RESEARCH DATA CENTRE (FDZ)
of the German Federal Employment Agency (BA)
at the Institute for Employment Research (IAB)

FDZ-DATENREPORT

Documentation of labour market data

06|2019 EN Linked Employer-Employee Data from the IAB: LIAB Cross-Sectional Model 2 (LIAB QM2) 1993-2017

Lisa Schmidtlein, Stefan Seth, Matthias Umkehrer

Bundesagentur für Arbeit

Linked Employer-Employee Data from the IAB: LIAB Cross-Sectional Model 2 (LIAB QM2) 1993-2017

Lisa Schmidtlein, Stefan Seth, Matthias Umkehrer
(Institute for Employment Research (IAB))

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 reports can be used to prepare evaluations.

Contents

Abstract

This data report describes the LIAB Cross-Sectional Model 2 1993 – 2017 (LIAB QM2 9317).

Zusammenfassung

Dieser Datenreport beschreibt das LIAB-Querschnittmodell 2 1993 – 2017 (LIAB QM2 9317).

Keywords

German administrative micro data, labour market data, linked employer-employee data, data manual

Acknowledgements

We would like to thank our colleagues in the Research Data Centre (FDZ), the department Establishments and Employment (D1) as well as the department Data and IT-Management (DIM) of the Institute for Employment Research for their cooperation and support. This Datenreport is a revision of the FDZ-Datenreport 05/2016 (authors: Klosterhuber, Lehnert, Seth). Additionally, individual passages were adopted from IAB-internal data documentations by IAB-DIM.

Data availability

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

1 Introduction and outline

1.1 Introduction

The LIAB Cross-Sectional Model 2 (QM2) 9317 is one of the linked employer-employee datasets from the Institute for Employment Research (IAB) provided for research purposes by the Research Data Centre (FDZ) of the Federal Employment Agency (BA) at the IAB. It links information on establishments from the IAB Establishment Panel, an annual establishment survey, with information on individuals employed at those establishments from the process-generated data of the BA.

The IAB Establishment Panel is an annual representative survey on various topics such as the determinants of labour demand. It has been conducted by the IAB since 1993 in West Germany and since 1996 in East Germany. The IAB Establishment Panel is the central basis for the analysis of labour demand in Germany. **This data report focuses on the description of data on individuals in the LIAB QM2 9317. More detailed information on the IAB Establishment Panel can be found on the FDZ's websites** (https://fdz.iab.de/en/FDZ_Establishment_Data/IAB_Establishment_Panel.aspx).

Data on individuals come from the Integrated Employment Biographies (IEB) of the IAB. The IEB comprise all individuals showing 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 jobseeker (recorded from 1997 onwards)
- participation in an employment or training measure (recorded from 2000 onwards)

These data originate from different sources. They are merged in the IEB and depict the statuses 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 History (Leistungshistoriken Grundsicherung - LHG and XLHG). The Jobseeker Histories (Arbeitsuchendenhistoriken - 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 Participation-in-Measures History File (Maßnahmeteilnahmehistoriken – MTH and XMTH). **Data on individuals in the LIAB QM2 9317 include information from all of these sources**.

The establishments surveyed in the IAB Establishment Panel build the basis for sampling the individual data from the IEB (see Chapter 3 for more details). In a first step, all of the establishments in the IAB Establishment Panel with a valid interview in the respective year (1993 to 2017) are selected. In a second step, all observations of individuals are drawn from the IEB that have been

employed at one of these establishments on June 30 of the respective survey year for at least one day. For these individuals, all observations are made available which contain June 30 of the respective survey year in a third step. Consequently, the LIAB QM2 9317 includes the survey data of all establishments surveyed in the IAB Establishment Panel between 1993 and 2017. Additionally, all workers of an establishment can be observed in the individual data on June 30 of a survey year. Furthermore, it is possible to merge administrative establishment-level information from the BHP for all establishments covered by the LIAB QM2 9317 (see https://fdz.iab.de/en/FDZ_Establishment_Data/Establishment_History_Panel.aspx).

The LIAB QM2 9317 is produced at the Research Data Centre (FDZ) of the Federal Employment Agency (BA) at the IAB. The data set covers information on 12,451,266 individuals whose labor market biographies are documented in a total of 56,150,704 lines of data. This Datenreport describes the different variables, which are largely based on the original and uncoarsened data. In order to protect the anonymity of the data subjects, some variables are classified as particularly sensitive and are only made available upon special application (see Section 1.2).

Other than the longitudinal model of the LIAB, the cross-sectional model includes employment biographies as of the reference date June 30 only. However, for all establishments surveyed in the IAB Establishment Panel the employment biographies of all individuals employed at one of these establishments on June 30 of a year in which a survey took place are included. The longitudinal model, in turn, contains comprehensive employment biographies of individuals of a subsample of establishments repeatedly interviewed in the IAB Establishment Panel. Therefore, employment biographies can still be tracked in the events of e. g. establishment changes or transitions into unemployment within this subgroup. **An overview of the different LIAB models can be found on the FDZ's websites** (https://fdz.iab.de/en/Integrated_Establishment_and_Individual_Data/LIAB.aspx).

This Datenreport 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 different data sources can be found in Section 2. Sections 3 and 4 discuss data preparation and data quality, whilst Section 5 describes the individual variables.

1.2 Data use

1.2.1 Data access

The LIAB data are weakly anonymous data and therefore may only be analysed on-site at the FDZ and via subsequent remote data execution.

In order to be able to use the data, it is necessary to submit an application to the FDZ. The FDZ decides on the approval of the research project on behalf of and, if necessary, in coordination with the Federal Ministry of Labour and Social Affairs (Bundesministerium für Arbeit und Soziales – BMAS). When approval has been granted, a data use agreement is concluded with the researcher's scientific institution. Details on applying for the dataset and possibilities for data processing are available on the FDZ's website.

1.2.2 Data management

The LIAB data have a modular structure and include labels both in German and English language.[1] The LIAB data are stored in several files (see Figure 1). A first module contains all waves of the IAB Establishment Panel in separate datasets. These waves include year-specific information on employment trends, business policies and development, company investments, innovations in business, public funding, personnel structure, vocational trainings and training posts, worker inflow and outflow, personnel search, wages and salaries, working hours, further trainings or general data of the establishment.

The individual data are stored in one dataset per year in a second module. These files contain identifiers (for both individuals and establishments), personal variables, information on employment, benefit receipt and job-search activity, variables regarding place of residence as well as technical and biographical variables.

A third module, the Basis Establishment File, contains the establishment number, the year, and variables regarding the place of work and economic activity as well as other establishment information as of the reference date June 30. The Basis Establishment File is extracted from the Establishment History Panel (BHP). The BHP comprises all establishments in Germany with at least one employee liable to social security on June 30 of a year. Since 1999, marginal part-time employees are included in this definition, too. The source of the BHP is the Employee History (BeH) of the IAB. In the BHP, individual data of the BeH is aggregated at the establishment-year level via the establishment numbers. Further information on the BHP can be found on the FDZ's websites (https://fdz.iab.de/en/FDZ_Establishment_Data/Establishment_History_Panel.aspx). Variables marked with an '*' in the list of variables (see Table 4), are included in the Basis Establishment File. Upon reasoned request, further establishment characteristics in variable blocks from the BHP can be provided (see Figure 1). In addition, the extension files Worker flows (inflows/outflows) and Establishment dynamics (entries/exits) can be applied for with a justified application.

Establishment variables and individual variables are organised separately, which makes the structure of the data clear and saves storage space. There are different program-specific commands for data formatting and analysis to link the files. For example, linkages in Stata can be conducted via the '`merge`' command. The linkage between the waves of the IAB Establishment Panel and the Individual Files is based on the Panel's establishment numbers, which are included in both modules (*idnum*; see Box 1). In contrast, the Basis Establishment File and the Individual Files are linked by an artificial establishment number (*betnr*) and year (*jahr*) (see Box 2). The same applies for the extension files of the BHP.

[1] With the Stata commands `label language de` or `label language en` labels can be switched to English or German, respectively.

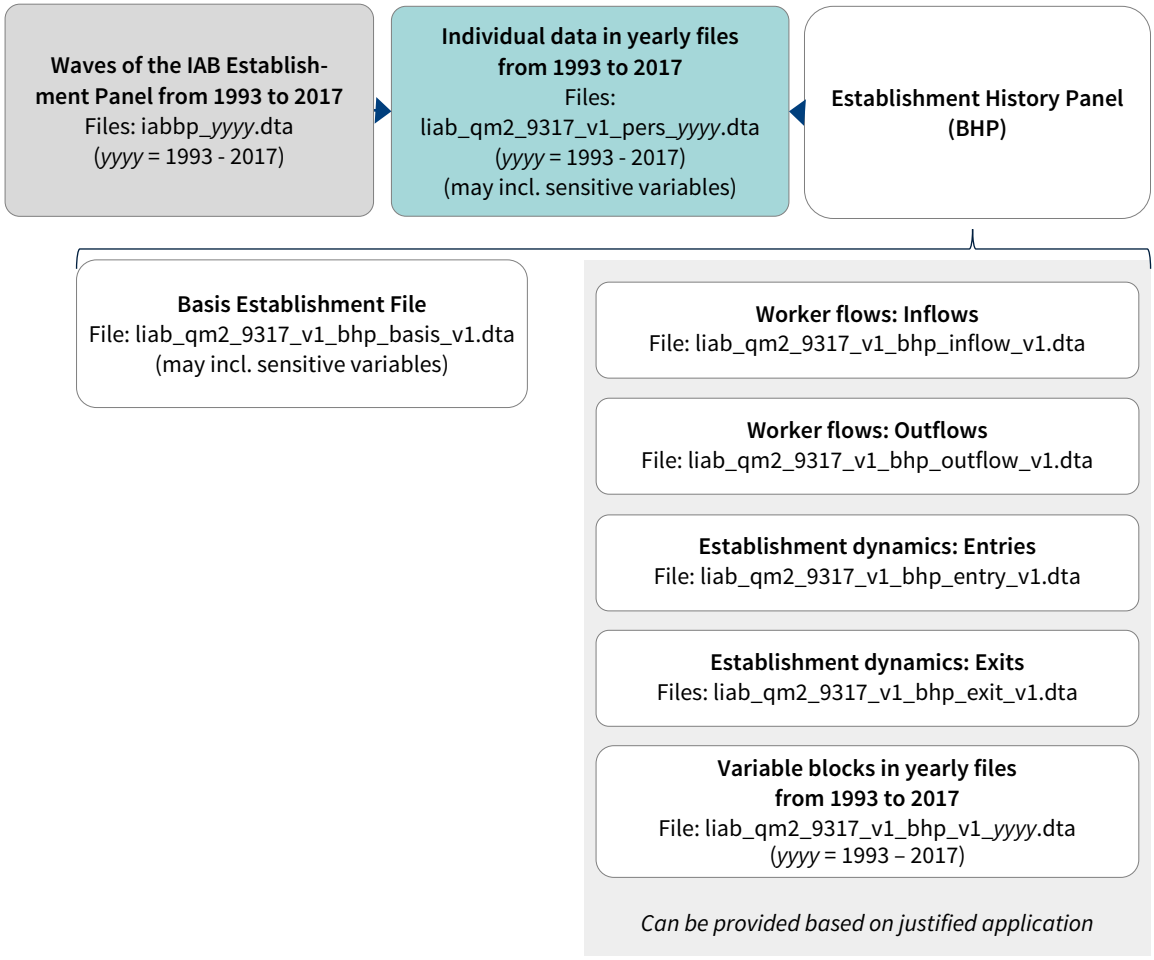

Figure 1: Data management of the Linked Employer-Employee Data in the LIAB QM2 9317

```
use liab_qm2_9317_v1_pers_2000.dta
sort idnum
merge m:1 idnum using iabbp_2000.dta
```

Box 1: Example for Stata 14; Linkage between an individual file and a single wave of the IAB Establishment Panel

```
use liab_qm2_9317_v1_pers_2000.dta
sort betnr jahr
merge m:1 betnr jahr using LIAB_qm2_9317_v1_bhp_basis_v1.dta
```

Box 2: Example for Stata 14; Linkage between an Individual file and the Basis Establishment File

Certain variables, which make identification of individuals or establishments easier, 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 following variables are particularly sensitive from a data protection point of view:

Individual Files:

- nationality (*nation*)
- month of birth (*gebmon*)
- 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 establishment number (*grd_dat*)
- date of last appearance establishment number (*lzt_dat*)

All sensitive variables listed in this section are located in the corresponding files, as long as the files have been requested and the request has been approved (see Figure 1).

