

FDZ-Datenreport

Documentation of labour market data

01/2017
EN

Sample of Integrated Labour Market Biographies – Regional File 1975–2014 (SIAB-R 7514)

Andreas Ganzer,
Alexandra Schmucker,
Philipp vom Berge,
Anja Wurdack



Sample of Integrated Labour Market Biographies – Regional File 1975-2014 (SIAB-R 7514)

Andreas Ganzer, Alexandra Schmucker, Philipp vom Berge, Anja Wurdack
(Institute for Employment Research, IAB)

Documentation: Version 1 (SIAB_Regional_7514_EN_v1_dok1)

Data: Version 1 (SIAB Regional 7514 v1)

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

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

Contents

1	INTRODUCTION AND OUTLINE.....	7
1.1	Introduction	7
1.2	Data use	8
1.3	Changes as compared to SIAB Regional 7510	8
1.3.1	Corrected territorial allocation	8
1.3.2	Introduction of the occupation code 2010.....	8
1.3.3	Set of variables	9
1.3.4	Anonymisation	10
1.4	Outline	10
1.5	List of variables	12
1.6	Volume Structure.....	14
2	DATA SOURCES	14
2.1	Employee History (BeH)	15
2.2	Benefit Recipient History (LeH)	16
2.3	Unemployment Benefit II Recipient History (LHG)	16
2.4	Jobseeker History (ASU / XASU)	17
3	DATA PREPARATION AND SAMPLING PROCEDURE	18
3.1	Corrections and validation procedures.....	18
3.1.1	Employee History (BeH).....	18
3.1.2	Benefit Recipient History (LeH).....	19
3.1.3	Unemployment Benefit II Recipient History (LHG).....	19
3.1.4	Jobseeker History (ASU / XASU).....	20
3.1.5	SGB II anonymisation	21
3.2	Episode splitting.....	21
3.3	Sampling procedure.....	22
3.4	Anonymisation of the SIAB-R 7514.....	22
3.4.1	Merging sources and deletion of sources	23
3.4.2	Periods of validity of the sources	23
3.4.3	Censoring of data on age.....	26
3.4.4	Aggregation and deletion of variables.....	26
3.4.5	Consequences of the anonymisation	28
3.5	Missing values	28
4	DATA QUALITY AND PROBLEMS.....	29

4.1	Entire IEB	29
4.2	Employee History (BeH)	30
4.3	Benefit Recipient History (LeH)	32
4.4	Unemployment Benefit II Recipient History (LHG)	32
4.5	Jobseeker History (ASU/XASU)	34
4.5.1	ASU	34
4.5.2	XASU	34
5	DESCRIPTION OF VARIABLES	35
5.1	Identifiers	35
5.1.1	Individual ID (persnr)	35
5.1.2	Establishment counter (bnn)	36
5.2	Generated technical variables	36
5.2.1	Observation counter per person (spell)	36
5.2.2	Source of spell (quelle_gr)	36
5.3	Period of validity	37
5.3.1	Original start date of observation (begorig)	37
5.3.2	Original end date of observation (endorig)	37
5.3.3	Start date of split episode (begepi)	38
5.3.4	End date of split episode (endepepi)	38
5.4	Personal information	38
5.4.1	Gender (frau)	38
5.4.2	Year of birth (gebjahr)	39
5.4.3	Nationality (deutsch)	39
5.4.4	Vocational training (ausbildung)	39
5.4.5	School leaving qualification (schule)	41
5.5	Information on employment, benefit receipt and job search	42
5.5.1	Daily wage, daily benefit rate (tentgelt_gr)	42
5.5.2	Occupation – current/most recent (KldB 1988) (beruf_gr)	43
5.5.3	Occupational group – current/most recent (KldB 2010) (beruf2010_gr)	44
5.5.4	Level of requirement – current/most recent (KldB 2010) (niveau)	45
5.5.5	Part-time (teilzeit)	46
5.5.6	Employment status (erwstat_gr)	46
5.5.7	Transition zone (gleitz)	48
5.5.8	Temporary agency work (leih)	48
5.5.9	Fixed-term contract (befrist)	48
5.5.10	Reason of cancellation/notification/termination (grund_gr)	49
5.5.11	Employment before the age of 17 (tage_jung)	50
5.5.12	Employment after the age of 62 (tage_alt)	50
5.5.13	Duration of unemployment (alo_dau)	50
5.6	Location data	51
5.6.1	Place of work: region (ao_region)	51
5.6.2	Commuter status (pendler)	52
5.7	Establishment variables	53

5.7.1	Economic activity 93 generated – completed by extrapolation / imputation (w93_3_gen_gr)...	53
6	REFERENCES	54
7	APPENDIX.....	56
7.1	Overview of date adjustments and spell deletions	56
7.2	Comparison of SIAB 7514 and SIAB-Regionalfile 7514.....	57
7.3	Recoding tables of aggregated variables in comparison with the weakly anonymous version of the SIAB 7510	57
7.4	Frequency tables	80
8	LIST OF ABBREVIATIONS.....	80

Abstract

This FDZ data report describes the “Regionalfile” of the Sample of Integrated Labour Market Biographies 1975-2014 (SIAB-R 7514). This dataset represents the factual anonymous version of the Sample of Integrated Labour Market Biographies (SIAB) and may be delivered to scientific research institutions after concluding a use agreement with the IAB.

Zusammenfassung

Dieser FDZ Datenreport beschreibt das „Regionalfile“ der Stichprobe der Integrierten Arbeitsmarktbiografien 1975-2014 (SIAB-R 7514). Der Datensatz stellt die faktisch anonymisierte Version der Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) dar und wird nach Abschluss eines Nutzungsvertrages mit dem IAB an Forschungseinrichtungen übermittelt.

Keywords: German administrative micro data, labour market data, data manual, factual anonymized data, Scientific Use File

Disclaimer

We would like to thank our colleagues in the Research Data Centre (FDZ) as well as the division for IT Services and Information Management (ITM) of the Institute for Employment Research for their cooperation and support. Our special thanks go to Casandra Munkert and Hanna Walch. This data report is a revision of the FDZ data report 09/2013 (authors: vom Berge, Burghardt, Trenkle). Furthermore, parts of this data report are taken from the Data report on the weakly anonymous version of the Sample of Integrated Labour Market Biographies (SIAB) (FDZ data report 04/2016). Additionally, individual passages from IAB-internal data documentations by IAB ITM were adopted.

Data availability

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

1 Introduction and outline

1.1 Introduction

The regional file of the Sample of Integrated Labour Market Biographies 1975-2014 (referred to as SIAB-R 7514 in the following) is a factually anonymous¹ scientific use file (SUF) of the Sample of Integrated Labour Market Biographies 1975-2014 (SIAB 7514)². The regional file is a specific type of SUF which distinguishes itself from other types of SUFs by preserving a comparatively detailed regional data structure. However, it provides only a limited range of variables and values at the level of individuals and establishments for reasons of data confidentiality. Altogether the SIAB-Regional file in its current version covers the employment histories of 1.707.228 individuals. Their employment biographies are documented in a total of 51.987.959 lines of data.

The starting point for the production of the SIAB-Regional File is the weakly anonymous version of the SIAB 7514, which is a 2 % sample of the Integrated Employment Biographies (Integrierte Erwerbsbiografien - IEB) of the Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung - IAB).

The IEB contains information which comes from different sources, and the individual's statuses are described exact to the day. They comprise all individuals who showed one of the following statuses at least once during the observation period (data source and the period covered in square brackets):

- employment subject to social security [Employee History (BeH); recorded from 1975 onwards]
- marginal part-time employment [Employee History (BeH); recorded from 1999 onwards]
- receipt of benefits in accordance with Social Code Book III [Benefit Recipient History (LeH); recorded from 1975 onwards] or Social Code Book II [Unemployment Benefit II Recipient History (LHG/XLHG); recorded from 2005 onwards]
- registered with the Federal Employment Agency [Bundesagentur für Arbeit – BA] as a jobseeker [Jobseeker History (ASU/XASU); recorded from 2000 onwards]
- participation in an employment or training measure [Participation-in-Measures History (MTH); recorded from 2000 onwards]

This FDZ data report describes the preparation of the SIAB-Regional File and the variables of the SUF. The data report is structured as follows: besides an introduction, Chapter 1 contains information on data access as well as an outline of the data, the volume structure and a list of

¹ For further information on the concept of factual anonymisation see Chapter 3.4.

² For the weakly anonymous version of the SIAB 7510 see vom Berge et al. (2013)

variables. A description of the individual data sources can be found in Chapter 2. Data preparation and the anonymisation concept of the SUFs are described in Chapter 3. Chapter 4 discusses the data quality and Chapter 5 concludes with a description of the individual variables contained in the data set.

1.2 Data use

Upon request, the SIAB-Regional File is made available to researchers for a limited period of time for a project with defined contents, and applications may be processed on the premises of their research institution (§ 282 para. 7 SGB III).

