# Sample of Integrated Labour Market Biographies (SIAB) 1975-2014

Manfred Antoni, Andreas Ganzer, Philipp vom Berge
(Institute for Employment Research - Institut für Arbeitsmarkt- und Berufsforschung)

Documentation: Version 1 (SIAB 7514_EN_v1_dok1)
Data: Version 1 (SIAB 7514 v1)

Die FDZ-Datenreporte beschreiben die Daten des FDZ im Detail. Diese Reihe hat somit eine doppelte Funktion: zum einen stellen Nutzerinnen und Nutzer fest, ob die angebotenen Daten für das Forschungsvorhaben geeignet sind, zum anderen dienen sie zur Vorbereitung der Auswertungen.

FDZ-Datenreporte (FDZ data reports) describe FDZ data in detail. As a result, this series of reports has a dual function: on the one hand, those using the reports can ascertain whether the data offered is suitable for their research task; on the other, the data can be used to prepare evaluations.

## Contents

## Zusammenfassung

Dieser Datenreport beschreibt die Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) 1975 - 2014.

## Abstract

This data report describes the Sample of Integrated Labour Market Biographies (SIAB) 1975 - 2014.

**Keywords:** German administrative micro data, labour market data, data manual

We would like to thank our colleagues in the Research Data Centre (FDZ) who were involved in the working group on the Sample of Integrated Labour Market Biographies as well as the ITM division of the Institute for Employment Research for their cooperation and support. This Datenreport is a revision of the FDZ-Datenreport 01/2013 (authors: vom Berge, König, Seth). Additionally, individual passages were adopted from IAB-internal data documentations by the ITM division of the IAB.

## Data availability

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

# 1 Introduction and outline

## 1.1 Introduction

The Sample of Integrated Labour Market Biographies (Stichprobe der integrierten Arbeitsmarktbiographien - SIAB) is a 2 % sample of the population of the Integrated Employment Biographies (IEB) of the Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung - IAB). The IEB comprises all individuals who showed one of the following statuses at least once during the observation period:

- employment subject to social security (recorded from 1975 onwards)
- marginal part-time employment (recorded from 1999 onwards)
- receipt of benefits in accordance with Social Code Book III (recorded from 1975 onwards) or Social Code Book II (recorded from 2005 onwards)
- registered with the Federal Employment Agency (Bundesagentur für Arbeit - BA) as a jobseeker (recorded from 2000 onwards)
- participation in an employment or training measure (recorded from 2000 onwards)

These data, which come from different sources, are merged in the IEB and the statuses are depicted exact to the day.

Whilst the Employee History (Beschäftigtenhistorik - BeH) is the origin of the information on employment subject to social security and marginal part-time employment, the receipt of benefits in accordance with Social Code Book III (SGB III) and Social Code Book II (SGB II) is recorded in the Benefit Recipient History (Leistungsempfängerhistorik - LeH) and the Unemployment Benefit II Recipient Histories (Leistungshistoriken Grundsicherung - LHG and XLHG). The Jobseeker Histories (Arbeitsuchenden-Historiken - ASU and XASU) are the data source for the periods of job search recorded by the BA, whilst participation in employment and training measures is recorded in the Participants-in-Measures History File (Maßnahmeteilnahmehistorik - MTH). Unlike in previous versions of the SIAB, information about participation in employment and training measures from the MTH is also contained in the current version of the SIAB.

The SIAB is produced at the Research Data Centre (FDZ) of the Federal Employment Agency (BA) at the IAB. The SIAB covers the employment histories of 1,757,925 individuals, and their employment biographies are documented in a total of 58,220,255 lines of data.

This Datenreport describes the variables of the weakly anonymous version of the SIAB, which largely comprises the original data, i.e. data which have not undergone anonymisation processes. In order to protect the anonymity of the data subjects, some variables are classified as particularly sensitive and are only disclosed on submission of a special application (see Section 1.2).

The Datenreport on the SIAB is structured as follows. Besides the introduction, Section 1 contains information on data access as well as an outline of the data, the volume structure and a list of variables. A description of the individual data sources can be found in Section 2. Data preparation and data quality are discussed in Sections 3 and 4, whilst the individual variables are described in Section 5.

## 1.2 Data use

### 1.2.1 Data access and data management

The SIAB data in the weakly anonymous version may only be analysed in the context of a research visit at the FDZ and subsequent remote data access.

In order to be able to use the data, it is first necessary to submit an application to the Research Data Centre (FDZ). The Federal Ministry of Labour and Social Affairs (Bundesministerium für Arbeit und Soziales – BMAS) makes the decision regarding the approval of the research project. When approval has been granted, a data use agreement is concluded with the researcher's institution. Details on applying for the dataset and possibilities for data processing can be found on the FDZ homepage.

The SIAB data, which include both German and English labels[1], have a modular structure and are stored in several files. One module, which is henceforth called the Individual File, contains identifiers (individual IDs and establishment IDs), the personal variables, the information on employment, benefit receipt and job search activity, the variables regarding place of residence, and technical variables. A second module, the Basis Establishment File, contains the establishment number, the year, and variables regarding the place of work and economic activities in aggregate form as well as other establishment information as of the reference date of 30 June. The datasets for eastern Germany are available from 1992 onwards.

Establishment variables and individual variables are therefore organised separately, which makes the structure of the data clear and saves storage space in the data management system (see Figure 1). The Basis Establishment File and the Individual File are linked via the establishment number and the year of the dataset. The variables which are marked with a "*" in the List of variables (see p. 14 et seq.) are contained in the Basis Establishment File. Please note that the variable which is necessary for linking the two modules, "establishment number", is available in both the Individual File and the Basis Establishment File.

[1] With the Stata command label language en / label language de labels can be switched to English or German.

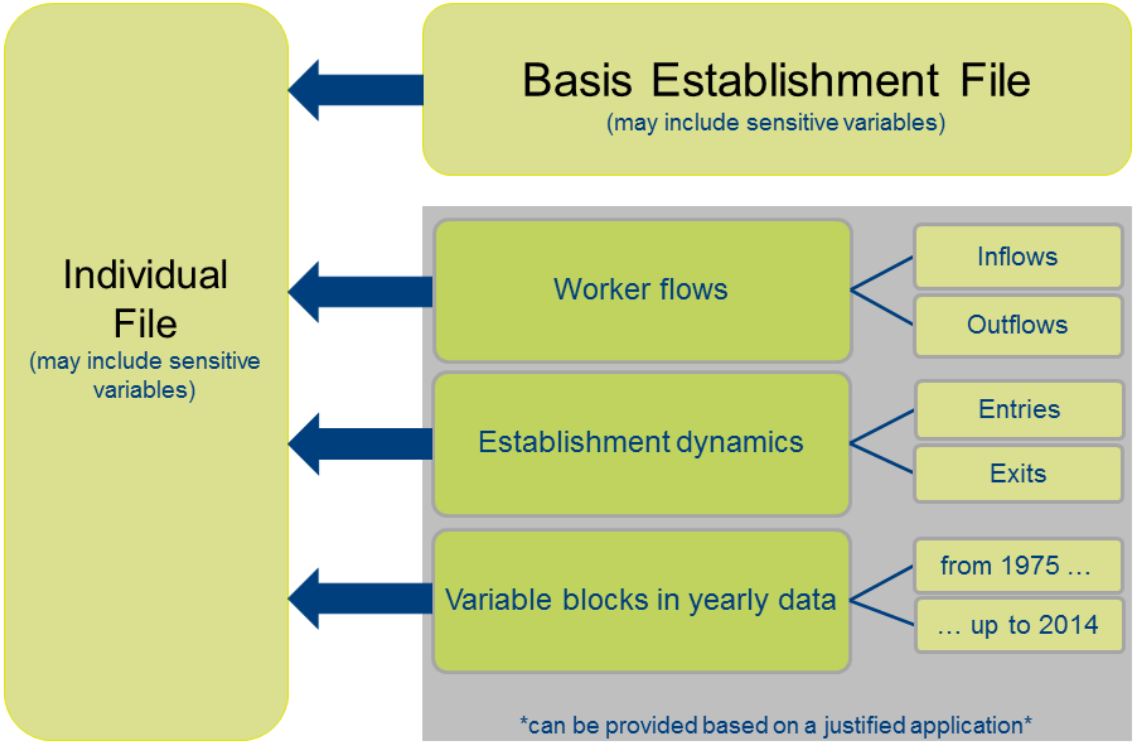


Figure 1: Data storage of SIAB data

Certain variables which make it possible to identify individuals or establishments are only disclosed in their original form if this is necessary for the analysis objective and is justified explicitly in the application for data access. The variables which are particularly sensitive from the viewpoint of data protection legislation are:

Individual File:

- nationality (nation)
- occupational sub-group (beruf2010_4)
- place of residence: employment agency (wo_aa)
- place of residence: district (Kreis) (wo_kreis)

Basis Establishment File:

- place of work: district (Kreis) (ao_kreis)
- economic activity 93 – sub-class of economic activity (five-digit code) (w93_5)
- economic activity 03 - sub-class of economic activity (five-digit code) (w03_5)
- economic activity 08 - sub-class of economic activity (five-digit code) (w08_5)
- date of first appearance (grd_dat)
- date of last appearance (lzt_dat)

Following a justified application further establishment variables can be provided from the Establishment History Panel (Betriebs-Historik-Panel - BHP) (see Figure 1). Further information on the BHP can be found under http://fdz.iab.de/en.aspx.

### 1.2.2 File names in the SIAB 7514

Individual File

siab_7514_v1.dta

Basis Establishment File

siab_7514_v1_bhp_basis_v1.dta

Extension Files BHP

Core dataset with variable blocks
siab_7514_v1_bhp_v1_jjjj.dta, jjjj = 1975 - 2014

Worker flows
siab_7514_v1_bhp_inflow_v1.dta

siab_7514_v1_bhp_outflow_v1.dta

Establishment dynamics
siab_7514_v1_bhp_entry_v1.dta

siab_7514_v1_bhp_exit_v1.dta

## 1.3 Changes as compared to SIAB 7510

### 1.3.1 Introduction of the occupation code 2010

Information on vocational training, the occupation/activity performed and the occupational status is transmitted by means of notifications made by the employer in accordance with the Data Collection and Transmission Regulation (DEÜV) (see Section 2.1) using a so-called occupation code. The new occupation code 2010 was adopted for periods with an end date later than 30.11.2011 (for further details see Bertat et al., 2013). The decision to switch to the new occupation code was made by the central organisations of the social security agencies as a number of facts could no longer be recorded in a way that was up-to-date and realistic using the occupation code 2003.

As the notifications made by employers in accordance with DEÜV go into the Employee History (BeH), the change of the occupation code impacts almost exclusively on observations from this source. The measurement of the following characteristics already reported using the occupation code 2003 is affected by this change: working hours, occupation, occupational status and school and vocational qualification level. In addition, since the switch to the new occupation code, details are also available about whether an employment relationship is fixed-term and whether a person is employed by a temporary work agency to be hired out to other firms. The most important consequence is the switchover to a new occupational classi-

fication. Instead of the previous Classification of Occupations 1988 (Klassifikation der Berufe 1988 (KldB 1988)), the more highly differentiated KldB 2010 is reported with the new occupation code.[2] Another important change is the discontinuation of the variable 'occupational status and working hours' (stib), as most of the values given in this variable can no longer be traced using the new occupation code. The categories of the previous stib variable that are available consistently across the entire observation period (employment status as an apprentice, distinction between part-time and full-time) can be traced in other variables in future. Section 1.3.3 shows which new variables are added as a result of the switch to the new occupation code, which variables are discontinued and which are contained in the data in a different form in future.

Section 4.2 shows what consequences the switch to the new occupation code has on the quality of the data in the BeH observations. One of the most serious consequences of the switchover is a temporary increase in the number of missing values in the variables reported by the employers via the occupation code. This was also the case with the new 'working time' variable, which is why an imputed version of the variable was used for the SIAB. The problem of increased numbers of missing values mainly concerns the years 2011 and 2012.

### 1.3.2 Incorporation of the Participants-In-Measures History Files (MTH)

Unlike the previous version, the SIAB 7514 also contains information from the Participants-In-Measures History Files (Maßnahmeteilnahmehistoriken (MTH)), which are also contained in the IEB, though the categories of measures are only provided in aggregated form.

### 1.3.3 Set of variables

For reasons of data economy and due to lack of demand, some variables that were contained in the SIAB 7510 have since been discontinued. In addition, some variables had to be removed as they would no longer have been available for the entire observation period as a result of the switch to the new occupation code. The adoption of the new occupation code made it possible to include a number of variables that were not contained previously. In some cases this was possible for the entire period, by means of recoding, but in other cases it was only possible for observations from the time of the switchover at the end of 2011 onwards. Table 1 provides an overview of the changes in the set of variables across all sources.

[2] See Paulus/Matthes, 2013, for details regarding the Classification of Occupations 2010.

| Variable | Type of change | Explanation |
|---|---|---|
| Start date of unemployment (alo_beg) | N | |
| Duration of unemployment (alo_dau) | N | |
| Place of work: employment agency (ao_aa) | D | Discontinued without replacement; relevant information contained in wo_aa |
| Place of work: regional directorate (ao_rd) | D | Discontinued without replacement; relevant information contained in wo_rd |
| Contract period (befrist) | N | Introduced with new occupation code; available from 2011 onwards |
| Occupational group – current/most recent (KldB 2010), 3-digit (beruf2010_3)/ 4-digit (beruf2010_4) | N | Introduced with new occupation code; available for entire period by recoding old variable. |
| School education and vocational training (bild) | D | Replaced by vocational education and training |
| Employment status after job search (estatnach) | N | |
| Health problems (gesund_ein) | D | Discontinued without replacement |
| BA client group (kunden_gr) | D | Replaced by client profile |
| Type of benefit (lart) | D | Discontinued without replacement |
| Temporary work (leih) | N | Introduced with new occupation code; available from 2011 onwards |
| Level of requirement – current/most recent (KldB 2010) (niveau) | N | Introduced with new occupation code; available for entire period by recoding old variable. |
| Client profile (profil) | N | Replaces old variable kunden_gr |
| Skills level (quali) | D | Now contained in vocational education and training |
| School leaving qualification (schbild) | D | Now contained in variable ‘schule’ |
| Severe disability status (schweb) | D | Discontinued without replacement |
| Occupational status and working hours (stib) | D | Discontinued because no longer recorded in notification procedure; distinction between full-time and part-time can be traced in variable ‘part-time’; employment status “apprentice” recorded in variable ‘erwstat’. |

Table 1: Changes in the set of variables; N = new, D = dropped

With regard to the discontinuation of the variables ‘place of work: employment agency / regional directorate’ (ao_aa / ao_rd) it should be noted that the local employment agency responsible is defined by the place of residence and not by the place of work. The distinction is therefore not useful anyway, so discontinuing the two variables does not lead to any loss of information.

