FDZ·Datenreport

Documentation of labour market data

10/2016

EN

Linked Employer–Employee Data from the IAB: LIAB Longitudinal Model 1993–2014

(LIAB LM 9314)

Jörg Heining,
Wolfram Klosterhuber,
Patrick Lehnert,
Stefan Seth

Bundesagentur für Arbeit

Linked Employer-Employee Data from the IAB: LIAB Longitudinal Model 1993-2014 (LIAB LM 9314)

Jörg Heining, Wolfram Klosterhuber, Patrick Lehnert, Stefan Seth

(Institut für Arbeitsmarkt- und Berufsforschung (IAB))

Documentation: Version 1 (LIAB_LM_9314_EN_v1_dok1)

Dataset: Version 1 (LIAB LM 9314 v1)

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

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

Contents

Abstract

This data report describes the LIAB Longitudinal Model 1993-2014 of the Linked Employer-Employee Data from the IAB (LIAB LM 9314).

Data availability

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

1 Introduction and outline

1.1 Introduction

. The LIAB Longitudinal Model 1993-2014 (LIAB LM 9314) is one of the linked employer-employee datasets from the Institute for Employment Research (IAB) provided for research analyses by the Research Data Centre (FDZ) of the Federal Employment Agency (BA) at the IAB. The Linked Employer-Employee Data from the IAB (LIAB) contain both establishment data from the annual waves of the IAB Establishment Panel and information on individuals from the process-generated data of the BA. The LIAB data are generated by linking these two data sources.

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, too. The IAB Establishment Panel is the central basis for the analysis of labour demand in Germany. The establishments of the IAB Establishment Panel are the basis for the sampling of the individual data from the Integrated Employment Biographies (IEB) of the IAB.

Data on individuals were taken from the Integrated Employment Biographies (IEB) of the IAB. The IEB comprises all individuals who showed one of the following statuses at least once during the observation period:

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

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

Whilst the Employee History (Beschäftigtenhistorik - BeH) is the origin of the information on employment subject to social security and marginal part-time employment, the receipt of benefits in accordance with Social Code Book III (SGB III) and Social Code Book II (SGB II) is recorded in the Benefit Recipient History (Leistungsempfängerhistorik - LeH) and the Unemployment Benefit II Recipient 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 (Maßnahmeteilnahmehistorik - MTH).

Data on individuals in the LIAB LM 9314 only include information from the sources Employee History (BeH) and Benefit Recipient History (LeH). Therefore the receipt of benefits under the sphere of Social Code Book II, which covers Unemployment Benefit II Recipient History (LHG and XLHG) as well as information from the Jobseeker Histories (ASU and XASU) and the Participation-in-Measures History File (MTH) are not included in the individual data of the LIAB LM 9314.

The data report focuses on the description of data on individuals in the LIAB LM 9314. Information on the IAB Establishment Panel as well as on the different LIAB datasets can be found on FDZ's homepage: http://fdz.iab.de.

The data report on the LIAB LM 9314 is structured as follows. Besides the introduction, Chapter 1 contains information on data access, as well as, an outline of the data, the volume structure and a list of variables. A description of the individual data sources can be found in Chapter 2. Data preparation and data quality are discussed in Chapters 3 and 4, whilst the individual variables are described in Chapter 5.

1.2 Data use

1.2.1 Data access

The weakly anonymous LIAB data may only be analysed on-site at the FDZ and through subsequent remote data access.

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

Certain variables which make it possible to identify individuals or establishments are only disclosed in their original form (e.g. *wo_bula*). If it is necessary for the analysis objective and justified explicitly in the application for data access, these variables can be provided, too. The

variables in the LIAB LM 9314, which are particularly sensitive from the viewpoint of data protection legislation are:

Individual data:
- nationality (nation)
- date of birth (gebdat)
- occupational sub-group (beruf2010_4)
- place of residence: employment agency area (wo_aa)
- place of residence: district (wo_kreis)

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

IAB Establishment Panel:
- type of local authority (BIK-type of municipality, since 1998)
- district key number (since 1998)
- economic activity – sub-class of economic activity (wz93)

  sub-class of economic activity for 2000, 2001, 2003; economic group for 2002
- economic activity – sub-class of economic activity (wz03)
- economic activity – sub-class of economic activity (wz08)

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). The first module contains the waves of the IAB Establishment Panel. In these waves you can find information about job trends, business policies and development of business, company investments, innovations in business, public funding, personnel structure, vocational trainings and training posts, worker

---

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

inflow and outflow, personnel search, wages and salaries, working hours, further trainings or general data of the company.

The individual data is stored in the second module. This file contains identifiers (artificial individual IDs and artificial establishment numbers), the personal variables, the information on employment and benefit receipt, the variables regarding place of residence as well as technical and biographical variables.

The third module, the Establishment File, contains the establishment number, the year, and variables regarding the place of work and economic activities in aggregate form as well as other non-sensitive establishment information as of the reference date of 30 June. The Establishment File is extracted from the Establishment History Panel (BHP). The BHP is a 50% sample of all establishments in Germany, which cover at least one employee liable to social security on the 30 June of a year. Since 1999 marginal part-time employees are included, too. The source of the BHP is the Employee History (BeH) of the Institute for Employment Research (IAB). In the BHP individual data of the BeH is aggregated on the establishment level via the establishment numbers. Further information on the BHP can be found on the homepage of FDZ. Variables, which are marked with a "*" in the list of variables (see Table 4), are stored in the Establishment File. Extension files from the BHP can be linked to the Establishment File. Following a justified application further establishment variables from the BHP can be provided.
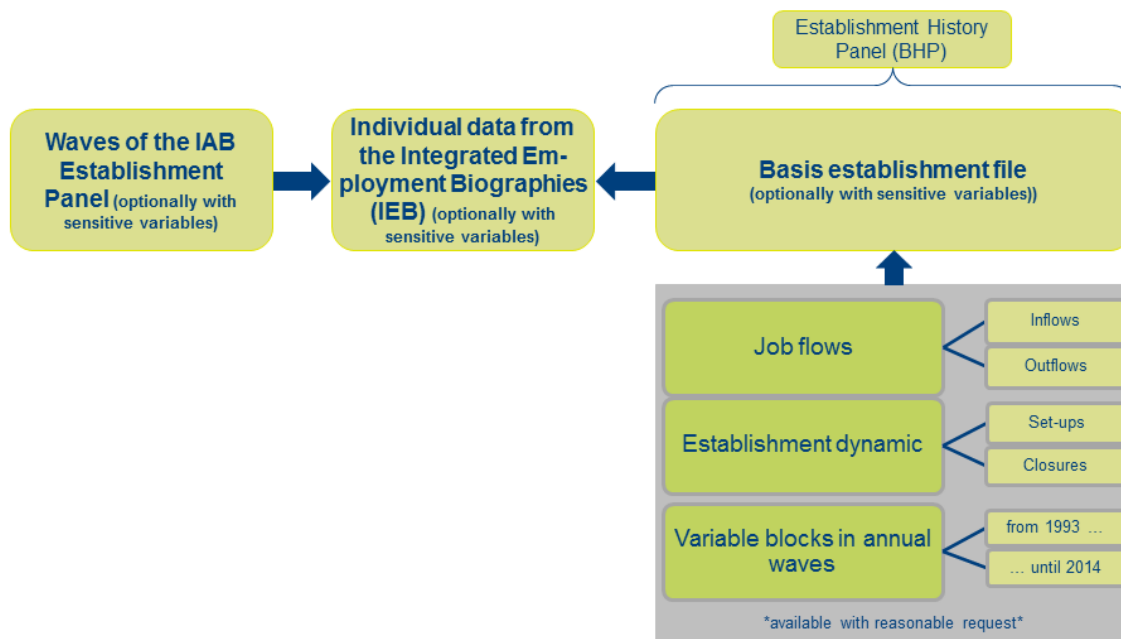
Figure 1: Data management of the Linked Employer-Employee Longitudinal Model dataset.

Establishment variables and individual variables are therefore organized separately, which makes the structure of the data clear and saves storage space in the data management system. The linkage between the waves of the IAB Establishment Panel and the individual data is based on the establishment number, which are included in both modules (*idnum*; see Box 1). In contrast, the Establishment File and individual data are linked by an artificial establishment number (*betnr*; see Box 2) and the year of the dataset. The same applies for the extension files of the BHP. The linkage between single files is provided by different program-specific commands for data formatting – and analysis of the used software packages. For example, the linkage in Stata can be conducted by the "`merge`"-command.

All sensitive variables listed in the previous section are located in the corresponding files, as long as the files have been requested and approved.

```
use LIAB_lm_9310_v1_pers.dta
sort idnum
merge m:1 idnum using iabbp_1993.dta
```

Box 1: Example for Stata 14; Linkage between the individual data and a single wave of the IAB Establishment Panel (here 2000)

1.2.3   File names of LIAB LM 9314

Waves of the IAB Establishment Panel

iabbp_*jjjj*.dta, *jjjj* = 1993 – 2014

Individual data

LIAB_lm_9314_v1_pers.dta

Establishment File

LIAB_lm_9314_v1_bhp_basis_v1.dta

Extension Files BHP

Core dataset

LIAB_lm_9314_v1_bhp_v1_*jjjj*.dta,   *jjjj* = 1993 - 2014

Worker flows

LIAB_lm_9314_v1_bhp_inflow_v1.dta

LIAB_lm_9314_v1_bhp_outflow_v1.dta

Entry and exit

LIAB_lm_9314_v1_bhp_entry_v1.dta

LIAB_lm_9314_v1_bhp_exit_v1.dta

1.3 Developments compared to former versions

1.3.1   Observation period

The LIAB LM 9314 differs from the LIAB LM 93310 data through the covered time periods of the individual and establishment data.

1.3.2   Sampling Procedure

A more recent time period was chosen for the selection to the so-called panel cases from the IAB Establishment Panel. The LIAB LM 9310 was based on the panel cases of the period 2000 to 2008. In contrast, the panel cases from 2003 to 2011 have been chosen for the LIAB LM 9314.