In order to be able to use the data a project application must first be submitted to the FDZ. After approval by the FDZ a project specific data use agreement is concluded with the researcher's institution.

Details about different options of data access, our requirements for data use and application procedures are available on the FDZ homepage at <http://fdz.iab.de/en.aspx>.

1.3 Changes as compared to SIAB Regional 7510

In comparison to the SIAB Regional 7510 some changes have been made. These changes are listed below.

1.3.1 Corrected territorial allocation

All locations stated in SIAB 7514 incorporate the territorial allocation as of 31 December 2014.

1.3.2 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.³ Another important change is the discontinuation of the variable 'occupational status and working hours' (stib), as most of the values given in this variable can no longer be traced using the new occupation code. The categories of the previous stib variable that are available consistently across the entire observation period (employment status as an apprentice, distinction between part-time and full-time) can be traced in other variables in future. Section 1.3.3 shows which new variables are added as a result of the switch to the new occupation code, which variables are discontinued and which are contained in the data in a different form in future. Section 4.2 shows what consequences the switch to the new occupation code has on the quality of the data in the BeH observations. One of the most serious consequences of the switchover is a temporary increase in the number of missing values in the variables reported by the employers via the occupation code. This was also the case with the new 'working time' variable, which is why an imputed version of the variable was used for the SIAB. The problem of increased numbers of missing values mainly concerns the years 2011 and 2012.

1.3.3 Set of variables

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

Table 1: Changes in set of variables

Variable	Type of change	Explanation
Duration of unemployment (alo_dau)	N	
Contract period (befrist)	N	Introduced with new occupation code; available from 2011 onwards

³ See Paulus/Matthes, 2013, for details regarding the Classification of Occupations 2010.

Occupational group- current/ most recent (KldB 2010) (beruf2010_gr)	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
Temporary work (leih)	N	Introduced with new occupation code; available from 2011 onwards
Level of requirement – current/most recent (KldB 2010) (niveau)	N	Introduced with new occupation code; available for entire period by recoding old variable.
Skills level (quali)	D	Now contained in vocational education and training
School leaving qualification (schbild)	D	Now contained in variable 'schule'
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'.

N = new, D = dropped

1.3.4 Anonymisation

It has been attempted to keep the anonymization procedures as close as possible to that of the previous version. However, some differences occur between the two versions, as aggregations or deletions were necessary. Thus, some variables may no longer contain certain values compared to the SIAB Regional 7510. In particular, users of the old dataset versions should become familiar with changes in the variables ,erwstat_gr' and ,grund_gr'.

1.4 Outline

Table 2: Outline

Topics/ groups of variables	<p>Employee History (Beschäftigten-Historik - 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.</p> <p>Benefit Recipient History (Leistungsempfänger-Historik - LeH): Information on benefit receipt in accordance with Social Code Book III (SGB III) for recipients of unemployment benefit, unemployment assistance and maintenance allowance.</p> <p>Unemployment Benefit II Recipient History (Leistungs-Historik-Grundsicherung - LHG): Data on individuals in receipt of basic social security benefits in accordance with Social Code Book II (SGB II) (Types of institution: cooperation of employment agencies and municipalities/joint facilities, separated responsibilities/municipalities exercising their duties separately, authorised municipalities).</p>
-----------------------------	--

	Jobseeker History (Arbeitsuchenden-Historik - ASU) as well as XSozial-BA-SGB II (XASU): Information on job-search activity (applicant pool of the Federal Employment Agency) including information on job-search activity reported via the transmission standard XSozial-BA-SGB II to the BA by authorised municipalities or separated responsibilities.
Data unit	Employees covered by social security (including marginal part-time employees from 1999 onwards), benefit recipients, jobseekers
Number of cases	1,707,228 individuals 51,987,959 non-overlapping observations (after episode splitting)
Period covered	The period covered depends on the data source. BeH: 1 January 1975 - 31 December 2014 (2012: 30-months file, 2013: 18-months file; 2014: 6-months file) LeH: 1 January 1975 - 31 December 2014 (X)ASU 1 January 1997 – 31 December 2004 and 1 January 2007 – 31 December 2014 LHG 1 January 2007 – 31 December 2014
Time reference	exact to the day
Regional structure	326 districts and aggregated regions based on the place of work, categorical variable for commuter status/place of residence.
Date of territorial allocation	Territorial allocation updated as of 31.12.2014
Survey design	2 % random sample
Institutions involved	Social security agencies, Federal Employment Agency (Bundesagentur für Arbeit)
Update frequency	Continuous
File format and size	STATA (approx. 2.7GB), SPSS (approx. 4.3GB)
Data access	Data transmission solely to public research institutes or institutes which verifiably perform tasks defined as independent scientific research.
Degree of anonymisation	Factually anonymous data (scientific use file)
Adjustments - outline	The data of the SIAB-Regional File went through several preparation and anonymisation procedures such as the recoding of individual variables, aggregations and deletions of variables and time periods after they have been drawn from the IEB.
Citation of data and data documentation	Sample of Integrated Labour Market Biographies – Regional File 1975-2014 (SIAB-R 7514), Nuremberg 2016 Data: “The data basis of this paper is the weakly anonymous Sample of Integrated Labour Market Biographies (SIAB) 1975 - 2014. The data were accessed on-site at the Research Data Centre (FDZ) of the Federal Employment Agency (BA) at the Institute for Employment Research (IAB) and/or via remote data access at the FDZ.”

	Data documentation: - Ganzer, Andreas; Schmucker, Alexandra; vom Berge, Philipp; Wurdack, Anja (2016): Sample of Integrated Labour Market Biographies- Regional File 1975-2014 (SIAB-R 7514). FDZ data report, 01/2017 (en), Nuremberg - Ganzer, Andreas; Schmucker, Alexandra; vom Berge, Philipp; Wurdack, Anja (2016): Stichprobe der Integrierten Arbeitsmarktbiografien Regionalfile 1975-2014 (SIAB-R 7514). FDZ Datenreport, 01/2017 (de), Nürnberg
Dataset version	Sample of Integrated Labour Market Biographies – Regional File 1975-2014 (SIAB-R 7514)

1.5 List of variables

The overview of variables in Table 4 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 ‘daily wage, daily benefit rate’ is only available for BeH and LeH observations; the observations of the other data sources contain the missing value “.n” for this variable. Another characteristic is that some variables have different contents depending on the data source. For instance, for BeH observations the ‘employment status’ variable contains the person group of the employment notification procedure, for LeH observations it contains the type of benefit, for LHG and XLHG observations it contains the SGB II status, for ASU and XASU observations the job search status. 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	Page	BeH	LeH	LHG	(X)ASU
		1975-2010	1975-2010	1997-2004 2007-20010	2007-2010
Identifiers	35				
Individual ID (persnr)					

List of variables	Page	BeH	LeH	LHG	(X)ASU
		1975-2010	1975-2010	1997-2004 2007-20010	2007-2010
Establishment counter (bnn)	36				
Generated technical variables	36				
Observation counter per person (spell)	36				
Source of spell (quelle_gr)	36				
Period of validity	37				
Original start date of observation (be-gorig)	37				
Original end date of observation (en-dorig)	37				
Start date of split episode (begepi)	38				
End date of split episode (ende-pi)	38				
Personal information	38				
Gender (frau)	38				
Year of birth (gebjahr)	39				
Nationality (deutsch)	39				
Vocational training (ausbildung)	39				
School leaving qualification (schule)	41				
Information on employment, benefit receipt and job search	42				
Daily wage, daily benefit rate (tentgelt_gr)	42				
Occupation – current/most recent (KldB 1988) (beruf_gr)	43				
Occupational group – current/most recent (KldB 2010) (beruf2010_gr)	44				
Level of requirement – current/most recent (KldB 2010) (niveau)	45				
Part-time (teilzeit)	46				
Employment status (erwstat_gr)	46				
Transition zone (gleitz)	48				
Temporary agency work (leih)	48				
Fixed-term contract (befrist)	48				
Reason of cancellation/notification/termination (grund_gr)	49				
Duration of Employment before the age of 17 (tage_jung)	50				
Duration of Employment after the age of 62 (tage_alt)	50				

List of variables	Page	BeH	LeH	LHG	(X)ASU
		1975-2010	1975-2010	1997-2004 2007-20010	2007-2010
Duration of unemployment (alo_dau)	50				
Location data	51				
Place of work: region (ao_region)	51				
Commuter status (pendler)	52				
Establishment variables	53				
Economic activity 93 generated – completed by extrapolation / imputation (w93_3_gen_gr)	53				