## 1.4 Outline

| | |
|---|---|
| Topics/ groups of variables | **Employee History (Beschäftigten-Historik - BeH):**<br>Annual notifications and end-of-employment notifications submitted to the social security agencies for employees covered by social security and employees in marginal part-time employment.<br>**Benefit Recipient History (Leistungsempfänger-Historik - LeH):**<br>Information on benefit receipt in accordance with Social Code Book III (SGB III) for recipients of unemployment benefit, unemployment assistance and maintenance allowance.<br>**Unemployment Benefit II Recipient History drawn from A2LL (Leistungs-Historik-Grundsicherung aus A2LL - LHG):**<br>Data on individuals in receipt of basic social security benefits in accordance with Social Code Book II (SGB II) (Types of institution: cooperation of employment agencies and municipalities/joint facilities, separated responsibilities/municipalities exercising their duties separately, authorised municipalities)<br>**Jobseeker History (Arbeitsuchenden-Historik - ASU):**<br>Information on job search activity<br>**Jobseeker History drawn from XSozial-BA-SGB II (XASU):**<br>Information on job search activity reported via the transmission standard XSozial-BA-SGBII to the BA by authorised municipalities.<br>**Participants-in-Measures History File (MTH):**<br>Information on participation in employment and training measures (not including measures of the authorised municipalities) |
| Data unit | Employees covered by social security (including marginal part-time employees from 1999 onwards), benefit recipients, jobseekers, participants in measures |
| Number of cases | 1,757,925 individuals<br>42,892,240 original observations<br>58,220,255 non-overlapping observations (after episode splitting) |
| Period covered | The period covered depends on the data source.<br>BeH: 1 January 1975 - 31 December 2014 (2012: 30-months file, 2013: 18-months file; 2014: 6-months file)<br>LeH: 1 January 1975 - 31 December 2014<br>ASU: 1 January 1997 - 31 December 2014<br>LHG: 1 January.2005 - 31 December 2014<br>XASU: 1 January 2005 - 31 December 2014<br>MTH: 1 January 2000 - 31 December 2014 |
| Time reference | exact to the day |
| Regional structure | German federal states (Bundesländer), districts (Kreise) |
| Date of territorial allocation | Territorial allocation updated as of 31.12.2014 |
| Survey design | 2 % random sample |

| | |
|---|---|
| Institutions involved | Social security agencies, Federal Employment Agency (Bundesagentur für Arbeit) |
| Update frequency | Continuous |
| File format and size | Stata, individual file approx. 3 GB, establishment file: 550 MB |
| File organisation | The data are stored in two files. One contains individual-level information and the other establishment-related information. Further files with additional information on establishments can be provided following a justified application. |
| Data access | On-site usage at the FDZ of the BA at the IAB and subsequent remote data access |
| Degree of anonymisation | Weakly anonymous |
| Sensitive variables: | Nationality (nation), occupational sub-group (beruf2010_4), place of residence: employment agency (wo_aa), place of residence: district (Kreis) (wo_kreis) place of work: district (Kreis) (ao_kreis), economic activity - sub-class of economic activity (five-digit code) (w93_5), economic activity - sub-class of economic activity (five-digit code) (w03_5), economic activity - sub-class of economic activity (five-digit code) (w08_5), first appearance of establishment (grd_dat), last appearance of establishment (lzt_dat) |
| Citation of data and data documentation | Data:<br>“The data basis of this paper is the weakly anonymous Sample of Integrated Labour Market Biographies (SIAB) 1975 - 2014. The data were accessed on-site at the Research Data Centre (FDZ) of the Federal Employment Agency (BA) at the Institute for Employment Research (IAB) and/or via remote data access at the FDZ.”<br>Data documentation:<br>Antoni, Manfred; Ganzer, Andreas; vom Berge, Philipp (2016): Sample of Integrated Labour Market Biographies (SIAB) 1975-2014. FDZ-Datenreport, 04/2016 (en), Nuremberg. |
| Dataset version | Sample of Integrated Labour Market Biographies (SIAB) 1975 - 2014 (7514) |

Table 2: Outline

## 1.5 List of variables

The overview of variables (Table 4) lists the variable names and the longer descriptions of the variables. It also shows which variables are available for which data sources:

| | |
|---|---|
| | Variable is available for the data source.<br>Degree of completeness always > 0.85. |
| | Variable is available for the data source.<br>Lower or varying degree of completeness, see description of variable and frequency count. |
| | Variable is not available for this data source.<br>Degree of completeness always < 0.05. |

Table 3: Degrees of completeness of the Variables

Example: the variable ‘daily wage, daily benefit rate’ is only available for BeH and LeH observations; the observations of the other data sources contain the missing value “.n” for this variable. Another characteristic is that some variables have different contents depending on the data source. For instance, for BeH observations the ‘employment status’ variable contains the person group of the employment notification procedure, for LeH observations it contains the type of benefit, for LHG and XLHG observations it contains the SGB II status, for ASU and XASU observations the job search status and for MTH observations it is the measure category. These differences are not immediately obvious from the variable name for every variable.

| List of variables | Module | Page | BeH | LeH | LHG | ASU | XASU | MTH |
|---|---|---|---|---|---|---|---|---|
| **Identifiers** | | **35** | | | | | | |
| Individual ID (persnr) | | 35 | | | | | | |
| Establishment ID (betnr) | * | 35 | | | | | | |
| **Generated technical variables** | | **37** | | | | | | |
| Observation counter per person (spell) | | 37 | | | | | | |
| Source of spell (quelle) | | 37 | | | | | | |
| Year (jahr) | * | 37 | | | | | | |
| **Period of validity** | | **37** | | | | | | |
| Original start date of observation (begorig) | | 37 | | | | | | |
| Original end date of observation (endorig) | | 38 | | | | | | |
| Start date of split episode (begepi) | | 39 | | | | | | |
| End date of split episode (endepi) | | 39 | | | | | | |
| **Personal information** | | **39** | | | | | | |
| Gender (frau) | | 39 | | | | | | |
| Year of birth (gebjahr) | | 39 | | | | | | |
| Nationality (nation) | | 40 | | | | | | |
| Nationality, aggregated (nation_gr) | | 40 | | | | | | |
| Marital status (famst) | | 40 | | | | | | |
| Number of children (kind) | | 40 | | | | | | |
| Vocational training (ausbildung) | | 41 | | | | | | |
| School leaving qualification (schule) | | 43 | | | | | | |
| **Information on employment, benefit receipt and job search** | | **44** | | | | | | |
| Daily wage, daily benefit rate (tentgelt) | | 44 | | | | | | |

| List of variables | Module | Page | BeH | LeH | LHG | ASU | XASU | MTH |
|---|---|---|---|---|---|---|---|---|
| Occupation – current/most recent (KldB 1988) (beruf) | | 45 | | | | | | |
| Occupational group – current/most recent (KldB 2010) (beruf2010_3) | | 46 | | | | | | |
| Occupational sub-group – current/most recent (KldB 2010) (beruf2010_4) | | 47 | | | | | | |
| Level of requirement – current/most recent (KldB 2010) (niveau) | | 48 | | | | | | |
| Part-time (teilzeit) | | 48 | | | | | | |
| Employment status (erwstat) | | 49 | | | | | | |
| Transition zone (gleitz) | | 51 | | | | | | |
| Temporary agency work (leih) | | 51 | | | | | | |
| Fixed-term contract (befrist) | | 51 | | | | | | |
| Reason of cancellation/notification/termination (grund) | | 52 | | | | | | |
| Employment status prior to job search (estatvor) | | 53 | | | | | | |
| Employment status after job search (estatnach) | | 53 | | | | | | |
| Client profile (profil) | | 54 | | | | | | |
| Type of termination of last job (art_kuend) | | 55 | | | | | | |
| Desired working hours of the job sought (arbzeit) | | 55 | | | | | | |
| Residual claim/planned duration (restanspruch) | | 55 | | | | | | |
| Type of institution (traeger) | | 56 | | | | | | |
| Start date of unemployment (alo_beg) | | 56 | | | | | | |
| Duration of unemployment (alo_dau) | | 57 | | | | | | |
| **Location data** | | **57** | | | | | | |
| Place of residence: district (Kreis) (wo_kreis) | | 57 | | | | | | |
| Place of residence: federal state (Bundesland) (wo_bula) | | 58 | | | | | | |
| Place of residence: employment agency (wo_aa) | | 58 | | | | | | |
| Place of residence: regional directorate (wo_rd) | | 59 | | | | | | |
| **Establishment variables** | | **59** | | | | | | |
| Economic activity 73 (w73_3) | * | 59 | | | | | | |
| Economic activity 93, 5-digit code (w93_5) | * | 60 | | | | | | |

| List of variables | Module | Page | BeH | LeH | LHG | ASU | XASU | MTH |
|---|---|---|---|---|---|---|---|---|
| Economic activity 93, 3-digit code (w93_3) | * | 60 | | | | | | |
| Economic activity 03, 5-digit code (w03_5) | * | 61 | | | | | | |
| Economic activity 03, 3-digit code (w03_3) | * | 61 | | | | | | |
| Economic activity 08, 5-digit code (w08_5) | * | 62 | | | | | | |
| Economic activity 08, 3-digit code (w08_3) | * | 62 | | | | | | |
| Economic activity 73 generated – completed by extrapolation / imputation (w73_3_gen) | * | 63 | | | | | | |
| Economic activity 73 generated – type of completion (group_w73_3) | * | 63 | | | | | | |
| Economic activity 93 generated – completed by extrapolation / imputation (w93_3_gen) | * | 64 | | | | | | |
| Economic activity 93 generated – type of completion (group_w93_3) | * | 64 | | | | | | |
| Year of first appearance of establishment number (grd_jahr) | * | 64 | | | | | | |
| First appearance of establishment number (grd_dat) | * | 65 | | | | | | |
| Year of last appearance of establishment number (lzt_jahr) | * | 65 | | | | | | |
| Last appearance of establishment number (lzt_dat) | * | 66 | | | | | | |
| Total number of employees (az_ges) | * | 66 | | | | | | |
| Number of full-time employees (regular workers + others) (az_vz) | * | 66 | | | | | | |
| Number of employees in marginal part-time employment (az_gf) | * | 67 | | | | | | |
| Mean imputed wage all full-time employees (te_imp_mw) | * | 67 | | | | | | |
| Place of work: district (ao_kreis) | * | 68 | | | | | | |
| Place of work: federal state (Bundesland) (ao_bula) | * | 68 | | | | | | |

Table 4: List of variables with degree of completeness

### 1.6 Volume structure

| No. of cases | before splitting | after splitting |
|---|---|---|
| BeH | 30,638,066 | 35,564,428 |
| LeH | 3,982,119 | 6,045,670 |
| LHG | 1,423,715 | 3,998,827 |
| ASU | 5,607,822 | 10,277,729 |
| XASU | 251,648 | 530,823 |
| MTH | 988,870 | 1,802,778 |
| Total number of observations | 42,892,240 | 58,220,255 |
| Individuals | 1,757,925 | |

Table 5: Volume structure

## 2 Data sources

The administrative individual data were drawn from the Integrated Employment Biographies (IEB) of the IAB. They unite data from five different data sources, each of which may contain information from different administrative procedures. In addition, some supplementary variables from these data sources which are not part of the IEB are incorporated into the administrative individual data. Figure 2 illustrates the data flows that lead to the SIAB and their relationship to other FDZ data products.

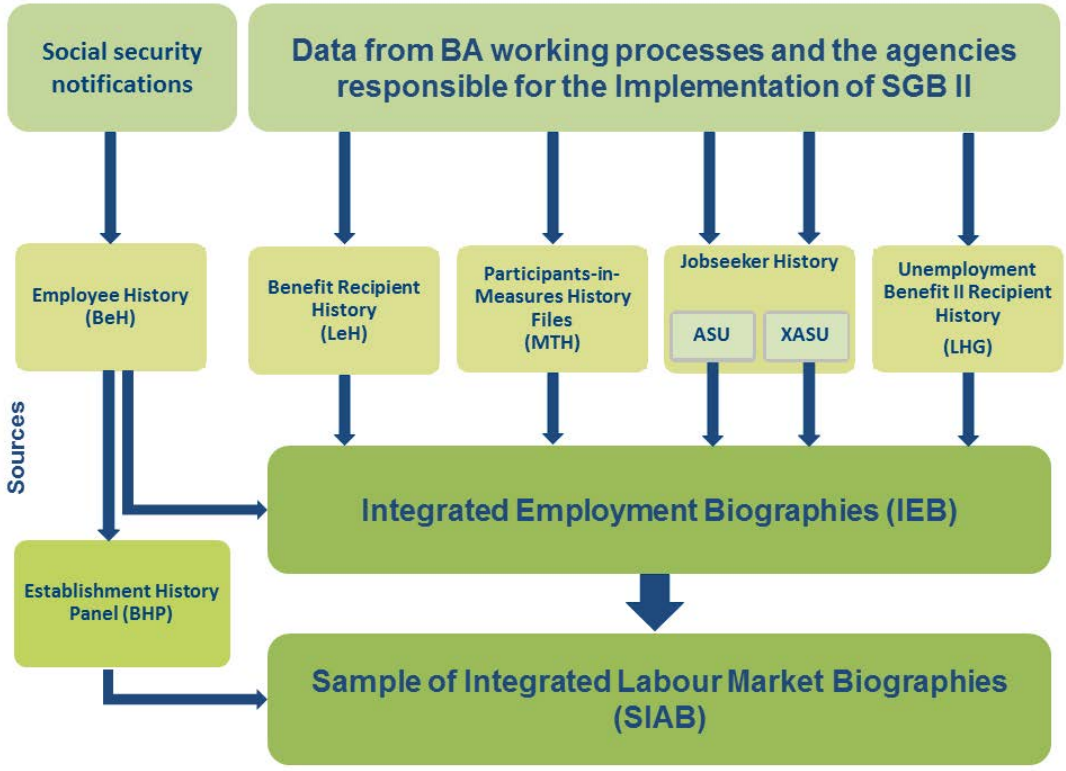


Figure 2: Data sources of the SIAB

## 2.1 Employee History (BeH)

The source of data regarding employment is the Employee History (Beschäftigten-Historik - BeH) of the IAB. The data basis is the integrated notification procedure for health, pension and unemployment insurance, which came into effect as of 1 January 1973 (and was extended to cover eastern Germany as of 1 January 1991) and is known by the abbreviation DEÜV (previously DEVO / DÜVO) (for further details see: Bender et al. 1996, p. 4 et seq.; Wermter /Cramer 1988). Under this procedure employers are required to submit notifications to the responsible social security agencies concerning all of their employees covered by social security at least once a year. The BeH covers all white- and blue-collar workers as well as apprentices as long as they are not exempt from social security contributions. This means that civil servants, self-employed persons and regular students[3] (see Cramer 1985) are in principle not recorded in the BeH. Since the notification procedure was changed on 1 January 1999, employees in marginal part-time employment and unpaid family workers have also been recorded (not contained in the data until 1 April 1999). The data are recorded by the health insurance companies, collected and edited by the Federal Employment Agency (BA) and subsequently integrated into the History File by the IAB.

The administrative individual data are supplemented by establishment data (Basis Establishment File and BHP Extension File). They are taken from the Establishment History Panel (Betriebs-Historik-Panel - BHP), which is also based on the BeH.

When linking individual data with establishment data it has to be taken into account that the variables in both the Basis Establishment File and the BHP Extension Files are aggregated on 30 June of a year. Unlike the data on individuals, the establishment variables are therefore not spell data but are only valid on 30 June precisely each year.[4]

The Basis Establishment File is linked with the Individual File via the programme-specific commands of the software packages used for preparing and analysing the data. In Stata, for instance, the two files can be linked using the "`merge`" command in connection with the relevant paths (see Box 1).

```
use siab_7514_v1.dta
gen jahr = year(begepi)
sort betnr jahr
merge m:1 betnr jahr using SIAB_7514_v1_bhp_basis_v1.dta
```

Box 1: Example code for Stata 14

[3] Students may still appear in the BeH if, for example, they had a marginal part-time job parallel to their degree course.

[4] An extreme example: an employment notification exists from 1 January 2006 to 30 May 2006; the establishment goes bankrupt in June 2006. There is then no information about this establishment in the BHP for 2006.

## 2.2 Benefit Recipient History (LeH)

The Benefit Recipient History (Leistungsempfänger-Historik - LeH) of the IAB covers periods during which individuals receive earnings replacement benefits from the Federal Employment Agency (sphere of Social Code Book III). The benefits comprise unemployment benefit, unemployment assistance and maintenance allowance, in other words not benefits under the sphere of Social Code Book II (e.g. unemployment benefit II). Since entitlement to benefits depends on meeting certain legal requirements, periods of unemployment in which the requirements are not met (e.g. no eligibility for unemployment assistance, or non-completion of the qualifying period for unemployment benefit) are not reported in the Benefit Recipient History. The earliest available data in the LeH are from 1 January 1975.

## 2.3 Unemployment Benefit II Recipient History (LHG)

The Unemployment Benefit II Recipient History (Leistungshistorik Grundsicherung - LHG) contains information about individuals who are eligible for benefit and capable of work, about the members of their benefit community (Bedarfsgemeinschaft) in accordance with § 7 SGB II and about certain individuals associated with the benefit community. In the SIAB it is not possible, however, to link individuals with benefit receipt in accordance with Social Code Book II (SGB II) at the level of benefit communities. The receipt of benefits in accordance with SGB II covers both basic social security benefits (e.g. Unemployment Benefit II) and supplements to unemployment benefit or additional benefits. The LHG does not contain any information about the benefit rates, however. As the amount of benefit received is not determined at the level of the individual but at the level of the benefit community in the case of Unemployment Benefit II, it is difficult to assign an individual benefit rate and this is not done in the SIAB.

Unlike the benefits in the sphere of Social Code Book III, the Federal Employment Agency (BA) is not the sole institution responsible for administering the benefits. The data therefore distinguish between the three possible types of institution responsible for implementing SGB II:

- Cooperation of employment agencies and municipalities (Arbeitsgemeinschaften – ARGE) until the end of 2010 / joint facilities (gemeinsame Einrichtungen) since 2011), in which the BA and the municipality deal with tasks jointly,
- separated responsibilities (getrennte Trägerschaft) / municipalities exercising their duties separately (until 2011) – here the tasks are divided between the BA and the municipality[5],

[5] The municipality pays the costs for housing and heating (Section 22 SGB II) and additional one-off benefit payments to cover extra costs (Section 23 (3) SGB II) and the additional benefits to support integration in accordance with Section 16 (2) Clause 2 No. 1 - 4 SGB II. The BA, on the other hand,

– authorised municipalities, which are also called opting local authorities or opting municipalities according to the initial experimental clause of Section 6a - here the local authority is responsible for all tasks in the sphere of SGB II.

