ISEI	MPS	Comments
9111	Street food vendors	24	29	38.3	
9112	Street vendors, non-food products	24	28	38.3	
9113	Door-to-door and telephone salespersons	26	29	38.3	
9120	Shoe cleaning and other street services elementary occupations	12	28	40.1	
9130	Domestic and related helpers, cleaners and launderers	21	16	30.1	
9131	Domestic helpers and cleaners	22	16	31.2	
9132	Helpers and cleaners in offices, hotels and other establishments	21	16	30.0	
9133	Hand-launderers and pressers	22	16	31.2	
9140	Building caretakers, window and related cleaners	23	23	44.5	
9141	Building caretakers	25	23	44.7	
9142	Vehicle, window and related cleaners	19	23	40.6	
9150	Messengers, porters, doorkeepers and related workers	20	27	35.6	
9151	Messengers, package and luggage porters and deliverers	22	25	32.4	
9152	Doorkeepers, watchpersons and related workers	20	27	36.8	
9153	Vending-machine money collectors, meter readers and related workers	21	27	35.6	
9160	Garbage collectors and related labourers	13	23	30.0	
9161	Garbage collectors	13	23	30.3	
9162	Sweepers and related labourers	13	23	28.6	
9169	Garbage collectors and related labourers not elsewhere classified	13	23	29.5	ZUMA special code
9200	Agricultural, fishery and related labourers	23	16	23.9	
9210	Agricultural, fishery and related labourers	23	16	23.9	
9211	Farm-hands and labourers	23	16	23.9	
9212	Forestry labourers	18	16	23.9	
9213	Fishery, hunting and trapping labourers	23	16	23.9	
9300	Labourers in mining, construction, manufacturing and transport	18	23	26.6	
9310	Mining and construction labourers	16	21	24.1	
9311	Mining and quarrying labourers	18	21	20.0	
9312	Construction and maintenance labourers: roads, dams and similar constructions	15	21	20.0	
9313	Building construction labourers	15	21	24.7	
9320	Manufacturing labourers	19	20	32.4	
9321	Assembling labourers	18	20	32.4	
9322	Hand packers and other manufacturing labourers	22	24	32.4	
9330	Transport labourers and freight handlers	20	29	26.9	
9331	Hand or pedal vehicle drivers	17	22	26.9	
9332	Drivers of animal-drawn vehicles and machinery	22	22	26.9	
9333	Freight handlers	20	30	26.9	

Appendix II: Brief description of the dataset

Content characteristics

Categories	Comments
Topics	**Socio-demographic characteristics:** Artificial individual ID, gender, 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; household income (incl. equivalised household income); basic information on assets and liabilities; household equipment (deprivation index); housing and residential environment **Employment-related characteristics:** Status of employment/ economic inactivity; mini-job; working hours; occupational status (detailed); occupation (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 01/2005 (from wave 2 onwards); fixed-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 **Benefit-related characteristics:** Unemployment Benefit I: start and end dates of the spell of benefit receipt since 01/2005 (wave 1 only); information on periods of Unemployment Benefit I receipt in the context of registered unemployment since 01/2005 (from wave 2 onwards) Unemployment Benefit II: start and end dates of the spell of benefit receipt since 01/2005; identification of household members receiving benefits; amount of benefits received; benefit cuts (start date, duration, reasons, which household members' benefit cut); Participation in measures: 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 **Subjective indicators:** Satisfaction; fears and problems; employment orientation; gender role orientation; subjective social position (top-bottom scale); subjective assessment of health state

Categories	Comments
Data unit	Individuals and households in receipt of Unemployment Benefit II in 7/2006 (sample I); individuals and households in the resident population of Germany (sample II); individuals and households in receipt of Unemployment Benefit II in 7/2006 but without receipt in 7/2006 (sample III; refreshment sample for sample I) Note: individuals aged 65 and above are interviewed using a shorter version of the questionnaire
Number of cases	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)
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 desired. Wave 1: N (CATI): 12,414 individuals (in 8,445 households) N (CAPI): 6,540 individuals (in 4,339 households) Wave 2: N (CATI): 7,888 individuals (in 5,378 households) N (CAPI): 4,599 individuals (3,051 households)
Interview languages	Wave 1: German: 18,231 individuals (12,346 households) Russian: 426 individuals (276 households) Turkish: 285 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 due to the low case numbers in wave 1
Response rates	Wave 1: Sample I: 35.1 % Sample II: 26.6 % Total: 30.5 % Wave 2: Sample I (re-interviewed households only): 51.1 % Sample II (re-interviewed households only): 64.7 % Sample III: 26.3 % Split-off households (from samples I and II): 13.4 % Total: 45.0 %

Categories	Comments
Response rates within households	Wave 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: 88.3 % Split-off households (from samples I and II): 88.3 % Total: 85.4 %
Fieldwork period	Wave 1: December 2006-June 2007 Wave 2: December 2007-July 2008
Period covered	Wave 1: fieldwork period and retrospective spell data from 01/2005 Wave 2: fieldwork period and retrospective spell data from 01/2005 or the respective reference period of the spell type
Time reference	Repeat interview (household panel)
Regional structure	German federal state, eastern/western Germany (Further regional information is available but is not contained in the scientific use file for data protection reasons. Detailed information available on request.)
Territorial allocation	As of survey date

Methodological characteristics

Categories	Comments
Survey design	Original sample, wave 1: two-stage random sample with two sub-populations 1st stage: selection of 300 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). 2nd stage, sample 1: drawing of benefit communities (Bedarfsgemeinschaften) from the register data of the Federal Employment Agency. The number of gross samples drawn per PSU was dependent 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 sector. 2nd stage, sample 2: for sample 2, first a sample of residential buildings was drawn from a commercial database (Microm mosaic). This was then stratified according to a stratification index contained in the database at a ratio of 4:2:1 for households with a low, medium or high status respectively. Interviewers from the surveying institute visited the selected buildings. In the event that a building accommodated several households, this 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 sector. Refreshment sample for sample I in wave 2: In addition to continuing the samples 1 and 2 which were drawn for wave 1, in the 2nd wave a refreshment sample was drawn from the register data of the Federal Employment Agency. For this, benefit communities which were in receipt of Unemployment Benefit II in July 2007 but not in July 2006 were selected. These benefit communities thus depict the inflows to benefit receipt. The sample was drawn in the postcode sectors selected for wave 1 following the procedure used in the 1st wave.
Institutions involved	Institute for Employment Research (IAB); TNS Infratest Sozialforschung
Frequency of data collection	Annually (panel)
File format and size	STATA, SPSS (several files)

Categories	Comments
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 (from wave 2 onwards) Spell data Employment: et_spells.dta/.sav (from wave 2 onwards) Spell data Gaps: lu_spells.dta/.sav (from wave 2 onwards) 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

Data access

Categories	Comments
Data access	Scientific use file (SUF)
Degree of anonymisation	Factually anonymous
Sensitive variables	None

Imprint

FDZ Datenreport
No. 6/2009

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

Editorial staff
Stefan Bender, Dana Singula

Technical production
Dana Singula

Download
http://doku.iab.de/fdz/reporte/2009/DR_06-09-I-EN.pdf

Internet
http://fdz.iab.de/

Corresponding authors
Daniel Gebhardt, Institute for Employment Research,
Regensburger Str. 104, D-90478 Nuremberg;
Phone: +49-(0)911/179-3479;
E-Mail: daniel.gebhardt@iab.de

Mark Trappmann, Institute for Employment Research,
Regensburger Str. 104, D-90478 Nuremberg;
Phone: +49-(0)911/179-3096
E-Mail: mark.trappmann@iab.de