1.6 Volume Structure

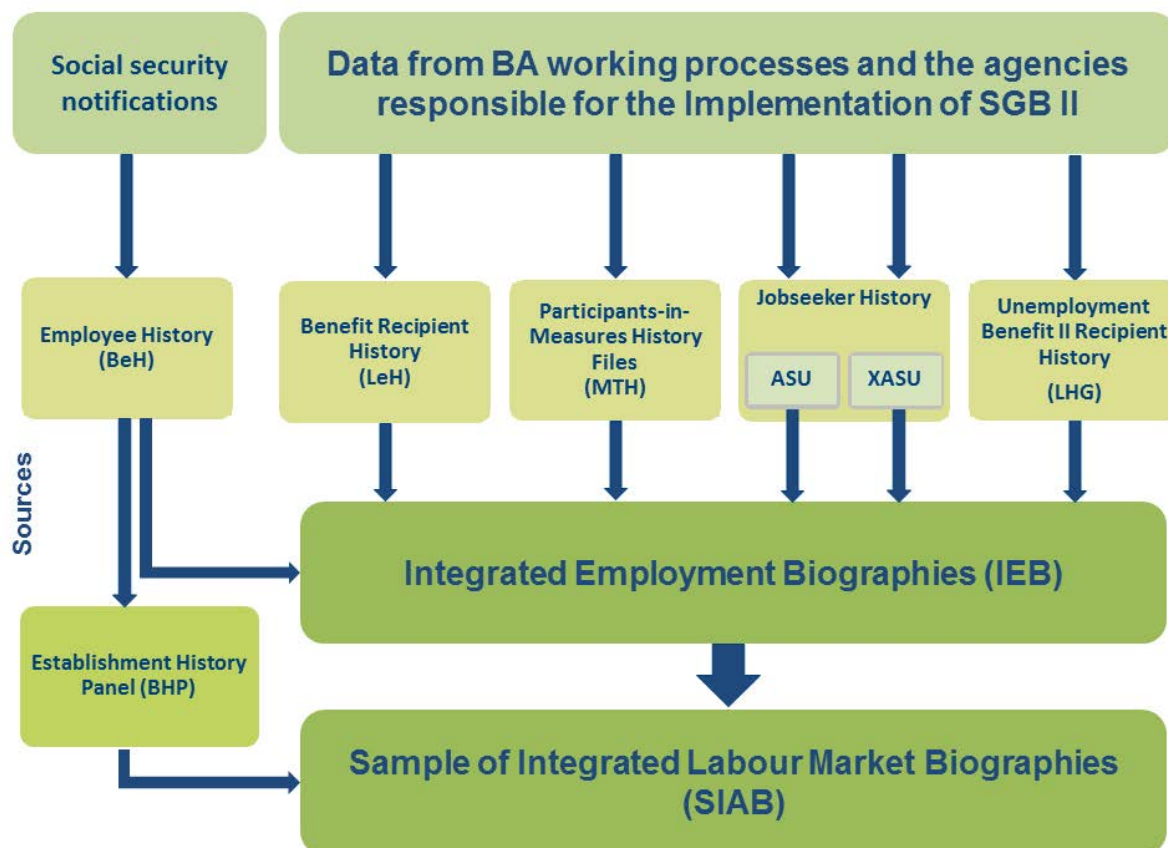
Table 5: Volume structure

No. of cases	before splitting	after splitting
BeH	29.625.506	34.482.587
LeH	3.957.163	6.079.643
LHG	1.372.823	3.069.942
ASU	6.164.261	8.355.787
Total number of observations	41.119.753	51.987.959
Individuals		1.707.228

2 Data sources

The administrative personal data were drawn from the Integrated Employment Biographies (IEB) of the IAB. The IEB unites data from five different data sources, each of which may contain information from different EDP procedures. In addition, some supplementary variables from these data sources which are not part of the IEB are incorporated into the administrative individual data. Figure 1 illustrates the data flows that lead to the SIAB and their relationship to other FDZ data products. Observations from all of the data sources shown, with the exception of participation (or planned participation) in employment and training measures (MTH), are incorporated into the SIAB Regional File from the IEB. Each data source of the IEB is specified below.

Figure 1: Data sources of the SIAB



2.1 Employee History (BeH)

The data source regarding employment is the Employment History (Beschäftigten-Historik - BeH) of the IAB. The data basis is the integrated notification procedure for health, pension and unemployment insurance, which came into effect as of 1 January 1973 (and was extended to cover East Germany as of 1 January 1991) and is known by the abbreviation DEÜV (previously DEVO/DÜVO) (for further details see: Bender et al. 1996, p. 4 ff.; 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. 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 and students (see Cramer 1985)⁴ are in principle not recorded in the BeH. Since the notification procedure was changed on 1 January 1999, employees in marginal part-time employment and unpaid family workers have also been recorded (not contained in the data until 1 April 1999). The data are

⁴ Students may still appear in the BeH if, for example, they had a marginal part-time job parallel to their degree course.

recorded by the health insurance companies, collected in a continuous file by the BA and subsequently integrated into the History File of the IAB.

The variables of the SIAB-Regional file related to establishments are taken from the Establishment-History-Panel (Betriebs-Historik-Panel - BHP) which itself is also based on the BeH (see Schmucker et al. 2016). The following modified variables based on the BHP are included in the scientific use file:

- Place of Work (ao_region)
- Industry Classification 93 (extrapolated 1975-2003 and 2008-2014) (w93_gen_gr)

Two issues have to be taken into account when linking individual data with establishment data:

1. Information on establishments is based on individual data and aggregated on the annual cut-off date of the BHP, June 30. Unlike the data on individuals, the establishment variables are not spell data but are only valid on 30 June precisely each year.
2. Secondly, the BHP is created directly from the BeH, whereas the SIAB Individual File is generated from the IEB (see Figure 1). The IEB however does not incorporate all BeH observations.

2.2 Benefit Recipient History (LeH)

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

2.3 Unemployment Benefit II Recipient History (LHG)

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

benefit rates, however. As the amount of benefit received is not determined at the level of the individual but at the level of the benefit community in the case of Unemployment Benefit II, it is difficult to assign an individual benefit rate and this is not done in the SIAB.

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

- Cooperation of employment agencies and municipalities (Arbeitsgemeinschaften – ARGE) until the end of 2010 / joint facilities (gemeinsame Einrichtungen) since 2011), in which the BA and the municipality deal with tasks jointly,
- separated responsibilities (getrennte Trägerschaft) / municipalities exercising their duties separately (until 2011) – here the tasks are divided between the BA and the municipality⁵,
- authorised municipalities, which are also called opting local authorities or opting municipalities according to the initial experimental clause of Section 6a - here the local authority is responsible for all tasks in the sphere of SGB II.

The data of the “Unemployment Benefit II Recipient History drawn from A2LL” (LHG) come from different reporting procedures. As a rule, the IT procedure A2LL was used in all ARGE cooperation projects until 2010, and in joint facilities from 2011 onwards⁶. Authorised municipalities use various IT procedures of their own and transmit their data to the BA by means of the XSozial-BA-SGB II standard. Both of the procedures are used by municipalities with separated responsibilities. The different data standards affect the scope and quality of the data supplied.

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

2.4 Jobseeker History (ASU / XASU)

Data about jobseekers are stored in the Jobseeker History (Arbeitsuchendenhistorik – ASU / XASU). The ASU data source contains information on jobseekers who are registered with employment agencies, and from 2005 onwards also includes ARGE cooperation projects and separated responsibilities for the implementation of SGB II. The XASU data source, on the

⁵ The municipality pays the costs for housing and heating (Section 22 SGB II) and additional one-off benefit payments to cover extra costs (Section 23 (3) SGB II) and the additional benefits to support integration in accordance with Section 16 (2) Clause 2 No. 1 - 4 SGB II. The BA, on the other hand, covers the costs for regular benefits, social security contributions and integration benefits (SGB III and SGB II) and specific benefits excluding the additional benefits to support integration cited above.

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

other hand, contains the data of jobseekers in receipt of Unemployment Benefit II (ALG-II) from authorised municipalities from 2005 onwards. These data are reported in accordance with the X-Sozial-BA-SGB II standard.

3 Data preparation and sampling procedure

3.1 Corrections and validation procedures

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

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

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

3.1.1 Employee History (BeH)

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

- Gender and date of birth are taken from the Data Warehouse (DWH) of the BA. This information is harmonised across data sources.
- The territorial allocations for place of work and place of residence are updated to the status as of 31 December 2014.

3.1.2 Benefit Recipient History (LeH)

- In cases of on-going 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. There is no information on the end of the period for which the benefit is allowed in the SIAB-Regional File because benefit receipt notifications still valid on 31.12.2014 are right-censored (see section 3.4.2). 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 and the spell was not deleted, then the end date is increased by one year.
- Between 2004 and 2006 the notification procedure from which the data originate was changed. Overlaps occurring between the old and the new procedures were corrected.
- Observations with no end date or an invalid end date are excluded, since in these cases it cannot be assumed that a benefit payment was made at all.
- The territorial allocations are corrected analogous to the BeH.