The data of the "Unemployment Benefit II Recipient History drawn from A2LL" (LHG) come from different reporting procedures. As a rule, the IT procedure A2LL was used in all ARGE cooperation projects until 2010, and in joint facilities from 2011 onwards[6]. Authorised municipalities use various IT procedures of their own and transmit their data to the BA by means of the XSozial-BA-SGB II standard. Both of the procedures are used by municipalities with separated responsibilities. The different data standards affect the scope and quality of the data supplied.

The earliest available data in the LHG are from 1 January 2005. However, the data source is incomplete until the beginning of 2007 (see Section 4).

## 2.4 Jobseeker History (ASU / XASU)

Data about jobseekers are stored in the Jobseeker History (Arbeitsuchendenhistorik – ASU / XASU). The ASU data source contains information on jobseekers who are registered with employment agencies, and from 2005 onwards also includes ARGE cooperation projects and separated responsibilities for the implementation of SGB II. The XASU data source, on the other hand, contains the data of jobseekers in receipt of Unemployment Benefit II (ALG-II) from authorised municipalities from 2005 onwards. These data are reported in accordance with the X-Sozial-BA-SGB II standard.

## 2.5 Participants-In-Measures History Files (MTH)

The Participants-In-Measures History Files (Maßnahmeteilnahmehistoriken - MTH) contain information that can be assigned to different legal spheres. First, they contain active labour market policy measures in accordance with Social Code Book III and participation in such measures. Second, the MTH contain measures in the legal sphere of Social Code Book II if these are recorded in BA administrative procedures. This means in particular that no measures implemented by the authorised municipalities are recorded in the MTH as these are reported via a different standard, XSozial. Information from these institutions is not included in the IEB due to a number of data problems. The earliest available data in the MTH are from 1 January 2000.

---

covers the costs for regular benefits, social security contributions and integration benefits (SGB III and SGB II) and specific benefits excluding the additional benefits to support integration cited above.

[6] In 2014 A2LL was gradually replaced by ALLEGRO as the new IT procedure for Unemployment Benefit II in the sphere of SGB II in joint facilities.

# 3 Data preparation and sampling procedure

## 3.1 Corrections and validation procedures

Before the data from the data sources specified in Section 2 are merged to form the IEB they undergo source-specific correction procedures (see the following Sections). The IEB as a whole undergo the following corrections:

- Observations in which the age is under 13 or over 75 are deleted.
- Observations whose end date precedes the start date are deleted.
- Inconsistent information on gender or date of birth within an account is corrected.
- Records with no information on the date of birth or on gender after the correction procedure are deleted.

Unlike in the IABS, no further corrections (such as the addition of presumably missing notifications, strike corrections) are performed.

### 3.1.1 Employee History (BeH)

- To capture a person group that is as constant as possible over time, some person groups for which data are not available throughout the entire observation period are excluded. From the reporting year 2011 onwards the BeH data originate from newly designed source data. As a result, a number of person groups have been introduced or reactivated as they are classified by the BA statistics as being subject to social security contributions. The person groups 101 - 107, 111 - 114, 118, 119, 120, 140, 141, 142, 143, 149, 201 and 203 - 205 are therefore contained from that time onwards as well as the two groups 109 and 209, which indicate people in marginal part-time employment. Groups that are not included are, for example, people in short-term employment, i.e. person groups 110, 202 and 210.
- Person groups 123, 124 and 127 have been newly introduced.
- For data protection reasons, the person groups 107, 111, 113, 114, 127 and 204 are combined to form the person group “other workers” (599).
- From the reporting year 2012 onwards apprentices were included as the new person groups 121 and 122.
  - Observations with earnings amounting to zero or with no details on earnings, and the value 101 for the person group variable, and the value 50 for the reason for notification (annual notification) are not incorporated into the IEB.
  - Gender and date of birth are taken from the Data Warehouse (DWH) of the BA. This information is harmonised across data sources.
  - The territorial allocations for place of work and place of residence are updated to the status as of 31 December 2014.

### 3.1.2 Benefit Recipient History (LeH)

- Observations without a valid start date are excluded.
- Observations whose end date precedes the start date are excluded.
- If the end date for the receipt of unemployment assistance precedes the start date by one day and the spell was not deleted, then the end date is increased by one year.
- Between 2004 and 2006 the notification procedure from which the data originate was changed. Overlaps occurring between the old and the new procedures were corrected.
- Observations with no end date or an invalid end date are excluded, since in these cases it cannot be assumed that a benefit payment was made at all.
- The territorial allocations are corrected in the same way as the BeH in the SIAB.

### 3.1.3 Unemployment Benefit II Recipient History (LHG)

- Observations without a BA client number are deleted.
- Observations without a valid date of birth are deleted.
- Cancelled data records are not used.
- It only contains observations of people who are capable of work and people under the age of 65.
- In each case non-overlapping periods of benefit entitlement of a person in a certain benefit community (BG) are depicted. New observations are begun for the following administrative reasons:
  - on certain birthdays of members of the BG that are stipulated by law and relevant for structural changes in the benefit community (14, 15, 18 and 25) and the individual retirement age of members of the BG (see Section 3.1.6),
  - when the structure of the benefit community changes (e.g. due to entries/exits),
  - when there are changes in a variable of the BG client and
  - at the beginning and the end of a case of benefit sanctions for observations from 1 April 2006 onwards. It must be taken into account, however, that it is not possible to identify the duration or type of sanction or the time when it was imposed or when it began on the basis of the data. The reason for this is that there is no corresponding variable or value that indicates the start, type or duration of the sanction.
- For the reason mentioned above, all individual-related variables that are available for the LHG source are valid for the entire duration of the observation.
- Double notifications due to the territorial reforms in 2009/2011 and the reorganisation of the institutions in 2012 were corrected as far as possible.
- The territorial allocations are corrected in the same way as the BeH in the SIAB.

### 3.1.4 Jobseeker History (ASU / XASU)

- Observations whose end date precedes the start date are not included in the ASU.
- There is no consolidation of the ASU observations for individual persons. Therefore, overlaps between ASU observations might occur.
- Individual-related variables that are only available for the (X)ASU sources always refer to the beginning of the spell.
- A new ASU spell is generated as soon as a change of status occurs (e.g. from seeking work to unemployed). This also applies if the type of institution (employment agency, cooperation of employment agency and municipality, joint facility, authorised municipalities, separated responsibilities) changes. The ASU data basis only distinguishes between observations with the status “unemployed” and “jobseeker”, and since 2006 “seeking advice” and “without status”. In the IEB, however, the additional status “ill / not able to work” is available. The employment status “ill / not able to work” is assigned to IEB spells when in the ASU data basis
  - a preceding observation with the status “unemployed” exists which joins the next observation without a gap and has “incapacitated for work” as the reason for exit, and
  - a subsequent observation with the status “unemployed” exists which also follows without a gap, and
  - the observation itself does not have the status “unemployed” but “jobseeker”.
  - In contrast to the ASU source, the XASU only distinguishes between the status
    - “not unemployed, but seeking work” or
    - “unemployed and simultaneously seeking work”.
  - Unlike in the ASU, periods of illness are not taken into account when generating the “employment status”, since no information about illness is available in the XASU data. When calculating the unemployment duration with XASU observations, gaps due to illness cannot be identified.
  - The XASU contains non-overlapping time periods for individuals. If one of the following variables changes, in each case a new data spell is generated for the XASU:
    - change of job search status
    - change of availability
    - change of SGB II institution (due to notification procedure)
    - change of place of residence
  - The territorial allocations are corrected in the same way as the BeH.

### 3.1.5 Participants-In-Measures History Files (MTH)

- Observations whose end date precedes the start date were excluded.
- Observations generated more than a year after the end of the measure are deleted if another observation exists that was generated within the year after completion of the measure.
- Only the most recent record of an individual case of participation in a measure is used.
- Only cases of participation in measures that are classified as “actually took place” are included in the IEB. Cases of participation that did not take place or have not yet taken place are deleted. Cases of participation are also classed as not having taken place when a deletion date is set during the participation in a measure.
- Certain types of measure are not included. These include services to support careers advice and job placement, mobility assistance and pure rehabilitation measures.

### 3.1.6 SGB II anonymisation

In order to reduce the risk of de-anonymisation, in general only the year of birth is indicated in the administrative data. However, in the LHG and (X)ASU sources there is the risk that the exact date of birth may still be obvious due to the chronological structure of the observations. Observations might end systematically on certain birthdays and/or the day before, or start again on the birthday.

In order to prevent an indirect determination of the exact date of birth, the following procedure is applied. Observations split on the 18$^{th}$, 25$^{th}$ or 65$^{th}$ birthday are merged into a single observation. Apart from the exact date of birth no other information is lost in this case. For observations beginning on the 15$^{th}$ birthday or on the first day of retirement (or ending on the first day of retirement), the start date (end date) of the observation is set to the beginning (the end) of the respective quarter. This correction results in a bias of the duration of the observation. Corrected observations are marked in the "employment status" variable.

## 3.2 Episode splitting

The administrative individual data are available with “split” episodes. If observations overlap within an account, these observations are replaced by artificial observations with new dates so that completely parallel periods and non-overlapping periods are created. This increases the number of observations (see Figure 3).

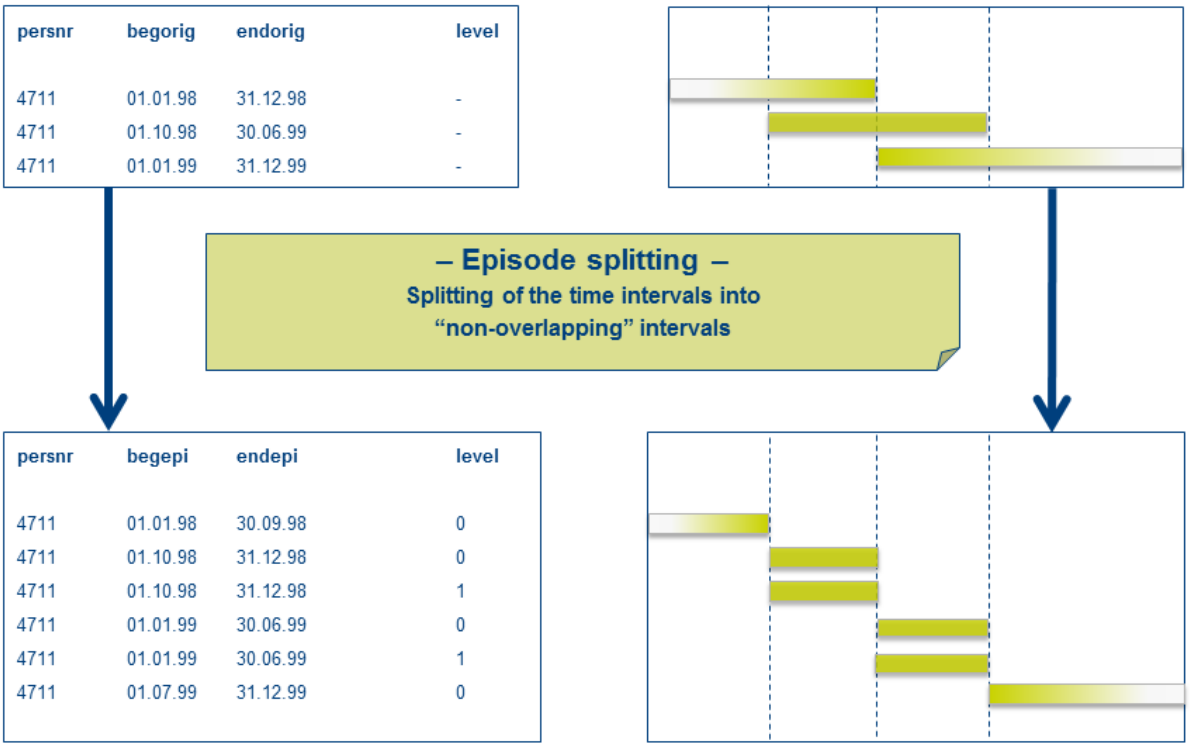


Figure 3: Episode splitting

The original date variables for the beginning and the end of the original observation (begorig and endorig) are retained, the variables 'start date of the split episode' and 'end date of the split episode' (begepi and endepi) mark the beginning and the end of the split episodes. It is possible to establish whether observations have been split by comparing the original period (begorig and endorig) with the episode period (begepi and endepi).
To restore the original data without the split episodes or to delete the episodes that were created artificially by means of episode splitting it is necessary to select all observations for which the start of the original observation is the same as the start of the split episode (`begepi == begorig`).

It is advisable to sort entirely parallel observations generated by the splitting procedure in a consistent manner. The variables ‘observation counter per episode’ (level2) and ‘observation counter per episode and source’ (level1) that were previously contained in the SIAB can be generated using the following Stata commands if required:

```
bysort persnr begepi quelle (spell): gen byte level1 = _n-1
```

```
bysort persnr begepi (spell): gen byte level2 = _n-1
```

### 3.3 Sampling procedure

The SIAB is a 2 % sample of the individuals for whom there is an entry in the IEB. This makes the SIAB representative of the person groups of all of the data sources, not only of people in employment. Accordingly, case numbers of the population can be estimated on the basis of the sample with a weighting factor of 50. This applies in principle to the person groups of all of the data sources; however, the differing degrees of completeness of the data sources have to be taken into account.

### 3.4 Missing values

In the SIAB, missing values are coded as follows:

| Term | Value | Description |
|---|---|---|
| No (valid) details available | .z | Values of a variable which are not systematically missing, i. e. the variable is available in principle for the data source, but no details are available for the value considered or cannot be interpreted reasonably. |
| Systematically not available | .n | A variable is not available in principle for a data source (dark grey cells in the overview of variables in Section 1.5) or is not available for a certain period. |

## 4 Data quality and problems

### 4.1 Entire IEB

The IEB contains employment histories. However, not every type of employment is included in the administrative data. Some individuals with certain life courses are not represented in the IEB at all.

For evaluation purposes, it is often relevant to know gaps in the included biographies (e.g. for creating control groups, analysing life courses, etc.). The gaps listed below are defined as periods of time after the end of school education for which no data are included in the IEB. These gaps can be divided into

- gaps with no information available at all, and
- gaps for which information may be available from the ‘reason for notification / reason for end of benefit receipt / reason for discontinuation of SGB II / reason for deregistration’ variable of the observation immediately preceding the gap (if a corresponding observation exists).