1.3.3 Introduction of the occupation code 2010

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

As the notifications made by employers in accordance with DEÜV go into the Employee History (BeH), the change of the occupation code impacts almost exclusively on observations from this source. The measurement of the following characteristics already reported using the occupation code 2003 is affected by this change: working hours, occupation, occupational status and school and vocational qualification level. In addition, since the switch to the new occupation code, details are also available about whether an employment relationship is fixed-term and whether a person is employed by a temporary work agency to be hired out to other firms. The most important consequence is the switchover to a new occupational 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.[2] Another important change is the discontinuation of the variable 'occupational status and working hours' (stib), as most of the values given in this variable can no longer be traced using the new occupation code. The categories of the previous stib variable that are available consistently across the entire observation period (employment status as an apprentice, distinction between part-time and full-time) can be traced in other variables in future. Section 1.3.4 shows which new variables are added as a result of the switch to the new occupation code, which variables are discontinued and which are contained in the data in a different form in future.

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

---

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

1.3.4 Set of variables

For reasons of data economy and due to lack of demand, some variables that were contained in the LIAB LM 9310 are no longer included.

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

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

Variable	Type of change	Explanation
	D	Discontinued without replacement
Place of work: employment agency (ao_aa)	D	Discontinued without replacement; relevant information contained in wo_aa
Place of work: regional directorate (ao_rd)	D	Discontinued without replacement; relevant information contained in wo_rd
Contract period (befrist)	N	Introduced with new occupation code; available from 2011 onwards
Occupational group – current/most recent (KldB 2010), 3-digit (beruf2010_3)/ 4-digit (beruf2010_4)	N	Introduced with new occupation code; available for entire period by recoding old variable.
School education and vocational training (bild)	D	Replaced by vocational education and training
Type of benefit (lart)	D	Discontinued without replacement
Employment status 2, 8 days before first day in establishment (estat2_8tv)	D	Discontinued without replacement
Employment status 8 days before first day in establishment (estat8tv)	D	Discontinued without replacement
Date of birth (gebdat)	N	In addition to the year of birth a new variable containing the year and the month of birth has been added to the LIAB LM 9314. This new variable is treated as sensitive information.
Type of benefit (lart)	D	Discontinued without replacement
Temporary work (leih)	N	Introduced with new occupation code; available from 2011 onwards

Variable	Type of change	Explanation
Daily wage 8 days before first day in establishment (lohn8tv)	D	Discontinued without replacement
Level of requirement – current/most recent (KldB 2010) (niveau)	N	Introduced with new occupation code; available for entire period by recoding old variable.
Occupational status and working hours (stib)	D	Discontinued because no longer recorded in notification procedure; distinction between full-time and part-time can be traced in variable 'part-time'; employment status "apprentice" recorded in variable 'erwstat'.
Number of days in previous job or unemployment (tage_estat2_8tv)	D	Discontinued without replacement
Number of days of benefit receipt (tage_lst)	N	Instead of two variables for the number of days with benefit receipt there is only one left (*tage_lst*). It corresponds to tage_lst2 in the LIAB LM 9310 with the difference, that only periods of unemployment benefits are included in the calculation. For example, unemployment assistance is no longer considered.

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

1.4 Outline

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. **IAB Establishment Panel** Information from the annual Establishment Panel survey with shifting priorities.
Data unit	Establishments, Employees covered by social security (including marginal part-time employees from 1999 onwards), benefit recipients
Number of cases	Linked establishments: 2.436 to 11.868 per year, 192.323 in total Individuals: 1.006.028 to 1.533.327 per year, 1.918.086 in total
Period covered	The period covered depends on the data source. BeH 01 January 1975 - 31 December 2014 (2012: 30-month file, 2013: 18-month file; 2014: 6-month file) LeH 01 January 1975 - 31 December 2014
Time reference	Daily exact information
Date of territorial allocation	31 December 2010
Institutions involved	Social security agencies, Federal Employment Agency (Bundesagentur für Arbeit), TNS Infratest Munich, SÖSTRA Institute Berlin (Institute for Socio-economic Structural Analysis)
Update frequency	Individual/Establishment data: continuous IAB Establishment Panel: annually

File format and size	Stata
	 Establishment data:
	      IAB Establishment Panel: about 1.5 to 11 MB per year
	      Establishment file: about 323 MB
	 Individual data: about 3.8 GB

File organisation	The LIAB data on individuals are stored in one file.

	A file with establishment information taken from the BHP (the Establishment File) is stored in long format and can be linked to the data on individuals by using the *betnr* and *year* variables.
	The IAB Establishment Panel files remain unchanged as one file per wave and can be linked to the data on individuals by using the *idnum* variable.
Data access	On-site use at the FDZ of the BA at the IAB and subsequent remote data access
Degree of anonymisation	Weakly anonymous
Sensitive variables	Individual file: Nationality (nation), Date of birth (gebdat), (beruf2010_4), Place of residence: district (wo_kreis), Place of residence: employment agency area (wo_aa) Establishment file: Place of work: district (ao_kreis), Economic activity - sub-class of economic activity (five-digit code) (w93_5), Economic activity - sub-class of economic activity (five-digit code) (w03_5), Economic activity - sub-class of economic activity (five-digit code) (w08_5), First appearance of establishment (grd_dat), Last appearance of establishment (lzt_dat) IAB-Betriebspanel: type of local authority (BIK-type of municipality, since 1998), district key number (since 1998), Economic activity 1993 (wave 2000, 2001, 2003 5-digit code, wave 2002 3-digit code of WZ93), Economic activity 2003 (up to wave 2004 5-digit code of WZ03), Economic activity 2008 (up to wave 2009 5-digit code of WZ08)
Citation method	Linked Employer-Employee Data from the IAB, Longitudinal Model 1993-2014 (LIAB LM 9314), Nuremberg 2016 Data: "The data basis of this paper is the Longitudinal Model (version 1993 – 2014) 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/or via remote data access at the FDZ." Data documentation: Heining, Jörg; Klosterhuber, Wolfram; Lehnert, Patrick; Seth, Stefan (2016): Linked Employer-Employee Data from the IAB: LIAB Cross-sectional model 2 1993 – 2014 (LIAB QM2 9314). FDZ-Datenreport, 10/2016 (en), Nuremberg.

Current data version	Linked Employer-Employee Data from the IAB, longitudinal model 1993 – 2014 (LIAB LM 9314)

1.5 List of Variables

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

Table 3: Degrees of completeness of the Variables

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

Example: The variable "Occupation – current/most recent (KldB 1988) (beruf) is only available for BeH observations. The observations of the data source LeH contain ".n" in this variable. Another characteristic is that some variables have different contents depending on the data source. For instance, for BeH observations the employment status variable (erwstat) contains the person group of the employment notification procedure, for LeH observations it contains the type of benefit. 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	Module	Page	BeH	LeH
Identifiers		**60**		
Artificial individual ID (persnr)		32		
Artificial establishment number (betnr)	*	32		
Establishment-ID acc. to Survey (idnum)		33		
Generated technical variables		34		
Observation counter per person (spell)		34		
Source of observation (quelle)		34		
Year (jahr)	*	34		
Status of establishment number (betr_st)		34		
Period of validity		35		
Original start date of observation (begorig)		35		
Original end date of observation (endorig)		36		
Start date of split episode (begepi)		36		

List of variables	Module	Page	BeH	LeH
End date of split episode (endepi)		36	▓	▓
Personal information		**37**		
Gender (frau)		37	▓	▓
Date of birth (gebdat)		37	▓	▓
Year of birth (gebjahr)		37	▓	▓
Nationality (nation)		38	▓	▓
Nationality, aggregated (nation_gr)		38	▓	▓
Marital status (famst)		38		▓
Number of children (kind)		38		▓
Vocational training (ausbildung)		39	░	
School leaving qualification (schule)		40	░	
Information on employment and benefit receipt		41		
Daily wage, daily benefit rate (tentgelt)		41	▓	▓
Occupation – current/most recent (KldB 1988) (beruf)		42	▓	
Occupational group – current/most recent (KldB 2010) (beruf2010_3)		42	▓	░
Occupational sub-group – current/most recent (KldB 2010) (beruf2010_4)		43	▓	░
Level of requirement – current/most recent (KldB 2010) (niveau)Level of requirement – current/most recent (KldB 2010) (niveau)Level of requirement – current/most recent (KldB 2010) (niveau)		44	▓	░
Part-time (teilzeit)		45	▓	
Employment status (erwstat)		45	▓	▓
Transition zone (gleitz)		46	░	
Temporary agency work (leih)		47	░	
Fixed-term contract (befrist)		47	░	
Reason for notification / reason for end of benefit receipt / reason for discontinuation of unemployment benefit II / reason for deregistration (grund)		47	▓	▓
Duration of remaining entitlement to unemployment benefit (restanspruch)		48		▓
Location data		49		
Place of residence: district (Kreis) (wo_kreis)		49	░	░
Place of residence: federal state (Bundesland) (wo_bula)		50	░	░
Place of residence: employment agency (wo_aa)		49	░	▓
Place of residence: regional directorate (wo_rd)		50	░	▓
Establishment variables		51		

List of variables	Module	Page	BeH	LeH
Economic activity 73, 3-digit code (w73_3)	*	51		
Economic activity 93, 5-digit code (w93_5)	*	51		
Economic activity 93, 3-digit code (w93_3)	*	52		
Economic activity 03, 5-digit code (w03_5)	*	52		
Economic activity 03, 3-digit code (w03_3)	*	53		
Economic activity 08, 5-digit code (w08_5)	*	53		
Economic activity 08, 3-digit code (w08_3)	*	54		
Economic activity 73 generated – completed by extrapolation / imputation (w73_3_gen)	*	54		
Economic activity 73 generated – type of completion (group_w73_3)	*	55		
Economic activity 93 generated – completed by extrapolation / imputation (w93_3_gen)	*	55		
Economic activity 93 generated – type of completion (group_w93_3)	*	55		
Year of first appearance of establishment number (grd_jahr)	*	56		
First appearance of establishment number (grd_dat)	*	56		
Year of last appearance of establishment number (lzt_jahr)	*	57		
Last appearance of establishment number (lzt_dat)	*	57		
Total number of employees (az_ges)	*	58		
Number of full-time employees (regular workers + others) (az_vz)	*	58		
Number of employees in marginal part-time employment (az_gf)	*	58		
Mean imputed wage all full-time employees (te_imp_mw)	*	59		
Place of work: district (ao_kreis)	*	59		
Place of work: federal state (ao_bula)	*	60		
Generated biographical variables		60		
First day in employment (ein_erw)		60		
Number of days in employment (tage_erw)		61		
First day in establishment (ein_bet)**First day in establish-ment (ein_bet)**		61		
Number of days in establishment (tage_bet)		62		
First day in job (ein_job)		62		
Number of days in job (tage_job)		63		
Number of days with benefit receipt (tage_lst)		64		