3.1.3 Unemployment Benefit II Recipient History (LHG)

- Observations without a BA client number are deleted.
- Observations without a valid date of birth are deleted.
- Cancelled data records are not used.
- It only contains observations of people who are capable of work and people under the age of 65.
- In each case non-overlapping periods of benefit entitlement of a person in a certain benefit community (BG) are depicted. New observations are begun for the following administrative reasons:
 - on certain birthdays of members of the BG that are stipulated by law and relevant for structural changes in the benefit community (14, 15, 18 and 25) and the individual retirement age of members of the BG,
 - when the structure of the benefit community changes (e.g. due to entries/exits),
 - when there are changes in a variable of the BG client and

- at the beginning and the end of a case of benefit sanctions for observations from 1 April 2006 onwards. It must be taken into account, however, that it is not possible to identify the duration or type of sanction or the time when it was imposed or when it began on the basis of the data. The reason for this is that there is no corresponding variable or value that indicates the start, type or duration of the sanction.
- For the reason mentioned above, all individual-related variables that are available for the LHG source are valid for the entire duration of the observation.
- Double notifications due to the territorial reforms in 2009/2011 and the reorganisation of the institutions in 2012 were corrected as far as possible.
- The territorial allocations are corrected in the same way as the BeH in the SIAB.

3.1.4 Jobseeker History (ASU / XASU)

- Observations whose end date precedes the start date are not included in the ASU.
- There is no consolidation of the ASU observations for individual persons. Therefore, overlaps between ASU observations might occur.
- Individual-related variables that are only available for the (X)ASU sources always refer to the beginning of the spell.
- A new ASU spell is generated as soon as a change of status occurs (e.g. from seeking work to unemployed). This also applies if the type of institution (employment agency, cooperation of employment agency and municipality, joint facility, authorised municipalities, separated responsibilities) changes. The ASU data basis only distinguishes between observations with the status “unemployed” and “jobseeker” and since 2006 between “seeking advice” and “without status”. In the IEB, however, the additional status “ill / not able to work” is available. The employment status “ill / not able to work” is assigned to IEB spells when in the ASU data basis
 - a preceding observation with the status “unemployed” exists which joins the next observation without a gap and has “incapacitated for work” as the reason for exit, and
 - a subsequent observation with the status “unemployed” exists which also follows without a gap, and
 - the observation itself does not have the status “unemployed” but “jobseeker”.
- In contrast to the ASU source, the XASU only distinguishes between the status
 - “not unemployed, but seeking work” or
 - “unemployed and simultaneously seeking work”.
- Unlike in the ASU, periods of illness are not taken into account when generating the “employment status”, since no information about illness is available in the XASU data. When

calculating the unemployment duration with XASU observations, gaps due to illness cannot be identified.

- The XASU contains non-overlapping time periods for individuals. If one of the following variables changes, in each case a new data spell is generated for the XASU:
 - change of job search status
 - change of availability
 - change of SGB II institution (due to notification procedure)
 - change of place of residence
- The territorial allocations are corrected in the same way as the BeH.

3.1.5 SGB II anonymisation

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

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

3.2 Episode splitting

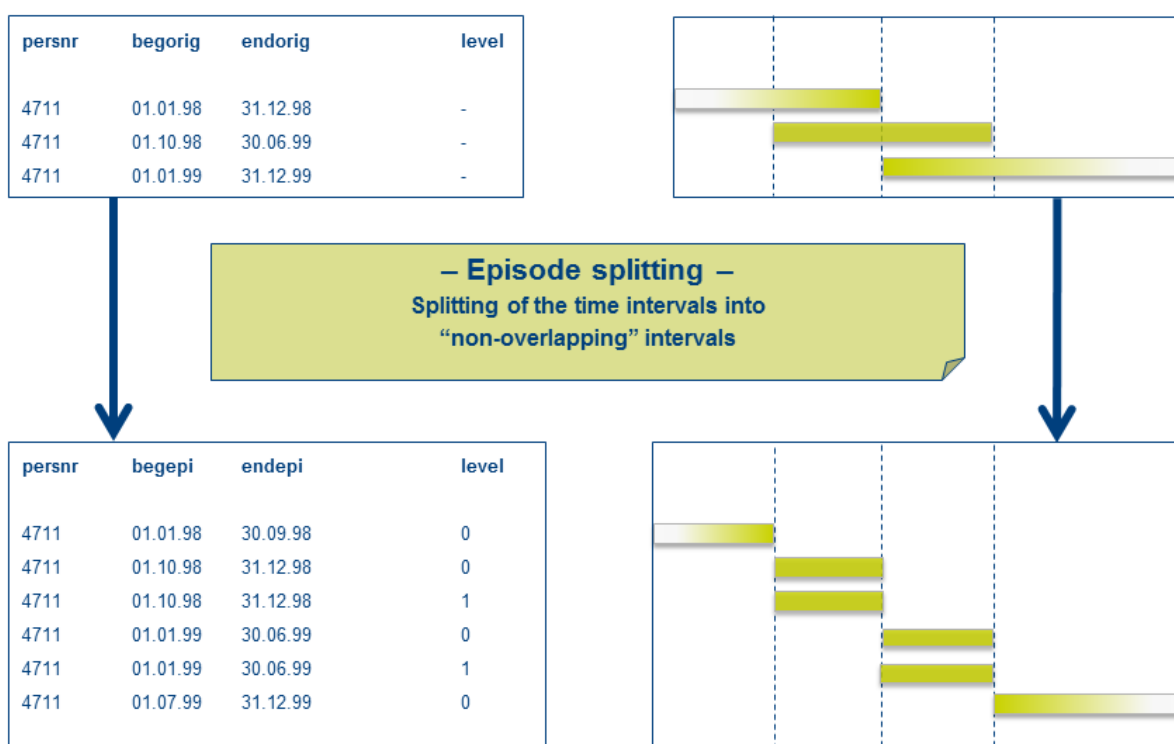
The administrative individual data are available with “split” episodes. If observations overlap within an account, these observations are replaced by artificial observations with new dates so that completely parallel periods and non-overlapping periods are created. This increases the number of observations (see Figure 2). The original date variables for the beginning and the end of the original observation (begorig and endorig) are retained, the variables 'start date of the split episode' and 'end date of the split episode' (begepi and endepi) mark the beginning and the end of the split episodes. It is possible to establish whether observations have been split by comparing the original period (begorig and endorig) with the episode period (begepi and endepi). To restore the original data without the split episodes or to delete the episodes that were created artificially by means of episode splitting it is necessary to select all observations for which the start of the original observation is the same as the start of the split episode (`begepi == begorig`).

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

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

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

Figure 2: Episode splitting



3.3 Sampling procedure

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

3.4 Anonymisation of the SIAB-R 7514

The Sample of Integrated Labour Market Biographies, which serves as the data basis for the SIAB-Regional File, contains a large amount of personal information on the status of a person

on the labour market. Additional establishment characteristics are available in the case of employment notifications from the social security notification procedure. The legislator stipulates that these social data (§ 67 SGB X) which are available at micro level at the FDZ may only be transmitted for scientific purposes (§ 282 para. 7 SGB III, § 75 SGB X) if both the personal and the establishment-related data can be regarded as factually anonymous when tested. This means that a disproportionately large amount of time, expense and effort would be required in order to identify an individual in the data set. The costs of deanonymisation must therefore exceed its benefits.

The guidelines provided by Müller et al. (1991) regarding the factual anonymity of micro data constitute the decisive orientation framework for the preparation of scientific use files at the FDZ (see for example Zimmermann et al. 2007; Drews 2008). The Sample of Integrated Labour Market Biographies was prepared and anonymised in the form of a regional file in order to guarantee the best possible continuity for users of previous scientific use files, in particular for users of the IAB Employment Samples, and in order to meet the constant demand for a regional file of employment history data from the IAB. The data were anonymised by deleting selected values and variables and aggregating information which is sensitive with regard to data protection legislation. The anonymisation procedures are described in the following.