These gaps were identified using the ‘reason for notification / reason for end of benefit receipt / reason for discontinuation of SGB II / reason for deregistration’ and ‘employment status’ variables in the various sources. The list makes no claims to be exhaustive.

| Biographical gap | Information on gap, identifiable using the details in the "grund" variable in the preceding observation of the source, if necessary |
|---|---|
| Civil servants, professional soldiers, judges, employees of bodies or foundations under public law | XASU |
| Self-employed persons without support | LeH, ASU |
| Students, persons in school-based further education | LeH, LHG, ASU, XASU |
| Persons who are ill / not able to work for more than 6 weeks (illness during unemployment, however, is represented in the ASU source under certain circumstances, see Section 3.1.4) | BeH, LeH, ASU |
| Persons receiving old-age pension without employment if not a member of a benefit community | LeH, LHG, ASU |
| Individuals on maternity leave / parental leave | XASU |
| Recipients of early retirement benefits | LeH, ASU |
| Trade professionals working from home | |
| Employees working short-time | ASU |
| Persons in youth welfare facilities, in vocational training centres, approved workshops or similar facilities for disabled persons | ASU |
| Participants in programmes to support participation in working life (people in rehabilitation) | ASU |
| (Sideline) farmers | |
| Caregivers according to Section 19 SBG XI | |
| Conscripts | BeH, LeH, LHG, ASU, XASU |
| Persons in reserve duty training | BeH, LeH, LHG, ASU, XASU |
| Persons fulfilling community service | BeH, LeH, LHG, ASU, XASU |
| Persons fulfilling a voluntary social or ecological year instead of community service | |
| Other people not registered with the statutory pension insurance or the Federal Employment Agency (e.g. sabbatical, funding from personal assets or pensions, emigration, employment abroad, voluntary work etc.) | BeH, LeH, ASU |
| Strikers in cases where the strike lasts more than a month | LeH |
| Social assistance recipients (prior to the introduction of SGB II in 2005), recipients of welfare payments (according to SGB II) | |
| Recipients of compensation according to FELEG (Gesetz zur Förderung der Einstellung der landwirtschaftlichen Erwerbstätigkeit, Act on Support in Case of Termination of Farming Activities) | |

Table 6: Biographical gaps and identification possibilities

## 4.2 Employee History (BeH)

- The introduction of the new occupation code in 2011 led to a number of problems. For example, during the transition period granted to employers in the social security notification procedure,[7] there was a temporary increase in the number of missing details. Analyses of the BA statistics (Bertat et al., 2013, p. 10) show that in 20 to 30 percent of cases no information was contained in the new or converted variables 'occupation - activity performed', 'working hours' and 'vocational education and training' after the switchover. This situation began to improve significantly in the first half of 2013. In order to improve the quality of the 'working time' variable in the transition period, Ludsteck/Thomsen (2016) developed an imputation procedure to replace the missing values by imputed values. The imputed data are already used in the SIAB 7514.
- Due to the introduction of the employment notification procedure in the federal states of eastern Germany, the notifications for eastern Germany can only be assumed to be sufficiently complete from 1993 onwards. For the same reason, a large number of spells for 1991 have missing values for several variables (such as employment status).
- The increase in the number of BeH observations from 1999 onwards is due to the introduction of the obligation to submit employment notifications for people in marginal part-time employment from 1 April 1999 onwards.
- Especially in 1999, observations of part-time employment increase significantly. This is caused by the actually observed increase in part-time work as well as by the fact that since 1999 employment notifications have been completed more correctly.
- Within the employment notification procedure, a certain time lag is unavoidable. Although changes in employment relationships have to be reported immediately, and existing employment relationships have to be confirmed annually by April (until the annual notification 2012) or mid-February (from the annual notification 2013 onwards) of the following year, some notifications actually arrive years later. The History File of the IAB is not updated continuously, however, but at certain intervals. This is done using files of employment notifications for one particular year which were submitted 36, 18, 12 or 6 months after the end of the reporting year (e.g. the 18-month file for 2013 can be created in July 2015 at the earliest). Notifications submitted more than three years late are not taken into account at the IAB, which means that a 36-month file shows a 100 % degree of completeness by definition.
- In the version of the IEB on which the SIAB data are based the degree of completeness of the BeH observations last stood at 100 % in 2012. 18-month files were used for 2013,

[7] The test programs used in the notification procedure permitted missing details in the occupation code 2010 until the end of May 2012.

and the observations for 2014 come from a 6-month file. It can therefore be assumed that the BeH observations for 2013 and in particular for 2014 are slightly underreported in the SIAB. However, this should not reduce the ability to analyse the data at individual level. The missing notifications occur more frequently in a few establishments, however. This means that in individual cases the establishment data, e.g. establishment size, are grossly incorrect and will change considerably in subsequent versions.

- In 1984, a change was made in the employment notification procedure. From that time onwards, one-off payments of gross earned income were reported as part of the annual earnings subject to social security contributions, which leads to an increase in the average daily wage. In particular, the proportion of wages and salaries above the upper earnings limit increased considerably from that year onwards (see Bender et al. 1996).
- For the years 1992 until 2000 noticeable decreases and increases in the number of notifications were observed. Decreases can be observed especially for the following 10 districts: Braunschweig (03101), Wolfsburg (03103), Emden (03402), Kassel (06633), Essen (05113), Neuss (05162), Erftkreis (05362), Hersfeld-Rotenburg (06632), Miltenberg (09676) and Kempten (Allgäu) (09763). This is due to notification problems of one or more establishments in these regions.
- Considerable decreases were also observed for the districts Salzgitter (03102) and Hoyerswerda (14264).
- Concerning the notifications for full-time employment, especially the districts Main-Taunus (06436) and Alzey-Worms (07331) are noteworthy. They feature above-average rises. Also in this case, the reasons are notification problems at one or more establishments in these regions.
- In the years 1996 to 1998, the values 841-844 (doctors and pharmacies) within the 'occupation – activity performed' variable are very rare compared to the neighbouring years.

## 4.3 Benefit Recipient History (LeH)

- For the states of eastern Germany, the LeH observations were not fully recorded until 1992.
- The benefit receipt data used to be saved on magnetic tapes. Owing to a fault in one magnetic tape, the benefit receipt data up to and including 1980 are only partially contained. Thus, in the present data product, too, it can be assumed that information on benefit receipt in that period is not available in full.
- Due to an internal change of systems, there is a break in the recording of periods of exclusion from benefits and of benefit suspension in 2004. Until 1 July 2004 periods of exclusion from benefits and of benefit suspension can only be identified via the 'reason for

end of benefit receipt' in the preceding LeH observation. After this date a separate observation is available with the daily benefit rate = 0 for periods of benefit exclusion and suspension.

## 4.4 Unemployment Benefit II Recipient History (LHG)

- With regard to the completeness of case numbers or benefit histories from the LHG data sources, there are substantial gaps in the years 2005 and 2006. We therefore strongly advise against analysing the data for this time period based merely on the LHG sources.
- Longitudinal analyses of individuals are affected by inaccuracies as it is not possible to distinguish between changes in the benefit entitlement status and relocations into and out of districts whose institutions had problems delivering data.
- Also from 2007 onwards, cases of underrecording occur at times. These generally last one month and occur mainly in the authorised municipalities.
- Underrecording and overrecording occur in connection with changes in the type of institution responsible for implementing SGB II:
  - In the context of the reform of the territories covered by the institutions, which came into force on 1 January 2011, cases of underreporting occurred in the districts covered by the employment agencies of Dessau-Roßlau, Halberstadt, Halle and Sangerhausen.
  - Double notifications due to the territorial reforms in 2009/2011 and the changes in the form of the institutions as of 1 January 2012 are already corrected as far as possible in the IEB. Nonetheless double notifications may still occur.
- In the following job centres there are inaccuracies with regard to the allocation of benefit cases:
  - between Emden and Norden between September and December 2009
  - between Döbeln and Mittelsachsen from October to December 2012
  - between Tirschenreuth and Wunsiedel from November 2012 to March 2013
- Some individuals for whom a (X)LHG spell exists are excluded entirely or partly from benefit receipt according to SGB II, for instance because they take part in a subsidised training programme, receive an old-age pension, live in an in-patient facility or a residential institution or receive insurance payments aimed at avoiding need. This affects on average 3 to 5 percent of all cases. In XSozial this person group is sometimes underrecorded by some institutions. Exclusion from benefits can not be identified in the SIAB.

## 4.5 Jobseeker History (ASU/XASU)

### 4.5.1 ASU

- The registered periods of job search activity in the ASU source are regarded as complete from the year 1997 onwards. Therefore, the analysis potential of the ASU spells before 1997 is limited.
- For the placement staff it is not always possible to record the allocation to the legal sphere immediately, since it is frequently only clear which institution is primarily responsible after a certain time due to a possible entitlement to SGB II benefits. Therefore, we recommend comparing the value of the ‘type of institution’ variable in the ASU with the value in the LHG and/or XLHG for the same period of time. Due to the recording gaps in the LHG and XLHG between 2005 and 2006 this is not always possible.
- For some individuals for whom an authorised municipality has been responsible since 2005, parallel "artificial" ASU datasets were created by the Federal Employment Agency. These can be identified via estatvor (transfer to an authorised municipality).
- From mid-2005 until mid-2006, the coArb IT procedure, from which the jobseeker and applicant pool data originate, was superseded by the VerBIS procedure at the Federal Employment Agency. In July 2005, coArb was first replaced by VerBIS in the employment agency in Wiesbaden as a pilot project. From December 2005 onwards, it was then gradually replaced by VerBIS in several stages in all employment agencies. The information for many of the variables recorded was gathered with different levels of differentiation and different qualitative weighting in the two systems. It is therefore very difficult to integrate these variables into the IEB, which is only possible using a special procedure (mapping). Unfortunately, a full conversion of the affected variables from coArb to VerBIS cannot be achieved by means of mapping, so for some variables there is an unusually large number of the values ‘no details available’, ‘other’ or ‘missing’. Moreover, striking differences may occur in frequency counts, depending on whether the original source of the data was coArb or VerBIS. Important limitations in the analysis potential are mentioned in the corresponding description of variables.
- The coArb procedure, which was used until June 2006, supported only the placement of unemployed persons and jobseekers. Some data were also collected about individuals who were only seeking advice, but these data are incomplete. The careers advice data were collected in a separate system. In VerBIS the attributes of the job-search status were extended to include ‘seeking advice’ and individuals ‘without status’. The latter group includes individuals eligible for Unemployment Benefit II who are only available for job placement to a limited degree. The recording of this group in VerBIS is only regarded as largely complete since January 2008.

- A change of the institution responsible for implementing SGB II or a change of place of residence does not lead to a new ASU observation, the value of the variable at the start of an episode is continued. The longer the observation becomes, the greater the risk is that the institution responsible or the place of residence is no longer correct.

### 4.5.2 XASU

- In contrast to the job search spells from the cooperation of employment agencies and municipalities (ARGE) and the separated responsibilities, systematic cases of under-recording have emerged for the authorised municipalities since 1 January 2005. Thus, data from the XASU source should only be analysed from 2007 onwards.
- A variety of variables sometimes have only a very low degree of completeness for the XASU. Variables which are affected by this include ‘school-leaving qualification’, ‘severe disability status’, ‘reason for notification’ as well as ‘employment status prior to job search’. Although the degree of completeness of these variables improves over time, some of them are still unsatisfactory. The ‘occupation – activity performed’ variable is not available in the XASU for almost the entire period available.
- For a number of institutions (districts), the proportion of registered recipients of unemployment benefit II who are also registered jobseekers is implausibly large at times or continuously in the IEB. One possible reason for this could be an incorrect determination of the status ‘not unemployed but seeking work’ by these institutions.
- The institution-related and period-related plausibility of the XASU data should be examined before use, taking the research question into account.

## 4.6 Participants-In-Measures History Files (MTH)

- The MTH is incomplete for measures with a start date before 01.01.2000.
- As of 01.01.2005 there is an inconsistency in the data as participants in measures were allocated to different institutions with the introduction of Social Code Book II (see Sections 2.2 and 2.4).
- The MTH contains only notifications that are recorded in BA procedures. The use of these procedures in cooperations of employment agencies and municipalities/separated responsibilities/municipalities exercising their duties separately increases continuously between 2005 and 2007. The notifications for these institutions are complete from March 2007 onwards.
- Measures that are reported via the XSozial standard are not contained in the MTH or in the SIAB.

- As a result of the reorganisation of the institutions responsible for implementing SGB-II, the documentation of participation in measures in the MTH may end or begin again when there is a change in the reporting procedure.[8]
- In the case of notifications regarding the bridging allowance (Überbrückungsgeld) the maximum permissible duration of six months is sometimes exceeded. In most cases this can be explained by a default setting in the input mask of the data recording system.
- The MTH is supplemented by applicant characteristics (e.g. vocational training) from other BA procedures. For these variables the administrative procedure was switched from co-Arb to VerBIS in 2006. The same quality limitations as for the ASU apply here.

# 5 Description of variables

Frequency counts and overviews of the individual values and labels of the variables can be found in separate files under http://fdz.iab.de/en.aspx.

## 5.1 Identifiers

### 5.1.1 Individual ID (persnr)

| | |
|---|---|
| Variable label | Individual ID |
| Variable name | persnr |
| Category | identifiers |
| Origin | BeH, LeH, LHG, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | The individual ID indicates which observations belong to the same person. Artificial means that it is not possible to infer any of the person's characteristics or any original identifiers from this individual ID.<br>As there is no uniform individual identifier in the different data sources, the allocation of the information from different data sources (e.g. employment and benefits) to individuals is not always unambiguous. In such cases, implausible employment histories may arise.<br>The formation of the individual identifier which spans all data sources is based on a heuristic developed by the BA. |

### 5.1.2 Establishment ID (betnr)

| | |
|---|---|
| Variable label | Establishment ID |
| Variable name | betnr |
| Category | identifiers |
| Origin | BeH |

[8] Further information concerning the territory structure of the institutions responsible for implementing Social Code Book II and relevant changes can be found at: http://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen/Regionale-Gliederungen/Gebietsstruktur-Traeger-Grundsicherung-Nav.html

| | |
|---|---|
| Data type | numerical |
| Hierarchy | none |
| Detailed description | The establishment ID indicates which observations belong to the same establishment. It is based on the establishment number allocated by the BA, which was replaced by an artificial number. (Further information on the allocation of establishment numbers by the BA can be found in Bender et al. 1996: p. 15 et seq. and pp. 27-30.) The establishment number and year specification can be used to merge individual and establishment information.<br><br>For the establishment number, the following should be observed in general:<br>a) If the company has one office only, or if the company has one office only in one municipality, this office is the establishment and is given an establishment number.<br>b) If the company has **several** branch offices in **one** municipality, these establishment premises / workplaces must be merged into a single establishment under one establishment number, if they belong to the same economic class. If they do **not** belong to the same economic class, each branch office is regarded as a separate establishment and is given its own establishment number.<br>c) If the company has **several** branch offices in **several** municipalities, each of these branch offices is an establishment and is given its own establishment number.<br><br>In this context, the following definitions with regards to the allocation of establishment numbers as part of the notification procedure for social security must be observed:<br>a) An **establishment** is a regionally and economically delimited unit in which employees work and which is allocated an establishment number according to the above-mentioned principles.<br>b) A **workplace** is a unit in which employees work and which is not allocated an establishment number according to the above-mentioned principles.<br>c) A **company** as a term combines establishment premises and workplaces belonging to the same employer.<br>d) An **employer** is any natural person or legal entity that employs at least one employee subject to social security contributions or in marginal part-time employment.<br>e) Establishment and establishment premises are synonyms; branch office is a synonym for subsidiary, district office, outsourced office, workplace etc. if it is not an establishment. |
| Notes on quality | The establishment ID is only missing in a very small number of cases. These observations are notifications for the person group “205” (earnings notifications for casual workers). As establishment variables (place of work, economic activity, establishment size etc.) are merged via the establishment ID, they are missing in these observations. |

## 5.2 Generated technical variables

### 5.2.1 Observation counter per person (spell)

| | |
|---|---|
| Variable label | counter per person |
| Variable name | spell |
| Category | generated technical variables |
| Origin | BeH, LeH, LHG, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | The observation counter per person counts a person's observations, beginning with 1. The variable is generated during the episode splitting procedure and refers to the split observations. Using the "observation counter per person" variable, it is easy to restore the original sorting order. The observations are sorted first by the start date of the split episode and then by the data source. |

### 5.2.2 Source of spell (quelle)

| | |
|---|---|
| Variable label | Source of spell |
| Variable name | quelle |
| Category | generated technical variables |
| Origin | BeH, LeH, LHG, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | The variable indicates the data source. |

### 5.2.3 Year (jahr)

| | |
|---|---|
| Variable label | Year |
| Variable name | jahr |
| Category | generated technical variables |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | This variable is only included in the Establishment File. It indicates the year of validity of the establishment data as of the reference date of 30 June.<br>This variable can be used together with the establishment number to link the Individual File and the Establishment File. |

## 5.3 Period of validity

### 5.3.1 Original start date of observation (begorig)

| | |
|---|---|
| Variable label | Original start date |
| Variable name | begorig |
| Category | period of validity |
| Origin | BeH, LeH, LHG, ASU, XASU, MTH |

| | |
|---|---|
| Data type | date |
| Hierarchy | none |
| Detailed description | The original start date of the observation corresponds to the original start date of the notification. This can differ from the start date of the episodes (begepi) (see also the comments on episode splitting in Section 3.2)<br>Since the notification logic might under certain circumstances permit re-identification of the exact day of birth, the original information on the date was changed in these cases by applying the anonymisation procedure described in Section 3.1.6.<br><br>**1) BeH**<br>Because of the rules of the notification procedure, in BeH observations the starting and ending year are always identical (obligation of the employer to submit annual employment notifications). A continuous employment relationship may therefore be distributed across several notifications.<br><br>**2) LHG, ASU, XASU**<br>Certain changes lead to the creation of a new observation (see Sections 3.1.1 and 3.1.4). begorig indicates the start date of the new period. |