1.6 Volume structure

Year	Individuals	Establishment	Establishments from the IAB Establishment Panel linked to the individual data		
	Total	Total	Total	Establishments selected for the LIAB LM 9314		Establish-ments not selected for the LIAB LM 9314
				betr_st == 1	betr_st == 2	betr_st == 3
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1993	1 006 028	255 434	2 642	406	293	2 231
1994	1 025 836	262 456	2 510	432	306	2 072
1995	1 061 647	265 788	2 436	487	349	1 941
1996	1 090 006	266 920	5 610	1 441	1 081	4 152
1997	1 128 578	270 833	5 390	1 639	1 210	3 718
1998	1 183 583	275 063	5 917	2 006	1 464	3 886
1999	1 285 338	350 838	6 605	2 545	2 053	4 034
2000	1 384 768	363 019	9 068	3 551	2 886	5 468
2001	1 440 295	352 286	10 121	4 278	3 588	5 801
2002	1 489 271	338 996	10 027	4 955	4 408	5 022
2003	1 510 280	347 889	10 662	6 163	5 737	4 401
2004	1 533 327	362 605	10 939	6 137	5 549	4 592
2005	1 481 661	367 426	10 925	6 053	5 372	4 653
2006	1 483 743	382 075	10 840	6 121	5 380	4 514
2007	1 504 035	395 536	10 898	6 281	5 502	4 421
2008	1 524 018	400 667	11 039	6 514	5 583	4 331
2009	1 521 673	407 249	11 225	6 801	5 873	4 212
2010	1 524 255	412 630	11 399	7 477	6 408	3 674
2011	1 529 924	418 379	11 868	8 115	6 905	3 490
2012	1 521 512	415 315	11 328	6 815	5 703	4 287
2013	1 487 753	414 211	10 757	5 988	4 834	4 597
2014	1 453 565	415 582	10 117	5 155	3 985	4 816

Table 5: Volume structure

Column (2) of table 5 indicates the number of individuals per year. In column (3) the establishment numbers (betnr) per year, contained in the individual data, are shown.

In contrast, columns (4) to (7) contain the total number of establishments from the IAB establishment panel that can be linked to the individual data. Column (4) displays the number of linked establishments in total. The number of linked establishments according to the variable "Status of establishment number" (*betr_st,* see section 5.2.4) is shown by columns (5) to (7).

Thereby column (5) contains all establishments, that have been originally selected for the LIAB LM 9314 (see also 3.3 Sampling procedure). These so-called panel cases have been selected

from the IAB Establishment Panel for the period 2003 to 2011 (see section 3.3). The displayed numbers refer to linkages of establishments which

- have been successfully interviewed (on 30 June of the respective year)

- and for which the observed episode includes this day of reference.

In contrast, column (6) presents the number of linked establishments which

- have been successfully interviewed (on 30 June of the respective year)

- but for which the observed episode excludes this day of reference.

Finally, column (7) shows the number of establishments from the IAB Establishment Panel which are not classified as panel cases in the survey for the period 2003 to 2011 but which can be linked to the individual data though.

2   Data sources

The administrative individual data were drawn from the Integrated Employment Biographies (IEB) of the IAB. They unite data from five different data sources, each of which may contain information from different administrative procedures.
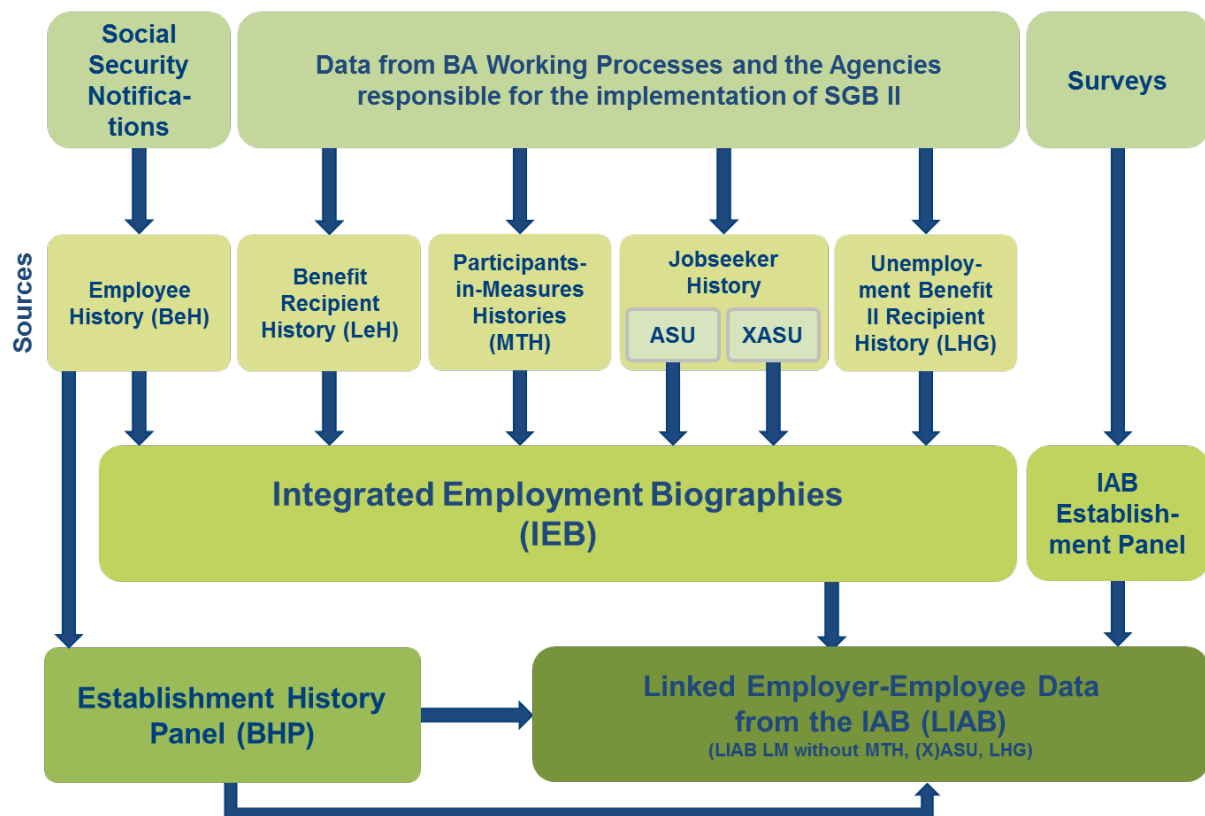
Figure 2: Data sources of the IEB

In the individual data of the LIAB Longitudinal Model 9314 only the datasets from the sources Employee History (BeH) and Benefit Recipient History (LeH) from the Integrated Employment Biographies (IEB) are adopted.

2.1 Employee History (BeH)

The source of data regarding employment is the Employee History (BeH) of the IAB. The data basis is the integrated notification procedure for health, pension and unemployment insurance, which came into effect as of 01 January 1973 (and was extended to cover Eastern Germany as of 01 January 1991) and is known by the abbreviation DEÜV (previously DEVO / DÜVO) (for further details see: Bender et al. 1996, p. 4 et seq.; Wermter /Cramer 1988). Under this procedure employers are required to submit notifications to the responsible social security agencies concerning all of their employees covered by social security at least once a year. The BeH covers all white- and blue-collar workers as well as apprentices as long as they are not exempt from social security contributions. This means that civil servants, self-employed persons and regular students (see Cramer 1985) are in principle not recorded in the BeH. Since the notification procedure was changed on 01 January 1999, employees in marginal part-time employment and unpaid family workers have also been recorded (not contained in the data until 01 April 1999). The data are recorded by the health insurance companies, collected and edited by the Federal Employment Agency (BA) and subsequently integrated into the History File by the IAB.

The administrative individual data are supplemented with establishment data (basic establishment data and BHP extension file). They are taken from the Establishment History Panel (Betriebs-Historik-Panel - BHP), which is also based on the BeH.

When linking individual data with establishment data it has to be taken into account that the variables in the Basis Establishment File as well as in the BHP Extension Files are aggregated on 30 June of a year. Unlike the data on individuals, the establishment variables are therefore not spell data but are only valid on 30 June precisely each year.[3]

The Establishment File is linked with the Individual File via the programme-specific commands of the software packages used for preparing and analysing the data. In Stata, for instance, the two files can be linked using the "`merge`" command in connection with the relevant paths (see Box 2).

---

[3] An extreme example: an employment notification exists from 01 January 2006 to 30 May 2006; the establishment goes bankrupt in June 2006. There is then no information about this establishment in the BHP for 2006.

```
use liab_lm_9314_v1_pers.dta
sort betnr jahr
merge m:1 betnr jahr using LIAB_lm_9314_v1_bhp_basis_v1.dta
```

Box 2: Example code for Stata14; Linkage of the individual data with the Establishment File

2.2 Benefit Recipient History (LeH)

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

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 without information on the date of birth are deleted. Further corrections, like in the IABS (Completion procedure of presumably missing notifications, strike corrections) do not happen.

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. To be precise, from 1999 onwards only the person groups 101, 102, 103, 105, 106, 112, 118, 119, 120, 140, 141, 142, 143, 149, 201, 203 and 205 are supplied, as well as the two groups 109 and 209, which indicate people in marginal part-

time employment. Groups that are not included from 1999 onwards include people in short-term employment, i.e. person groups 110, 202 and 210.

– Observations with earnings amounting to zero or with no details on earnings, and the value 101 for the person group variable, and the value 50 for the reason for notification (annual notification) are not incorporated into the IEB.

– Gender and date of birth are taken from the social security number. This information is harmonised across data sources.

– The territorial allocations for place of work and place of residence are updated to the status as of 31 December 2010.