3.4.1 Merging sources and deletion of sources

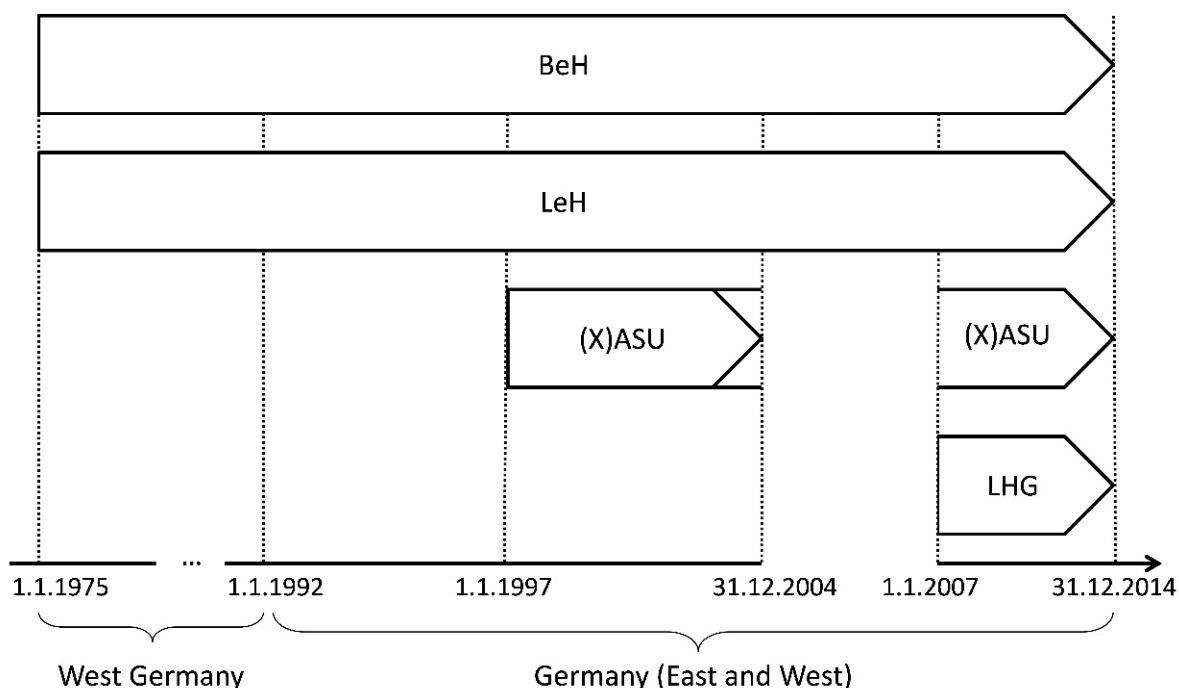
When developing the SIAB-Regional File from the SIAB 7514, details on the source of certain observations were merged or deleted. For example, information about job-search (ASU and XASU) were combined in each case to a joint data source, (X)ASU, in the SIAB-Regional File. The reason for this is that the information about whether a person in a region receives his/her support from an authorised municipality (zugelassener kommunaler Träger - zkT) constitutes a risk of the data subject being re-identified. This risk results from the fact that the information about the institution responsible for the person, in combination with regional information from other sources or from the employment history, can make it possible to identify the location at local authority level. This is also the reason why the source Participation-in-Measures History (MTH) was completely deleted: information on support by a zkT is systematically missing in the MTH.

3.4.2 Periods of validity of the sources

After aggregating the data sources (see Section 3.4.1) and in accordance with the data protection regulations mentioned above, information about the start and end dates of selected observations were adjusted. Figure 3 illustrates the periods for which the various sources are available in the SIAB-Regional File. For the Employee History (BeH) and the Benefit Recipient History (LeH) the SIAB-R 7514 only contains observations for the period between 1.1.1975

and 31.12.2014. Notifications submitted in the context of the social security notification procedure in East Germany are only assumed to be complete from 1993 onwards. In employment notifications with valid places of work in East Germany before the year 1992 the information on the district region is set to missing (.z). It is therefore only possible to distinguish whether a person is employed in East or West Germany in the SIAB-Regional File from 1992 onwards.

Figure 3: Periods of validity of the data sources



The (X)ASU is available in the SIAB-Regional File for the period between 1.1.1997 and 31.12.2004 and the period between 1.1.2007 and 31.12.2014. As the ASU original source is only assumed to be complete from the year 1997 onwards, observations with an end date before 1.1.1997 are not reported in the SIAB-Regional File. The deletion of observations in the years 2005 and 2006 is first due to the quality-relevant under-representation of notifications from the authorised municipalities responsible for implementing SGB II (zugelassene kommunale Träger) in the IEB original data (see Antoni et al. 2016). Second, revealing these data is problematic in terms of data protection legislation because of the risk of individuals reported by authorised municipalities being identified in combination with regional information. The outlined problems associated with notifications from authorised municipalities also apply to the LHG in 2005 and 2006. Accordingly the SIAB-Regional File only contains observations from the LHG with episode start dates from 1.1.2007 onwards.

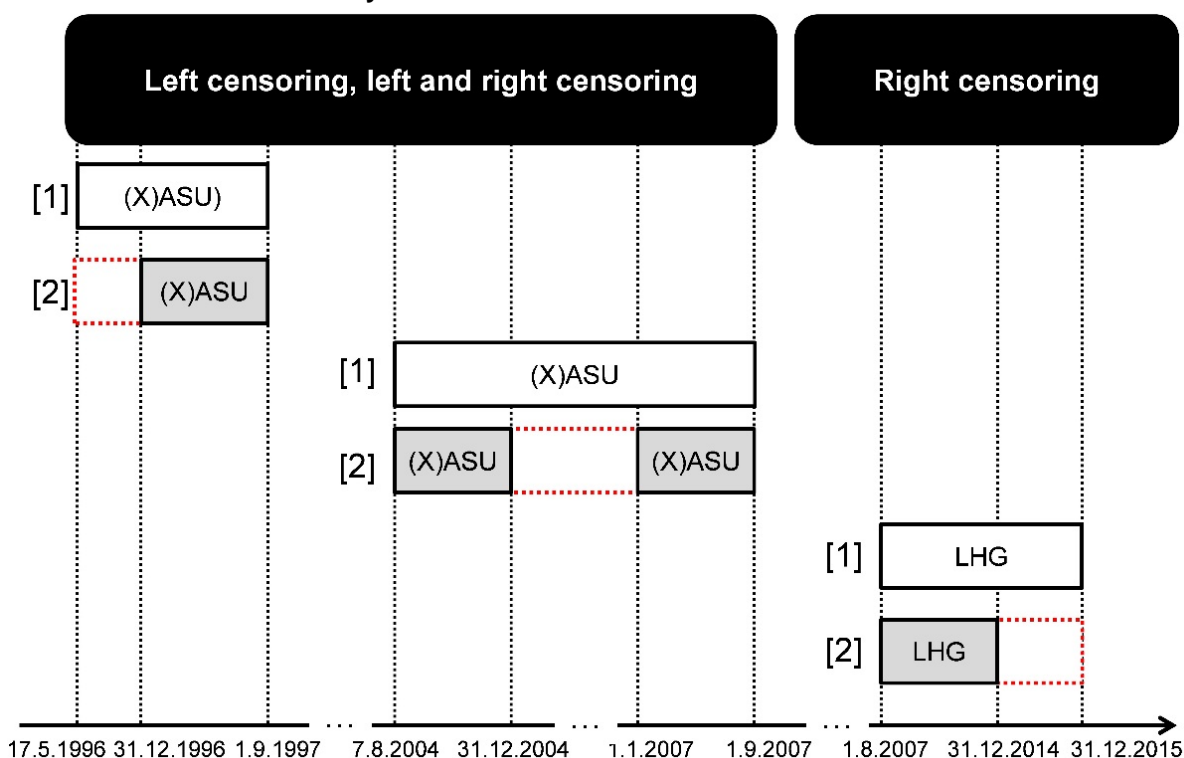
Observations whose start and end dates are outside these source-specific periods of validity are no longer contained in the data. In the case of observations from the LeH, (X)ASU and

(X)LHG sources which span more than one year and in which either the start date or the end date falls in one of the source-specific periods of validity, the dates are changed. This means that the original information regarding either the start or end date is replaced by the respective start or end date of the source-specific period of validity.

(X)ASU records constitute a peculiarity in the data. They can begin before the deleted period of 2005/2006 and can extend into 2007 without interruption. These observations were duplicated in the SIAB-Regional File and the start and end dates were adapted to the period of validity, so the data contain two observations – a right-censored record and a left-censored one.

The following fictitious examples in Figure 4 illustrate the procedure used for adjusting the dates. The white bar [1] always represents the original period and the grey bar [2] the period contained in the SIAB-Regional File after the date adjustment procedure.

Figure 4: Examples of the date adjustment procedure for observations spanning more than one year



In the example showing the left-censoring of a job-search observation from the (X)ASU to 1.1.1997 and in the example depicting the right-censoring of an observation from the LHG (also LeH or (X)ASU) to 31.12.2014, the number of data records in the data remains the same. Only the parts of the multi-year observation which are outside the observation period covered by the SIAB-Regional File are deleted.

The example in the middle shows the procedure used for multi-year observations which span the period deleted from the (X)ASU in 2005 and 2006. In this case two censored observations are created from the original notification. Accordingly the number of lines of data in the SIAB-Regional File increases compared with the original file.

3.4.3 Censoring of data on age

For data confidentiality reasons the scientific use file only contains those data records in which a person's age is between 17 and 62. Records that end before the age of 17 or begin after the age of 62 are excluded. In the case of records spanning more than one year which begin before or in the year of a person's 17th birthday but end after that date, the original start date ('begorig') was converted to 1st of Januar of the year in which the person reaches the age limit of 17. Similarly, the original end date ('endorig') was converted to 31 December of the year in which the person reaches the age limit of 62 if the record begins before or in the year of a person's 62nd birthday and end after that year. The variables 'tage_jung' and 'tage_alt' report the number of days in (regular) employment⁷ before the age of 17 and after the age of 62, respectively.