### 5.3.2 Original end date of observation (endorig)

| | |
|---|---|
| Variable label | Original end date |
| Variable name | endorig |
| Category | period of validity |
| Origin | BeH, LeH, LHG, ASU, XASU, MTH |
| Data type | date |
| Hierarchy | none |
| Detailed description | The original end date of the observation corresponds to the original end date of the notification. This can differ from the end date of the relevant line of data, the so-called end date of the split episode (see also the comments on episode splitting in Section 3.2).<br>Since the notification logic might under certain circumstances permit re-identification of the exact day of birth, the original information on the date was changed in these cases by applying the anonymisation procedure described in Section 3.1.6.<br><br>**1) BeH**<br>Because of the rules of the notification procedure, in BeH observations the starting and ending year are always identical (obligation of the employer to submit annual employment notifications). A continuous employment relationship may therefore be distributed across several notifications.<br><br>**2) LHG, ASU, XASU**<br>Certain changes lead to the creation of a new observation (see Sec- |

| | |
|---|---|
| | tions 3.1.3 and 3.1.4). endorig indicates the end date of the new period. |

### 5.3.3 Start date of split episode (begepi)

| | |
|---|---|
| Variable label | Episode start date |
| Variable name | begepi |
| Category | generated period of validity |
| Origin | BeH, LeH, LHG, ASU, XASU, MTH |
| Data type | date |
| Hierarchy | none |
| Detailed description | The start date of the split episode is always equal to or greater than the start date of the original observation (see also the comments on episode splitting in Section 3.2). |

### 5.3.4 End date of split episode (endepi)

| | |
|---|---|
| Variable label | Episode end date |
| Variable name | endepi |
| Category | generated period of validity |
| Origin | BeH, LeH, LHG, ASU, XASU, MTH |
| Data type | date |
| Hierarchy | none |
| Detailed description | The end date of the split episode is always equal to or smaller than the end date of the original observation (see also the comments on episode splitting in Section 3.2). |

## 5.4 Personal information

### 5.4.1 Gender (frau)

| | |
|---|---|
| Variable label | Gender |
| Variable name | frau |
| Category | personal variable |
| Origin | BeH, LeH, LHG, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | Gender dummy (0 - man, 1 - woman). The gender information is constant within one individual account. |

### 5.4.2 Year of birth (gebjahr)

| | |
|---|---|
| Variable label | Year of birth |
| Variable name | gebjahr |
| Category | personal variables |
| Origin | BeH, LeH, LHG, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |

| | |
|---|---|
| Detailed description | The year of birth is constant within one individual account. |
| Notes on quality | In the original data, it may happen that the year of birth changes between the data sources. This is corrected during the data preparation process. The information from the social security number is given highest priority here. |

### 5.4.3 Nationality (nation)

| | |
|---|---|
| Variable label | Nationality |
| Variable name | nation |
| Category | personal variables |
| Origin | BeH, LeH, LHG, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | The variable contains the nation codes used by the Federal Statistical Office (Statistisches Bundesamt). |
| Notes | Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. |

### 5.4.4 Nationality, aggregated (nation_gr)

| | |
|---|---|
| Variable label | Nationality, grouped |
| Variable name | nation_gr |
| Category | personal variables |
| Origin | BeH, LeH, LHG, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | The variable contains a grouped version of the nation codes used by the Federal Statistical Office. |

### 5.4.5 Marital status (famst)

| | |
|---|---|
| Variable label | Marital status |
| Variable name | famst |
| Category | personal variables |
| Origin | LeH, LHG, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | This variable describes the marital status. The characteristic in the LeH source has only two values (0 - not married, 1 - married), while in the LHG/ASU/XASU/MTH sources, a distinction is made between six values. The information from the sources was not compared. |

### 5.4.6 Number of children (kind)

| | |
|---|---|
| Variable label | Number of children |
| Variable name | kind |

| | |
|---|---|
| Category | personal variables |
| Origin | LeH, LHG, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | This variable has a different meaning depending on the data source.<br><br>**1) LeH**<br>The LeH source indicates the number of children aged under 16 at the time when the application was made. It only distinguishes between whether the number of children is zero or at least one. The following values are therefore available:<br>0 no children<br>100 one or more children<br>The variable is not updated when there are changes in the type of benefit or the approval of benefits, but only when a new case of benefit receipt occurs after a period of employment. This restricts the quality of the data.<br><br>**2) ASU, MTH**<br>Here, the value of the variable corresponds to the number of children.<br>Until 30 June 2006, only up to nine children could be recorded. The value 0 does not exist. For observations prior to 30 June 2006, the zero value was recoded to "missing", since it is not clear whether zero should be interpreted as "no children" or as "field not filled in". For observations after 30 June 2006, the variable is only recorded if children exist.<br><br>**3) XASU, LHG**<br>The variable shows the number of children aged under 16 in the benefit community (Bedarfsgemeinschaft). In the LHG sources, the value is valid for the entire original period. |

### 5.4.7 Vocational training (ausbildung)

| | |
|---|---|
| Variable label | Vocational training |
| Variable name | ausbildung |
| Category | personal variables |
| Origin | BeH, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |

| | |
|---|---|
| Detailed description | It must be taken into account that this variable has a different meaning depending on the data source:<br><br>**1) BeH**<br>For BeH observations the variable contains the vocational education and training reported by the employers in the employment notification procedure. The following values exist:<br><br>1 No vocational training<br>2 In-company voc. training/traineeship/external voc. training<br>11 Degree from a university of applied sciences<br>12 University degree<br><br>In notifications using the new occupation code (see Section 1.3.1) it is no longer possible to identify graduates of universities of applied sciences clearly, as the new occupation code no longer has a separate category for this vocational qualification. They are assigned to category 12.<br><br>“Changes in the vocational training status frequently occur at the same time as a change of establishment. This is because the notification data are compiled anew in the new firm. If, for example, an employee has gained a higher qualification via a part-time further training course while still working then this change of status is probably not recorded until he/she joins a new firm. It can generally be assumed that when a person is employed in a firm for a longer period, the personal data that they reported when they joined the firm is simply continued.” (Meinken / Koch 2004, p. 63).<br>A method for correcting missing values or temporal inconsistencies in the education and training data in the predecessor sample, the IABS, can be found in Fitzenberger et al. (2006) and in Drews (2006). Only the BeH data source is used for this, however, as this was the only data source with information on education and training in the IABS.<br><br>**2) ASU, XASU, MTH**<br>For these observations the vocational education or training completed most recently is reported. The following values exist:<br><br>1 no completed vocational training<br>2 in-firm vocational training/external vocational training<br>3 full-time vocational school (Berufsfachschule)<br>4 technical college (Fachschule)<br>5 university of applied sciences (Fachhochschule)<br>6 university<br>7 vocational education/training not recognised in Germany<br>8 university degree not recognised in Germany<br><br>Values 7 and 8 are only valid for the MTH data source. |

| | |
|---|---|
| Notes on quality | The number of missing values increases continuously over time. In the most recent data more than 40% of values are missing. Missing values occur especially frequently in the following person groups: persons in marginal part-time employment, persons working part-time, foreign employees and eastern German workers. The reason for this is that the variable is not of particular importance as regards social security (see Meinken/Koch, 2004, p. 63).<br>As a result of the switchover from coArb to VerBIS it is not possible to distinguish correctly between “no completed vocational training” and “no information available” in the ASU and MTH data sources between 2006 and 2008. A missing value in this period therefore does not necessarily mean that the person has no vocational training or that there is no information available on vocational education and training, but may also mean that it was not possible to apply the relevant data generation procedure. The degree of completeness in the XASU data source is generally low. |

### 5.4.8 School leaving qualification (schule)

| | |
|---|---|
| Variable label | School leaving qualification |
| Variable name | schule |
| Category | personal variables |
| Origin | BeH, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | This variable contains the school leaving qualification. Different values are possible depending on the source.<br><br>**1) BeH**<br>The possible values change when the new occupation code is adopted (see Section 1.3.1). The values with the old occupation code are:<br><br>5 Grade-/lower secondary school with or without leaving certificate, intermediate school leaving certificate or equivalent qualification<br>8 Upper secondary school leaving certificate from a specialised upper secondary school (Fachoberschule), general upper secondary school leaving certificate, A-level equivalent, qualification for university<br>9 General upper secondary school leaving certificate, A-level equivalent, qualification for university<br><br>With the new occupation code the values are:<br><br>1 No school leaving certificate<br>4 Lower secondary school certificate/ grade school certificate<br>6 Intermediate school leaving certificate<br>8 Upper secondary school leaving certificate from a specialised upper secondary school/general upper secondary school leaving certificate, A-level equivalent, qualification for university |

| | |
|---|---|
| | **2) ASU, XASU, MTH**<br>The following values are possible for these data sources:<br><br>1 No school leaving certificate<br>4 Lower secondary school certificate/ grade school certificate<br>6 Intermediate school leaving certificate<br>7 Upper secondary school leaving certificate from a specialised upper secondary school (Fachoberschule)<br>9 General upper secondary school leaving certificate, A-level equivalent, qualification for university<br>They are valid at the beginning of the period of job-search or participation in a measure. In the case of people seeking an apprenticeship position, the variable may also contain the school qualification they are working towards in the XASU data source. |
| Notes on quality | The degree of completeness has been decreasing continuously over time in the BeH and has stood at under 2/3 in recent years. In the XASU it has been increasing continuously and has stood at over 2/3 since 2012. In ASU and MTH the degree of completeness is good on the whole. |

## 5.5 Information on employment, benefit receipt and job search

### 5.5.1 Daily wage, daily benefit rate (tentgelt)

| | |
|---|---|
| Variable label | Daily wage/daily benefit |
| Variable name | tentgelt |
| Category | information on employment, benefit receipt and job search |
| Origin | BeH, LeH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | **1) BeH**<br>In BeH observations, this variable shows the employee's gross daily wage. It is calculated from the fixed-period wages reported by the employer and the duration of the (unsplit) original notification period in calendar days. The daily wage is shown in euros.<br>Until 1998, employers in principle only reported the earnings which were subject to social security contributions. Earnings below the marginal part-time income threshold were not reported. Earnings exceeding the upper earnings limit for statutory pension insurance are only reported up to this limit. There are two upper earnings limits in the statutory pension insurance scheme. The earnings limit of the miners' pension insurance is generally higher than the earnings limit of the pension insurance for wage and salary earners. However, it is not possible to differentiate between these two insurance providers in the data.<br>Since the inclusion of marginal part-time employees in the employment notification procedure on 1 April 1999, earnings below the marginal part-time income threshold have also been recorded; the upper earnings limit still applies as the upper ceiling. In some cases, however, the reported earnings nonetheless exceed the upper earnings limit. Generally, this can probably be attributed to the payment of annual bonus- |

| | |
|---|---|
| | es which the employer can add to the regular earnings in the annual, employment interruption or end of employment notifications. In this case, it is irrelevant whether the upper earnings limit in the statutory pension insurance which is decisive for the notification period is exceeded as a result of this addition. However, such earnings notifications could also be due to incorrect details in the employment period. (The earnings information, however, may be considered less error-prone due to its insurance relevance.)The marginal part-time income threshold and the upper earnings limit for statutory pension insurance differ from year to year as well as between eastern and western Germany (the decisive factor is the location of the establishment). An overview of these limits and thresholds can be found under http://fdz.iab.de.<br>A daily wage reported as 0 euros can be put down to “employment interruption notifications”. During these periods, the employment relationship continues to exist in legal terms, but without pay. This is the case for periods of illness after the end of continued payment of wages, for periods of maternity leave and for sabbaticals.<br>The daily wage is shown with two decimal places. All values greater than 0 and smaller than 0.01 were rounded up to 0.01. This makes it possible to identify the above-mentioned employment interruption notifications with the condition daily wage = 0.<br><br>**2) LeH**<br>For LeH observations, the variable shows the daily benefit rate, converted into euros in each case. It must be taken into account that for observations with an original start date prior to 1 January 1998 the daily benefit rate applies to working days, while for observations with an original start date from 1 January 1998 onwards it applies to calendar days.<br>Since 1 January 2005, a daily benefit rate reported as 0 euros can be put down to benefit suspension periods or interruptions of benefit payments. If a reason for end of benefit is reported for an observation with a daily benefit rate equal to 0, then it is a notification of interruption of benefit payments. In the case of observations that reflect a period of benefit suspension, the entitlement is the same as before the start of the benefit suspension period. |

### 5.5.2 Occupation – current/most recent (KldB 1988) (beruf)

| | |
|---|---|
| Variable label | Occupation – current/most recent (KldB 1988) |
| Variable name | beruf |
| Category | information on employment, benefit receipt and job search |
| Origin | BeH, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | **1) BeH**<br>The occupational title of the job performed by the employee during the notification period is a component of the 'employment details' submitted by the employer. If more than one job title with different classification codes applies to one employee, the employer is required to select |

| | |
|---|---|
| | the job title that best defines the main activity performed (see BA 2005, p. V).<br>For this, the employer encodes the employee's job in accordance with the "Classification of Occupations. Systematic and Alphabetical Directory of Job Titles" (published by the Federal Employment Agency, Nuremberg, 1988), which contains approx. 25,000 job titles. The occupational classification consists of a 3-digit code and comprises about 330 values.<br>Employment notifications with an end date later than 30.11.2011 are reported using the new occupation code 2010 (KldB2010) (see Section 1.3.1). These values are recoded to the KldB1988 by transferring the key area. This results in inaccuracies.<br><br>**2) ASU, XASU, MTH**<br>The variable contains the occupation of the last job. See 1) with regard to the occupation code. |
| Note on quality | There is a considerable increase in the number of missing values in 2011 due to the change in the reporting procedure.<br>In the XASU the occupation variable is not reported for almost the entire period available. |

### 5.5.3 Occupational group – current/most recent (KldB 2010) (beruf2010_3)

| | |
|---|---|
| Variable label | Occupational group - current/most recent (KldB 2010), 3-digit |
| Variable name | beruf2010_3 |
| Category | information on employment, benefit receipt and job search |
| Origin | BeH, LeH, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | **1) BeH**<br>The occupational title of the job performed by the employee during the notification period is a component of the 'employment details' submitted by the employer. If more than one job title with different classification codes apply for one employee, the employer is required to select the job title that best defines the main activity performed (see Bundesagentur für Arbeit, 2005, p. V).<br>For this the employer encodes the employee's job in accordance with the "Classification of Occupations 2010" (Klassifikation der Berufe 2010, KldB2010, Bundesagentur für Arbeit, 2011). The occupational class consists of a 5-digit code and comprises about 1300 values. The less detailed occupational group is recorded by the first three digits of the code. The skill level required for a job, which is recorded in the fifth digit of the codes in the KldB2010, is made available separately in the variable 'level of requirement' (niveau).<br>Employment notifications with an end date earlier than 30.11.2011 are reported using the old occupation code 1988 (KldB 1988) (see Section 1.3.1). These values are recoded to the KldB2010 by transferring the key area. As the new occupation code is considerably more detailed than the old one, this results in substantial inaccuracies. This must be taken into account when analysing the data. |

| | |
|---|---|
| | **2) LeH, ASU, XASU, MTH**<br>The variable contains the occupation of the last job. See 1) with regard to the occupation code. |
| Note on quality | There is a considerable increase in the number of missing values in 2011 due to the change in the reporting procedure. |

### 5.5.4 Occupational sub-group – current/most recent (KldB 2010) (beruf2010_4)

| | |
|---|---|
| Variable label | Occupational sub-group - current/most recent (KldB 2010), 4-digit |
| Variable name | beruf2010_4 |
| Category | information on employment, benefit receipt and job search |
| Origin | BeH, LeH, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | **1) BeH**<br>The occupational title of the job performed by the employee during the notification period is a component of the 'employment details' submitted by the employer. If more than one job title with different classification codes apply for one employee, the employer is required to select the job title that best defines the main activity performed (see Bundesagentur für Arbeit, 2005, p. V).<br>For this the employer encodes the employee's job in accordance with the "Classification of Occupations 2010" (Klassifikation der Berufe 2010, KldB2010, Bundesagentur für Arbeit, 2011). The occupational class consists of a 5-digit code and comprises about 1300 values. The less detailed occupational sub-group is recorded by the first four digits of the code. The skill level required for a job, which is recorded in the fifth digit of the codes in the KldB2010, is made available separately in the variable 'level of requirement' (niveau).<br>Employment notifications with an end date earlier than 30.11.2011 are reported using the old occupation code 1988 (KldB 1988) (see Section 1.3.1). These values are recoded to the KldB2010 by transferring the key area. As the new occupation code is considerably more detailed than the old one, this results in substantial inaccuracies. This must be taken into account when analysing the data.<br><br>**2) LeH, ASU, XASU, MTH**<br>The variable contains the occupation of the last job. See 1) with regard to the occupation code. |
| Note on quality | There is a considerable increase in the number of missing values in 2011 due to the change in the reporting procedure. |
| Notes | Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise the occupational group (beruf2010_3) is the only occupation-related information using KldB2010 that is made available. |