3.1.2 Benefit Recipient History (LeH)

– It must be taken into account that in cases of ongoing receipt the end date corresponds to the end of the period for which the benefit is allowed and not the end of the benefit receipt period. LeH observations which go beyond 14 July 2011 thus depict the period for which the benefit is allowed. Observations with no end date or an invalid end date are excluded from further processing, since in these cases it cannot be assumed that a benefit payment was made at all.

– Observations without a valid start date are excluded.

– Observations whose end date precedes the start date are excluded.

– If the end date for the receipt of unemployment assistance precedes the start date by one day, then the end date is increased by one year.

– The territorial allocations are corrected analogous to the BeH.

3.2 Episode splitting

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

persnr	begorig	endorig	level
4711	01.01.98	31.12.98	-
4711	01.10.98	30.06.99	-
4711	01.01.99	31.12.99	-

Episode Splitting
Splitting of the time intervals into "non-overlapping" intervals.

persnr	begepi	endepi	level
4711	01.01.98	30.09.98	0
4711	01.10.98	31.12.98	0
4711	01.10.99	31.12.98	1
4711	01.01.99	30.06.99	0
4711	01.01.99	30.06.99	1
4711	01.07.99	31.12.99	0
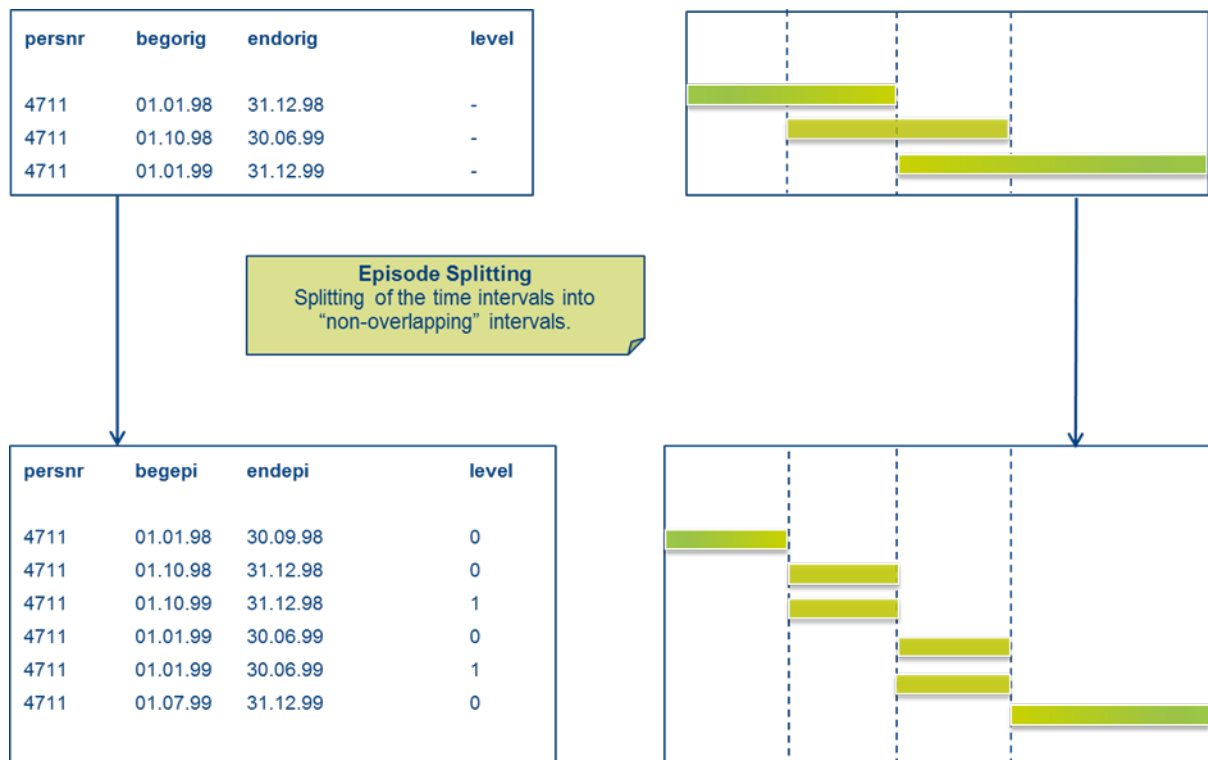
Figure 3: Episode splitting

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

3.3 Sampling Procedure

The LIAB is sampled on the basis of the establishments surveyed in 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.

In a first step all so-called panel cases for the time period 2003 to 2011 are selected for the LIAB LM 9310. Panel cases comprise

- the stock of establishments in the IAB Establishment Panel in the year 2003,

- "new" establishments in the subsequent waves,
- and establishments going out of operation during the panel period.

A prerequisite for the selection as a panel case is a valid interview in the following years or that the establishment goes out of operation.[4] It is important to bear in mind that not all cross-sectional cases of one wave of the IAB establishment panel are panel cases simultaneously. If one establishment was added in the course of the reserve sample, which is drawn from the stock records, or the same establishment unit wasn't surveyed as in the previous year, the corresponding case is not a panel case (see Fischer et al. (2008), p. 20f).

In a second step, all individuals, who were employed in one of these establishments at least one day during the time period 2002 to 2012, are identified. In the third step the employment biographies from the BeH and LeH for these individuals are drawn. The employment biographies cover the time period from 1993 to 2014. Figure 4 illustrates the sampling procedure.

Time period of establishments selected from the IAB Establishment Panel; 2003 - 2011

Period of employment of individuals in the selected establishments; 2002 - 2012

Time period of employment biographies; 1993 – 2014

1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014
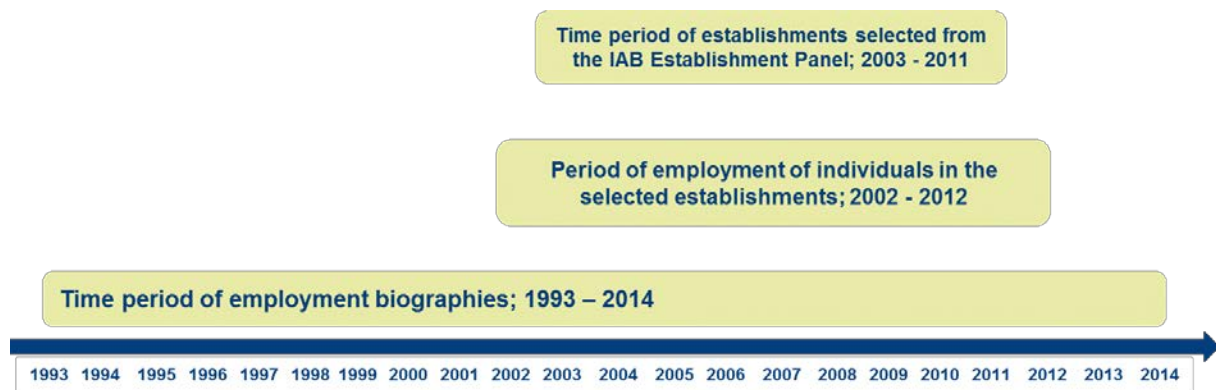
Figure 4: Sampling procedure of the LIAB LM 9314

3.4 Missing values

In the LIAB, missing values are coded as follows:

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

---

[4] See Fischer et al. (2008), p. 26f.

4 Data quality and problems

4.1 Entire IEB

The IEBs contain employment histories. Not each kind of employment is included in the administrative data, though. Certain persons with certain courses of life are not represented in the IEBs at all.

For evaluation purposes, it is often relevant to know gaps in the included biographies (e.g. for control group creation, course of life analyses, etc.) The gaps listed in the following are defined as periods of time after the end of school education for which no data is included in the IEBs. These gaps can be divided into

- gaps with no information available, and
- gaps for which information may be available from the reason for notification / reason for end of benefit receipts / reason for discontinuation of SGB II / reason for deregistration characteristic of the direct preceding dataset (if a corresponding dataset exists).

These gaps were identified using the manifestation of the reason for notification / reason for end of benefit receipts / reason for discontinuation of SGB II / reason for deregistration and employment status characteristics in the various sources. The list makes no claims of being complete.

Biographical gap	Information on gap, identifiable using the manifestation of the "grund" characteristic in the preceding data-set of the source, if necessary
Civil servants, professional soldiers, judges, employees of bodies or foundations under public law, unsupported self-employed persons	LeH
Students, people in school-based further education	LeH
People who are ill / not able to work for more than 6 weeks	BeH, LeH
People receiving old-age pension without employment if not member of a benefit community	LeH
Recipients of early retirement benefits	LeH
Trade professionals working from home	
(Sideline) farmers	
Caregivers according to Section 19 SBG XI	
Conscripts	BeH, LeH
Persons in reserve duty training	BeH, LeH

Biographical gap	Information on gap, identifiable using the manifestation of the "grund" characteristic in the preceding data-set of the source, if necessary
Persons fulfilling community service	BeH, LeH
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.)	LeH, BeH
Strikers in case of a duration of the strike of more than a month	LeH
Social benefit recipients (prior to the introduction of SGB II in 2005), social allowance recipients (according to SGB II)	
Compensation recipients according to FELEG (Gesetz zur Förderung der Einstellung der landwirtschaftlichen Erwerbstätigkeit, Act on the Support in Case of Termination of Farming Activities)	

Table 2: Biographical gaps and identification possibilities

4.2 Employee History (BeH)

–   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 great deal of spells in 1991 show missing values for several variables (such as employment status).

–   The increase in the number of BeH observations from 1999 onwards is due to the introduction of the obligation to submit employment notifications for people in marginal part-time employment from 01 April 1999 onwards.

–   Especially in 1999, observations with 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 properly.

–   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 of the following year, some notifications actually arrive years later. The History File

of the IAB is not updated continuously, however, but at certain intervals. This is done using files of employment notifications for one particular year which were submitted 36, 18, 12 or 6 months after the end of the reporting year (e.g. the 18-month file for 2007 can be created in July 2009 at the earliest). Notifications submitted more than three years late are not taken into account at the IAB, which means that a 36-month file shows a 100 % degree of completeness by definition.

– In the version of the IEB on which the LIAB data are based the degree of completeness of the BeH observations last stood at 100 % in 2007. 18-month files were used for 2008 and 2009, and the observations for 2010 come from a 6-month file. It can therefore be assumed that the BeH observations for 2008, 2009 and in particular for 2010 are slightly underreported. 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 has increased considerably from that year onwards (see Bender et al. 1996).