3.4.4 Aggregation and deletion of variables

While in the case of the sources with information on benefit receipt and job-search activity the spectrum of variables was strongly reduced compared with the weakly anonymous version of the SIAB for data protection reasons, the employment-related variables from the Employee History (BeH) have largely been retained and aggregated in the SIAB-Regional File.

The variables remaining in the SIAB-Regional File constitute the basis for the application of further anonymisation procedures in the form of the aggregation of information. Müller et al. (1991, p. 444 f.) suggest the following three rules for testing and creating a factually anonymous regional file:

1. *It must not be possible to identify any regional unit contained in the data which has fewer than 100,000 inhabitants by combining regional classifications (see Müller et al. 1991, p. 444)*

This demand was met by checking the number of inhabitants in the 402 autonomous municipal districts (kreisfreie Städte) and rural districts (Landkreise) as of 31.12.2014⁸. The regionalisation contained in the SIAB-Regional File combines the information on districts in the original

⁷ The calculation of periods of employment takes the following restrictions into account: `quelle_gr==1 & level2==0` (Stata Syntax).

⁸ Data basis: Population projection of the Federal Statistical Office (Statistisches Bundesamt) as of 31.12.2014 for the autonomous municipal authorities (kreisfreie Städte) and the administrative districts (Kreise) (NUTS 3-level) of the Federal Republic of Germany.

data in such a way that each of the district regions created has no fewer than 100,000 inhabitants as of the reference date. This regionalisation in the SIAB-Regional File is based on the procedure used to create the microcensus district regions (Mikrozensuskreisregionen - MZKR) (see Christians/Wirth 2009, p.22). The aggregation results in a total of 326 district regions which are shown in the variable 'place of work: region' (ao_region) (territorial allocation 31.12.2014). When defining the district regions superordinate spatial units such as spatial planning regions according to the Federal Institute for Research on Building, Urban Affairs and Spatial Development (Bundesinstitut für Bau-, Stadt- und Raumforschung - BBSR) and NUTS-II regions were taken into account as far as possible (see Table A 9 and Figure A 1).

In addition the variable 'commuter status' (pendler) was generated, which indicates whether the current place of residence differs from the current or most recently reported place of work.

2. *Variables that may potentially be suitable for identifying individuals must be so strongly aggregated that no values of variables are shown which*
 - a. *do not comprise at least 50,000 individuals in the population,*
 - b. *do not contain at least three individuals in the data set per regional unit transmitted; variable values which contain only one or two cases in the data set may only be shown in a more strongly aggregated form (see Müller et al. 1991, p. 444 f.).*
3. *If necessary for compliance with data protection legislation, all other variables contained in the regional file must be aggregated to a level at which every variable value disclosed covers at least 5000 persons in the population (see Müller et al. 1991, p. 445).*

On the basis of these criteria, further variables were aggregated when producing the SIAB-Regional File. Frequency counts of individuals as of 30th June of any year were taken as the basis. Attributes of variables with a high risk of deanonymisation were aggregated first if their univariate marginal distribution as of 30th June of the respective year did not comprise a minimum number of persons (2a). Second, they were aggregated if the cells of the contingency tables showing the relationship between the particular variable and the variable 'place of work: region' did not contain a minimum of three persons (2b). When checking the data on economic activity, the number of establishments was taken as the basis and not the number of persons as this is an establishment variable. All other variables are counted over the entire observation period of the data and are aggregated if necessary on the basis of (3) if at any time the minimum case numbers given above are not reached in the data cells of the SIAB 7514.

If parallel information from different sources is available for a person on the reference date, e.g. job-search and simultaneous benefit receipt, then the person was only taken into account once in a variable that is available in both sources (e.g. age or nationality) when conducting the frequency counts covering more than one data source.

Aggregations were conducted for the following variables: nationality, occupation – activity performed, classification of economic activity, occupational status, reason for notification/ reason for end of benefit receipt/ reason for discontinuation of unemployment benefit II/ reason for deregistration. For these variables the years 1975 up to and including 1984 were not used to assess the aggregation level required for the factual anonymity of the variables. Due to the structural change on the labor market, a strict interpretation of the anonymisation rules described above would have resulted in a very high level of aggregation in this period and therefore in a substantial restriction of the analysis potential for the data for certain groups of persons. The daily wage and the benefit rates were rounded to whole Euros.

3.4.5 Consequences of the anonymisation

The information on employment and benefit receipt in accordance with SGB III in the SIAB-Regional File are only affected to a small extent by deletions at the level of data records and individuals. The deletion of individual data records for anonymisation reasons in the process of adjusting the observation periods for the (X)ASU and (X)LHG sources resulted in some individuals being dropped from the SIAB-Regional File. As a result of the data preparation procedures mentioned above a total of 50.697 individual accounts were deleted with reference to the SIAB 7514. The SIAB-Regional File therefore still contains 97.1 percent of the persons included in the source file, SIAB 7514.

Some personal and establishment-related variables were either aggregated or deleted from the data during the anonymisation process across all sources. In the data sources with information on benefit receipt and job-search, the spectrum of variables was reduced considerably, while the employment-related variables were largely retained and were simply aggregated. More detailed information about deletions and aggregations at variable level due to anonymisation compared to the weakly anonymous version of the SIAB 7514 and compared to related FDZ data products can be found in the tables in the Appendix to the data report (see section 7.2).

3.5 Missing values

In the SIAB, missing values are coded as follows:

Term	Value	Description
No (valid) details available	.z	Values of a variable which are not systematically missing, i. e. the variable is available in principle for the data source, but no details are available for the value considered or cannot be interpreted reasonably. Please note that it is not possible to distinguish between anonymised and invalid values originating from the weakly anonymous version of the SIAB.

Systematically not available	.n	A variable is not available in principle for a data source (dark grey cells in the overview of variables in section 1.5) or is not available for a certain period.
------------------------------	----	--

4 Data quality and problems

4.1 Entire IEB

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

For evaluation purposes, it is often relevant to know gaps in the included biographies (e.g. for 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 IEB.

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 of SIAB 7514. The list makes no claims of being complete. The anonymization of this variable in the SIAB Regional File further reduces identifiability.

Table 6: Biographical gaps and identification possibilities

Biographical gap	Information on gap, identifiable using the details in the "grund" variable in the preceding observation of the source, if necessary
Civil servants, professional soldiers, judges, employees of bodies or foundations under public law	XASU
Self-employed persons without support	LeH, ASU
Students, persons in school-based further education	LeH, LHG, ASU, XASU
Persons who are ill / not able to work for more than 6 weeks (illness during unemployment, however, is represented in the ASU source under certain circumstances, see Section 3.1.4)	BeH, LeH, ASU

Persons receiving old-age pension without employment if not a member of a benefit community	LeH, LHG, ASU
Individuals on maternity leave / parental leave	XASU
Recipients of early retirement benefits	LeH, ASU
Trade professionals working from home	
Employees working short-time	ASU
Persons in youth welfare facilities, in vocational training centres, approved workshops or similar facilities for disabled persons	ASU
Participants in programmes to support participation in working life (people in rehabilitation)	ASU
(Sideline) farmers	
Caregivers according to Section 19 SGB XI	
Conscripts	BeH, LeH, LHG, ASU, XASU
Persons in reserve duty training	BeH, LeH, LHG, ASU, XASU
Persons fulfilling community service	BeH, LeH, LHG, ASU, XASU
Persons fulfilling a voluntary social or ecological year instead of community service	
Other people not registered with the statutory pension insurance or the Federal Employment Agency (e.g. sabbatical, funding from personal assets or pensions, emigration, employment abroad, voluntary work etc.)	BeH, LeH, ASU
Strikers in cases where the strike lasts more than a month	LeH
Social assistance recipients (prior to the introduction of SGB II in 2005), recipients of welfare payments (according to SGB II)	
Recipients of compensation according to FELEG (Gesetz zur Förderung der Einstellung der landwirtschaftlichen Erwerbstätigkeit, Act on Support in Case of Termination of Farming Activities)	

4.2 Employee History (BeH)

- The introduction of the new occupation code in 2011 led to a number of problems. For example, during the transition period granted to employers in the social security notification procedure,⁹ there was a temporary increase in the number of missing details. Analyses of the BA statistics (Bertat et al., 2013, p. 10) show that in 20 to 30 percent of cases no information was contained in the new or converted variables ‘occupation - activity performed’,

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

‘working hours’ and ‘vocational education and training’ after the switchover. This situation began to improve significantly in the first half of 2013. In order to improve the quality of the ‘working time’ variable in the transition period, Ludsteck/Thomsen (2016) developed an imputation procedure to replace the missing values by imputed values. The imputed data are already used in the SIAB 7514.