1.2.3 File names of the LIAB QM2 9317

Waves of the IAB Establishment Panel

iabbp_*yyyy*.dta, *yyyy* = 1993 – 2017

Individual Files

liab_qm2_9317_v1_pers_*yyyy*.dta, *yyyy* = 1993 – 2017

Basis Establishment File

liab_qm2_9317_v1_bhp_basis_v1.dta

Extension Files BHP

Stocks with variable blocks

liab_qm2_9317_v1_bhp_v1_*yyyy*.dta, *yyyy* = 1993 – 2017

Worker flows

liab_qm2_9317_v1_bhp_inflow_v1.dta

liab_qm2_9317_v1_bhp_outflow_v1.dta

Establishment dynamics

liab_qm2_9317_v1_bhp_entry_v1.dta

liab_qm2_9317_v1_bhp_exit_v1.dta

1.3 Changes as compared to LIAB QM2 9314

1.3.1 Observation period

The LIAB QM2 9317 differs from the LIAB QM2 9314 data in terms of time periods covered by both the individual and establishment data.

1.3.2 Set of variables

The LIAB QM2 9317 offers a number of additional variables as compared to the LIAB QM2 9314. The variable 'Occupational status and working hours' (*stib*), which was removed in the previous version, is included again in the LIAB QM2 9317. However, the restriction applies that the variable is only filled for notifications that were submitted before the introduction of the new occupation code (see Section 4.2). In addition, the variable 'Vocational training (imputed)' (*ausbildung_imp*, see Section 5.3.9) has been added. This variable contains imputed data for observations with originally missing information on the level of vocational training. Additionally, the LIAB QM2 9317 now also contains the sensitive variable 'Month of birth' (*gebmon*, see Section 5.3.3) as separate variable. Furthermore, a time-consistent version of the WZ08 industry classification was generated by extrapolation and imputation. Further information can be found in Sections 5.6.12 and 5.6.13.

The code for generating the following biographical variables was slightly adjusted:

- First day in employment (*ein_erw*)
- Number of days in employment (*tage_erw*)
- First day in establishment (*ein_bet*)
- Number of days in establishment (*tage_bet*)
- First day in job (*ein_job*)
- Numbers of days in job (*tage_job*)
- Number of days of benefit receipt (*tage_lst*)

Adjustment concerns corrections of how one-time payments are considered (BeH spells with reason for notification of '54'). Eberle and Schmucker (2019) offer the corresponding revised programmes for Stata.

Table 1 provides an overview of the changes in the set of variables across all sources.

Table 1: Changes in the set of variables

Variable	Type of modification	Explanation
Vocational training (imputed) (*ausbildung_imp*)	N	Newly included, yet with fewer categories than the variable 'Vocational training' (*ausbildung*)
Month of birth (*gebmon*)	N	Sensitive variable
Occupational status and working hours (*stib*)	N	Included again, yet only filled for notifications submitted before the introduction of the new occupation code
w08_3 completed by extrapolation/imputation (*w08_3_gen*)	N	Newly included in Basis Establishment File
Art der Vervollständigung w08_3 (*group_w08_3*)	N	Newly included in Basis Establishment File

N = new, D = dropped

1.4 Outline

Table 2: Outline

Topics/ groups of variables	**Employee History (BeH)**: 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 **Benefit Recipient History (LeH)**: Information on benefit receipt in accordance with Social Code Book III (SGB III) for recipients of unemployment benefit, unemployment assistance and maintenance allowance **Unemployment Benefit II Recipient History (Leistungshistorik Grundsicherung - LHG)**: 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) **Jobseeker History (Arbeitsuchendenhistorik - ASU)**: Information on job search activities that are recorded in BA procedures **Jobseeker History from XSozial-BA-SGB II (Arbeitsuchendenhistorik aus XSozial-BA-SGB II - XASU)**: Information on job search activity reported to the BA by authorised municipalities via the transmission standard XSozial-BA-SGB II **Participants-in-Measures History File (Maßnahmeteilnahmehistorik - MTH)**: Information on participation in employment and training measures (not including measures of authorised municipalities) **Participants-in-Measures History File from XSozial-BA-SGB II (Maßnahmeteilnahmehistorik aus XSozial-BA-SGB II - XMTH)**:

	Information on participation in employment and training measures reported to the BA by authorised municipalities via the transmission standard XSozial-BA-SGB II **IAB Establishment Panel**: Information from the annual Establishment Panel survey with shifting priorities.
Data units	Establishments, Employees covered by social security (including marginal part-time employees from 1999 onwards), benefit recipients, jobseekers, participants in measures
Number of cases	Linked establishments: 3,933 to 15,061 per year, 67,407 in total Individuals: 1,469,218 to 2,584,415 per year, 12,451,266 in total Number of intervals: 56,150,704
Period covered	The period covered depends on the data source. BeH January 1, 1993 – December 31, 2017 (2014: 36-Months-File 2015 and 2016: 18-Months-File, 2017: 6-Months-File) LeH January 1, 1993 – December 31, 2017 LHG January 1, 2005 – December 31, 2017 ASU January 1, 1997 – December 31, 2017 XASU January 1, 2005 – December 31, 2017 MTH January 1, 2000 – December 31, 2017 XMTH January 1, 2005 – December 31, 2017
Time reference	Establishments: referenced to June 30 Individuals: referenced to June 30
Regional structure	German federal states (Bundesländer), districts (Kreise)
Date of territorial allocation	Territorial allocation updated as of December 31, 2017
Data collection method	Linked employer-employee data; selection of all establishments from establishment survey and linkage with their employees' employment biographies from process data as of June 30 of a year in which a survey took place.
Institutions involved	Social security agencies, Federal Employment Agency (Bundesagentur für Arbeit), Kantar Public (formerly TNS Infratest Sozialforschung GmbH) in cooperation with SÖSTRA (Institute for Socio-economic Structural Analysis).
Frequency of data collection	Administrative data: continuous IAB Establishment Panel: annually
File format and size	Stata Establishment Files (without sensitive variables): IAB Establishment Panel: about 1.7 to 11 MB per year Basis Establishment File: about 108 MB Individual files (without sensitive variables): about 111 MB to 187 MB per year

File architecture	The data on individuals are stored in one file per year. Another file with establishment information taken from the BHP (the Basis Establishment File) is stored in long format and can be linked to the data files on individuals by using the variables *betnr* and *year*. Upon reasoned request, further information on establishments from the BHP can be provided. The IAB Establishment Panel files remain unchanged as one file per wave and can be linked to the data files on individuals by using the variable *idnum*.
Data access	On-site access and subsequent remote data execution
Degree of anonymisation	Weakly anonymous
Sensitive variables	Individual files: Nationality (*nation*), Month of birth (*gebmon*), Occupational sub-group (*beruf2010_4*), Place of residence – district (*wo_kreis*), Place of residence – employment agency area (*wo_aa*) Basis Establishment File: Place of work – district (*ao_kreis*), Economic activity - sub-class of economic activity (5-digit code) (*w93_5*), Economic activity - sub-class of economic activity (5-digit code) (*w03_5*), Economic activity - sub-class of economic activity (5-digit code) (*w08_5*), First appearance of establishment (*grd_dat*), Last appearance of establishment (*lzt_dat*)
Citation of data and data documentation	Dataset: Linked Employer-Employee Data from the IAB, Cross-Sectional Model 2 1993-2017 (LIAB QM2 9317), Nuremberg 2019 Data: 'The data basis of this paper is the Cross-Sectional Model 2 1993 – 2017 of the Linked Employer-Employee Data from the IAB. The data were accessed on-site at the Research Data Centre of the Federal Employment Agency at the Institute for Employment Research (FDZ) and via remote data execution at the FDZ.' DOI: 10.5164/IAB.LIABQM29317.de.en.v1 Data documentation: Schmidtlein, Lisa; Seth, Stefan; Umkehrer, Matthias (2019): Linked Employer-Employee Data from the IAB: LIAB Cross-Sectional Model 2 (LIAB QM2) 1993 – 2017. FDZ-Datenreport, 06/2019 (en), Nuremberg. DOI: 10.5164/IAB.FDZD.1906.en.v1
Dataset version	Linked Employer-Employee Data from the IAB, Cross-Sectional Model 2 1993 – 2017 (LIAB QM2 9317) – Version 9317 v1; DOI: 10.5164/IAB.LIABQM29317.de.en.v1

1.5 List of variables

The overview of variables in Table 4 lists the variable names and the longer descriptions of variables. It also provides an overview of whether and how well variables are filled in the individual data sources. Table 3 illustrates the meaning of the shading, which indicates the degree of completeness per variable and source in Table 4.

Example: The variable 'Daily wage/daily benefit' is only available for BeH and LeH observations; the observations of the other data sources contain the missing value ".n" for this variable.

Table 3: Degrees of completeness of variables

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

Another characteristic is that some variables have different contents depending on the data source. For instance, for BeH observations the 'Employment status' (*erwstat*) 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 and XMTH observations it is the measure category. These differences are not immediately obvious from the variable name for every variable.

Table 4: List of variables with degree of completeness

List of variables	In Basis Estab. File	p.	BeH	LeH	LHG	ASU	X ASU	MTH	X MTH
Identifiers		37							
Individual ID (*persnr*)		37	**h**	**h**	**h**	**h**	**h**	**h**	**h**
Establishment ID (*betnr*)	*	38	**h**	**l**	**l**	**l**	**l**	**l**	**l**
Establishment-ID acc. to Survey (*idnum*)		39	**h**	**l**	**l**	**l**	**l**	**l**	**l**
Generated technical variables		39							
Source of spell (*quelle*)		39	**h**	**h**	**h**	**h**	**h**	**h**	**h**
Year (*jahr*)	*	39	**h**	**h**	**h**	**h**	**h**	**h**	**h**
Status of establishment number (*betr_st*)		39	**h**	**l**	**l**	**l**	**l**	**l**	**l**
Personal information		40							
Gender (*frau*)		40	**h**	**h**	**h**	**h**	**h**	**h**	**h**
Year of birth (*gebjahr*)		40	**h**	**h**	**h**	**h**	**h**	**h**	**h**
Month of birth (*gebmon*)		40	**h**	**h**	**h**	**h**	**h**	**h**	**h**
Nationality (*nation*)		41	**h**	**h**	**h**	**h**	**h**	**h**	**h**
Nationality, grouped (*nation_gr*)		41	**h**	**h**	**h**	**h**	**h**	**h**	**h**
Marital status (*famst*)		41	**l**	**h**	**h**	**h**	**h**	**m**	**h**
Number of children (*kind*)		42	**l**	**h**	**h**	**m**	**m**	**m**	**m**
Vocational training (*ausbildung*)		42	**m**	**l**	**l**	**h**	**m**	**h**	**m**

Vocational training (imputed) (*ausbildung_imp*)		45	h	l	l	l	l	l	l
School leaving qualification (*schule*)		45	m	l	l	h	m	h	m
Information on employment, benefit receipt and job search		46							
Daily wage/daily benefit (*tentgelt*)		46	h	h	l	l	l	l	l
Occupation - current/most recent (KldB 1988) (*beruf*)		48	h	l	l	m	l	m	l
Occupational group - current/most recent (KldB 2010), 3-digit (*beruf2010_3*)		48	h	l	l	m	m	m	m
Occupational sub-group - current/most recent (KldB 2010), 4-digit (*beruf2010_4*)		49	h	l	l	m	m	m	m
Level of requirement - current/most recent job (KldB 2010) (*niveau*)		50	h	l	l	m	m	m	m
Part-time (*teilzeit*)		51	h	l	l	l	l	l	l
Occupational status and working hours (*stib*)		51	h	l	l	l	l	l	l
Employment status (*erwstat*)		52	h	h	h	h	h	h	h
Transition zone (*gleitz*)		54	h	l	l	l	l	l	l
Temporary agency work (*leih*)		54	h	l	l	l	l	l	l
Fixed-term contract (*befrist*)		55	h	l	l	l	l	l	l
Reason of cancellation/notification/termination (*grund*)		55	h	h	m	h	m	l	l
Client profile (*profil*)		56	l	l	l	m	l	m	l
Reason for end of previous employment (*art_kuend*)		57	l	l	l	m	l	m	l
Working hours of job application (*arbzeit*)		57	l	l	l	m	l	m	l
Residual claim/planned duration (*restanspruch*)		57	l	h	l	l	l	h	l
Type of provider (*traeger*)		58	l	l	h	h	h	h	h
Location data		59							
Place of residence - district (Kreis) (*wo_kreis*)		59	h	h	h	h	h	h	h
Place of residence - federal state (Bundesland) (*wo_bula*)		59	h	h	h	h	h	h	h
Place of residence - employment agency (Arbeitsagentur) (*wo_aa*)		60	h	h	h	h	h	h	h