– For the years 1992 until 2000 noticeable drops and rises in the number of notifications were noticed. Drops 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 single or several establishments in these regions.

– Considerable drops were also observed for the districts Salzgitter (03102) und Hoyerswerda (14264).

– Concerning the notifications for full-time employment, especially the districts Main-Taunus (06436) und Alzey-Worms (07331) are noteworthy. They feature above-average rises. Also in this case, the reasons are notification problems at single or several establishments in these regions.

– In the years 1996 to 1998, the values 841-844 (doctors and pharmacies) within the "occupation – activity performed" variable are very rare compared to the neighbouring years. The reasons for this are not known.

4.3 Benefit Recipient History (LeH)

- Caused by an internal change of systems, there is a break in the collection of periods of exclusion of benefits and of benefit suspension in 2004.

5 Description of variables

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

5.1 Identifiers

5.1.1 Artificial individual ID (persnr)

Variable label	Individual ID
Variable name	persnr
Category	identifiers
Origin	BeH, LeH
Data type	numerical
Hierarchy	none
Detailed description	The individual ID indicates which observations belong to the same person. Artificial means that it is not possible to infer any of the person's characteristics or any original identifiers from this individual ID. 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 Artificial establishment number (betnr)

Variable label	Establishment ID
Variable name	betnr
Category	identifiers
Origin	BeH
Data type	numerical
Hierarchy	none
Detailed description	The establishment ID indicates which observations belong to the same establishment. It is based on the establishment number allocated by the BA, which was replaced by an artificial number. (Further information on the allocation of establishment numbers by the BA can be found in Bender et al. 1996: p. 15 et seq. and pp. 27-30.) The establishment number and year specification can be used to merge individual and establishment information.

<table>
<tr><td></td><td>For the establishment number, the following should be observed in general:

a) If the company has one office only, or if the company has one office only in one municipality, this office is the establishment and is given an establishment number.

b) If the company has **several** branch offices in **one** municipality, these establishment premises / workplaces must be merged into a single establishment under one establishment number, if they belong to the same economic class. If they do **not** belong to the same economic class, each branch office is regarded as a separate establishment and is given its own establishment number.

c) If the company has **several** branch offices in **several** municipalities, each of these branch offices is an establishment and is given its own establishment number.

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:

a) An **establishment** is a regionally and economically delimited unit in which employees work and which is allocated an establishment number according to the above-mentioned principles.

b) A **workplace** is a unit in which employees work and which is not allocated an establishment number according to the above-mentioned principles.

c) A **company** as a term combines establishment premises and workplaces belonging to the same employer.

d) An **employer** is any natural person or legal entity that employs at least one employee subject to social security contributions or in marginal part-time employment.

e) Establishment and establishment premises are synonyms; branch office is a synonym for subsidiary, district office, outsourced office, workplace etc. if it is not an establishment.</td></tr>
<tr><td>Note on quality</td><td>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.</td></tr>
</table>

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	Numeric
Hierarchy	None
Detailed description	The Establishment ID acc. to Survey is adopted from the IAB Establishment Panel. The variable is only filled in the Individual File, 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 Observation counter per person (spell)

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

5.2.2 Source of observation (quelle)

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

5.2.3 Year (jahr)

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

5.2.4 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
Hierarchy	none
Detailed description	The variable indicates for the current episode in the individual data if there is information from the IAB Establishment Panel for this episode.

1 – Employment in establishment of Establishment Panel with interview in same year, establishment selected for LIAB model
2 – Employment in establishment of Establishment Panel without interview in same year, establishment selected for LIAB model
Establishment chosen for LIAB-model
3 – Employment in establishment of Establishment Panel, establishment not selected for LIAB model
4 – Employment in establishment not of Establishment Panel

The values 1-3 indicate employments in establishments, which have taken part in the IAB Establishment Panel at least once, whereas value 4 marks the employment in establishments, which have never taken part in the IAB Establishment Panel.
1 and 2 describe employments in establishments, which are in the IAB Establishment Panel, which has been chosen for the respective LIAB model. Number 3 defines employments in firms, which have not been chosen for the respective LIAB model.

The difference between 1 and 2 depends on the fact, whether the episode includes the 30 June of the respective year, for a firm that has been interviewed, or not.

Value 3 does not appear by definition in a Cross-sectional Model of the LIAB, because all firms of the IAB Establishment Panel have been interviewed at least once and due to construction chosen for the LIAB Cross-sectional Model. |

5.3 Period of validity

5.3.1 Original start date of observation (begorig)

Variable label	Original start date
Variable name	begorig
Category	period of validity
Origin	BeH, LeH
Data type	date
Hierarchy	none
Detailed description	The original start date of the observation corresponds to the original start date of the notification. This can differ from the start date of the episodes (begepi) (see also the comments on episode splitting in Section 3.2)

	Because of the rules of the notification procedure, in BeH observations the starting and ending year are always identical (obligation of the employer to submit annual employment notifications). A continuous employment relationship may therefore be distributed across several notifications.

5.3.2   Original end date of observation (endorig)

Variable label	Original end date
Variable name	endorig
Category	period of validity
Origin	BeH, LeH
Data type	date
Hierarchy	none
Detailed description	The original end date of the observation corresponds to the original end date of the notification. This can differ from the end date of the relevant line of data, the so-called end date of the split episode (see also the comments on episode splitting in Section 3.2).
	Because of the rules of the notification procedure, in BeH observations the starting and ending year are always identical (obligation of the employer to submit annual employment notifications). A continuous employment relationship may therefore be distributed across several notifications.

5.3.3   Start date of split episode (begepi)

Variablenlabel	Episode start date
Variablenname	begepi
Kategorie	generated period of validity
Herkunft	BeH, LeH
Datentyp	date
Hierarchie	none
Detailbeschreibung	The start date of the split episode is always equal to or greater than the start date of the original observation (see also the comments on episode splitting in Section 3.2).

5.3.4   End date of split episode (endepi)

Variablenlabel	Episode end date
Variablenname	endepi
Kategorie	generated period of validity
Herkunft	BeH, LeH
Datentyp	date
Hierarchie	none
Detailbeschreibung	The end date of the split episode is always equal to or smaller than the end date of the original observation (see also the comments on episode splitting in Section 3.2).

5.4 Personal information

5.4.1 Gender (frau)

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

5.4.2 Date of birth (gebdat)

Variable label	Date of birth
Variable name	gebdat
Category	personal variables
Origin	BeH, LeH
Data type	numerical
Hierarchy	none

Detailed description	The variable contains year and month of the date of birth. The format is YYYYmM (e.g. 1984m6). It is constant within one individual account. If you want to extract the year and/or month of the date of birth, you can use the following commands:
	
	`gen int temp = dofm(gebdat)`
	`gen int year = year(temp)`
	`gen byte month = month(temp)`
	`drop temp`

Note 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. Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available on application and only in well-founded cases. Otherwise, only the year of birth (*gebjahr*) is provided.

5.4.3 Year of birth (gebjahr)

Variable label	Year of birth
Variable name	gebjahr
Category	personal variables
Origin	BeH, LeH
Data type	numerical
Hierarchy	none
Detailed description	The year of birth is constant within one individual account.

Note on quality	In the original data, it may happen that the year of birth changes between the data sources. This is corrected during the data preparation process. The information from the social security number is given highest priority here.

5.4.4 Nationality (nation)

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

5.4.5 Nationality, aggregated (nation_gr)

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

5.4.6 Marital status (famst)

Variable label	Marital status
Variable name	famst
Category	personal variables
Origin	LeH
Data type	numerical
Hierarchy	none
Detailed description	This variable describes the marital status. The characteristic in the LeH source has only two manifestations (0 - not married, 1 – married).

5.4.7 Number of children (kind)

Variable label	Number of children
Variable name	kind
Category	personal variables
Origin	LeH

Data type	numerical
Hierarchy	none
Detailed description	The LeH source indicates the number of children aged under 16 at the time when the application was made. It only distinguishes between whether the number of children is zero or at least one. The following values are therefore available: 0 no children 100 one or more children The variable is not updated when there are changes in the type of benefit or the approval of benefits, but only when a new case of benefit receipt occurs after a period of employment. This restricts the quality of the data.

5.4.8 Vocational training (ausbildung)

Variable label	Vocational training
Variable name	ausbildung
Category	personal variables
Origin	BeH
Data type	numerical
Hierarchy	none
Detailed description	For BeH observations the variable contains the vocational education and training reported by the employers in the employment notification procedure. The following values exist: 1 No vocational training 2 In-company voc. training/traineeship/external voc. training 11 Degree from a university of applied sciences 12 University degree In notifications using the new occupation code (see Section 1.3.3) 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. "Changes in the vocational training status frequently occur at the same time as a change of establishment. This is because the notification data are compiled anew in the new firm. If, for example, an employee has gained a higher qualification via a part-time further training course while still working then this change of status is probably not recorded until he/she joins a new firm. It can generally be assumed that when a person is employed in a firm for a longer period, the personal data that they reported when they joined the firm is simply continued." (Meinken / Koch 2004, p. 63).

	A method for correcting missing values or temporal inconsistencies in the education and training data in the predecessor sample, the IABS, can be found in Fitzenberger et al. (2006) and in Drews (2006). Only the BeH data source is used for this, however, as this was the only data source with information on education and training in the IABS.
Notes on quality	The number of missing values increases continuously over time. In the most recent data more than 40% of values are missing. Missing values occur especially frequently in the following person groups: persons in marginal part-time employment, persons working part-time, foreign employees and eastern German workers. The reason for this is that the variable is not of particular importance as regards social security (see Meinken/Koch, 2004, p. 63).