- Due to the introduction of the employment notification procedure in the federal states of eastern Germany, the notifications for eastern Germany can only be assumed to be sufficiently complete from 1993 onwards. For the same reason, a large number of spells for 1991 have missing values for several variables (such as employment status).
- The increase in the number of BeH observations from 1999 onwards is due to the introduction of the obligation to submit employment notifications for people in marginal part-time employment from 1 April 1999 onwards.
- Especially in 1999, observations of part-time employment increase significantly. This is caused by the actually observed increase in part-time work as well as by the fact that since 1999 employment notifications have been completed more correctly.
- Within the employment notification procedure, a certain time lag is unavoidable. Although changes in employment relationships have to be reported immediately, and existing employment relationships have to be confirmed annually by April (until the annual notification 2012) or mid-February (from the annual notification 2013 onwards) of the following year, some notifications actually arrive years later. The History File of the IAB is not updated continuously, however, but at certain intervals. This is done using files of employment notifications for one particular year which were submitted 36, 18, 12 or 6 months after the end of the reporting year (e.g. the 18-month file for 2013 can be created in July 2015 at the earliest). Notifications submitted more than three years late are not taken into account at the IAB, which means that a 36-month file shows a 100 % degree of completeness by definition.
- In the version of the IEB on which the SIAB data are based the degree of completeness of the BeH observations last stood at 100 % in 2012. 18-month files were used for 2013, and the observations for 2014 come from a 6-month file. It can therefore be assumed that the BeH observations for 2013 and in particular for 2014 are slightly underreported in the SIAB. However, this should not reduce the ability to analyse the data at individual level. The missing notifications occur more frequently in a few establishments, however. This means that in individual cases the establishment data, e.g. establishment size, are grossly incorrect and will change considerably in subsequent versions.
- In 1984, a change was made in the employment notification procedure. From that time onwards, one-off payments of gross earned income were reported as part of the annual

earnings subject to social security contributions, which leads to an increase in the average daily wage. In particular, the proportion of wages and salaries above the upper earnings limit increased considerably from that year onwards (see Bender et al. 1996).

- For the years 1992 until 2000 noticeable decreases and increases in the number of notifications were observed. Decreases can be observed especially for the following 10 districts: Braunschweig (03101), Wolfsburg (03103), Emden (03402), Kassel (06633), Essen (05113), Neuss (05162), Erftkreis (05362), Hersfeld-Rotenburg (06632), Miltenberg (09676) and Kempten (Allgäu) (09763). This is due to notification problems of one or more establishments in these regions.
- Considerable decreases were also observed for the districts Salzgitter (03102) and Hoyer-swerda (14264).
- Concerning the notifications for full-time employment, especially the districts Main-Taunus (06436) and Alzey-Worms (07331) are noteworthy. They feature above-average rises. Also in this case, the reasons are notification problems at one or more establishments in these regions.
- In the years 1996 to 1998, the values 841-844 (doctors and pharmacies) within the 'occupation – activity performed' variable are very rare compared to the neighbouring years.

4.3 Benefit Recipient History (LeH)

- For the states of eastern Germany, the LeH observations were not fully recorded until 1992.
- The benefit receipt data used to be saved on magnetic tapes. Owing to a fault in one magnetic tape, the benefit receipt data up to and including 1980 are only partially contained. Thus, in the present data product, too, it can be assumed that information on benefit receipt in that period is not available in full.
- Due to an internal change of systems, there is a break in the recording of periods of exclusion from benefits and of benefit suspension in 2004. Until 1 July 2004 periods of exclusion from benefits and of benefit suspension can only be identified via the 'reason for end of benefit receipt' in the preceding LeH observation. After this date a separate observation is available with the daily benefit rate = 0 for periods of benefit exclusion and suspension.

4.4 Unemployment Benefit II Recipient History (LHG)

- With regard to the completeness of case numbers or benefit histories from the LHG data sources, there are substantial gaps in the years 2005 and 2006. These years are deleted in the SIAB-Regional file.

- Longitudinal analyses of individuals are affected by inaccuracies as it is not possible to distinguish between changes in the benefit entitlement status and relocations into and out of districts whose institutions had problems delivering data.
- Also from 2007 onwards, cases of underrecording occur at times. These generally last one month and occur mainly in the authorised municipalities.
- Underrecording and overrecording occur in connection with changes in the type of institution responsible for implementing SGB II:
 - In the context of the reform of the territories covered by the institutions, which came into force on 1 January 2011, cases of underreporting occurred in the districts covered by the employment agencies of Dessau-Roßlau, Halberstadt, Halle and Sangerhausen.
 - Double notifications due to the territorial reforms in 2009/2011 and the changes in the form of the institutions as of 1 January 2012 are already corrected as far as possible in the IEB. Nonetheless double notifications may still occur.
- In the following job centres there are inaccuracies with regard to the allocation of benefit cases:
 - between Emden and Norden between September and December 2009
 - between Döbeln and Mittelsachsen from October to December 2012
 - between Tirschenreuth and Wunsiedel from November 2012 to March 2013
- Some individuals for whom a (X)LHG spell exists are excluded entirely or partly from benefit receipt according to SGB II, for instance because they take part in a subsidised training programme, receive an old-age pension, live in an in-patient facility or a residential institution or receive insurance payments aimed at avoiding need. This affects on average 3 to 5 percent of all cases. In XSozial this person group is sometimes underrecorded by some institutions. Exclusion from benefits can not be identified in the SIAB.
- The degree of completeness of the variable “Reason for discontinuation of unemployment benefit II” is comparatively low in the SIAB data. However, it must be taken into consideration here that the “Reason for discontinuation of unemployment benefit II” is a purely person-related variable. Thus if an observation of an individual in receipt of Unemployment Benefit II is split due to a change in the structure of his/her benefit community, the variable “Reason for discontinuation of unemployment benefit II” (grund_gr) shows the value “missing” (.z).
- Although the beginning and the end of a benefit sanction lead to the observations being split, it is not possible to determine from the data either the duration or the type of the sanction or the time when the sanction was imposed or when it began.

4.5 Jobseeker History (ASU/XASU)

4.5.1 ASU

- The registered periods of job search activity in the ASU source are regarded as complete from the year 1997 onwards.
- From mid-2005 until mid-2006, the coArb IT procedure, from which the jobseeker and applicant pool data originate, was superseded by the VerBIS procedure at the Federal Employment Agency. In July 2005, coArb was first replaced by VerBIS in the employment agency in Wiesbaden as a pilot project. From December 2005 onwards, it was then gradually replaced by VerBIS in several stages in all employment agencies. The information for many of the variables recorded was gathered with different levels of differentiation and different qualitative weighting in the two systems. It is therefore very difficult to integrate these variables into the IEB, which is only possible using a special procedure (mapping). Unfortunately, a full conversion of the affected variables from coArb to VerBIS cannot be achieved by means of mapping, so for some variables there is an unusually large number of the values 'no details available', 'other' or 'missing'. Moreover, striking differences may occur in frequency counts, depending on whether the original source of the data was coArb or VerBIS. Important limitations in the analysis potential are mentioned in the corresponding description of variables.
- The coArb procedure, which was used until June 2006, supported only the placement of unemployed persons and jobseekers. Some data were also collected about individuals who were only seeking advice, but these data are incomplete. The careers advice data were collected in a separate system. In VerBIS the attributes of the job-search status were extended to include 'seeking advice' and individuals 'without status'. The latter group includes individuals eligible for Unemployment Benefit II who are only available for job placement to a limited degree. The recording of this group in VerBIS is only regarded as largely complete since January 2008.
- A change of the institution responsible for implementing SGB II or a change of place of residence does not lead to a new ASU observation, the value of the variable at the start of an episode is continued. The longer the observation becomes, the greater the risk is that the institution responsible or the place of residence is no longer correct.

4.5.2 XASU

- In the SIAB-Regional file the XASU notifications from the 69 authorised municipalities can be assumed to be complete from the year 2007 onwards. Before that, systematic underreporting can be observed. These years are therefore deleted in the SIAB Regional file.

- As there is no central allocation of identification numbers, inconsistencies or implausible job-search histories are possible in the IEB and therefore also in the SIAB data.
- A variety of variables sometimes have only a very low degree of completeness for the XASU. Variables which are affected by this include 'school-leaving qualification', 'severe disability status', 'reason for notification' as well as 'employment status prior to job search'. Although the degree of completeness of these variables improves over time, some of them are still unsatisfactory. The 'occupation – activity performed' variable is not available in the XASU for almost the entire period available. Thus, this variable is set to missing in the (X)ASU data source.
- For a number of institutions (districts), the proportion of registered recipients of unemployment benefit II who are also registered jobseekers is implausibly large at times or continuously in the IEB. One possible reason for this could be an incorrect determination of the status 'not unemployed but seeking work' by these institutions.
- The institution-related and period-related plausibility of the XASU data should be examined before use, taking the research question into account.