Place of residence - regional directorate (Regionaldirektion) (*wo_rd*)		60	h	h	h	h	h	h	h
Establishment variables		61							
Classification of economic activities 73 (*w73_3*)	*	61	h	l	l	l	l	l	l
Classification of economic activities 93, sub-classes (*w93_5*)	*	61	h	l	l	l	l	l	l
Classification of economic activities 93, groups (*w93_3*)	*	62	h	l	l	l	l	l	l
Classification of economic activities 03, sub-classes (*w03_5*)	*	62	h	l	l	l	l	l	l
Classification of economic activities 03, groups (*w03_3*)	*	63	h	l	l	l	l	l	l
Classification of economic activities 08, sub-classes (*w08_5*)	*	63	h	l	l	l	l	l	l
Classification of economic activities 08, groups (*w08_3*)	*	63	h	l	l	l	l	l	l
w73_3 completed by extrapolation/imputation (*w73_3_gen*)	*	64	h	l	l	l	l	l	l
Type of imputation w73_3 (*group_w73_3*)	*	64	h	l	l	l	l	l	l
w93_3 completed by extrapolation/imputation (*w93_3_gen*)	*	64	h	l	l	l	l	l	l
Type of imputation w93_3 (*group_w93_3*)	*	65	h	l	l	l	l	l	l
w08_3 completed by extrapolation/imputation (*w08_3_gen*)	*	65	h	l	l	l	l	l	l
Type of imputation w08_3 (*group_w08_3*)	*	65	h	l	l	l	l	l	l
Year of first appearance (*grd_jahr*)	*	66	h	l	l	l	l	l	l
First appearance (*grd_dat*)	*	66	h	l	l	l	l	l	l
Year of last appearance (*lzt_jahr*)	*	67	h	l	l	l	l	l	l
Last appearance (*lzt_dat*)	*	67	h	l	l	l	l	l	l
Total number of employees (*az_ges*)	*	67	h	l	l	l	l	l	l
Number of full-time employees (regular workers + others) (*az_vz*)	*	68	h	l	l	l	l	l	l

Number of employees in marginal part-time employment (*az_gf*)	*	68	h	l	l	l	l	l	l
Mean imputed wage all full-time employees (*te_imp_mw*)	*	68	h	l	l	l	l	l	l
Place of work - district (Kreis) (*ao_kreis*)	*	69	h	l	l	l	l	l	l
Place of work - federal state (Bundesland) (*ao_bula*)	*	69	h	l	l	l	l	l	l
Generated biographical variables		70							
First day in employment (*ein_erw*)		70	h	h	h	h	h	h	h
Number of days in employment (*tage_erw*)		70	h	h	h	h	h	h	h
First day in establishment (*ein_bet*)		70	h	l	l	l	l	l	l
Number of days in establishment (*tage_bet*)		71	h	l	l	l	l	l	l
First day in job (*ein_job*)		71	h	l	l	l	l	l	l
Number of days in job (*tage_job*)		72	h	l	l	l	l	l	l
Number of days with benefit receipt (*tage_lst*)		73	h	h	h	h	h	h	h

1.6 Volume structure

The second column in Table 5 shows the number of establishments with a valid interview in the IAB Establishment Panel. Not all of the surveyed establishments can be linked to the individual data.[2] The number of establishments, which can be linked to the individual data, is listed in the third column. The number of individuals per year is presented in the fourth column. It should be taken into account that the number of data lines in the individual data is higher than the explicit number of individuals. Parallel employment contracts, benefit receipt and job search are documented respectively in separate lines of data. The last column refers to the number of different artificial establishment numbers in the individual data. Because of parallel employment spells in establishments not participating in the IAB Establishment Panel, these numbers are much higher than the numbers in the second and third columns.

It is apparent that the number of establishments has quadrupled whereas the number of employees remained roughly constant. This pattern is caused by a decline in the fraction of large establishments in the IAB Establishment Panel.

[2] For instance, there is no consolidation of establishment numbers in the IAB Establishment Panel in case of repeated interviews. Therefore, if the establishment number changes due to e.g. a change in ownership this will not be captured by the interview. From the respective wave onwards, no or incorrect individual data will be merged.

Table 5: Volume structure

Year	Establishments surveyed in IAB Establishment Panel	Establishments surveyed in IAB Establishment Panel and linkage to individual data	Individuals	Establishments in LIAB QM2 9317 overall
(1)	(2)	(3)	(4)	(5)
1993	4,265	4,188	2,584,415	10,742
1994	4,154	4,041	2,223,867	9,259
1995	4,114	3,933	1,884,526	8,985
1996	8,604	8,292	2,472,562	13,577
1997	8,917	8,385	2,121,115	12,979
1998	9,334	8,787	2,080,445	13,778
1999	9,915	9,433	2,019,135	54,124
2000	14,083	13,462	2,343,731	69,424
2001	15,782	14,981	2,536,432	73,523
2002	15,682	14,698	2,396,807	67,113
2003	16,165	15,061	2,184,581	74,190
2004	16,063	14,832	2,385,610	98,720
2005	16,280	14,870	2,396,903	105,490
2006	15,976	14,460	2,229,543	104,292
2007	16,181	14,590	2,016,979	101,300
2008	16,058	14,388	1,961,701	102,519
2009	16,145	14,414	1,909,679	105,651
2010	16,296	14,515	1,640,566	96,685
2011	15,967	14,171	1,763,918	102,783
2012	16,270	14,470	1,852,233	110,851
2013	16,495	14,676	1,785,162	110,461
2014	16,385	14,572	1,598,457	102,328
2015	16,358	14,539	1,576,906	100,748
2016	16,214	14,341	1,516,222	99,769
2017	16,255	14,314	1,469,218	96,648

2 Data sources

The administrative individual data were drawn from the Integrated Employment Biographies (IEB) of the IAB. The IEB combine data from 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 LIAB QM2 and their relationship to other FDZ data products.

2.1 Employee History (BeH)

The source of data regarding employment is the Employee History (Beschäftigtenhistorik - BeH) of the IAB. The data basis is the integrated notification procedure for health, pension and unemployment insurance, which came into effect as of January 1, 1973 (and was extended to cover East Germany as of January 1, 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 not recorded in the BeH in principle. As the notification procedure was modified on January 1, 1999, employees in marginal part-time employment and unpaid family workers have also been recorded (not contained in the data until April 1, 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 with administrative establishment data (Basis Establishment File and Extension Files). 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 the Basis Establishment File as well as in the BHP Extension Files are aggregated as of June 30 of a year. Unlike the data on individuals, the establishment variables are therefore not spell data but are only valid on June 30 precisely each year.[4]

The Basis Establishment File is linked with the Individual Files 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 2).

[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 January 1, 2006 to May 30, 2006; the establishment goes bankrupt in June 2006. There is then no information on this establishment in the BHP for 2006.

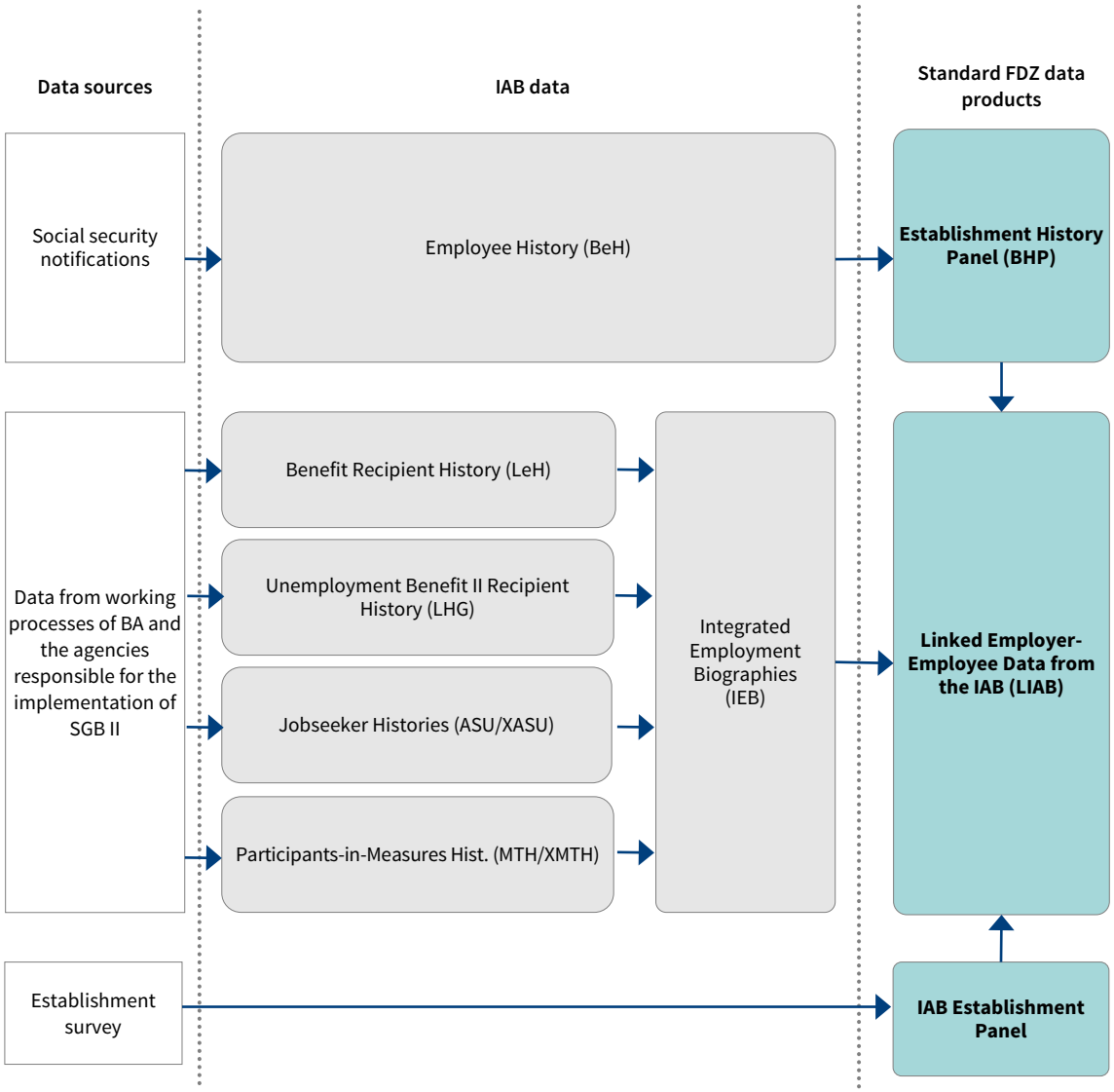

Figure 2: Data sources of the LIAB

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, maintenance allowance, and contributions paid by the BA to private health or care insurance while benefits are being drawn. Benefits in the context of Social Code Book II (e.g., unemployment benefit II) are not included in the data. As 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 LeH. The earliest available data in the LeH are from January 1, 1975.

2.3 Unemployment Benefit II Recipient History (LHG)

The Unemployment Benefit II Recipient History (Leistungshistorik Grundsicherung - LHG) contains information on individuals who are eligible for benefits and capable of work, on the members of their benefit unit (Bedarfsgemeinschaft) in accordance with Section 7 SGB II and on certain individuals associated with the benefit unit. In the LIAB, however, it is not possible to link individuals with benefit receipt in accordance with Social Code Book II (SGB II) within benefit units. 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 unit in the case of Unemployment Benefit II, it is difficult to assign an individual benefit rate. Therefore, the LIAB also does not contain information on SGB-II-benefit rates.

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 - gt) / municipalities exercising their duties separately (until 2011) – here the tasks are divided between the BA and the municipality,[5]
- 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 once a month. 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 January 1, 2005. However, the data source is incomplete until the beginning of 2007 (see Section 4).

2.4 Jobseeker Histories (ASU/XASU)

Data about jobseekers are stored in the Jobseeker Histories (Arbeitsuchendenhistoriken – ASU/XASU). The ASU data source contains information on jobseekers who are registered with em-

[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, 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.

ployment 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. The earliest available data in the ASU are from January 1, 1997, and in the XASU from January 1, 2005, respectively.

2.5 Participants-In-Measures History Files (MTH/XMTH)

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. Measures implemented by authorised municipalities or opting local authorities are recorded in the XMTH from 2005 onwards. The earliest available data in the MTH are from January 1, 2000, those in the XMTH are from January 1, 2005.

3 Data preparation and sampling procedure

3.1 Corrections and validation procedures

Before the data from the data sources specified in Chapter 2 are merged to form the IEB they undergo source-specific correction procedures (see the following chapters). 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.
- Datasets lacking information on the date of birth and gender after correction are deleted.

No further corrections (such as the addition of presumably missing notifications, strike corrections etc.) 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 in 2011.
- 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 no details on earnings, 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 December 31, 2017.

3.1.2 Benefit Recipient History (LeH)

- Observations without a valid 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, as 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 in the BeH.

3.1.3 Unemployment Benefit II Recipient History (LHG)

- Observations without a BA client number 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 unit (BG) are created. 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 unit (14, 15, 18 and 25) and the individual retirement age of members of the BG (see Section 3.1.7),
  - when the structure of the benefit unit 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 January 1, 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 (see Section 4.1.2) were corrected as far as possible.
- The territorial allocations are corrected in the same way as in the BeH.

3.1.4 Jobseeker Histories (ASU/XASU)

- Observations with an end date before January 1, 1995 are not included.
- 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 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 in the BeH.

3.1.5 Participants-In-Measures History File (MTH)

- 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.
- The territorial allocations are corrected in the same way as in the BeH.