5.4.9 School leaving qualification (schule)

Variable label	School leaving qualification
Variable name	schule
Category	personal variables
Origin	BeH
Data type	numerical
Hierarchy	none
Detailed description	The possible values change when the new occupation code is adopted (see Section 1.3.3). The values with the old occupation code are: 5 Grade-/lower secondary school with or without leaving certificate, intermediate school leaving certificate or equivalent qualification 8 Upper secondary school leaving certificate from a specialised upper secondary school (Fachoberschule), general upper secondary school leaving certificate, A-level equivalent, qualification for university 9 General upper secondary school leaving certificate, A-level equivalent, qualification for university 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 Upper secondary school leaving certificate from a specialised upper secondary school/general upper secondary school leaving certificate, A-level equivalent, qualification for university
Notes on quality	The degree of completeness has been decreasing continuously over time in the BeH and has stood at under 2/3 in recent years

5.5 Information on employment and benefit receipt

5.5.1 Daily wage, daily benefit rate (tentgelt)

Variable label	Daily wage/daily benefit
Variable name	tentgelt
Category	information on employment and benefit receipt
Origin	BeH, LeH
Data type	numerical
Hierarchy	none
Detailed description	**1) BeH** In BeH observations, this variable shows the employee's gross daily wage. It is calculated from the fixed-period wages reported by the employer and the duration of the (unsplit) original notification period in calendar days. The daily wage is shown in euros. 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 1 April 1999, earnings below the marginal part-time income threshold have also been recorded; the upper earnings limit still applies as the upper ceiling. In some cases, however, the reported earnings nonetheless exceed the upper earnings limit. Generally, this can probably be attributed to the payment of annual 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 information, however, may be considered less error-prone due to its insurance relevance).The marginal part-time income threshold and the upper earnings limit for statutory pension insurance differ from year to year as well as between eastern and western Germany (the decisive factor is the location of the establishment). An overview of these limits and thresholds can be found under http://fdz.iab.de. 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. 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.

<table>
<tr><td></td><td>

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 start date prior to 1 January 1998 the daily benefit rate applies to working days, while for observations with an original start date from 1 January 1998 onwards it applies to calendar days.

Since 1 January 2005, a daily benefit rate reported as 0 euros can be put down to benefit suspension periods or interruptions of benefit payments. If a reason for end of benefit is reported for an observation with a daily benefit rate equal to 0, then it is a notification of interruption of benefit payments. In the case of observations that reflect a period of benefit suspension, the entitlement is the same as before the start of the benefit suspension period.

</td></tr>
</table>

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

Variable label	Occupation - current/most recent (KldB 1988)
Variable name	beruf
Category	employment / benefit receipt
Origin	BeH
Data type	numerical
Hierarchy	none
Detailed description	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 30.11.2011 are reported using the new occupation code 2010 (KldB2010) (see Section 1.3.3). These values are recoded to the KldB1988 by transferring the key area. This results in inaccuracies.
Note on quality	There is a considerable increase in the number of missing values in 2011 due to the change in the reporting procedure.

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

Variable label	Occupational group - current/most recent (KldB 2010), 3-digit
Variable name	beruf2010_3
Category	information on employment, benefit receipt and job search
Origin	BeH, LeH
Data type	numerical
Hierarchy	none

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 1300 values. The less detailed occupational group is recorded by the first three digits of the code. The skill level required for a job, which is recorded in the fifth digit of the codes in the KldB2010, is made available separately in the variable 'level of requirement' (*niveau*). Employment notifications with an end date earlier than 30.11.2011 are reported using the old occupation code 1988 (KldB 1988) (see Section 1.3.3). 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** The variable contains the occupation of the last job. See 1) with regard to the occupation code.
Note on quality	There is a considerable increase in the number of missing values in 2011 due to the change in the reporting procedure.

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

Variable label	Occupational sub-group - current/most recent (KldB 2010), 4-digit
Variable name	beruf2010_4
Category	information on employment, benefit receipt and job search
Origin	BeH, LeH
Data type	numerical
Hierarchy	none
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 1300 values. The less detailed occupational sub-group is recorded by the first four digits of the code. The skill level required for a job, which is recorded in the fifth digit

of the codes in the KldB2010, is made available separately in the variable 'level of requirement' (*niveau*).

Employment notifications with an end date earlier than 30.11.2011 are reported using the old occupation code 1988 (KldB 1988) (see Section 1.3.3). 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
The variable contains the occupation of the last job. See 1) with regard to the occupation code.

Note on quality	There is a considerable increase in the number of missing values in 2011 due to the change in the reporting procedure.
Notes	Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise the occupational group (*beruf2010_3*) is the only occupation-related information using KldB2010 that is made available.

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

Variable label	Level of requirement - current/most recent job (KldB 2010)
Variable name	niveau
Category	information on employment, benefit receipt and job search
Origin	BeH, LeH
Data type	numerical
Hierarchy	none
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 30.11.2011 are reported using the old occupation code 1988 (KldB 1988) (see Section 1.3.3). 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
	The variable contains the occupation of the last job. See 1) with regard to the occupation code.
Note on quality	There is a considerable increase in the number of missing values in 2011 due to the change in the reporting procedure.

5.5.6 Part-time (teilzeit)

Variable label	Part-time
Variable name	teilzeit
Category	information on employment, benefit receipt and job search
Origin	BeH
Data type	numerical
Hierarchy	none
Detailed description	The employee's occupational status during the notification period is reported by the employer in the "employment details".
	The variable "occupational status" distinguishes between full-time and part-time employees. The decisive factor here is the ratio between the contracted hours and the usual working hours in the establishment. For part-time employees the variable only records whether their working hours exceed a certain limit or not. Until 1978 this limit was 20 hours of work per week, between 1979 and 1987 it was 15 hours per week and since 1988 it has been 18 hours per week.
	The variable only provides actual details regarding the occupational status for full-time employees, distinguishing among other things between blue-collar and white-collar employees in full-time employment and trainees/apprentices. When the new occupation code was introduced (see Section 1.3.3), however, this distinction was no longer available. The variable *teilzeit* therefore only distinguishes between full-time and part-time employment in the entire reporting period. No further information about the occupational status is used.
Note on quality	There is a considerable increase in the number of missing values in 2011 due to the change in the reporting procedure. In order to reduce this problem, the working hours were imputed at the IAB for the period in question. Further information about the procedure can be found in Ludsteck/Thomsen (2016).

5.5.7 Employment status (erwstat)

Variable label	Employment status
Variable name	erwstat
Category	personal variables
Origin	BeH, LeH
Data type	numerical
Hierarchy	none
Detailed description	This variable assumes different values with different meanings for each data source.

1) BeH

For BeH observations, the variable 'employment status' corresponds to the person group recorded in the new notification procedure (DEÜV) from 1 January 1999 onwards. It indicates contribution- or benefit-related particularities of the employment relationship.

If multiple codes apply to an employment notification, the smallest must be indicated by the reporting employer. The majority of these cases are employment relationships subject to social security contributions without any distinctive characteristics, which are recorded under code number 101. Accordingly, it is possible that these employment relationships are slightly overestimated.

The notification procedure stipulates that changes in the employment status - e.g. when an apprentice is taken on by his/her training company after completing his/her vocational training - must be indicated by a new notification (cf. Deutsche BKK 2012, p. 31).

The person group can be contained in employment notifications that refer to the years prior to 1999 but were not received until 1999 or later. For notifications which were received before 1999, an attempt is made to allocate the notifications to the person groups on the basis of certain rules and with the aid of the vocational education and training', 'occupational status and working hours' and 'occupation' variables as well as other information. In many cases, however, conclusive allocations are not possible.

Since 1 April 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' 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.

5.5.8 Transition zone (gleitz)

Variable label	Transition zone
Variable name	gleitz
Category	information on employment and benefit receipt
Origin	BeH
Data type	numerical
Hierarchy	none

Detailed description	This variable is only available from 2003 onwards and only for BeH observations. It indicates whether the employment notification relates to employment in the low-wage sector, within the so-called transition zone. Jobs in the transition zone have a gross monthly wage of € 400.01 to € 800.00 (so-called midi jobs) for which the employee only has to pay a reduced overall social security contribution. As employees with earnings in the transition zone can voluntarily pay the "regular" social security contribution, not all employees with corresponding earnings are automatically classified as being in the transition zone. The corresponding legislation has been in force since 1. April 2003.

5.5.9 Temporary agency work (leih)

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

5.5.10 Fixed-term contract (befrist)

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

5.5.11 Reason for notification / reason for end of benefit receipt / reason for discontinuation of unemployment benefit II / reason for deregistration (grund)

Variable label	Reason of cancellation/notification/termination
Variable name	grund
Category	information on employment and benefit receipt
Origin	BeH, LeH
Data type	numerical

Hierarchy	none
Detailed description	**1) BeH** In BeH observations, the 'reason for notification' variable indicates the reason why the employer submitted the employment notification in question to the social security agencies. However, not all of the possible reasons for submitting a notification that may occur in the context of the notification procedure are available in the IEB. For instance, the IEB only includes notifications that have information on earnings (i.e. annual, employment interruption and end of employment notifications), whilst initial registrations are not contained as they contain no information on earnings. However, this does not involve a loss of information, as the details from a registration are transmitted again with the following annual, employment interruption or end of employment notification. The reasons for submitting employment notifications are encoded according to the regulations of the notification procedure which has been in effect since 1 January 1999 (in accordance with DEÜV). **2) LeH** In LeH observations, the variable indicates the reason for the end of receipt of unemployment benefit, unemployment assistance or maintenance allowance. In contrast, the LeH contains no information about the reasons for the start of benefit receipt, as the information for the LeH comes from the notifications submitted by the employment agencies to the health insurance institutions regarding finished periods of benefit receipt.

5.5.12 Duration of remaining entitlement to unemployment benefit (restanspruch)

Variable label	Residual claim/planned duration
Variable name	restanspruch
Category	information on employment, benefit receipt and job search
Origin	LeH
Data type	numerical
Hierarchy	none
Detailed description	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 1 January 1998, the remaining entitlement is reported in working days, after this date it is reported in calendar days. This information does not refer to the start date of the episode, but to the start of the original time period.

5.6 Location data

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

Variable label	Place of residence: district (Kreis)
Variable name	wo_kreis
Category	location data
Origin	BeH, LeH
Data type	numerical
Hierarchy	federal state | district
Detailed description	In BeH and LeH observations, the place of residence at district level is only available for the years from 1999 onwards. The variable indicates the district (urban district or rural district) in which the social security contributor lives. The first two digits of the 5-digit district code (Kreisschlüssel) show the code for the federal state (Bundesland), positions 1-3 indicate the regional authority (Regierungsbezirk), and positions 1-5 show the district authority (Kreis). Federal states without a regional authority have a 0 in the third position. 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 LeH, the variable contains the place of residence at the beginning of the original period of time. In order to guarantee consistent regional allocations across the entire observation period, the information on the district was recoded with reference to the territorial allocation of 31 December 2014 for all sources, i.e. in all calendar years, a place of residence is assigned to a district in accordance with the boundaries that the district had on 31 December 2014. As the district boundaries have changed over time, cases would occur in which the district code changes without the individual concerned having relocated if the territorial allocations of the districts were not updated.
Notes	Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, only the federal state (*wo_bula*) is shown as regional information.