### 5.5.5 Level of requirement – current/most recent (KldB 2010) (niveau)

| | |
|---|---|
| Variable label | Level of requirement - current/most recent (KldB 2010) |
| Variable name | niveau |
| Category | information on employment, benefit receipt and job search |
| Origin | BeH, LeH, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | **1) BeH**<br>The occupational title of the job performed by the employee during the notification period is a component of the 'employment details' submitted by the employer. If more than one job title with different classification codes apply for one employee, the employer is required to select the job title that best defines the main activity performed (see Bundesagentur für Arbeit, 2005, p. V).<br>For this the employer encodes the employee's job in accordance with the "Classification of Occupations 2010" (Klassifikation der Berufe 2010, KldB2010, Bundesagentur für Arbeit, 2011). The occupational class consists of a 5-digit code and comprises about 1300 values. The less detailed occupational sub-group is recorded by the first four digits of the code. The skill level required for a job, which is recorded in the fifth digit of the codes in the KldB2010, is made available separately in the variable 'level of requirement' (niveau).<br>Employment notifications with an end date earlier than 30.11.2011 are reported using the old occupation code 1988 (KldB 1988) (see Section 1.3.1). These values are recoded to the KldB2010 by transferring the key area. As the new occupation code is considerably more detailed than the old one, this results in substantial inaccuracies. This must be taken into account when analysing the data.<br><br>**2) LeH, ASU, XASU, MTH**<br>The variable contains the occupation of the last job. See 1) with regard to the occupation code. |
| Note on quality | There is a considerable increase in the number of missing values in 2011 due to the change in the reporting procedure. |

### 5.5.6 Part-time (teilzeit)

| | |
|---|---|
| Variable label | Part-time |
| Variable name | teilzeit |
| Category | information on employment, benefit receipt and job search |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | The employee's occupational status during the notification period is reported by the employer in the "employment details".<br>The variable "occupational status" distinguishes between full-time and part-time employees. The decisive factor here is the ratio between the contracted hours and the usual working hours in the establishment. |

| | |
|---|---|
| | For part-time employees the variable only records whether their working hours exceed a certain limit or not. Until 1978 this limit was 20 hours of work per week, between 1979 and 1987 it was 15 hours per week and since 1988 it has been 18 hours per week.<br>The variable only provides actual details regarding the occupational status for full-time employees, distinguishing among other things between blue-collar and white-collar employees in full-time employment and trainees/apprentices. When the new occupation code was introduced (see Section 1.3.1), however, this distinction was no longer available. The variable 'teilzeit' therefore only distinguishes between full-time and part-time employment in the entire reporting period. No further information about the occupational status is used. |
| Note on quality | There is a considerable increase in the number of missing values in 2011 due to the change in the reporting procedure. In order to reduce this problem, the working hours were imputed at the IAB for the period in question. Further information about the procedure can be found in Ludsteck/Thomsen (2016). |

### 5.5.7 Employment status (erwstat)

| | |
|---|---|
| Variable label | Employment status |
| Variable name | erwstat |
| Category | information on employment, benefit receipt and job search |
| Origin | BeH, LeH, LHG, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | This variable assumes different values with different meanings for each data source.<br><br>**1) BeH**<br>For BeH observations, the variable 'employment status' corresponds to the person group recorded in the new notification procedure (DEÜV) from 1 January 1999 onwards. It indicates contribution- or benefit-related particularities of the employment relationship.<br>If multiple codes apply to an employment notification, the smallest must be indicated by the reporting employer. The majority of these cases are employment relationships subject to social security contributions without any distinctive characteristics, which are recorded under code number 101. Accordingly, it is possible that these employment relationships are slightly overestimated.<br>The notification procedure stipulates that changes in the employment status - e.g. when an apprentice is taken on by his/her training company after completing his/her vocational training - must be indicated by a new notification (cf. Deutsche BKK 2012, p. 31).<br>The person group can be contained in employment notifications that refer to the years prior to 1999 but were not received until 1999 or later. For notifications which were received before 1999, an attempt is made to allocate the notifications to the person groups on the basis of certain rules and with the aid of the vocational education and training', 'occupational status and working hours' and 'occupation' variables as well as other information. In many cases, however, conclusive alloca- |

| | |
|---|---|
| | tions are not possible.<br>Since 1 April 1999, employees in marginal part-time employment have also been recorded in the DEÜV notification procedure. This persongroup can be distinguished via the values 109 and 209. For employees in marginal part-time employment, no data prior to the introduction of the notification obligation in 1999 could be collected.<br><br>**2) LeH**<br>For LeH observations, the variable 'employment status' contains the grouped benefit type. Thus, it is possible to distinguish whether a person receives unemployment benefit, unemployment assistance or maintenance allowance or whether contributions to private long-term care insurance are paid by the BA.<br><br>**3) ASU/XASU**<br>For ASU observations, the 'employment status' variable reports the job search status. Recipients of unemployment benefits (Unemployment Benefit I or II) over the age of 58 who receive benefits under the relaxed conditions according to Section 428 of Social Code Book III (or Section 65 Para. 4 of Social Code Book II) and individuals aged over 58 who are not benefit recipients and are not willing to be placed in employment in the sense of Section 252 Para. 8 Social Code Book VI are recorded as individuals seeking advice.<br>The status 'without status' (statistics: 'not set') mainly implies individuals who cannot be expected to be activated or placed in employment in accordance with Section 10 SGB II. Individuals who are classed as unfit for work for more than 42 days but continue to receive Unemployment Benefit II are also recorded under this status in the system.<br>In XASU observations, the variable 'employment status' has so far also contained the values 'not unemployed, but seeking work' as well as 'unemployed and simultaneously seeking work'.<br>Since the notification logic might in many cases permit re-identification of the exact day of birth, the original information on the date was changed by applying the anonymisation procedure described in Section 3.1.6.<br><br>**4) LHG**<br>For LHG datasets, the 'employment status' variable shows whether the person is registered as an employable minor, an employable person of full age or not employable from the old-age pension threshold.<br>Since the notification logic might in many cases permit re-identification of the exact day of birth, the original information on the date was changed by applying the anonymisation procedure described in Section 3.1.6.<br><br>**5) MTH**<br>For MTH observations the 'employment status' variable indicates the measure-type category. This is the highest level in the hierarchy of the measure-type classifications of the BA. |

| | |
|---|---|
| Note on quality | In the LHG it can be observed that there is an above-average number of 15-year-olds and to a lesser extent 16- and 17-year-olds classed as unable to work. 15- and 16-year-old benefit recipients of the authorised municipalities may therefore be under-represented, as 'individuals who are unable to work' are not included in the IEB. |

### 5.5.8 Transition zone (gleitz)

| | |
|---|---|
| Variable label | Transition zone |
| Variable name | gleitz |
| Category | information on employment, benefit receipt and job search |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | This variable is only available from 2003 onwards and only for BeH observations. It indicates whether the employment notification relates to employment in the low-wage sector, within the so-called transition zone. Jobs in the transition zone have a gross monthly wage of € 400.01 to € 800.00 (so-called midi jobs) for which the employee only has to pay a reduced overall social security contribution. As employees with earnings in the transition zone can voluntarily pay the "regular" social security contribution, not all employees with corresponding earnings are automatically classified as being in the transition zone. The corresponding legislation has been in force since 1 April 2003. |

### 5.5.9 Temporary agency work (leih)

| | |
|---|---|
| Variable label | Temporary agency work |
| Variable name | leih |
| Category | information on employment, benefit receipt and job search |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | The variable reports whether the person's employment is a temporary job via an employment agency. The variable is derived from the occupation code 2010 and is only available for notifications with an end date later than 30.11.2011. |
| Note on quality | There is a considerable increase in the number of missing values in 2011 due to the change in the reporting procedure. |

### 5.5.10 Fixed-term contract (befrist)

| | |
|---|---|
| Variable label | Fixed-term contract |
| Variable name | befrist |
| Category | information on employment, benefit receipt and job search |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | none |

| | |
|---|---|
| Detailed description | The variable reports whether the person's employment relationship is fixed-term or permanent. The variable is derived from the occupation code 2010 and is only available for notifications with an end date later than 30.11.2011. |
| Note on quality | There is a considerable increase in the number of missing values in 2011 due to the change in the reporting procedure. |

### 5.5.11 Reason of cancellation/notification/termination (grund)

| | |
|---|---|
| Variable label | Reason of cancellation/notification/termination |
| Variable name | grund |
| Category | information on employment, benefit receipt and job search |
| Origin | BeH, LeH, LHG, ASU, XASU |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | **1) BeH**<br>In BeH observations, the 'reason for notification' variable indicates the reason why the employer submitted the employment notification in question to the social security agencies. However, not all of the possible reasons for submitting a notification that may occur in the context of the notification procedure are available in the IEB. For instance, the IEB only includes notifications that have information on earnings (i.e. annual, employment interruption and end of employment notifications), whilst initial registrations are not contained as they contain no information on earnings. However, this does not involve a loss of information, as the details from a registration are transmitted again with the following annual, employment interruption or end of employment notification.<br>The reasons for submitting employment notifications are encoded according to the regulations of the notification procedure which has been in effect since 1 January 1999 (in accordance with DEÜV).<br><br>**3) LHG**<br>The LHG observations contain the 'reason for discontinuation of Unemployment Benefit II' and indicate the reason why current benefits have been discontinued. The 'reason for discontinuation of Unemployment Benefit II' variable **refers to the individual**, not to the benefit community. If the Unemployment Benefit II receipt of a different member of the benefit community is discontinued, new observations for all members of the benefit community are started on this date, but the reason for discontinuation of Unemployment Benefit II is only available for the individual whose benefit is discontinued. This variable is valid exactly at the end of the original observation.<br><br>**4) ASU**<br>In the case of ASU observations, the variable contains the deregistration or exit reason. In the case of a change of legal sphere, the observation is split artificially and 'generated by data splitting' is entered as the reason for deregistration. In order to depict the reasons for deregistration correctly it is also necessary to take into account the variable |

| | |
|---|---|
| | ‘status after job search’.<br>The number of values of the variable was reduced from 26 April 2003 onwards. For analyses over long periods of time, the old values can be recoded to the currently valid ones using the table below:<br><br>**old -> new  old -> new  old -> new  old -> new**<br>29 -> 60  36 -> 61  44 -> 74  51 -> 74<br>30 -> 60  37 -> 66  45 -> 77  52 -> 76<br>31 -> 61  38 -> 66  46 -> 67  53 -> 68<br>32 -> 60  39 -> 71  47 -> 67  54 -> 78<br>33 -> 60  40 -> 69  48 -> 78<br>34 -> 60  42 -> 65  49 -> 69<br>35 -> 60  43 -> 70  50 -> 75<br><br>**5) XASU**<br>In the case of XASU observations, the variable contains the deregistration or exit reason. In the case of a change of legal sphere, the observation is split artificially and ‘generated by data splitting’ is entered as the reason for deregistration. |
| Note on quality | The proportion of valid values (degree of completeness) for the reason for notification in the LHG data sources is very small (< 20%) across all years. The degree of completeness is also low in the XASU, but improves over time. |

### 5.5.12 Employment status prior to job search (estatvor)

| | |
|---|---|
| Variable label | Employment status prior to job search |
| Variable name | estatvor |
| Category | information on employment, benefit receipt and job search |
| Origin | ASU, XASU |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | For ASU and XASU observations, this variable shows the employment status prior to the job search activity. From 12/2002 onwards the number of values of the variable was reduced substantially. The values of older observations were recoded to the currently valid values, which are thus valid for the entire period.<br>If an observation within the ASU/XASU is split artificially (for instance in the case of a change of legal sphere from SGB III to SGB II), the reason for registration is reported as ‘generated by data splitting’. However, this does not apply for episode splitting within the SIAB.<br>This information does not refer to the start date of the episode, but to the start of the original time period |
| Note on quality | The proportion of valid values (degree of completeness) increases for XASU observations from approx. 8 % to approx. 78 % between 2005 and 2013. |

### 5.5.13 Employment status after job search (estatnach)

| | |
|---|---|
| Variable label | Employment status after job search |

| | |
|---|---|
| Variable name | estatnach |
| Category | information on employment, benefit receipt and job search |
| Origin | ASU |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | The variable contains the person's status after leaving unemployment. Longer periods of illness can be identified via this variable.<br>The values are classified as follows:<br>• 1000s: measure (assisted employment)<br>• 2000s: non-assisted employment<br>• 3000s: training etc.<br>• 4000s: self-employment<br>• 5000s: exclusion<br>• 6000s: other |

### 5.5.14 Client profile (profil)

| | |
|---|---|
| Variable label | Client profile |
| Variable name | profil |
| Category | information on employment, benefit receipt and job search |
| Origin | ASU, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | The variable reports the client profile assigned to the client in the profiling process. The profiling process serves to create a client profile, i.e. a list of the client's skills, experiences and interests with labour-market relevance, in order to identify the client's position in the labour market more easily. Towards the end of the profiling process, the items are summarised to create a client profile. To this end, the client's overall integration prospects are first ascertained. The following options are available:<br>• good integration prospects (integration into the regular labour market within 12 months is realistic)<br>• complex (integration into the regular labour market within 12 months is not realistic)<br>• other<br>The allocation of the client profile depends on the identification of the integration prospects. Clients whose integration prospects are classed as good can be assigned the client profiles 'market profile', 'activation profile' and 'assistance profile', while clients with complex prospects are assigned the client profiles 'development profile', 'stabilisation profile' or 'support profile'. The selection of the specific client profile is based on the need for action as assessed by the placement officer. If the client's prospects are classed as 'other', the option 'assignment not required' or – only for SGB II clients – 'integrated but in receipt of benefits' may be selected as the client profile. |

| | |
|---|---|
| Note on quality | The variable was introduced in 2009 but was mapped back to 2006 using other variables. The quality has been assessed as reliable by the BA statistics department since 2010. |

### 5.5.15 Type of termination of last job (art_kuend)

| | |
|---|---|
| Variable label | Reason for end of previous employment |
| Variable name | art_kuend |
| Category | information on employment, benefit receipt and job search |
| Origin | ASU, MHT |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | This variable describes how the last employment or training relationship was terminated before a new period of job search. It can therefore be used to identify job-to-job placements. |

### 5.5.16 Desired working hours of the job sought (arbzeit)

| | |
|---|---|
| Variable label | Working hours of job application |
| Variable name | arbzeit |
| Category | information on employment, benefit receipt and job search |
| Origin | ASU, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | During the placement procedure, jobseekers indicate how many working hours the job they are seeking should have. |

### 5.5.17 Residual claim/planned duration (restanspruch)

| | |
|---|---|
| Variable label | Residual claim/planned duration |
| Variable name | restanspruch |
| Category | information on employment, benefit receipt and job search |
| Origin | LeH, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | The variable has a different meaning depending on the data source.<br><br>**1) LeH**<br>The variable contains the residual entitlement to unemployment benefit that remains after the end of the current benefit receipt period. If the period of benefit receipt ends before the maximum duration of entitlement has been reached (e.g. due to taking up employment again), a residual entitlement remains which is equivalent to the duration of benefit entitlement that was not used up. If new entitlement is acquired within five years, the duration of the residual entitlement is added to the new duration of entitlement. However, the maximum duration of entitlement for the client's age is the upper limit. If no new entitlement is acquired, the residual entitlement can be used for benefits within |

| | |
|---|---|
| | four years on application.<br>If the end date of the benefit receipt is before 1 January 1998, the remaining entitlement is reported in working days, after this date it is reported in calendar days. This information does not refer to the start date of the episode, but to the start of the original time period.<br><br>**2) MTH**<br>The variable contains the planned duration of the measure. |

### 5.5.18 Type of institution (traeger)

| | |
|---|---|
| Variable label | Type of provider |
| Variable name | traeger |
| Category | information on employment, benefit receipt and job search |
| Origin | LHG, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | The variable contains the type of institution responsible for implementing Unemployment Benefit II (LHG), the type of institution providing the measure (MTH) or the institution responsible for managing the applicant profile (ASU/XASU). The variable contains not only the "Bundesagentur für Arbeit (BA)" as the type of institution responsible for implementing Social Code Book III (SGB III), but also three types of institution responsible for implementing Social Code Book II (SGB II). For further information see Section 2.2. The type of institution responsible for Unemployment Benefit II may change over time (e.g. from a cooperation of an employment agency and a municipality to a municipality excercising its duties separately, or from joint facilities to an authorised municipality). It must also be taken into account that the district territories covered by the institutions are not always distinct and thus may not necessarily correspond to the boundaries of the districts. The same applies for the employment agencies.<br>ASU and XASU observations contain the type of institution that holds the records of the Applicant Pool Data. |