3.1.6 Participants-In-Measures History File from XSozial-BA-SGB II (XMTH)

- For XMTH, the particular challenge is to identify and handle multiple notifications of participation that are caused by technical or organisational issues. The causes of these duplicate notifications are, on the one hand, a missing identification number for participation in measures until April 2009 or improper handling of the subsequently introduced promotion ID and, on the other hand, a change of provider numbers in connection with mergers or other reorganisations of institutions responsible for implementing SGB II due to the reorganisation of the employment agencies in 2012/2013. In addition, provider-specific problems arose with version or product changes of the municipal software or with version changes of the XSozial standard. At times, this results in a considerable number of reports of participations in measures of the same type, which partly or completely overlap in time per participant. Real duplicates are identified and sorted out because they are highly likely to be technical duplicates. Hidden duplicates can also be identified and sorted out. In principle, the rule always applies that the most up-to-date information is retained.
- In addition, notifications with a starting date before 2005 are excluded.
- Overlapping and immediately adjacent notifications of the same type of measure (from XSozial) are combined to one observation. Measure-specific characteristics of the combined periods are set to system missings.
- Subsequently, the remaining characteristics are compiled and calculated. For this, consolidated individual data from XSozial-histories and LHG are used.
- Finally, the following measures are excluded:
  - one-off benefits (like UBV/Mobi-/Vermittlungsbudget /LES) and
  - specific rehabilitation measures
- The territorial allocations are corrected in the same way as in the BeH.

3.1.7 SGB II anonymisation

In order to reduce the risk of de-anonymisation, only the year of birth is available in the LIAB by default. The month of birth can only be requested as a sensitive variable if there is a justified need for it. However, in the LHG and (X)ASU 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 18th, 25th or 65th 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 15th 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 variable ‘Employment status’ (*erwstat*).

3.2 Sampling procedure

The foundation of the sampling procedure in the LIAB are the establishments from the IAB Establishment Panel.

The IAB Establishment Panel is an annually conducted survey. The sample is drawn from the population of all German establishments with at least one employee liable to social security. The sample is stratified according to industry, firm size, and federal state.[7]

The data on individuals in the LIAB QM2 9317 are taken from the IEB according to the following procedure. In a first step, all of the establishments in the IAB Establishment Panel with a valid interview in the respective year (1993 to 2017) are selected. In a second step, all observations of individuals are drawn from the IEB that have been employed at one of these establishments on June 30 of the respective survey year for at least one day. For these individuals, all observations are made available which contain June 30 of the respective survey year in a third step.

3.3 Missing values

In the LIAB, missing values are coded as follows:

Term	Value	Description
No (valid) details available	.z	Values of a variable that 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 (cells shaded in dark grey in the overview of Table 4) or is not available for a certain period.

4 Data quality and problems

4.1 Entire IEB

4.1.1 Gaps in employment histories

The IEB contain comprehensive 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 variable reason for notification/reason for end of benefit receipt/reason for deregistration of the observation immediately preceding the gap (if a corresponding observation exists).

[7] See Fischer et al. (2008),

These gaps were identified using the variables 'Reason of cancellation/notification/termination' (*grund*) and 'Employment status' (*erwstat*) in the various sources. The list in Table 6 makes no claims to be exhaustive.

Table 6: Biographical gaps and possible ways of identifying them

Biographical gap	**Information on gap, potentially identifiable using the details in the '*grund*' variable in the preceding observation of the source**
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)	BeH, LeH, ASU
Persons receiving old-age pension without employment if not a member of a benefit unit	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)	
SGB-II recipients whose providers have experienced delivery failures	

Recipients of compensation according to FELEG (Gesetz zur Förderung der Einstellung der landwirtschaftlichen Erwerbstätigkeit, Act on Support in Case of Termination of Farming Activities)	

4.1.2 Introduction of SGB II and subsequent institutional changes

With the introduction of the SGB II on January 1, 2005, the responsibilities for the care of jobseekers were redesigned, so that tasks in the sphere of the SGB II can be taken over by different types of providers:

- Usually, the Arbeitsgemeinschaften (ARGE) between BA and the district took over the tasks according to the SGB II. These were replaced in 2011 by the joint facilities (gE) – also known as 'Job Centers'.
- In addition, it was possible until the end of 2011 for the BA and the district to perform the tasks assigned to them in their own (separate) responsibility (gT and gAw).
- Since 2005, it is also possible for authorised municipalities (zkT; also: opting municipalities) to take over the tasks. Initially, 69 administrative districts or independent towns took sole responsibility for the basic provision for jobseekers until December 31, 2010. With a constitutional amendment, the option was extended to 110 providers from 2012 onwards.

While the ARGEn/gE maintain benefit and case management via the BA procedures and transfer the recorded data to the statistics department of the BA, the municipal institutions responsible for implementing SGB II each use their own software systems. The transfer to the statistics department of the BA takes place via the XSozial standard. A division of responsibilities existed for separate providers. Any given provider was only responsible for the collection and delivery of the data falling within its area of responsibility. For this reason, there were reduced data requirements for the municipal side of the separate providers. Transmission via the XSozial standard takes place once a month at a fixed time window. In the following week, there is the possibility of repeating failed deliveries on a second reporting day. Over the years, numerous quality assurance instruments were developed. Nevertheless, version changes of the standard or the software used on site regularly increase the risk of data quality problems.

As part of the reorganisation of SGB II institutions on January 1 of the years 2011 to 2014, various ARGEn/gE and separate providers were transferred to municipal providers or converted from zkT to gE. This change of responsibility and the associated change of the software used locally also lead to some breaks in the data of the SGB-II sources. Detailed information on the resulting quality problems in the data can be found in the individual chapters on the data sources.

4.2 Employee History (BeH)

- 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 notifications with an end date later than November 30, 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 only enter the Employee History (BeH), the change of the occupation code only affects observations from this source. The measurement of the following characteristics previously reported using the occupation code 2003 is affected by this change: working hours, occupation, occupational status and school and vocational qualification levels. 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 switch to a new occupational classification. 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.[8]

- In contrast to LIAB QM2 9314, the LIAB QM2 9317 contains the variable 'Occupational status and working hours' (*stib*) again. However, the variable is only filled for reports that were submitted before the introduction of the new occupation code, and the FDZ does not extrapolate or impute for later reports. The categories of the variable *stib* that can be consistently observed over the whole observation period (occupational status as a trainee, distinction between part-time and full-time) are to be reproduced in the variables *erwstat* and *teilzeit*.
- 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,[9] 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 available in the new or converted variables 'Occupation/activity performed', 'Working time' and 'Vocational education and training' after the switch. 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 included in the LIAB QM2 9317. No imputation is performed regarding the gaps in the other variables.
- 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 'Vocational training', 'Employment status', and 'Daily wage').
- 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 April 1, 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.

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

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

- 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-months 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-months file shows a 100 percent degree of completeness by definition.
- In the version of the IEB on which the LIAB QM2 9317 is based, the year 2014 is the last year with a degree of completeness of BeH observations of 100 percent.[10] For the years 2015 and 2016, the 18-months files were used, and the observations for 2017 originate from a 6-months file. It can therefore be assumed that employment notifications for 2015 and 2016 are slightly underreported in the LIAB, and that those for 2017 are underreported to a slightly higher degree. 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' (*beruf*) variable are very rare compared to the neighbouring years. The reasons for this are not known.

[10] Due to a redesign of the data basis of the BeH, at the time of the preparation of the BeH as an exception only the 30-months files were available instead of the 36-months files. However, analyses with earlier data versions have shown that, as a rule, the 18-months file already has a degree of completeness of around 99 percent. This means that in the next 18 months there will only be extremely few follow-up or correction notifications.

- In the years 1975 and 1977, there were so far considerably too many employees with a place of work municipality of Lahn (district of Emsland, Lower Saxony). Instead of the expected double-digit figure, there were up to 90,000 employee registrations per year with this place of work in the population of the BeH. The reason for this is a historical misclassification of employment reports from the city of Lahn. This was a merger of the Hessian municipalities Wetzlar, Gießen, Heuchelheim, Wettenberg and Lahnau (districts of Gießen and Lahn-Dill-Kreis) which was dissolved after a short time. On the basis of these findings and assuming that there were no real establishment relocations between these regions, the following adjustment rule was implemented at the level of the establishment number: As soon as the establishment location changes from the municipality of Lahn to one of the listed Hessian municipalities in the years 1975 to 1978, the former specification of the municipality of Lahn is overwritten with the later correct specification. This rule significantly reduces the overhang and the municipality of Lahn in Lower Saxony then only has less than 2,000 employees in the population of the BeH in the critical 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 July 1, 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 January 1, 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 January 1, 2012 and of January 1, 2013 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 cannot be identified in the LIAB.
- Due to the reporting logic, information from the XSozial transmission standard can only be updated monthly.
- In the official performance statistics of the BA, reporting gaps are supplemented by a statistical estimation procedure at an aggregated regional level. No supplementary data sets are provided in the LHG. A comparison of the IEB with the performance statistics is therefore only possible to a limited extent.

4.5 Jobseeker Histories (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.
- With the introduction of SGB II on January 1, 2005, jobseekers are no longer fully covered by BA procedures. From this date, the ASU only covers persons who are supported by the BA in the sphere of the SGB III (employment promotion) or by ARGEn, gE or gT in the sphere of the SGB II (basic security).
- For the placement staff, it is not always possible to record the allocation to the legal sphere immediately, as 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 variable ‘Type of provider’ 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.
- 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. It can be assumed that integration agreements are under-recorded. 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.
- The job search status is hardly corrected afterwards, which is why the case numbers largely coincide with the BA statistics until 2005. Since spring 2011, the jobseeker data from BA procedures and XSozial have been consolidated in the integrated unemployment statistics. This may lead to larger deviations. At the current margin, however, the stock data are identical.

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 underrecording have emerged for the authorised municipalities since January 1, 2005. Thus, data from the XASU source should only be analysed from 2007 onwards.
- From 2007 onwards, individual months are also repeatedly affected by delivery failures, analogous to the benefit receipt data (LHG).
- Over-reporting may also occur at certain points. A large number of technical problems can be responsible for this, including systematically missing deregistrations, incorrect reversals of cancellations and parallel job search notifications in several districts.
- 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', 'Reason of 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

- current/most recent' 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.
- Due to the reporting logic, information from the XSozial reporting procedure can only be updated monthly.
- Differences in consolidation rules, time references and regional assignments may result in differences to the published BA statistics.

4.6 Participants-In-Measures History Files (MTH/XMTH)

4.6.1 MTH

- The MTH is incomplete for measures with a start date before January 1, 2000.
- As of January 1, 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 2.5 and 4.1.2).
- The MTH only contains 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 by authorised municipalities via the XSozial standard are contained in the XMTH.
- Because of the reorganisation of the institutions responsible for implementing SGB-II in 2011 to 2014, a split of the documentation of participations in measures in the MTH and in the XMTH may occur when there is a change in the reporting procedure.[11] This might result in a split or a duplication of the spell of the measure (see Section 3.1.6).
- 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 coArb to VerBIS in 2006. The same quality limitations as for the ASU apply here.

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

4.6.2 XMTH

- In the years 2005 to 2007, the notifications of participation in measures are incomplete. The degree of under-reporting is unclear due to a lack of or inaccurate comparative figures.
- Between 2005 and the beginning of 2017, about 13% of all institutions responsible for implementing SGB II report almost no municipal integration benefits (formerly accompanying benefits), which are included in the summarised category 'other support'. Many other job centers report only temporarily and/or only selected types. This leads to an under-recording of the 'other support' category.
- Over-registration of participations: Total stocks are considered to be stable from the end of 2008. However, the majority of the providers still register notification profiles which are temporarily conspicuous and which raise doubts about the quality. Only the introduction of the XSozial promotion-ID in 2009 and the meanwhile several years of experience of all participants stabilise the reporting process and the subsequent data processing. Also the new zkT, introduced in2012, report inconspicuously for the most part, so that the scope and duration of presumed over-recording decrease noticeably.
- The reorganisation of SGB II providers in 2011-2014 (see Section 4.1.2) lead to a split in the documentation of participation in measures in MTH and XMTH if the reporting procedure was changed. This could result in a split of the measure spell, but also in duplications (see Section 3.1.6).
- The figures or person counts received in the XMTH from monthly key date counts differ in several respects from the statistics published by the BA.

5 Description of variables

Frequency counts and overviews of the individual values and labels of the variables can be found in separate files under https://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, XMTH
Data type	numerical
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. 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.

	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
Data type	numerical
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) as well as directly on the website of the establishment number service of the BA at https://www.arbeitsagentur.de/betriebsnummern-service/alles-wichtige). The establishment number and year specification can be used to merge individual and establishment information. For the establishment number, the following should be observed in general: If the company has only one office, or if the company has only one office in one municipality, this office is the establishment and is given an establishment number. 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. 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. 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: 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. A **workplace** is a unit in which employees work and which is not allocated an establishment number according to the above-mentioned principles. A **company** as a term combines establishment premises and workplaces belonging to the same employer. 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. Establishment and establishment premises are synonyms; branch office is a synonym for subsidiary, district office, out-sourced 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.