5.6.2 Place of residence: employment agency (wo_aa)

Variable label	Place of residence: employment agency (Arbeitsagentur)
Variable name	wo_aa
Category	location data
Origin	BeH, LeH
Data type	numerical
Hierarchy	regional directorate | employment agency

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

5.6.3 Place of residence: regional directorate (wo_rd)

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

5.6.4 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
Data type	numerical

Hierarchy	federal state	district
Detailed description	This variable is an aggregation of the "district" variable to the 16 German federal states. In BeH observations, the federal state of the place of residence is only available for the years from 1999 onwards. Further information on the district of the place of residence can be found under Location data Place of residence: district (Kreis) (*wo_kreis*).

5.7 Establishment variables

5.7.1 Economic activity 73, 3-digit code (w73_3)

Variable label	classification of economic activities 73
Variable name	w73_3
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	division (1-digit code)	group (2-digit code)	class (3-digit code) of economic activity
Detailed description	This variable indicates the economic activity as a 3-digit code in accordance with the WS73 classification and is available from 1975 up to and including 2002. 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.

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

Variable label	classification of economic activities 93, sub-classes
Variable name	w93_5
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	section (1-digit code)	division (2-digit code)	group (3-digit code)	class (4-digit code)	sub-class (5 digit code) of economic activity

Detailed description	This variable indicates the economic activity as a 5-digit code in accordance with the WZ93 classification and is available from 1999 up to and including 2003. WZ93 stands for the "Classification of Economic Activities for the Statistics of the Federal Employment Services, edition 1993" ("Klassifikation der Wirtschaftszweige für die Statistik der Bundesanstalt für Arbeit, Ausgabe 1993"). The WZ93 is based on the Statistical Classification of Economic Activities in the European Community NACE Rev. 1 („Nomenclature génerale des activités économiques dans les communautés européennes") which has four levels the first two of which are based on the international standard ISIC Rev. 3 („International Standard Industrial Classification of All Economic Activities"). Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected.
Notes	Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, the economic activity is only shown as the 3-digit code (*w93_3*).

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

Variable label	classification of economic activities 93, groups
Variable name	w93_3
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	section (1-digit code) | division (2-digit code) | group (3-digit code) | class (4-digit code) | sub-class (5 digit code) of economic activity
Detailed description	This variable indicates the economic activity as a 3-digit code in accordance with the WZ93 classification and is available from 1999 up to and including 2003. WZ93 stands for the "Classification of Economic Activities for the Statistics of the Federal Employment Services, edition 1993" ("Klassifikation der Wirtschaftszweige für die Statistik der Bundesanstalt für Arbeit, Ausgabe 1993"). The WZ93 is based on the Statistical Classification of Economic Activities in the European Community NACE Rev. 1 („Nomenclature génerale des activités économiques dans les communautés européennes") which has four levels the first two of which are based on the international standard ISIC Rev. 3 („International Standard Industrial Classification of All Economic Activities"). Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected.

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

Variable label	classification of economic activities 03, sub-classes
Variable name	w03_5
Category	establishment variables
Origin	BeH
Data type	numerical

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

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

Variable label	classification of economic activities 03, groups
Variable name	w03_3
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	section (1-digit code) | division (2-digit code) | group (3-digit code) | class (4-digit code) | sub-class (5 digit code) of economic activity
Detailed description	This variable indicates the economic activity as a 3-digit code in accordance with the WZ03 classification and is available from 2003 up to 2008. WZ03 stands for the "Classification of Economic Activities, Edition 2003" ("Klassifikation der Wirtschaftszweige Ausgabe 2003") of the Federal Statistical Office (eds.). Like the WZ93, the WZ03 is based on the Statistical Classification of Economic Activities in the European Community NACE Rev. 1 (see description of variables *w93_3*, *w93_5*). The classifications of the economic activity have been updated, but the structure of the WZ93 has been largely retained.
	Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected.

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

Variable label	classification of economic activities 08, sub-classes
Variable name	w08_5
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	section (1-digit code) | division (2-digit code) | group (3-digit code) | class (4-digit code) | sub-class (5-digit code) of economic activity

Detailed description	This variable indicates the economic activity as a 5-digit code in accordance with the WZ08 classification and is available from 2008 onwards. WZ08 stands for the "Classification of Economic Activities, Edition 2008" ("Klassifikation der Wirtschaftszweige Ausgabe 2008") of the Federal Statistical Office (eds.). The WZ08 is based on the Statistical Classification of Economic Activities in the European Community NACE Rev. 2. Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected.
Notes	Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise the economic activity is only shown as the 3-digit code (*w08_3*).

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

Variable label	classification of economic activities 08, groups
Variable name	w08_3
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	section (1-digit code)	division (2-digit code)	group (3-digit code)	class (4-digit code)	sub-class (5 digit code) of economic activity
Detailed description	This variable indicates the economic activity as a 3-digit code in accordance with the WZ08 classification and is available for the years since 2008. WZ08 stands for the "Classification of Economic Activities, Edition 2008" ("Klassifikation der Wirtschaftszweige Ausgabe 2008") of the Federal Statistical Office (eds.). The WZ08 is based on the Statistical Classification of Economic Activities in the European Community NACE Rev. 2. Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected.

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

Variable label	w73_3 completed by extrapolation/imputation
Variable name	w73_3_gen
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	division (1-digit code)	group (2-digit code)	class (3-digit code)

Detailed description	This variable indicates the economic activity as a 3-digit code in accordance with the WZ73 classification. From 1975 up to and including 2002, the variable contains the original values from *w73_3*. From 2003 onwards, the information is either continued or replaced with the help of recoding tables. Thus the variable provides time-consistent information on the economic activity based on the economic activity classification WS73. A detailed description can be found in Eberle et al. (2011). Further information on the WS73 classification can be found in the description of variable *w73_3*.

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

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

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

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

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

Variable label	Type of imputation w93_3
Variable name	group_w93_3

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

5.7.12 Year of first appearance of establishment number (grd_jahr)

Variable label	year of first appearance
Variable name	grd_jahr
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	none
Detailed description	This variable indicates the first appearance of the establishment number in the dataset. If an establishment number in western Germany is only determined for the first time after 1975 (or after 1992 in eastern Germany), this variable could indicate the date when the respective establishment was founded. However, it could also be an establishment that has been in existence for a longer time but has been allocated a new establishment number following a change of owner or a change in the legal form of the establishment (For the allocation of establishment numbers see Bundesagentur für Arbeit 2007, pp. 9-11). It is also possible that the establishment already existed before, but had no employees subject to social security, or from 1999 onwards, no marginal part-time workers.

5.7.13 First appearance of establishment number (grd_dat)

Variable label	first appearance
Variable name	grd_dat
Category	establishment variables
Origin	BeH
Data type	date
Hierarchy	none

Detailed description	This variable indicates the first appearance of the establishment number in the BeH to the day. If an establishment number in western Germany is only determined for the first time after 1975, or after 1992 in eastern Germany, this variable could indicate the date when the respective establishment was founded. However, it could also be an establishment that has been in existence for a longer time but has been allocated a new establishment number following a change of ownership or a change in the legal form of the establishment (for the allocation of establishment numbers see Bundesagentur für Arbeit 2007, pp. 9-11). It is also possible that the establishment already existed before, but had no employees subject to social security, or from 1999 onwards, no marginal part-time workers.
Notes	Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, only the year when the establishment number first appeared is shown (*grd_jahr*).

5.7.14 Year of last appearance of establishment number (lzt_jahr)

Variable label	year of last appearance
Variable name	lzt_jahr
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	none
Detailed description	This variable indicates the last appearance of the establishment number in the dataset (see Bender et. al. 1996). If the existence of an establishment number in the BHP already ends before 2008, it could indicate the closure of the establishment. However, other possible reasons for this are an "arbitrary change of the establishment number following a change of owner or a change in the legal form of the establishment", the "outsourcing of parts of the firm under a new establishment number" or other administrative changes (see Bender et. al. 1996 or Bundesagentur für Arbeit 2007, pp. 9-11).

5.7.15 Last appearance of establishment number (lzt_dat)

Variable label	last appearance
Variable name	lzt_dat
Category	establishment variables
Origin	BeH
Data type	date
Hierarchy	none

Detailed description	This variable indicates the last appearance of the establishment number in the dataset to the day (see Bender et. al. 1996). If the existence of an establishment number in the BHP already ends before 2008, it could indicate the closure of the establishment. However, other possible reasons for this are an "arbitrary change of the establishment number following a change of ownership or a change in the legal form of the establishment", the "outsourcing of parts of the firm under a new establishment number" or other administrative changes (see Bender et al. 1996 or Bundesagentur für Arbeit 2007, pp. 9-11).
Notes	Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, only the year when the establishment number last appeared is shown (*lzt_jahr*).

5.7.16 Total number of employees (az_ges)

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

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

Variable label	no. full-time (regular workers + others)
Variable name	az_vz
Category	generated establishment variables
Origin	BeH
Hierarchy	none
Detailed description	This variable contains the number of people in the establishment who are reported on 30 June of a year as full-time employees under the person group codes 101, 140, 143, 105, 106, 112, 118, 119, 120, 149, 201, 203, 205, 999 and YYY. Apprentices, marginally part-time employees and individuals participating in partial retirement schemes are not considered.

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

Variable label	no. marginal part-time workers
Variable name	az_gf
Category	generated establishment variables
Origin	BeH
Hierarchy	none

Detailed description	The number of employees in marginal part-time employment is generated using the person group code – values 109 and 209. This variable has only been contained in the dataset since 1999 as it has only been included in the social security notification procedure since that year.