### 5.5.19 Start date of unemployment (alo_beg)

| | |
|---|---|
| Variable label | Start date of unemployment |
| Variable name | alo_beg |
| Category | information on employment, benefit receipt and job search |
| Origin | LeH, LHG, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |

| | |
|---|---|
| Detailed description | The variable reports the start date of an uninterrupted sequence of periods of unemployment and is valid at the beginning of the observation.<br><br>The following gaps do not result in an interruption of the period of unemployment:<br>• any gap lasting seven days or less<br>• periods of illness lasting up to 42 days (ASU)<br><br>No information about illnesses is contained in XASU observations, which is why it cannot be taken into account in the calculations. |

### 5.5.20 Duration of unemployment (alo_dau)

| | |
|---|---|
| Variable label | Duration of unemployment |
| Variable name | alo_dau |
| Category | information on employment, benefit receipt and job search |
| Origin | LeH, LHG, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | The variable reports the duration (in days) of an uninterrupted sequence of periods of unemployment and is valid at the beginning of the observation.<br><br>The following gaps do not result in an interruption of the period of unemployment:<br>• any gap lasting seven days or less<br>• periods of illness lasting up to 42 days (ASU)<br><br>When calculating the duration these gaps are not added, however. No information about illnesses is contained in XASU observations, which is why it cannot be taken into account in the calculations.<br><br>Prior to 1997 the value “0” does not mean that the individual was not unemployed, as the ASU/XASU sources are not available here. |

## 5.6 Location data

### 5.6.1 Place of residence: district (Kreis) (wo_kreis)

| | |
|---|---|
| Variable label | Place of residence: district (Kreis) |
| Variable name | wo_kreis |
| Category | location data |
| Origin | BeH, LeH, LHG, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | federal state \| district |
| Detailed description | In BeH and LeH observations, the place of residence at district level is only available for the years from 1999 onwards. The variable indicates the district (urban district or rural district) in which the social security |

| | |
|---|---|
| | contributor lives. The first two digits of the 5-digit district code (Kreisschlüssel) show the code for the federal state (Bundesland), positions 1-3 indicate the regional authority (Regierungsbezirk), and positions 1-5 show the district authority (Kreis). Federal states without a regional authority have a 0 in the third position.<br>In the BeH, the place of residence is determined at the end of each year and added consistently to all datasets of a year. For the LHG, XLHG and XASU sources, the place of residence applies to the period of the original observation. For the ASU and LeH, the variable contains the place of residence at the beginning of the original period of time.<br>In order to guarantee consistent regional allocations across the entire observation period, the information on the district was recoded with reference to the territorial allocation of 31 December 2014 for all sources, i.e. in all calendar years, a place of residence is assigned to a district in accordance with the boundaries that the district had on 31 December 2014. As the district boundaries have changed over time, cases would occur in which the district code changes without the individual concerned having relocated if the territorial allocations of the districts were not updated. |
| Notes | Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, only the federal state (wo_bula) is shown as regional information. |

### 5.6.2 Place of residence: federal state (Bundesland) (wo_bula)

| | |
|---|---|
| Variable label | Place of residence: federal state (Bundesland) |
| Variable name | wo_bula |
| Category | location data |
| Origin | BeH, LeH, LHG, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | federal state \| district |
| Detailed description | This variable is an aggregation of the “district” variable to the 16 German federal states.<br>In BeH observations, the federal state of the place of residence is only available for the years from 1999 onwards. Further information on the district of the place of residence can be found under Location data<br>Place of residence: district (Kreis) (wo_kreis). |

### 5.6.3 Place of residence: employment agency (wo_aa)

| | |
|---|---|
| Variable label | Place of residence: employment agency (Arbeitsagentur) |
| Variable name | wo_aa |
| Category | location data |
| Origin | BeH, LeH, LHG, XLHG, ASU, XASU |
| Data type | numerical |
| Hierarchy | regional directorate \| employment agency |
| Detailed description | This variable contains the employment agency of the employee’s / BA client’s place of residence from 1999 onwards. This information is determined from the residence address. For the LHG and XASU data |

| | |
|---|---|
| | sources, the place of residence is valid for the period of the original observation. In the case of the LeH, ASU and MTH, the variable contains the place of residence at the start of the period of unemployment or job search. For the BeH and LeH, it is available from 1999 onwards.<br>In order to guarantee consistent regional allocations across the entire observation period, the information on the employment agency was recoded to the territorial allocation of 31 December 2014 for all data sources, i.e. in all calendar years, a place of residence is assigned to an employment agency area in accordance with the boundaries that the employment agency area had on 31 December 2014. As the boundaries of the employment agency areas have changed over time, cases would occur in which a person's employment agency area changes without him/her having relocated if the territorial allocations were not updated. Berlin constitutes a problematic case with regard to updating territorial allocations, however: The boundaries of Berlin's employment agency areas have been changed repeatedly over the years, which could not be fully corrected even by recoding the territorial allocations. Prior to 1999 the variable contains the employment agency that last processed the case of benefit receipt. It is not determined from the residence address and is only available for the LeH data source. Moreover, it is not possible to update the territorial allocations. This must be taken into account when conducting calculations over time (e.g. regional unemployment figures). |
| Note on quality | Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, only the area of the regional directorate in which the social security contributor's place of residence is located is shown. |

### 5.6.4 Place of residence: regional directorate (wo_rd)

| | |
|---|---|
| Variable label | Place of residence: regional directorate (Regionaldirektion) |
| Variable name | wo_rd |
| Category | location data |
| Origin | BeH, LeH, LHG, ASU, XASU, MTH |
| Data type | numerical |
| Hierarchy | regional directorate \| employment agency |
| Detailed description | This variable is an aggregation of the variable 'place of residence: employment agency' at the level of the regional directorates. Further information can be found there. |

## 5.7 Establishment variables

### 5.7.1 Economic activity 73 (w73_3)

| | |
|---|---|
| Variable label | classification of economic activities 73 |
| Variable name | w73_3 |
| Category | establishment variables |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | division (1-digit code) \| group (2-digit code) \| class (3-digit code) of |

| | |
|---|---|
| | economic activity |
| Detailed description | This variable indicates the economic activity as a 3-digit code in accordance with the WS73 classification and is available from 1975 up to and including 2002.<br>WS73 stands for the “Classification of Economic Activities for the Statistics of the Federal Employment Services, edition 1973” (“Klassifikation der Wirtschaftszweige für die Statistik der Bundesanstalt für Arbeit, Ausgabe 1973”). Using a 3-digit code, the classification distinguishes between 269 classes of economic activity, whereby the first digit of the code defines the division of economic activity of a total of 10, and the first two digits together define the particular group of economic activity of a total of 95.<br>Each establishment is only assigned one code. The assignment to the relevant class of economic activity is carried out under consideration of the institutional orientation of the establishment. |

### 5.7.2 Economic activity 93, 5-digit code (w93_5)

| | |
|---|---|
| Variable label | classification of economic activities 93, sub-classes |
| Variable name | w93_5 |
| Category | establishment variables |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | section (1-digit code) \| division (2-digit code) \| group (3-digit code) \| class (4-digit code) \| sub-class (5 digit code) of economic activity |
| Detailed description | This variable indicates the economic activity as a 5-digit code in accordance with the WZ93 classification and is available from 1999 up to and including 2003. WZ93 stands for the “Classification of Economic Activities for the Statistics of the Federal Employment Services, edition 1993” (“Klassifikation der Wirtschaftszweige für die Statistik der Bundesanstalt für Arbeit, Ausgabe 1993”). The WZ93 is based on the Statistical Classification of Economic Activities in the European Community NACE Rev. 1 („Nomenclature génerale des activités économiques dans les communautés européennes“) which has four levels the first two of which are based on the international standard ISIC Rev. 3 („International Standard Industrial Classification of All Economic Activities“).<br>Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected. |
| Note on quality | Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, the economic activity is only shown as the 3-digit code (w93_3). |

### 5.7.3 Economic activity 93, 3-digit code (w93_3)

| | |
|---|---|
| Variable label | classification of economic activities 93, groups |
| Variable name | w93_3 |
| Category | establishment variables |

| | |
|---|---|
| Origin | BeH |
| Data type | numerical |
| Hierarchy | section (1-digit code) \| division (2-digit code) \| group (3-digit code) \| class (4-digit code) \| sub-class (5 digit code) of economic activity |
| Detailed description | This variable indicates the economic activity as a 3-digit code in accordance with the WZ93 classification and is available from 1999 up to and including 2003. WZ93 stands for the "Classification of Economic Activities for the Statistics of the Federal Employment Services, edition 1993" ("Klassifikation der Wirtschaftszweige für die Statistik der Bundesanstalt für Arbeit, Ausgabe 1993"). The WZ93 is based on the Statistical Classification of Economic Activities in the European Community NACE Rev. 1 („Nomenclature génerale des activités économiques dans les communautés européennes") which has four levels the first two of which are based on the international standard ISIC Rev. 3 („International Standard Industrial Classification of All Economic Activities").<br>Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected. |

### 5.7.4 Economic activity 03, 5-digit code (w03_5)

| | |
|---|---|
| Variable label | classification of economic activities 03, sub-classes |
| Variable name | w03_5 |
| Category | establishment variables |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | section (1-digit code) \| division (2-digit code) \| group (3-digit code) \| class (4-digit code) \| sub-class (5-digit code) of economic activity |
| Detailed description | This variable indicates the economic activity as a 5-digit code in accordance with the WZ03 classification and is available from 2003 up to 2008. WZ03 stands for the "Classification of Economic Activities, Edition 2003" ("Klassifikation der Wirtschaftszweige Ausgabe 2003") of the Federal Statistical Office (eds.). Like the WZ93, the WZ03 is based on the Statistical Classification of Economic Activities in the European Community NACE Rev. 1 (see description of variables w93_3, w93_5). The classifications of the economic activity have been updated, but the structure of the WZ93 has been largely retained.<br>Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected. |
| Note on quality | Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, the economic activity is only shown as the 3-digit code (w03_3). |

### 5.7.5 Economic activity 03, 3-digit code (w03_3)

| | |
|---|---|
| Variable label | classification of economic activities 03, groups |
| Variable name | w03_3 |

| | |
|---|---|
| Category | establishment variables |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | section (1-digit code) \| division (2-digit code) \| group (3-digit code) \| class (4-digit code) \| sub-class (5 digit code) of economic activity |
| Detailed description | This variable indicates the economic activity as a 3-digit code in accordance with the WZ03 classification and is available from 2003 up to 2008. WZ03 stands for the "Classification of Economic Activities, Edition 2003" ("Klassifikation der Wirtschaftszweige Ausgabe 2003") of the Federal Statistical Office (eds.). Like the WZ93, the WZ03 is based on the Statistical Classification of Economic Activities in the European Community NACE Rev. 1 (see description of variables w93_3, w93_5). The classifications of the economic activity have been updated, but the structure of the WZ93 has been largely retained.<br>Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected. |

### 5.7.6 Economic activity 08, 5-digit code (w08_5)

| | |
|---|---|
| Variable label | classification of economic activities 08, sub-classes |
| Variable name | w08_5 |
| Category | establishment variables |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | section (1-digit code) \| division (2-digit code) \| group (3-digit code) \| class (4-digit code) \| sub-class (5-digit code) of economic activity |
| Detailed description | This variable indicates the economic activity as a 5-digit code in accordance with the WZ08 classification and is available from 2008 onwards. WZ08 stands for the "Classification of Economic Activities, Edition 2008" ("Klassifikation der Wirtschaftszweige Ausgabe 2008") of the Federal Statistical Office (eds.). The WZ08 is based on the Statistical Classification of Economic Activities in the European Community NACE Rev. 2.<br>Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected. |
| Notes | Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise the economic activity is only shown as the 3-digit code (w08_3). |

### 5.7.7 Economic activity 08, 3-digit code (w08_3)

| | |
|---|---|
| Variable label | classification of economic activities 08, groups |
| Variable name | w08_3 |
| Category | establishment variables |
| Origin | BeH |
| Data type | numerical |

| | |
|---|---|
| Hierarchy | section (1-digit code) \| division (2-digit code) \| group (3-digit code) \| class (4-digit code) \| sub-class (5 digit code) of economic activity |
| Detailed description | This variable indicates the economic activity as a 3-digit code in accordance with the WZ08 classification and is available for the years since 2008. WZ08 stands for the "Classification of Economic Activities, Edition 2008" ("Klassifikation der Wirtschaftszweige Ausgabe 2008") of the Federal Statistical Office (eds.). The WZ08 is based on the Statistical Classification of Economic Activities in the European Community NACE Rev. 2.<br>Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected. |

### 5.7.8 Economic activity 73 generated – completed by extrapolation / imputation (w73_3_gen)

| | |
|---|---|
| Variable label | w73_3 completed by extrapolation/imputation |
| Variable name | w73_3_gen |
| Category | establishment variables |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | division (1-digit code) \| group (2-digit code) \| class (3-digit code) |
| Detailed description | This variable indicates the economic activity as a 3-digit code in accordance with the WZ73 classification. From 1975 up to and including 2002, the variable contains the original values from w73_3. From 2003 onwards, the information is either continued or replaced with the help of recoding tables. Thus the variable provides time-consistent information on the economic activity based on the economic activity classification WS73. A detailed description can be found in Eberle et al. (2011).<br>Further information on the WS73 classification can be found in the description of variable w73_3. |

### 5.7.9 Economic activity 73 generated – type of completion (group_w73_3)

| | |
|---|---|
| Variable label | Type of imputation w73_3 |
| Variable name | group_w73_3 |
| Category | establishment variables |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | None |
| Detailed description | This variable indicates the type of completion for the w73_3_gen variable. It reports whether the respective value in w73_3_gen is consistent with the original value from w73_3, still missing / extrapolated or imputed based on recording tables.<br>A detailed description of the procedure can be found in Eberle et al. (2011). |

### 5.7.10 Economic activity 93 generated – completed by extrapolation / imputation (w93_3_gen)

| | |
|---|---|
| Variable label | w93_3 completed by extrapolation/imputation |
| Variable name | W93_3_gen |
| Category | establishment variables |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | section (1-digit code) \| division (2-digit code) \| group (3-digit code) \| class (4-digit code) \| sub-class (5 digit code) of economic activity |
| Detailed description | This variable indicates the economic activity as a 3-digit code in accordance with the WZ93 classification. From 1998 up to and including 2003, the variable contains the original values from w93_3. Before 1998 and after 2003, the information is either written back / continued or replaced with the help of recoding tables, so that the variable contains time-consistent information on the economic activity based on the economic activity classification WS93. A detailed description can be found in Eberle et al. (2011).<br>Further information on the WS93 classification can be found in the description of variable w93_3. |

### 5.7.11 Economic activity 93 generated – type of completion (group_w93_3)

| | |
|---|---|
| Variable label | Type of imputation w93_3 |
| Variable name | group_w93_3 |
| Category | establishment variables |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | None |
| Detailed description | This variable indicates the type of completion for the w93_3_gen variable. It reports whether the respective value in w93_3_gen is consistent with the original value from w93_3, still missing / extrapolated or imputed based on recording tables.<br>A detailed description of the procedure can be found in Eberle et al. (2011). |

### 5.7.12 Year of first appearance of establishment number (grd_jahr)

| | |
|---|---|
| Variable label | year of first appearance |
| Variable name | grd_jahr |
| Category | establishment variables |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | none |

| | |
|---|---|
| Detailed description | This variable indicates the first appearance of the establishment number in the dataset.<br>If an establishment number in western Germany is only determined for the first time after 1975 (or after 1992 in eastern Germany), this variable could indicate the date when the respective establishment was founded. However, it could also be an establishment that has been in existence for a longer time but has been allocated a new establishment number following a change of owner or a change in the legal form of the establishment. (For the allocation of establishment numbers see Bundesagentur für Arbeit 2007, pp. 9-11). It is also possible that the establishment already existed before, but had no employees subject to social security, or from 1999 onwards, no marginal part-time workers. |