5.1.3 Establishment-ID acc. to Survey (idnum)

Variable label	Establishment-ID acc. to Survey
Variable name	idnum
Category	identifiers
Origin	IAB Establishment Panel
Data type	numerical
Detailed description	The Establishment ID acc. to Survey is adopted from the IAB Establishment Panel. The variable is only filled in the Individual Files if the person is employed in an establishment which is in the IAB Establishment Panel on the reference date with a valid interview.

5.2 Generated technical variables

5.2.1 Source of spell (quelle)

Variable label	Source of spell
Variable name	quelle
Category	generated technical variables
Origin	BeH, LeH, LHG, ASU, XASU, MTH, XMTH
Data type	numerical
Detailed description	The variable indicates the data source (see Section 2).

5.2.2 Year (jahr)

Variable label	Year
Variable name	jahr
Category	generated technical variables
Origin	BeH
Data type	numerical
Detailed description	In the Individual Files, this variable indicates the calendar year at the beginning of each episode. In the Establishment Files, it indicates the calendar year on which the establishment information is valid, referenced to June 30. This variable can be used together with the establishment number to link the Individual and the Establishment Files, see Box 2 on page 10 for an example code with the '`merge`' command in Stata 14.

5.2.3 Status of establishment number (betr_st)

Variable label	Status of establishment number
Variable name	betr_st
Category	generated technical variables
Origin	BeH
Data type	numerical
Detailed description	For the current episode in the individual data, the variable indicates if there is information from the IAB Establishment Panel for this episode. The following characteristics exist: 1. Estab. surveyed in current year

	2. Estab. surveyed in other year 4. Estab. never surveyed The values 1-2 indicate employment in establishments which have taken part in the IAB Establishment Panel survey at least once, whereas value 4 marks employment in establishments which have never taken part in the IAB Establishment Panel. Value 1 marks establishments with an interview in the current year, while value 2 marks establishments surveyed in different year(s).

5.3 Personal information

5.3.1 Gender (frau)

Variable label	Gender
Variable name	frau
Category	personal variable
Origin	BeH, LeH, LHG, ASU, XASU, MTH, XMTH
Data type	numerical
Detailed description	Gender dummy (0 - man, 1 - woman). The gender information is constant within one individual account.

5.3.2 Year of birth (gebjahr)

Variable label	Year of birth
Variable name	gebjahr
Category	personal variables
Origin	BeH, LeH, LHG, ASU, XASU, MTH, XMTH
Data type	numerical
Detailed description	The year of birth is constant within one individual account.
Notes on quality	In the original data, it may happen that the date 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.

5.3.3 Month of birth (gebmon)

Variable label	Month of birth
Variable name	gebmon
Category	personal variables
Origin	BeH, LeH, LHG, ASU, XASU, MTH, XMTH
Data type	numerical

Detailed description	The month of birth is constant within one individual account.
	One can use the variables 'Year of birth' (*gebjahr*) and 'Month of birth' (*gebmon*) to generate a variable in the date format JJJJmM (e.g., 1984m6) with the following syntax in Stata:
	`gen int gebdat = ym(gebjahr, gebmon)`
	`format gebdat %tm`

Notes on quality	In the original data, it may happen that the date 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.

Anonymisation	Due to its particular sensitivity with regard to data privacy, this sensitive variable is only made available on application and only in well-founded cases. By default, only the coarsened variable 'Year of birth' (*gebjahr*) is provided.

5.3.4 Nationality (nation)

Variable label	Nationality
Variable name	nation
Category	personal variables
Origin	BeH, LeH, LHG, ASU, XASU, MTH, XMTH
Data type	numerical
Detailed description	The variable contains the nation codes used by the Federal Statistical Office (Statistisches Bundesamt, 2019).
Notes on quality	The variable is not filled well in the LeH before 1983.
Anonymisation	Due to its particular sensitivity with regard to data privacy, this sensitive variable is only made available on application and only in well-founded cases. By default, only the coarsened variable 'Nationality, grouped' (*nation_gr*) is provided.

5.3.5 Nationality, grouped (nation_gr)

Variable label	Nationality, grouped
Variable name	nation_gr
Category	personal variables
Origin	BeH, LeH, LHG, ASU, XASU, MTH, XMTH
Data type	numerical
Detailed description	The variable contains a grouped version of the nation codes used by the Federal Statistical Office (Statistisches Bundesamt, 2019).
Notes on quality	The variable is not filled well in the LeH before 1983.

5.3.6 Marital status (famst)

Variable label	Marital status
Variable name	famst
Category	personal variables
Origin	LeH, LHG, ASU, XASU, MTH, XMTH
Data type	numerical
Detailed description	This variable describes the marital status. **1) LeH** In the LeH, the variable has only two values (0 - not married, 1 - married). **2) LHG, ASU, XASU, MTH, XMTH** In the sources LHG, ASU, XASU, MTH and XMTH, a distinction is made between six values (values 11-16). The information from the different sources was not compared.
Notes on quality	**1) LeH** The quality of the information originating from the LeH is classified as poor until 2005, as it is not quite clear how the marital status was maintained. The quality is

	also questionable since the beginning of 2014. Between 2013 and 2017, for example, the proportion of married people drops from 37% to 14%. **2) LHG, XASU, XMTH** Due to deviating reporting standards in XSozial, the information from the zkT has to be classified as less valid until December 2009.

5.3.7 Number of children (kind)

Variable label	Number of children
Variable name	kind
Category	personal variables
Origin	LeH, LHG, ASU, XASU, MTH, XMTH
Data type	numerical
Detailed description	This variable has a different meaning depending on the data source. **1) LeH** In the LeH, the variable indicates the number of children aged under 16 living in the household 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: 0 no children 100 one or more children **2) ASU, MTH** In these sources, the value of the variable corresponds to the actual number of children under 15 living in the household. **3) LHG, XASU, XMTH** The variable reports the actual number of children aged under 15 in the benefit unit (Bedarfsgemeinschaft). In the LHG, the value is valid for the entire original period.
Notes on quality	**1) LeH** In the LeH, 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 strongly restricts the quality of the data. **2) ASU, MTH** Until June 30, 2006, only up to nine children could be recorded. The value zero does not exist. For observations prior to June 30, 2006, the value zero was recoded to 'missing', as it is not clear whether zero should be interpreted as 'no children' or as 'field not filled in'. For observations after June 30, 2006, the variable is only recorded if children exist.

5.3.8 Vocational training (ausbildung)

Variable label	Vocational training
Variable name	ausbildung
Category	personal variables
Origin	BeH, ASU, XASU, MTH, XMTH
Data type	numerical

Detailed description	It must be taken into account that this variable has a different meaning depending on the data source. **1) BeH** For observations obtained from the BeH, the variable contains the vocational education reported by the employers as part of the employment notification procedure. The following values exist: 1 Without vocational training 2 In-company voc. training/traineeship/external voc. training 11 University of applied sciences without further specifications 12 University without further specifications In notifications that rely on the new occupation code (see Section 4.2) 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. **2) ASU, MTH** For these observations the vocational education completed most recently is reported. The following values exist in spells with a start date until June 30, 2006: 1 Without vocational training 2 In-company voc. training/traineeship/external voc. training 3 Technical school (voc. training) 4 Technical school (advanced voc. training) 5 University of applied sciences (FH) 6 University In 2006, the IT procedure from which the jobseeker data originate was switched from coArb to VerBIS. Many variables, such as training, were reported with different levels of differentiation in the two systems. This means that in spells from ASU and MTH which have a start date from July 1, 2006 onwards, the following differentiated categories are available: 7 Voc. training not accepted in Germany 8 University degree not accepted in Germany 9 In-company voc. training/traineeship/in-school voc. training 10 Other exams 11 University of applied sciences without further specifications 12 University without further specifications 13 Doctorate 14 Bachelor (BA) 15 Bachelor (FH) 16 Bachelor (University) 17 Master (FH) 18 Master (University) 19 Diploma (BA) 20 Diploma (FH)

21 Diploma (University)
23 Undergraduate studies
24 Secondary/additional studies
25 Other (promotion-) advanced training for graduates

3) XASU, XMTH

For spells that originate from these sources, the vocational education completed most recently is reported. The following values exist:

1 Without vocational training
2 In-company voc. training/traineeship/external voc. training
3 Technical school (voc. training)
4 Technical school (advanced voc. training)
5 University of applied sciences (FH)
6 University
7 Voc. training not accepted in Germany
8 University degree not accepted in Germany

For analyses that cover a longer period of time, the values can be aggregated as follows:

ausbildung	***ausbildung_agg***	**value label**
1, 22	1	Without (recognised) vocational training
7	2	Vocational training not accepted in Germany
8	3	University degree not accepted in Germany
2, 3, 4, 9	4	In-company / school-based training
5, 11, 14, 15, 17, 19, 20	5	University of applied sciences
6, 12, 13, 16, 18, 21	6	University
10	6	Other exams
23	7	Undergraduate studies
24	7	Secondary/additional studies
25	7	Other (promotion-) advanced training for graduates

Notes on quality

1) BeH

'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 establishment. If, for example, an employee obtained 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 establishment. It can generally be assumed that when a person is employed in an establishment for a longer period, the personal data that they reported when they joined the establishment is simply carried forward' (own translation of Meinken / Koch 2004, p. 63).

In the BeH, the share of missing values increases almost continuously over time. Due to the introduction of the new occupation code in 2011, the share even temporarily strongly increased to around 51%. Since 2014, however, the proportion of missing values in the BeH has levelled off at around 40%.

	Missing values occur particularly frequently in the following groups: marginal part-time employees, part-time workers, foreign employees and workers from Eastern German. The reason for this is that the variable is not of particular importance as regards social security contributions (see Meinken/Koch, 2004, p. 63). For the variable 'Vocational training (imputed)' (*ausbildung_imp*, see Section 5.3.9), a method was applied to correct missing values or inconsistent changes of the training variable. However, this variable is only filled in the source BeH and has different categories than the variable *ausbildung*. **2) ASU, MTH** 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. In the source MTH, older categories partially still occur even after July 1, 2006. **3) XASU, XMTH** The degree of completeness in the XASU and the XMTH is generally low.

5.3.9 Vocational training (imputed) (ausbildung_imp)

Variable label	Vocational training (imputed)
Variable name	ausbildung_imp
Category	personal variables
Origin	BeH
Data type	numerical
Detailed description	The variable 'Vocational training (imputed)' (*ausbildung_imp*) is a supplement to the variable 'Vocational training' (*ausbildung*) and contains additional and harmonised information on the vocational training of employees for BeH spells. The variable thus offers a solution to the problems concerning the variable *ausbildung* described in Section 5.3.8. The imputation procedure is described in Thomsen et al (2018). As the variable only uses the training information from BeH notifications and because the educational categories of the old and the new occupation codes had to be harmonised for the variable *ausbildung*, the variable *ausbildung_imp* has other categories than the variable *ausbildung*.

5.3.10 School leaving qualification (schule)

Variable label	School leaving qualification
Variable name	schule
Category	personal variables
Origin	BeH, ASU, XASU, MTH, XMTH
Data type	numerical
Detailed description	This variable contains the school leaving qualification. Different values are possible depending on the source. **1) BeH**

	With the switch to the new occupation code (see Section 4.2), the possible values of the variable change. In the BeH, the values from the old occupation code are: 5 Grade-/lower school certificate, intermediate school or equivalent qualification 8 Completion of education at a specialised upper secondary school/completion of higher education at a specialised college or upper secondary school leaving certificate, A-level equivalent, qualification for university; 13 years of schooling 9 Upper secondary school leaving certificate, A-level equivalent, qualification for university; 13 years of schooling With the new occupation code the values are: 1 No school leaving certificate 4 Lower secondary school certificate/ grade school certificate 6 Intermediate school leaving certificate 8 Completion of education at a specialised upper secondary school/completion of higher education at a specialised college or upper secondary school leaving certificate, A-level equivalent, qualification for university; 13 years of schooling **2) ASU, XASU, MTH, XMTH** The following values are possible for these data sources: 1 No school leaving certificate 4 Lower secondary school certificate/ grade school certificate 6 Intermediate school leaving certificate 7 Completion of education at a specialised upper secondary school/completion of higher education at a specialised college 9 Upper secondary school leaving certificate, A-level equivalent, qualification for university; 13 years of schooling 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 in the BeH has been decreasing continuously over time and seems to have levelled off at under two thirds in recent years. In the XASU and the XMTH it has been increasing continuously and has levelled off at over two thirds since 2013 (XASU) and 2012 (XMTH), respectively. In the ASU and the MTH the degree of completeness is generally high.