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

Variable label	mean imp. Wage all full-time employees
Variable name	te_imp_mw
Category	generated establishment variables
Origin	BeH
Data type	numerical
Hierarchy	none
Detailed description	This variable contains the mean imputed gross daily wage of the full-time employees in an establishment. It does not include the wages of marginally part-time staff, apprentices or individuals participating in partial retirement schemes. 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 cut. In order to calculate the gross daily wage, the fixed period wage is divided by the number of calendar days in the period. To calculate the mean, these censored wages were imputed (see Section 8.2 in Schmucker et al. 2016). These data were then aggregated at establishment level.

5.7.20 Place of work: district (ao_kreis)

Variable label	Place of work: district (Kreis)
Variable name	ao_kreis
Category	location data
Origin	BeH
Data type	numerical
Hierarchy	federal state	district
Detailed description	The variable indicates the district (urban district or rural district) in which the employee's establishment is located. The first two digits of the 5-digit district code (Kreisschlüssel) show the code for the federal state

	(Bundesland), positions 1-3 indicate the regional authority (Regierungs-bezirk), and positions 1-5 show the district authority (Kreis). Federal states without a regional authority have a 0 in the third position.
	In order to guarantee consistent regional allocations across the entire observation period, the information on the district was recoded to the territorial allocation of 31 December 2014, i.e. in all calendar years, a place of work is assigned to a district in accordance with the boundaries that the district had on 31 December 2014. As the district boundaries have changed over time, cases would occur in which the district code of the location of the establishment would change without the establishment concerned having relocated, if the territorial allocations were not updated.
Notes	Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, only the federal state (*ao_bula*) is shown as regional information.

5.7.21 Place of work: federal state (ao_bula)

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

5.8 Generated biographical variables

5.8.1   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 1.1.1975. For East Germans the censoring is not so clear. Entries on 1.1.1990 are censored for sure, but often also entries on 1.1.1991 and 1.1.1992 may be affected because in 1990 and 1991 many employment notifications are missing.

5.8.2 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, which an individual has worked till the 30 June of a year (Cross-sectional Model) or to the end date of the respective episode (Longitudinal Model). Training periods are not taken into account (*erwstat* == 102, 121, 122, 141). If an individual was just in training, the variable adopts 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 work force until the 30 June of the respective year (Cross-sectional Model) or still not entered the work force until the end of the episode (Longitudinal Model). |
| Note on quality | For West Germans the variable is left censored on 1.1.1975. For East Germans the censoring is not so clear. Entries on 1.1.1990 are censored for sure, but often also entries on 1.1.1991 and 1.1.1992 may be affected because in 1990 and 1991 many employment notifications are missing. |

5.8.3 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 or if the *grund* variable indicates a lump sum payment (*grund* == 54). |

Note on quality	For West Germans the variable is left censored on 1.1.1975. For East Germans the censoring is not so clear. Entries on 1.1.1990 are censored for sure, but often also entries on 1.1.1991 and 1.1.1992 may be affected because in 1990 and 1991 many employment notifications are missing.

5.8.4 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 the 30 June of this year, in the Longitudinal Model until the end of the episode. Training periods are also included, meanwhile employment gaps are not.
	
	If the number of days in the establishment is alternatively calculated with the variable First day in establishment (ein_bet)
	`in QM: gen tag_bet_neu = mdy(6,30,jahr) – ein_bet + 1`
	`in LM: gen tag_bet_neu = endepi – ein_bet + 1`
	
	the values, which are obtained, are larger than or equal to the variable tage_bet, because tage_bet does not include interruptions of employment.
	Number of days in establishment (*tage_bet*) can be larger than Duration of a working life (*tage_erw*). The variable *tage_bet* includes periods of education, *tage_erw* doesn't include these periods.
	
	This variable is coded missing (.n) if the establishment ID is missing or if the *grund* variable indicates a lump sum payment (*grund* == 54).

Note on quality	For West Germans the variable is left censored on 1.1.1975. For East Germans the censoring is not so clear. Entries on 1.1.1990 are censored for sure, but often also entries on 1.1.1991 and 1.1.1992 may be affected because in 1990 and 1991 many employment notifications are missing.

5.8.5 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 first day in employment (*ein_erw*).
	This variable is coded missing (.n) if the establishment ID is missing or if the *grund* variable indicates a lump sum payment (*grund* == 54).
Note on quality	For West Germans the variable is left censored on 1.1.1975. For East Germans the censoring is not so clear. Entries on 1.1.1990 are censored for sure, but often also entries on 1.1.1991 and 1.1.1992 may be affected because in 1990 and 1991 many employment notifications are missing.

5.8.6 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 relates to 30 June of the year, in all other LIAB data it relates to the end of the episode. Training periods in the establishment are included, employment gaps not. If the number of days in the current job is alternatively calculated with the first day in job variable (*ein_job*),
	
	`in QM:gen tage_job_neu = mdy(6,30,jahr) – ein_job+1`
	`in LM:gen tage_job_neu = endepi – ein_job+1`
	
	the values obtained are larger or equal than the variable *tage_job* because it does not include interruptions of employment.
	
	 This variable is coded missing (.n) if the establishment ID is missing or if the *grund* variable indicates a lump sum payment (*grund* == 54).

Note on quality	For West Germans the variable is left censored on 1.1.1975. For East Germans the censoring is not so clear. Entries on 1.1.1990 are censored for sure, but often also entries on 1.1.1991 and 1.1.1992 may be affected because in 1990 and 1991 many employment notifications are missing.

5.8.7 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 LeH
Data type	numerical
Hierarchy	none
Detailed description	The variable contains the number of days of benefit receipts of a person until the 30 June of each year (Cross-sectional Model) or the end date of the respective episode (Longitudinal Model). If there are no benefit receipts, the variable adopts the value 0. Other payments (e.g. benefits according to SGB II, the former unemployment assistance, etc.) than unemployment benefits (Arbeitslosengeld) have not considered for this variable. This allows a consistent evaluation of this variable over time. Moreover, periods of parallel employment have not been considered for this variable.
Note on quality	For West Germans the variable is left censored on 1.1.1975. For East Germans the censoring is not so clear. Entries on 1.1.1990 are censored for sure, but often also entries on 1.1.1991 and 1.1.1992 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, Nürnberg.

Bundesagentur für Arbeit (Hg.) 1988: Klassifizierung der Berufe – Systematisches und Alphabetisches Verzeichnis der Berufsbenennungen, Nürnberg.

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

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

vom Berge, Philipp; König, Marion; Seth, Stefan (2013): Sample of Integrated Labour Market Biographies (SIAB) 1975-2010. (FDZ-Datenreport, 01/2013 (en)), Nürnberg, 65 S.

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

Deutsche BKK (2012): Ratgeber Sozialversicherung 2012, Wolfsburg, 56 S., URL: http://www.deutschebkk.de/fileadmin/user_upload/Service/Download-Center/Ratgeber_Sozialversicherung_2012_01.pdf, (Zugriff: 28.03.2012).

Drews, Nils 2006: Qualitätsverbesserung der Bildungsvariable in der IAB-Beschäftigtenstichprobe 1975-2001. (FDZ Methodenreport, 05/2006 (de)), Nürnberg, 16 S.

Eberle, Johanna; Jacobebbinghaus, Peter; Ludsteck, Johannes; Witter, Julia 2011: Generation of time-consistent industry codes in the face of classification changes * Simple heuristic based on the Establishment History Panel (BHP). (FDZ Methodenreport, 05/2011 (en)), Nürnberg, 21 S.

Fischer, Gabriele; Janik, Florian; Müller, Dana; Schmucker, Alexandra (2008): The IAB Establishment Panel – from Sample to Survey to Projection (FDZ Methodenreport, 01/2008 (en)), Nürnberg, 42 S.

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

Meinken, Holger; Koch, Iris 2004: BA-Beschäftigtenpanel 1998-2002. Codebuch, Nürnberg.

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, S. 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 under http://fdz.iab.de.

.

8   List of abbreviations

AA	Agentur für Arbeit/ Arbeitsamt	Employment agency / employment office
ALG	Arbeitslosengeld	unemployment benefit
ALH	Arbeitslosenhilfe	unemployment support
ASU	Arbeitsuchendenhistorik	Jobseeker History
BA	Bundesagentur für Arbeit	Federal Employment Agency
BeH	Beschäftigtenhistorik	Employee History
BfA	Bundesversicherungsanstalt für Angestellte	Federal Social Insurance Office for Salaried Employees
BHP	Betriebs-Historik-Panel	Establishment History Panel
DEÜV	Verordnung über die Erfassung und Übermittlung von Daten für die Träger der Sozialversicherung – Datenerfassungs- und –übermittlungs-Verordnung –	Data Collection and Transmission Regulation - regulation on the collection and transmission of data for the social security agencies
DEVO	Zweite VO über die Erfassung von Daten für die Träger der Sozialversicherung und für die BA – Datenerfassungs-Verordnung –	Data Collection Regulation - second regulation on the collection of data for the social security agencies and for the Federal Employment Agency
DÜVO	Zweite VO über die Datenübermittlung auf maschinell verwertbaren Datenträgern im Bereich der Sozialversicherung und der BA – Datenübermittlungs-Verordnung –	Data Transmission Regulation - second regulation on the transfer of data on machine-readable data media in the field of social security and the BA
FDZ	Forschungsdatenzentrum	Research Data Centre
IAB	Institut für Arbeitsmarkt- und Berufsforschung	Institute for Employment Research
IEB	Integrierte Erwerbsbiographien	Integrated Employment Biographies
LeH	Leistungsempfängerhistorik	Benefit Recipient History
LIAB	Linked-Employer-Employee-Daten des IAB	Linked Employer-Employee Data of the IAB
LHG	Leistungshistorik Grundsicherung	Unemployment Benefit II Recipient History
LM	Längsschnittmodell	Longitudinal Model
LVA	Landesversicherungsanstalt	Land Social Insurance Office
MTH	Maßnahmeteilnehmerhistorik	Participants-in-Measures History
SGB	Sozialgesetzbuch	German Social Code
SIAB	Stichprobe der Integrierten Arbeitsmarktbiographien	Sample of Integrated Labour Market Biographies
UHG	Unterhaltsgeld	subsistence payment
XASU	Arbeitsuchendenhistorik 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

Imprint

Corresponding author:
Wolfram Klosterhuber
Institute for Employment Research (IAB)
Research Data Centre (FDZ)
Regensburger Str. 104
D-90478 Nürnberg
Phone: +49-911-179-7007
Email: Wolfram.Klosterhuber@iab.de