### 5.7.13 First appearance of establishment number (grd_dat)

| | |
|---|---|
| Variable label | first appearance |
| Variable name | grd_dat |
| Category | establishment variables |
| Origin | BeH |
| Data type | date |
| Hierarchy | none |
| Detailed description | This variable indicates the first appearance of the establishment number in the BeH to the day. If an establishment number in western Germany is only determined for the first time after 1975, or after 1992 in eastern Germany, this variable could indicate the date when the respective establishment was founded. However, it could also be an establishment that has been in existence for a longer time but has been allocated a new establishment number following a change of ownership or a change in the legal form of the establishment. (For the allocation of establishment numbers see Bundesagentur für Arbeit 2007, pp. 9-11). It is also possible that the establishment already existed before, but had no employees subject to social security, or from 1999 onwards, no marginal part-time workers. |
| Notes | Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, only the year when the establishment number first appeared is shown (grd_jahr). |

### 5.7.14 Year of last appearance of establishment number (lzt_jahr)

| | |
|---|---|
| Variable label | year of last appearance |
| Variable name | lzt_jahr |
| Category | establishment variables |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | none |

| | |
|---|---|
| Detailed description | This variable indicates the last appearance of the establishment number in the dataset (see Bender et. al. 1996).<br>If the existence of an establishment number in the BHP already ends before 2008, it could indicate the closure of the establishment. However, other possible reasons for this are an “arbitrary change of the establishment number following a change of owner or a change in the legal form of the establishment”, the “outsourcing of parts of the firm under a new establishment number” or other administrative changes (see Bender et. al. 1996 or Bundesagentur für Arbeit 2007, pp. 9-11). |

### 5.7.15 Last appearance of establishment number (lzt_dat)

| | |
|---|---|
| Variable label | last appearance |
| Variable name | lzt_dat |
| Category | establishment variables |
| Origin | BeH |
| Data type | date |
| Hierarchy | none |
| Detailed description | This variable indicates the last appearance of the establishment number in the dataset to the day (see Bender et. al. 1996). If the existence of an establishment number in the BHP already ends before 2008, it could indicate the closure of the establishment. However, other possible reasons for this are an “arbitrary change of the establishment number following a change of ownership or a change in the legal form of the establishment”, the “outsourcing of parts of the firm under a new establishment number” or other administrative changes (see Bender et al. 1996 or Bundesagentur für Arbeit 2007, pp. 9-11). |
| Notes | Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, only the year when the establishment number last appeared is shown (lzt_jahr). |

### 5.7.16 Total number of employees (az_ges)

| | |
|---|---|
| Variable label | no. employees |
| Variable name | az_ges |
| Category | generated establishment variables |
| Origin | BeH |
| Detailed description | This variable contains the total number of an establishment’s employees reported to the social security agencies as of 30 June of a year. Since the introduction of the new notification regulations in 1999, people in marginal part-time employment have also been recorded. Part-time employees and dormant employment relationships (daily wage of zero) are also included. |

### 5.7.17 Number of full-time employees (regular workers + others) (az_vz)

| | |
|---|---|
| Variable label | No. full-time (regular workers + others) |
| Variable name | az_vz |
| Category | generated establishment variables |

| | |
|---|---|
| Origin | BeH |
| Hierarchy | none |
| Detailed description | This variable contains the number of people in the establishment who are reported on 30 June of a year as full-time employees under the person group codes 101, 140, 143, 105, 106, 112, 118, 119, 120, 149, 201, 203, 205, 999 and YYY. Apprentices, marginally part-time employees and individuals participating in partial retirement schemes are not considered. |

### 5.7.18 Number of employees in marginal part-time employment (az_gf)

| | |
|---|---|
| Variable label | no. marginal part-time workers |
| Variable name | az_gf |
| Category | generated establishment variables |
| Origin | BeH |
| Hierarchy | none |
| Detailed description | The number of employees in marginal part-time employment is generated using the person group code – values 109 and 209. This variable has only been contained in the dataset since 1999 as it has only been included in the social security notification procedure since that year. |

### 5.7.19 Mean imputed wage all full-time employees (te_imp_mw)

| | |
|---|---|
| Variable label | mean imp. wage all full-time employees |
| Variable name | te_imp_mw |
| Category | generated establishment variables |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | none |
| Detailed description | This variable contains the mean imputed gross daily wage of the full-time employees in an establishment. It does not include the wages of marginally part-time staff, apprentices or individuals participating in partial retirement schemes.<br>The values are reported in euros for all years.<br>According to the social security notification regulations, employers must indicate the employee's gross wage subject to social security contributions for a certain period of time (fixed period wage). Until the end of 1998, employers had to report the gross wage subject to social security contributions only. So only wages above the marginal part-time income threshold and below the contribution assessment ceiling were recorded. Since 1999, wages below the marginal part-time income threshold have also been recorded as part of the new notification procedure. Gross wages above the contribution assessment ceiling, however, are still cut.<br>In order to calculate the gross daily wage, the fixed period wage is divided by the number of calendar days in the period. To calculate the mean, these censored wages were imputed (see Section 8.2 in Schmucker et al. 2016). These data were then aggregated at establishment level. |

### 5.7.20 Place of work: district (ao_kreis)

| | |
|---|---|
| Variable label | Place of work: district (Kreis) |
| Variable name | ao_kreis |
| Category | location data |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | federal state \| district |
| Detailed description | The variable indicates the district (urban district or rural district) in which the employee's establishment is located. The first two digits of the 5-digit district code (Kreisschlüssel) show the code for the federal state (Bundesland), positions 1-3 indicate the regional authority (Regierungsbezirk), and positions 1-5 show the district authority (Kreis). Federal states without a regional authority have a 0 in the third position.<br>In order to guarantee consistent regional allocations across the entire observation period, the information on the district was recoded to the territorial allocation of 31 December 2014, i.e. in all calendar years, a place of work is assigned to a district in accordance with the boundaries that the district had on 31 December 2014. As the district boundaries have changed over time, cases would occur in which the district code of the location of the establishment would change without the establishment concerned having relocated, if the territorial allocations were not updated. |
| Notes | Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, only the federal state (ao_bula) is shown as regional information. |

### 5.7.21 Place of work: federal state (Bundesland) (ao_bula)

| | |
|---|---|
| Variable label | Place of work: federal state (Bundesland) |
| Variable name | ao_bula |
| Category | location data |
| Origin | BeH |
| Data type | numerical |
| Hierarchy | federal state \| district |
| Detailed description | The variable indicates the federal state in which the establishment is located. This variable is generated from the district code (ao_kreis). The first two positions of the district code indicate the federal state. |

# 6 References


**Bender, Stefan / Hilzendegen, Jürgen / Rohwer, Götz / Rudolph, Helmut** (1996): Die IAB-Beschäftigtenstichprobe 1975-1990. Beiträge zur Arbeitsmarkt- und Berufsforschung 197, Nürnberg.

**Bertat, Thomas / Dundler, Agnes / Grimm, Christopher / Kiewitt, Jochen / Schomaker, Christine / Schridde, Dr. Henning / Zemann, Dr. Christian** (2013): Neue Erhebungsinhalte ‚Arbeitszeit', ‚ausgeübte Tätigkeit' sowie ‚Schul- und Berufsabschluss' in der Beschäftigungsstatistik. Methodenbericht, Bundesagentur für Arbeit – Statistik, URL: http://statistik.arbeitsagentur.de/Statischer-Content/Grundlagen/Methodenberichte/Beschaeftigungsstatistik/Generische-Publikationen/Methodenbericht-Neue-Erbebungsinhalte-Arbeitszeita-ausgeuebte-Taetigkeit-sowie-Schul-und-Berufsabschluss-in-der-Beschaeftigungsstatistik.pdf, (31 March 2016).

**Bundesagentur für Arbeit** (Hrsg.) (2005): Schlüsselverzeichnis für die Angaben zur Tätigkeit in den Meldungen zur Sozialversicherung. Ausgabe Januar 2005, Nürnberg.

**Bundesagentur für Arbeit** (Hrsg.) (2007): Handbuch für die Betriebsnummernvergabe und –pflege im Rahmen des Meldeverfahren zur Sozialversicherung. Ausgabe Dezember 2007, Nürnberg.

**Bundesagentur für Arbeit** (2009): Klassifikation der Wirtschaftszweige 1973, Nürnberg. URL: http://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen/Klassifikation-der-Wirtschaftszweige/Klassifikation-der-Wirtschaftszweige-1973-2003/Klassifikationen-der-Wirtschaftszweige-1973-2003-Nav.html, (21 April 2016).

**Bundesagentur für Arbeit** (2011): Klassifikation der Berufe 2010. Band 1: Systematischer und alphabetischer Teil mit Erläuterungen, Nürnberg. URL: http://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen/Klassifikation-der-Berufe/KldB2010/Printausgabe-KldB2010/Printausgabe-KldB-2010-Nav.html, (21 April 2016).

**Bundesanstalt für Arbeit** (1988): Klassifikation der Berufe – Systematisches und Alphabetisches Verzeichnis der Berufsbenennung, Nürnberg. URL: http://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen/Klassifikation-der-Berufe/KldB1975-1992/KldB1975-1992-Nav.html, (21 April 2016).

**Cramer, Ulrich** (1985): Probleme der Genauigkeit der Beschäftigtenstatistik. In: Allgemeines Statistisches Archiv 69: 56-68.

**Deutsche BKK** (2016): Ratgeber Sozialversicherung 2016, Wolfsburg, URL: https://www.deutschebkk.de/fileadmin/user_upload/microsites/arbeitgeber/medien/pdf/ratgeber-sozialversicherung-2016.pdf, (31 March 2016).

**Drews, Nils** (2006): Qualitätsverbesserung der Bildungsvariable in der IAB-Beschäftigtenstichprobe 1975-2001. FDZ Methodenreport, 05/2006 (de), Nürnberg.

**Eberle, Johanna/ Jacobebbinghaus, Peter/ Ludsteck, Johannes/ Witter, Julia** (2011): Generation of time-consistent industry codes in the face of classification changes * Simple heuristic based on the Establishment History Panel (BHP). FDZ Methodenreport, 05/2011 (en), Nürnberg.

**Fitzenberger, Bernd/ Osikominu, Aderonke/ Völter, Robert** (2006): Imputation rules to improve the education variable in the IAB employment subsample. In: Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften, Jg. 126, H. 3, 405-436.

**Ludsteck, Johannes/ Thomsen, Ulrich** (2016): Imputation of the Working Time Information for the Employment Register Data. FDZ Methodenreport 01/2016 (en), Nürnberg.

**Meinken, Holger/ Koch, Iris** (2004): BA-Beschäftigtenpanel 1998-2002. Codebuch, Nürnberg.

**Paulus, Wiebke/ Matthes, Britta** (2013): Klassifikation der Berufe * Struktur, Codierung und Umsteigeschlüssel. FDZ-Methodenreport, 08/2013 (de), Nürnberg.

**Schmucker, Alexandra; Seth, Stefan; Ludsteck, Johannes; Eberle, Johanna; Ganzer, Andreas** (2016): Establishment History Panel 1975-2014. FDZ-Datenreport, 03/2016 (en), Nürnberg.

**Statistisches Bundesamt** (2002): Klassifikation der Wirtschaftszweige, Ausgabe 1993 (WZ 93), Wiesbaden. URL: https://www.destatis.de/DE/Methoden/Klassifikationen/GueterWirtschaftklassifikationen/Content75/KlassifikationWZ93.html, (21 April 2016).

**Statistisches Bundesamt** (2003): Klassifikation der Wirtschaftszweige, Ausgabe 2003 (WZ 2003), Wiesbaden. URL: https://www.destatis.de/DE/Methoden/Klassifikationen/GueterWirtschaftklassifikationen/Content75/KlassifikationWZ2003.html, (21 April 2016).

**Statistisches Bundesamt** (2008): Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008), Wiesbaden. URL: https://www.destatis.de/DE/Methoden/Klassifikationen/GueterWirtschaftklassifikationen/Content75/KlassifikationWZ08.html, (21 April 2016).

**Wermter, Winfried/ Cramer, Ulrich** (1988): Wie hoch war der Beschäftigtenanstieg seit 1983? – Ein Diskussionsbeitrag aus der Sicht der Beschäftigtenstatistik der Bundesanstalt für Arbeit. In: Mitteilungen aus der Arbeitsmarkt – und Berufsforschung 4/88, 468-482.

# 7 Appendix

## 7.1 Frequency tables

Frequency tables and overviews of the individual values and labels of the variables can be found in separate files at http://fdz.iab.de/en.aspx.

# 8 List of abbreviations

| | | |
|---|---|---|
| AA | Agentur für Arbeit / Arbeitsamt | employment agency / employment office |
| ALG | Arbeitslosengeld | unemployment benefit |
| ARGE | Arbeitsgemeinschaft | cooperation of employment agencies and municipalities |
| ASU | Arbeitsuchende-Historik | Jobseeker History |
| A2LL | Arbeitslosengeld II – Leistungen zum Lebensunterhalt | unemployment benefit II - benefits to secure a livelihood |
| BA | Bundesagentur für Arbeit | Federal Employment Agency |
| BeH | Beschäftigten-Historik | Employee History |
| BfA | Bundesversicherungsanstalt für Angestellte | Federal Social Insurance Office for Salaried Employees |
| BHP | Betriebs-Historik-Panel | Establishment History Panel |
| BMAS | Bundesministerium für Arbeit und Soziales | Federal Ministry of Labour and Social Affairs |
| coArb | Computerunterstützte Arbeitsvermittlung (operatives Verfahren zur Verwaltung der Vermittlung (Altverfahren)) | computer-aided job placement (procedure for the administration of job placements – old procedure) |
| DEÜV | Verordnung über die Erfassung und Übermittlung von Daten für die Träger der Sozialversicherung – Datenerfassungs- und –übermittlungsverordnung | Data Collection and Transmission Regulation - regulation on the collection and transmission of data for the social security agencies |
| DEVO | Zweite VO über die Erfassung von Daten für die Träger der Sozialversicherung und für die BA – Datenerfassungs-Verordnung – | Data Collection Regulation - second regulation on the collection of data for the social security agencies and for the Federal Employment Agency |
| DÜVO | Zweite VO über die Datenübermittlung auf maschinell verwertbaren Datenträgern im Bereich der Sozialversicherung und der BA – Datenübermittlungs-Verordnung – | Data Transmission Regulation - second regulation on the transfer of data on machine-readable data media in the field of social security and the BA |
| EDV | Elektronische Datenverarbeitung | Electronic data processing |
| FDZ | Forschungsdatenzentrum | Research Data Centre |
| FELEG | Gesetz zur Förderung der Einstellung der landwirtschaftlichen Erwerbstätigkeit | Act on the Support in Case of Termination of Farming Activities |
| gAw | Träger mit getrennter Aufgabenwahrnehmung | Municipalities exercising their duties separately |
| gE | Gemeinsame Einrichtung | Joint facility |
| gT | Getrennte Trägerschaft | Separated responsibilities |

| | | |
|---|---|---|
| IAB | Institut für Arbeitsmarkt- und Berufsforschung | Institute for Employment Research |
| IABS | IAB-Beschäftigtenstichprobe | IAB Employment Samples |
| IEB | Integrierte Erwerbsbiographien | Integrated Employment Biographies |
| IEBS | Stichprobe der Integrierten Erwerbsbiographien | Integrated Employment Biographies Sample |
| ISIC | International Standard Industrial Classification of All Economic Activities | International Standard Industrial Classification of All Economic Activities |
| ITM | IT- und Informationsmanagement des Instituts für Arbeitsmarkt- und Berufsforschung | IT and information management of the Institute for Employment Research |
| LIAB | Linked-Employer-Employee-Daten des IAB | Linked employer-employee data of the IAB |
| LeH | Leistungsempfänger-Historik | Benefit Recipient History |
| LHG | Leistungs-Historik Grundsicherung | Unemployment Benefit II Recipient History |
| LVA | Landesversicherungsanstalt | Land Social Insurance Office |
| MTH | Maßnahmeteilnehmer-Historik | Participants-in-Measures History File |
| NACE | Nomenclature génerale des activités économiques dans les communautés européennes | Nomenclature génerale des activités économiques dans les communautés européennes |
| SchwbG | Gesetz zur Sicherung der Eingliederung Schwerbehinderter in Arbeit, Beruf und Gesellschaft – Schwerbehindertengesetz- | law to guarantee the integration of persons with severe disabilities into employment and society – Severely Disabled Persons Act |
| SGB | Sozialgesetzbuch | German Social Code |
| SIAB | Stichprobe der Integrierten Arbeitsmarktbiographien | Sample of Integrated Labour Market Biographies |
| VerBIS | Vermittlungs- und Beratungsinformationssystems | Information System for Placement and Counselling |
| XASU | Arbeitsuchenden-Historik aus XSozial-BA-SGB II | Jobseeker History from XSozial-BA-SGB II |
| zkT | Zugelassener kommunaler Träger | Authorised municipalities |