5.4 Information on employment, benefit receipt and job search

5.4.1 Daily wage/daily benefit (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
Detailed description	**1) BeH**

	In BeH observations, this variable shows the employee's average gross daily wage. It is calculated from the fixed-period earnings reported by the employer and the duration of the (unsplit) original notification period in calendar days. The daily wage is shown in Euros. 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. Since the inclusion of marginal part-time employees in the employment notification procedure on April 1, 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 bonuses 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 in-formation, 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 https://fdz.iab.de/en.aspx. 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. From 2013 onwards, the number of notifications with a reason for deregistration of 54 (notification of a one-off wage) increases sharply (see Section 5.4.12). It is likely that special payments which were reported with the annual declarations before 2013 are now reported separately. It is therefore advisable to add these variable one-time payments to the corresponding wages for simultaneous employment episodes within the same establishment when analysing wages over time. 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. **2) LeH** 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 end date prior to January 1, 1998 the daily benefit rate applies to working days (i.e., including Saturdays but excluding Sundays and public holidays), while for observations with an original end date from January 1, 1998 onwards it applies to calendar days. Since January 1, 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.

5.4.2 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
Detailed description	**1) BeH** In the BeH, 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). 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. Employment notifications with an end date later than November 30, 2011 are reported using the new occupation code 2010 (KldB2010) (see Section 4.2). These values are transcoded to the KldB1988 via a priority switch. This results in inaccuracies. **2) ASU, XASU, MTH** The variable contains the occupation of the last job. See 1) with regard to the occupation code.
Notes on quality	**1) BeH** In the BeH, there is a considerable increase in the number of missing values in 2011 due to the change in the occupation code. A similar accumulation of missing values occurs in 1991 due to reunification. **2) ASU, MTH** The degree of completeness decreases significantly from 2014 to 2017. **3) XASU** The occupation variable is not filled for almost the entire period available.

5.4.3 Occupational group - current/most recent (KldB 2010), 3-digit (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, XMTH
Data type	numerical

Detailed description	**1) BeH** 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). 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 1,300 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*). Employment notifications with an end date earlier than November 30, 2011 are reported using the old occupation code 1988 (KldB 1988) (see Section 4.2). 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, XMTH** The variable contains the occupation of the last job. See 1) with regard to the occupation code.
Notes on quality	**1) BeH** There is a considerable increase in the number of missing values in 2011 due to the change in the occupation code. A similar accumulation of missing values occurs in 1991 due to reunification. **2) LeH** This variable was previously not filled in the LeH because the occupational data in that source was of poor quality. By now, the occupation in the LeH is adopted from the jobseeker history, leading to an increase in quality. **3) XASU, XMTH** The variable is not filled for almost the entire period available.

5.4.4 Occupational sub-group - current/most recent (KldB 2010), 4-digit (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, XMTH
Data type	numerical

Detailed description	**1) BeH** 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). 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 1,300 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*). Employment notifications with an end date earlier than November 30, 2011 are reported using the old occupation code 1988 (KldB 1988) (see Section 4.2). 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, XMTH** The variable contains the occupation of the last job. See 1) with regard to the occupation code.
Notes on quality	**1) BeH** There is a considerable increase in the number of missing values in 2011 due to the change in the occupation code. A similar accumulation of missing values occurs in 1991 due to the reunification. **2) LeH** This variable was previously not filled in the LeH because the occupational data in that source was of poor quality. By now, the occupation in the LeH is adopted from the jobseeker history, leading to an increase in quality. **3) XASU, XMTH** The variable is not filled for almost the entire period available.
Anonymisation	Due to its particular sensitivity with regard to data privacy, this sensitive variable is only made available on application and only in well-founded cases. By default, only the coarsened variable *beruf2010_3* is provided.

5.4.5 Level of requirement - current/most recent job (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, XMTH
Data type	numerical
Detailbeschreibung	**1) BeH** 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).

	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 1,300 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*). Employment notifications with an end date earlier than November 30, 2011 are reported using the old occupation code 1988 (KldB 1988) (see Section 4.2). 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, XMTH** The variable contains the occupation of the last job. See 1) with regard to the occupation code.
Notes on quality	**1) BeH** There is a considerable increase in the number of missing values in 2011 due to the change in the occupation code. A similar accumulation of missing values occurs in 1991 due to the reunification. **2) LeH** This variable was previously not filled in the LeH because the occupational data in that source was of poor quality. By now, the occupation in the LeH is adopted from the jobseeker history, leading to an increase in quality. **3) XASU, XMTH** The variable is not filled for almost the entire period available.

5.4.6 Part-time (teilzeit)

Variable label	Part-time
Variable name	teilzeit
Category	information on employment, benefit receipt and job search
Origin	BeH
Data type	numerical
Detailed description	The variable 'Part-time' (*teilzeit*) distinguishes between full-time and part-time employees. The decisive factor 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.
Notes 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). A similar accumulation of missing values occurs in 1991 due to the reunification. No imputation is performed here.

5.4.7 Occupational status and working hours (stib)

Variable label	Occupational status and working hours

Variable name	stib
Category	information on employment, benefit receipt and job search
Origin	BeH
Data type	numerical
Detailed description	The employee's occupational status during the notification period is reported by the employer as part of the 'employment details'. The variable 'Occupational status and working hours' (*stib*) distinguishes between full-time and part-time employees. The decisive factor for this differentiation is the ratio between the contracted hours and the usual working hours in the establishment. For part-time employees, the variable only indicates whether their working hours exceed a certain threshold. Until 1978, this threshold was 20 hours of work per week, between 1979 and 1987 it was 15 hours per week and since 1988 it is 18 hours per week. The variable only provides information regarding the occupational status for full-time employees, distinguishing among other things between blue-collar and white-collar employees in full-time employment and apprentices. The distinction between (full-time) blue-collar and white-collar employees is solely based on the type of compulsory pension insurance (Federal Social Insurance Office for Salaried Employees – BfA – for white-collar workers, and State Social Insurance Office – LVA – for blue-collar workers). Master craftsmen and foremen are only included in a separate category if they are compulsorily insured in the workers' pension insurance. The assignment of a master craftsman or foreman to the blue-collar or white-collar employees can only be made via the respective pension provider. The 'Employees in vocational training' category covers not only apprentices, volunteers and interns but also semi-skilled trainees, students at colleges for health occupations and participants in subsidised further vocational training, retraining and induction training. If more than one code is eligible for an employee, the employer is required to classify the job according to the activity which is predominantly performed. If this cannot be determined clearly, the code of the higher occupational status is to be entered (see BA 2005, p. VI). Owing to the introduction of the new occupation code (see Section 4.2), however, this distinction is no longer possible. The variable *stib* is therefore only filled for notifications which date back to before the introduction of the new occupation code.
Notes on quality	There is a considerable number of missing values in 1991 due to the German reunification.

5.4.8 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, XMTH
Data type	numerical
Detailed description	This variable takes on different values with different meanings for each data source. **1) BeH** For BeH observations, the variable 'Employment status' (*erwstat*) corresponds to the person group recorded in the new notification procedure (DEÜV) from January 1, 1999 onwards. It indicates contribution- or benefit-related particularities of the employment relationship. If multiple codes apply to an employment notification, the smallest must be indicated by the reporting employer. The majority of these cases are employment rela-

	tionships 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. 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. 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 variables 'Vocational education and training', 'Occupational status and working hours' and 'Occupation' as well as other information. In many cases, however, conclusive allocations are not possible. Since April 1, 1999, employees in marginal part-time employment have also been recorded in the DEÜV notification procedure. This person group 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. **2) LeH** For LeH observations, the variable 'Employment status' (*erwstat*) 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. **3) LHG** For LHG spells, the variable shows whether the person is underage and able to work, adult and able to work, or unable to work and beyond the retirement pension limit. As the reporting logic would make it possible to re-identify the exact date of birth in many cases, the original dates were changed by means of the anonymization procedure described in Section 3.1.7. **4) ASU, XASU** For ASU/XASU observations, the 'employment status' variable reports the job search status. A distinction is made between those who are 'unemployed and seeking work', 'not unemployed and seeking work', 'seeking advice' and 'not seeking work'. Applicants who only want advice from the BA are considered as 'seeking advice'. In addition, there are rehabilitants and, before 2008, persons aged 58 or older who are not fully available for placement. As of August 1, 2016, this may also include persons who no longer require support, but who are nevertheless still under the job center's care. 'Not seeking work' mainly subsumes persons of whom activation or placement cannot be reasonably expected according to § 10 SGB II. Similarly, persons with an incapacity to work of more than 42 days who continue to receive ALG II are listed in the system under this status. As the reporting logic would make it possible to re-identify the exact date of birth in many cases, the original dates were changed by the anonymization procedure described in Section 3.1.7. **5) MTH, XMTH**

	For observations in the MTH and the XMTH, 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.
Notes on quality	**1) LHG** 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 classified 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. In some cases, there may be conflicting information on a person's ability to work if he or she is a member of different benefit units (BGs) at the same time. A possible reason for this is the determination of the ability to work by the job center specialist supervising the BG or by the respective responsible medical services. They can come to different results or document these at different times. A further reason may be different birthday records for simultaneous BG customers, which may affect the status of their ability to work. **2) ASU, XASU** The categories 'seeking advice' and 'not seeking work' have only existed since the introduction of Verbis (see Section 4.5). Due to late notifications, however, entries before 2006 can also be found. The characteristic 'not seeking work' is considered to be under-recorded before 2008. In the XASU, the quality of status determination is partly limited for some institutions, especially in the first years.

5.4.9 Transition zone (gleitz)

Variable label	Transition zone
Variable name	gleitz
Category	information on employment, benefit receipt and job search
Origin	BeH
Data type	numerical
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 April 1, 2003.

5.4.10 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
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 November 30, 2011.

Notes on quality	There is a considerable increase in the number of missing values in 2011 due to the change in the reporting procedure. By 2012, the share of missing values is down to about 3%.

5.4.11 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
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 November 30, 2011.
Notes on quality	There is a considerable increase in the number of missing values in 2011 due to the change in the reporting procedure. By 2012, the share of missing values is down to about 3%.

5.4.12 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
Detailed description	**1) BeH** In BeH observations, the 'Reason of 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), while 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. The reasons for submitting employment notifications are encoded according to the regulations of the notification procedure, which has been in effect since January 1, 1999 (in accordance with DEÜV). **2) LeH** In the LeH, this variable specifies the reason for the end of the receipt of unemployment benefits, unemployment assistance, or maintenance allowance. There is no information in the LeH on the reasons for the start of the benefit receipt because the LeH is filled with the notifications from the employment agencies to the health insurance about completed benefit receipt durations. **3) LHG** 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 unit. If the Unemployment Benefit II receipt of a different

	member of the benefit unit is discontinued, new observations for all members of the benefit unit 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. **4) ASU** 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'. The number of values of the variable was reduced from April 26, 2003 onwards. For analyses over long periods of time, the old values can be recoded to the currently valid ones using the table below (in the 5,000s number range): **5) XASU** 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.
Notes on quality	**1) BeH** In the BeH, from 2013 onwards the number of notifications with a reason for deregistration of 54 (notification of a one-off payment) increases sharply. As long as an employment relationship exists, special payments that are paid out by March of the following year can be included into the usual notifications (mostly annual notifications) for the previous year. A separate notification with a reason for deregistration of 54 is then not required. Until 2012, the annual notifications could be submitted until mid-April; since 2013, they must now be submitted by mid-February at the latest. Special payments made in February and March must now be reported separately. **2) LHG** The degree of completeness for the reason for notification in the LHG data sources is very small (< 20%) across all years.

29->60	33->60	37->66	42->65	46->67	50->75	54->78
30->60	34->60	38->66	43->70	47->67	51->74	
31->61	35->60	39->71	44->74	48->78	52->76	
32->60	36->61	40->69	45->77	49->69	53->68	

5.4.13 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
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:

	• good integration prospects (integration into the regular labour market within 12 months is realistic) • complex (integration into the regular labour market within 12 months is not realistic) • other 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.
Notes 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.

5.4.14 Reason for end of previous employment (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, MTH
Data type	numerical
Detailed description	This variable describes how the last employment or training relationship was terminated before a period of job search. It can therefore be used to identify job-to-job placements.

5.4.15 Working hours of job application (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
Detailed description	During the placement procedure, jobseekers indicate how many working hours the job they are seeking should have.

5.4.16 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
Detailed description	The variable has a different meaning depending on the data source. **1) LeH**

	In the LeH, 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. If the end date of the benefit receipt is before January 1, 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. **2) MTH** The variable contains the planned duration of the measure.
Notes on quality	In the LeH, over the course of a benefit receipt biography a systematic development of the remaining entitlement is to be expected. This means that the remaining entitlement at the end of a benefit period minus the duration of the subsequent period should give the remaining entitlement after the subsequent period as long as no claims have arisen or expired in the meantime. However, there may be unexpected increases or decreases in the remaining entitlement. These are presumably mainly caused by corrections during the processing of benefit cases. Such corrections occur, for example, if a claim was not correctly determined at the start of the benefit case due to incomplete information or if a remaining claim that has not yet expired was not taken into account. However, these corrections are usually only administered for the current record, i.e. there is no correction of the complete case. In addition, in individual cases the information on entitlement durations may exceed the individual upper limit according to the statutory requirements.

5.4.17 Type of provider (traeger)

Variable label	Type of provider
Variable name	traeger
Category	information on employment, benefit receipt and job search
Origin	LHG, ASU, XASU, MTH, XMTH
Data type	numerical
Detailed description	The variable contains the type of institution responsible for implementing Unemployment Benefit II (LHG), the type of institution providing the measure (MTH/XMTH) 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 2.3. The type of institutions responsible for implementing SGB II may change over time (e.g., from a cooperation of an employment agency and a municipality to a municipality exercising 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. ASU and XASU observations contain the type of institution that holds the records of the applicant pool data.

5.5 Location data

5.5.1 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, XMTH
Data type	numerical
Hierarchy	federal state | district
Detailed description	For BeH and LeH observations, the place of residence at the 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, NUTS 1), and positions 1-3 indicate the regional authority (Regierungsbezirk, NUTS 2). Federal states without a regional authority have a '0' in the third position. In the NUTS classification (Nomenclature des unités territoriales statistiques) of the European Union, districts correspond to the level NUTS 3. 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 and XASU sources, the place of residence applies to the whole period of the original observation. For the ASU, LeH, MTH and XMTH, the variable contains the place of residence at the beginning of the original period of time. This means that the longer the spell lasts, the higher the risk that the place of residence will become obsolete and that the information given for later dates will be incorrect. 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 December 31, 2017 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 December 31, 2017. 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 on quality	There are inaccuracies in the information provided for some employees with regard to where they live. The reporting requirement does not clarify which residence - main or secondary residence with predominant residence - is to be reported by the employer. In the employment history, this can lead to the impression of 'long-distance commuters' between the registered main residence and the place of work, even though the person is in fact employed at the secondary residence, i.e. does not actually commute. In the year 2015, the data show a reduction in the change of residence of approx. 10% to 15% compared to the usual level of the surrounding years, with regional differences occurring. It has not yet been possible to determine the exact reason for this deviation.
Anonymisation	Due to its particular sensitivity with regard to data privacy, this sensitive variable is only made available on application and only in well-founded cases. By default, only the coarsened variable 'Place of residence – federal state (Bundesland)' (*wo_bula*) is provided.

5.5.2 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, XMTH
Data type	numerical
Hierarchy	federal state | district
Detailed description	This variable is an aggregation of the variable 'Place of residence - district (Kreis)' (*wo_kreis*) to the 16 German federal states. For BeH and LeH 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 in Section 5.5.1.

5.5.3 Place of residence - employment agency (Arbeitsagentur) (wo_aa)

Variable label	Place of residence - employment agency (Arbeitsagentur)
Variable name	wo_aa
Category	location data
Origin	BeH, LeH, LHG, ASU, XASU, MTH, XMTH
Data type	numerical
Hierarchy	regional directorate | employment agency
Detailed description	From 1999 onwards, this variable contains the agency district of the employment agency that is responsible for the employee's / BA client's place of residence. 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, MTH and XMTH, the variable contains the place of residence at the start of the period of unemployment or job search. Accordingly, the longer the spell lasts, the higher the risk that the place of residence will become obsolete and that the information given for later dates will be incorrect. For the BeH and LeH, it is available from 1999 onwards. In order to guarantee consistent regional allocations across the entire observation period, the information on the agency district was recoded to the territorial allocation of December 31, 2017 for all data sources, i.e. in all calendar years, a place of residence is assigned to an agency district in accordance with the boundaries that the agency district had on December 31, 2017. As the boundaries of the agency district 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. Berlin should preferably be analysed as a unit and not separately for East and West, because since the territorial reform of the employment offices in Berlin on July 1, 1997 their clear allocation to East and West is no longer possible. The BA statistics uniformly assigns Berlin to the East. 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).
Anonymisation	Owing to its particular sensitivity with regard to data privacy, 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.

5.5.4 Place of residence - regional directorate (Regionaldirektion) (wo_rd)

Variable label	Place of residence - regional directorate (Regionaldirektion)

Variable name	wo_rd
Category	location data
Origin	BeH, LeH, LHG, ASU, XASU, MTH, XMTH
Data type	numerical
Hierarchy	regional directorate | employment agency
Detailed description	This variable is an aggregation of the variable 'Place of residence: employment agency (Arbeitsangentur)' (*wo_aa*) at the level of the regional directorates. Further information can be found in Section 5.5.3.

5.6 Establishment variables

5.6.1 Classification of economic activities 73 (w73_3)

Variable label	classification of economic activities 73
Variable name	w73_3
Category	establishment variables
Origin	BHP
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. 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. 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.

5.6.2 Classification of economic activities 93, sub-classes (w93_5)

Variable label	classification of economic activities 93, sub-classes
Variable name	w93_5
Category	establishment variables
Origin	BHP
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’). Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected.
Anonymisation	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*).

5.6.3 Classification of economic activities 93, groups (w93_3)

Variable label	classification of economic activities 93, groups
Variable name	w93_3
Category	establishment variables
Origin	BHP
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’). Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected.

5.6.4 Classification of economic activities 03, sub-classes (w03_5)

Variable label	classification of economic activities 03, sub-classes
Variable name	w03_5
Category	establishment variables
Origin	BHP
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. Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected.

Anonymisation	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*).

5.6.5 Classification of economic activities 03, groups (w03_3)

Variable label	classification of economic activities 03, groups
Variable name	w03_3
Category	establishment variables
Origin	BHP
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. Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected.

5.6.6 Classification of economic activities 08, sub-classes (w08_5)

Variable label	classification of economic activities 08, sub-classes
Variable name	w08_5
Category	establishment variables
Origin	BHP
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. Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected.
Anonymisation	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*).

5.6.7 Classification of economic activities 08, groups (w08_3)

Variable label	classification of economic activities 08, groups
Variable name	w08_3
Category	establishment variables

Origin	BHP
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. Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected.

5.6.8 w73_3 completed by extrapolation/imputation (w73_3_gen)

Variable label	w73_3 completed by extrapolation/imputation
Variable name	w73_3_gen
Category	establishment variables
Origin	BHP
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). Further information on the WS73 classification can be found in the description of the variable *w73_3*.

5.6.9 Type of imputation w73_3 (group_w73_3)

Variable label	Type of imputation w73_3
Variable name	group_w73_3
Category	establishment variables
Origin	BHP
Data type	numerical
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. A detailed description of the procedure can be found in Eberle et al. (2011).

5.6.10 w93_3 completed by extrapolation/imputation (w93_3_gen)

Variable label	w93_3 completed by extrapolation/imputation
Variable name	w93_3_gen
Category	establishment variables
Origin	BHP
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). Further information on the WS93 classification can be found in the description of the variable *w93_3*.

5.6.11 Type of imputation w93_3 (group_w93_3)

Variable label	Type of imputation w93_3
Variable name	group_w93_3
Category	establishment variables
Origin	BHP
Data type	numerical
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. A detailed description of the procedure can be found in Eberle et al. (2011).

5.6.12 w08_3 completed by extrapolation/imputation (w08_3_gen)

Variable label	w08_3 completed by extrapolation/imputation
Variable name	w08_3_gen
Category	establishment variables
Origin	BHP
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 from 2008 onwards. From 2008 up to and including 2017, the variable contains the original values from *w08_3*. Before 2008, the information is either written back 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 WZ08. A detailed description can be found in Eberle et al. (2011). Further information on the WZ08 classification can be found in the description of the variable *w08_3*.

5.6.13 Type of imputation w08_3 (group_w08_3)

Variable label	Type of imputation w08_3
Variable name	group_w08_3
Category	establishment variables
Origin	BHP
Data type	numerical

Detailed description	This variable indicates the type of completion for the *w08_3_gen* variable. It reports whether the respective value in *w08_3_gen* is consistent with the original value from *w08_3*, still missing / extrapolated or imputed based on recording tables. A detailed description of the procedure can be found in Eberle et al. (2011).

5.6.14 Year of first appearance (grd_jahr)

Variable label	year of first appearance
Variable name	grd_jahr
Category	establishment variables
Origin	BHP
Data type	numerical
Detailed description	This variable indicates the first appearance of the establishment number in the dataset. 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. An establishment does not necessarily have to be included in the BHP in the year of its first appearance, as only the key date June 30 is relevant for this inclusion. If the establishment has no employees on June 30 of its year of foundation, it consequently does not appear in the BHP in that year.

5.6.15 First appearance (grd_dat)

Variable label	first appearance
Variable name	grd_dat
Category	establishment variables
Origin	BHP
Data type	date
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. An establishment does not necessarily have to be included in the BHP in the year of its first appearance, as only the key date June 30 is relevant for this inclusion. If the establishment has no employees on June 30 of its year of foundation, it consequently does not appear in the BHP in that year.
Anonymisation	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*).

5.6.16 Year of last appearance (lzt_jahr)

Variable label	year of last appearance
Variable name	lzt_jahr
Category	establishment variables
Origin	BHP
Data type	numerical
Detailed description	This variable indicates the last appearance of the establishment number in the dataset (see Bender et. al. 1996). If the existence of an establishment number in the BHP already ends before 2017, this 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). An establishment does not necessarily have to be included in the BHP in the year of its last appearance, as only the key date June 30 is relevant for this inclusion. If the establishment has no employees on June 30 of its year of closure, it consequently does not appear in the BHP in that year.

5.6.17 Last appearance (lzt_dat)

Variable label	last appearance
Variable name	lzt_dat
Category	establishment variables
Origin	BHP
Data type	date
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 2017, 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). An establishment does not necessarily have to be included in the BHP in the year of its last appearance, as only the key date June 30 is relevant for this inclusion. If the establishment has no employees on June 30 of its year of closure, it consequently does not appear in the BHP in that year.
Anonymisation	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*).

5.6.18 Total number of employees (az_ges)

Variable label	no. employees
Variable name	az_ges
Category	generated establishment variables
Origin	BHP
Data type	numerical

Detailed description	This variable contains the total number of an establishment's employees reported to the social security agencies as of June 30 of a year. Since the introduction of the new notification regulations in 1999, people in marginal part-time employment have also been recorded. Dormant employment relationships (daily wage of zero) are not included.

5.6.19 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	BHP
Data type	numerical
Detailed description	This variable contains the number of people in the establishment who are reported on June 30 of a year as full-time employees under the person group codes 101, 140, 143, 105, 106, 112, 113, 114, 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.

5.6.20 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	BHP
Data type	numerical
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.

5.6.21 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	BHP
Data type	numerical

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. The values are reported in Euros for all years. 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 capped. 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 3.1.3.3 in Schmucker et al. 2018). These data were then aggregated at establishment level. The values are rounded to two decimal places. However, due to the 'storage type' in Stata, additional decimal places are displayed that are not correct.

5.6.22 Place of work - district (Kreis) (ao_kreis)

Variable label	Place of work - district (Kreis)
Variable name	ao_kreis
Category	location data
Origin	BHP
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, NUTS 1), positions 1-3 indicate the regional authority (Regierungsbezirk, NUTS 2). Federal states without a regional authority have a 0 in the third position. In the NUTS classification (Nomenclature des unités territoriales statistiques) of the European Union, districts correspond to the level NUTS 3. In order to guarantee consistent regional allocations across the entire observation period, the information on the district was recoded to the territorial allocation of December 31, 2017, 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 December 31, 2017. 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.
Anonymisation	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.

5.6.23 Place of work - federal state (Bundesland) (ao_bula)

Variable label	Place of work - federal state (Bundesland)
Variable name	ao_bula
Category	location data
Origin	BHP

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 (NUTS 1).

5.7 Generated biographical variables

5.7.1 First day in employment (ein_erw)

Variable label	First day in employment
Variable name	ein_erw
Category	generated biographical variables
Origin	generated from BeH
Data type	date
Hierarchy	none
Detailed description	This variable specifies the date of start of employment subject to social security in the IEB. Training periods are not included (*erwstat* == 102, 121, 122, 141). Persons have always a missing value if they pass a training period in the IEB but do not have an employment covered by the social security system. The 'First day in employment' (*ein_erw*) can occur a long time after the 'First day in establishment' (*ein_bet*) and the 'First day in job' (*ein_job*) because in the latter cases training periods are included.
Note on quality	For West Germans the variable is left censored on January 1, 1975. For East Germans the censoring is not so clear. Entries on January 1, 1990 are censored for sure, but often also entries on January 1, 1991 and 1January 1, 1992 may be affected because in 1990 and 1991 many employment notifications are missing.

5.7.2 Number of days in employment (tage_erw)

Variable label	Number of days in employment
Variable name	tage_erw
Category	generated biographical variables
Origin	generated from BeH
Data type	numerical
Hierarchy	none
Detailed description	The variable contains the number of days that an individual has worked until June 30 of a year. Training periods are not taken into account (*erwstat* == 102, 121, 122, 141). If an individual was just in training, the variable takes on the value '0'. The 'Number of days in employment' (*tage_erw*) can be zero, even though the variable 'First day in employment' (*ein_erw*) is filled. This happens when individuals have not entered the workforce until June 30 of the respective year.
Note on quality	For West Germans the variable is left censored on January 1, 1975. For East Germans the censoring is not so clear. Entries on January 1, 1990 are censored for sure, but often also entries on January 1, 1991 and January 1, 1992 may be affected because in 1990 and 1991 many employment notifications are missing.

5.7.3 First day in establishment (ein_bet)

Variable label	First day in establishment
Variable name	ein_bet
Category	generated biographical variables

Origin	generated from BeH
Data type	date
Hierarchy	none
Detailed description	This variable contains the start date of the first employment notification in the current establishment in the IEB. Training periods are also considered. An interruption of the employment in the establishment does not change the start date, i.e. it is constant for each combination of individual and establishment number. The 'first day in employment' (*ein_erw*) can occur a long time after the 'First day in establishment' (*ein_bet*) and the 'First day in job' (*ein_job*) because in the latter cases training periods are included. This variable is coded missing (.n) if the establishment ID is missing.
Note on quality	For West Germans the variable is left censored on January 1, 1975. For East Germans the censoring is not so clear. Entries on January 1, 1990 are censored for sure, but often also entries on January 1, 1991 and January 1, 1992 may be affected because in 1990 and 1991 many employment notifications are missing.

5.7.4 Number of days in establishment (tage_bet)

Variable label	Number of days in establishment
Variable name	tage_bet
Category	Generated biographical variables
Origin	Generated from BeH
Data type	numeric
Hierarchy	none
Detailed description	The variable contains the number of days a person has been working in the respective establishment. In the Cross-Sectional Model, the duration is calculated until June 30 of this year. Training periods are included, too, while employment gaps are not. If the number of days in the establishment is alternatively calculated with the variable 'First day in establishment' (*ein_bet*) `gen tag_bet_neu = mdy(6,30, jahr) - ein_bet + 1` the values obtained are larger than or equal to the variable *tage_bet* because *tage_bet* does not include interruptions of employment. The 'Number of days in establishment' (*tage_bet*) can be larger than the 'Duration of a working life' (*tage_erw*). The variable *tage_bet* includes periods of education while *tage_erw* does not. This variable is coded missing (.n) if the establishment ID is missing. In some rare cases, the variable is coded 0 if the date of entry into the establishment is characterized by a lump sum payment (*grund* == 54).
Note on quality	For West Germans the variable is left censored on January 1, 1975. For East Germans the censoring is not so clear. Entries on January 1, 1990 are censored for sure, but often also entries on January 1, 1991 and January 1, 1992 may be affected because in 1990 and 1991 many employment notifications are missing.

5.7.5 First day in job (ein_job)

Variable label	First day in job
Variable name	ein_job
Category	generated biographical variables
Origin	generated from BeH
Data type	numerical

Hierarchy	none
Detailed description	This variable contains the start date of the first employment notification in the current job. Training periods (*erwstat* == 102, 121, 122, 141) in the same establishment are treated as separate jobs, even if they follow directly or are followed directly by a job in the same establishment. An employment in the same establishment after a gap is considered a new job if • the reason for notification of the last employment record before the gap indicates the end of the last job (*grund* = 30, 34, 40 or 49) and the gap is longer than 92 days or • the reason for notification of the last employment record before the gap does not indicate the end of the last job and the gap is longer than 366 days. The 'First day in job' (*ein_job*) cannot occur before the 'First day in establishment' (*ein_bet*), but it can occur before the 'First day in employment' (*ein_erw*). This variable is coded missing (.n) if the establishment ID is missing. This is also true in the case of a lump sum payment (*grund* == 54) that cannot be assigned to a specific job.
Note on quality	For West Germans the variable is left censored on January 1, 1975. For East Germans the censoring is not so clear. Entries on January 1, 1990 are censored for sure, but often also entries on January 1, 1991 and January 1, 1992 may be affected because in 1990 and 1991 many employment notifications are missing.

5.7.6 Number of days in job (tage_job)

Variable label	Numbers of days in job
Variable name	tage_job
Category	generated biographical variables
Origin	generated from BeH
Data type	numerical
Hierarchy	none
Detailed description	The variable counts how many days a person has been working in the current job. Training periods (*erwstat* == 102, 121, 122, 141) in the same establishment are treated as separate jobs, even if they follow directly or are followed directly by a job in the same establishment. An employment in the same establishment after a gap is considered a new job if • the reason for notification of the last employment record before the gap indicates the end of the last job (*grund* = 30, 34, 40, or 49) and the gap is longer than 92 days or • the reason for notification of the last employment record before the gap does not indicate the end of the last job and the gap is longer than 366 days. In the Cross-Sectional Model, this duration is referenced to June 30 of the year. Training periods in the establishment are included while employment gaps are not. If the number of days in the current job is alternatively calculated with the variable 'First day in job' (*ein_job*) `gen tage_job_neu = mdy(6,30, jahr) - ein_job + 1` the values obtained are larger than or equal to the variable *tage_job* because it does not include interruptions of employment. This variable is coded missing (.n) if the establishment ID is missing. This is also true in the case of a lump sum payment (*grund* == 54) that cannot be assigned to a specific job.

Note on quality	For West Germans the variable is left censored on January 1, 1975. For East Germans the censoring is not so clear. Entries on January 1, 1990 are censored for sure, but often also entries on January 1, 1991 and January 1, 1992 may be affected because in 1990 and 1991 many employment notifications are missing.

5.7.7 Number of days with benefit receipt (tage_lst)

Variable label	Number of days of benefit receipt
Variable name	tage_lst
Category	generated biographical variables
Origin	generated from BeH, LeH
Data type	numerical
Hierarchy	none
Detailed description	The variable contains the number of days in benefit receipt of a person until June 30 of each year. If there is no benefit receipt the variable takes on the value '0'. Other payments (e.g. benefits according to SGB II, the former unemployment assistance, etc.) than unemployment benefits (Arbeitslosengeld) are not considered in this variable. This allows a consistent evaluation of this variable over time. Periods of parallel employment are not considered in this variable, too.
Note on quality	For West Germans the variable is left censored on January 1, 1975. For East Germans the censoring is not so clear. Entries on January 1, 1990 are censored for sure, but often also entries on January 1, 1991 and January 1, 1992 may be affected because in 1990 and 1991 many employment notifications are missing.

6 References

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

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: https://statistik.arbeitsagentur.de/Statischer-Content/Grundlagen/Methodenberichte/Beschaeftigungsstatistik/Generische-Publikationen/Methodenbericht-Neue-Erbebungsinhalte-Arbeitszeita-ausgeuebte-Taetigkeit-sowie-Schul-und-Berufsabschluss-in-der-Beschaeftigungsstatistik.pdf (accessed: June 12, 2019).

Bundesagentur für Arbeit (ed.) (2005): Schlüsselverzeichnis für die Angaben zur Tätigkeit in den Meldungen zur Sozialversicherung. Issue January 2005, Nuremberg.

Bundesagentur für Arbeit (ed.) (2007): Handbuch für die Betriebsnummernvergabe und –pflege im Rahmen des Meldeverfahren zur Sozialversicherung. Issue December 2007, Nuremberg.

Bundesagentur für Arbeit (2009): Klassifikation der Wirtschaftszweige 1973, Nuremberg. URL: http://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen/Klassifikation-der-Wirtschaftszweige/Klassifikation-der-Wirtschaftszweige-1973-2003/Klassifikationen-der-Wirtschaftszweige-1973-2003-Nav.html (accessed: June 12, 2019)

Bundesagentur für Arbeit (2011): Klassifikation der Berufe 2010. Band 1: Systematischer und alphabetischer Teil mit Erläuterungen, Nuremberg. URL: http://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen/Klassifikation-der-Berufe/KldB2010/Printausgabe-KldB2010/Printausgabe-KldB-2010-Nav.html (accessed: June 12, 2019).

Bundesanstalt für Arbeit (1988): Klassifikation der Berufe – Systematisches und Alphabetisches Verzeichnis der Berufsbenennung, Nuremberg. URL: http://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen/Klassifikation-der-Berufe/KldB1975-1992/KldB1975-1992-Nav.html (accessed: June 12, 2019)

Cramer, Ulrich (1985): Probleme der Genauigkeit der Beschäftigtenstatistik. In: Allgemeines Statistisches Archiv 69: pp. 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 (accessed: June 13, 2016)

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

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), Nuremberg.

Eberle, Johanna/ Schmucker, Alexandra (2019): Creating cross-sectional data and biographical variables with the Sample of Integrated Labour Market Biographies 1975-2017 * Programming examples for Stata. FDZ Methodenreport, 04/2019 (en), Nuremberg.

Fischer, Gabriele/ Janik, Florian/ Müller, Dana/ Schmucker, Alexandra (2008): The IAB Establishment Panel – from Sample to Survey to Projection. FDZ Methodenreport, 01/2008 (en), Nuremberg.

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, vol. 126, issue 3, pp. 405-436.

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

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

Paulus, Wiebke/ Matthes, Britta (2013): The German Classification of Occupations 2010 – Structure, Coding and Conversion Table. FDZ-Methodenreport, 08/2013 (en), Nuremberg.

Schmucker, Alexandra/Eberle, Johanna/ Ganzer, Andreas/Stegmaier, Jens/ Umkehrer, Matthias (2018): Establishment History Panel 1975-2016. FDZ-Datenreport, 01/2018 (en), Nuremberg.

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

Statistisches Bundesamt (2002): Klassifikation der Wirtschaftszweige, Ausgabe 1993 (WZ 93), Wiesbaden. URL: https://www.destatis.de/DE/Methoden/Klassifikationen/Gueter-Wirtschaftsklassifikationen/Downloads/klassifikation-wz-1993.html (accessed: June 12, 2019).

Statistisches Bundesamt (2003): Klassifikation der Wirtschaftszweige, Ausgabe 2003 (WZ 2003), Wiesbaden. URL: https://www.destatis.de/DE/Methoden/Klassifikationen/GueterWirtschaftklassifikationen/Content75/KlassifikationWZ2003.html (accessed: June 12, 2019).

Statistisches Bundesamt (2008): Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008), Wiesbaden. URL: https://www.destatis.de/DE/Methoden/Klassifikationen/Gueter-Wirtschaftsklassifikationen/klassifikation-wz-2008.html (accessed: June 12, 2019).

Statistisches Bundesamt (2019): Staats- und Gebietssystematik, Wiesbaden. URL: https://www.destatis.de/DE/Methoden/Klassifikationen/Staat-Gebietsystematik/Staatsangehoerigkeitsgebietsschluessel_pdf.pdf?__blob=publicationFile&v=5 (accessed: June 12, 2019).

Thomsen, Ulrich/ Ludsteck, Johannes/ Schmucker, Alexandra (2018): Skilled or unskilled - Improving the information on qualification for employee data in the IAB Employee Biography. FDZ-Methodenreport, 09/2018 (en), Nuremberg.

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, pp. 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 https://fdz.iab.de/en.aspx.

7.2 List of abbreviations

AA	Agentur für Arbeit / Arbeitsamt	employment agency / employment office
ALG	Arbeitslosengeld	unemployment benefit
ALH	Arbeitslosenhilfe	unemployment support
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
BG	Bedarfsgemeinschaft	Benefit unit
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
DIM	Daten- und IT-Management	Data and IT Management
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
IEB	Integrierte Erwerbsbiographien	Integrated Employment Biographies
ISIC	International Standard Industrial Classification of All Economic Activities	International Standard Industrial Classification of All Economic Activities
KldB	Klassifikation der Berufe	Classification of occupations
LeH	Leistungsempfängerhistorik	Benefit Recipient History
LHG	Leistungshistorik Grundsicherung	Unemployment Benefit II Recipient History
LIAB	Linked-Employer-Employee-Daten des IAB	Linked Employer-Employee Data of the IAB
LM	Längsschnittmodell	Longitudinal Model
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
NUTS	Nomenclature des unités territoriales statistiques	Nomenclature des unités territoriales statistiques
QM	Querschnittmodell	Cross-sectional model
SGB	Sozialgesetzbuch	German Social Code
UHG	Unterhaltsgeld	subsistence payment
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
XLHG	Leistungshistorik Grundsicherung aus XSozial-BA-SGB II	Unemployment Benefit II Recipient History from XSozial-BA-SGB II
zkT	Zugelassene kommunale Träger	Authorised municipalities

Imprint

FDZ-Datenreport 06|2019 EN

Date of publication

October 21, 2019

Publisher

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

Download

http://doku.iab.de/fdz/reporte/2019/DR_06-19_EN.pdf

Documentation version

LIAB_QM2_9317_EN_v1_dok1, DOI: 10.5164/IAB.FDZD.1906.en.v1

Dataset version

LIAB_QM2_9317_v1, DOI: 10.5164/IAB.LIABQM29317.de.en.v1

All publications in the series 'FDZ-Datenreport' can be downloaded from

https://fdz.iab.de/en/FDZ_Publications/FDZ_Publication_Series/FDZ-Datenreporte.aspx

Website

https://fdz.iab.de/en.aspx

Corresponding author

Matthias Umkehrer
Phone: 0911 179-6211
Email: Matthias.Umkehrer@iab.de