5 Description of variables

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

5.1 Identifiers

5.1.1 Individual ID (persnr)

Variable label	Individual ID
Variable name	persnr
Category	identifiers
Origin	BeH, LeH, LHG, ASU, XASU
Data type	numerical
Hierarchy	none
Detailed description	<p>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.</p> <p>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.</p> <p>The formation of the individual identifier which spans all data sources is based on a heuristic developed by the BA.</p>

5.1.2 Establishment counter (bnn)

Variable label	Counter of different establishment IDs
Variable name	bnn
Category	identifiers
Origin	BeH
Data type	numerical
Hierarchy	none
Detailed description	<p>Numbers the establishments in a person's working life in ascending order.</p> <p>Example: The first establishment in which a person was employed is given the value 1. If the person moves to a different establishment, this establishment is given the value 2, etc. If the person returns to an establishment in which he or she was previously employed, then this establishment is given the value that applied for the first period of employment there (e.g. 2). If a person returns to the first establishment after just one change of establishment, this would result in the sequence 1-2-1 for the variable 'bnn' over time.</p> <p>Missing establishment numbers in the original data were also set to missing (.z) in the SUF.</p>

5.2 Generated technical variables

5.2.1 Observation counter per person (spell)

Variable label	Spell counter per person
Variable name	spell
Category	generated technical variables
Origin	BeH, LeH, LHG, ASU, XASU
Data type	numerical
Hierarchy	none
Detailed description	<p>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.</p>

5.2.2 Source of spell (quelle_gr)

Variable label	Source of spell, grouped
Variable name	quelle_gr
Category	generated technical variables
Origin	BeH, LeH, LHG, ASU, XASU
Data type	numerical
Hierarchy	none

Detailed description	<p>The variable indicates the data source. The original sources ASU and XASU were each aggregated to one source for data confidentiality reasons.</p> <p>Hence, it is not possible to distinguish between the types of institution. Because of the adjustments made, the designated time periods of the sources differ from those provided in the SIAB 7514. For further information on data preparation see Chapter 3. For further information on aggregation see Table A 3.</p>
----------------------	---

5.3 Period of validity

5.3.1 Original start date of observation (begorig)

Variable label	Original start date
Variable name	begorig
Category	period of validity
Origin	BeH, LeH, LHG, ASU, XASU
Data type	date
Hierarchy	none
Detailed description	<p>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)</p> <p>Since the notification logic might under certain circumstances permit re-identification of the exact day of birth, the original information on the date was changed in these cases by applying the anonymisation procedure described in Section 3.1.5.</p> <p>1) BeH</p> <p>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.</p> <p>2) LHG, ASU, XASU</p> <p>Certain changes lead to the creation of a new observation (see Sections 3.1.3 and 3.1.4). begorig indicates the start date of the new period.</p>

5.3.2 Original end date of observation (endorig)

Variable label	Original end date
Variable name	endorig
Category	period of validity
Origin	BeH, LeH, LHG, ASU, XASU
Data type	date
Hierarchy	none

Detailed description	<p>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).</p> <p>Since the notification logic might under certain circumstances permit re-identification of the exact day of birth, the original information on the date was changed in these cases by applying the anonymisation procedure described in Section o.</p> <p>1) BeH</p> <p>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.</p> <p>2) LHG, ASU, XASU</p> <p>Certain changes lead to the creation of a new observation (see Sections 3.1.3 and 3.1.4). endorig indicates the end date of the new period.</p>
----------------------	---

5.3.3 Start date of split episode (begepi)

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

5.3.4 End date of split episode (endepe)

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

5.4 Personal information

5.4.1 Gender (frau)

Variable label	Gender
----------------	--------

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

5.4.2 Year of birth (gebjahr)

Variable label	Year of birth
Variable name	gebjahr
Category	personal variables
Origin	BeH, LeH, LHG, ASU, XASU, MTH
Data type	numerical
Hierarchy	none
Detailed description	The year of birth is constant within one individual account.
Notes on quality	In the original data, it may happen that the year of birth changes between the data sources. This is corrected during the data preparation process. The information from the social security number is given highest priority here. For further information on the anonymisation of the year of birth see Section 3.4.3.

5.4.3 Nationality (deutsch)

Variable label	nationality
Variable name	deutsch
Category	personal variables
Origin	BeH, LeH, LHG, XLHG, ASU, XASU
Data type	numerical
Hierarchy	none
Detailed description	Reports whether the person is a German citizen (1) or whether he or she is not (0). The notification may change over time. For data confidentiality reasons the variable is set to missing (.z) for individuals working in Eastern Germany or East/West border-commuters.

5.4.4 Vocational training (ausbildung)

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

Detailed description	<p>It must be taken into account that this variable has a different meaning depending on the data source:</p> <p>1) BeH</p> <p>For BeH observations the variable contains the vocational education and training reported by the employers in the employment notification procedure. The following values exist:</p> <ul style="list-style-type: none"> 1 No vocational training 2 In-company voc. training/traineeship/external voc. training 11 Degree from a university of applied sciences 12 University degree <p>In notifications using the new occupation code (see Section 1.3.2) it is no longer possible to identify graduates of universities of applied sciences clearly, as the new occupation code no longer has a separate category for this vocational qualification. They are assigned to category 12.</p> <p>“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).</p> <p>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.</p> <p>2) ASU, XASU</p> <p>For these observations the vocational education or training completed most recently is reported. The following values exist:</p> <ul style="list-style-type: none"> 1 no completed vocational training 2 in-firm vocational training/external vocational training 3 full-time vocational school (Berufsfachschule) 4 technical college (Fachschule) 5 university of applied sciences (Fachhochschule) 6 university <p>The data source XASU partly contains degrees which are not recognised in Germany. These observations were set to missing (.z) due to data confidentiality.</p>
----------------------	---

Notes on quality	<p>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).</p> <p>As a result of the switchover from coArb to VerBIS it is not possible to distinguish correctly between “no completed vocational training” and “no information available” in the ASU and MTH data sources between 2006 and 2008. A missing value in this period therefore does not necessarily mean that the person has no vocational training or that there is no information available on vocational education and training, but may also mean that it was not possible to apply the relevant data generation procedure. The degree of completeness in the XASU data source is generally low.</p>
------------------	---

5.4.5 School leaving qualification (schule)

Variable label	School leaving qualification
Variable name	schule
Category	personal variables
Origin	BeH, ASU, XASU
Data type	numerical
Hierarchy	none
Detailed description	<p>This variable contains the school leaving qualification. Different values are possible depending on the source.</p> <p>1) BeH</p> <p>The possible values change when the new occupation code is adopted (see Section 1.3.2). The values with the old occupation code are:</p> <ul style="list-style-type: none"> 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 <p>With the new occupation code the values are:</p> <ul style="list-style-type: none"> 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 certify

	<p>cate, A-level equivalent, qualification for university</p> <p>2) ASU, XASU</p> <p>The following values are possible for these data sources:</p> <ul style="list-style-type: none"> 1 No school leaving certificate 4 Lower secondary school certificate/ grade school certificate 6 Intermediate school leaving certificate 7 Upper secondary school leaving certificate from a specialised upper secondary school (Fachoberschule) 9 General upper secondary school leaving certificate, A-level equivalent, qualification for university <p>They are valid at the beginning of the period of job-search or participation in a measure. In the case of people seeking an apprenticeship position, the variable may also contain the school qualification they are working towards in the XASU data source.</p>
Notes on quality	<p>The degree of completeness has been decreasing continuously over time in the BeH and has stood at under 2/3 in recent years. In the XASU it has been increasing continuously and has stood at over 2/3 since 2012. In ASU the degree of completeness is good on the whole.</p>

5.5 Information on employment, benefit receipt and job search

5.5.1 Daily wage, daily benefit rate (tentgelt_gr)

Variable label	Daily wage/daily benefit, rounded
Variable name	Tentgelt_gr
Category	information on employment, benefit receipt and job search
Origin	BeH, LeH
Data type	numerical
Hierarchy	none
Detailed description	<p>1) BeH</p> <p>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.</p> <p>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.</p> <p>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</p>

	<p>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.</p> <p>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.</p> <p>The daily wage is shown with two decimal places. All values greater than 0 and smaller than 0.5 were rounded up to 0.1. This makes it possible to identify the above-mentioned employment interruption notifications with the condition daily wage = 0. Earnings exceeding the upper earnings limit for statutory pension insurance are censored.</p> <p>2) LeH</p> <p>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.</p> <p>In the SUF, the variable is rounded to whole Euros. All values greater than 0 and smaller than 0.5 were rounded up to 0.1.</p>
--	---

5.5.2 Occupation – current/most recent (KIdB 1988) (beruf_gr)

Variable label	Occupation – current/most recent (KIdB 1988) , grouped
Variable name	beruf_gr
Category	information on employment, benefit receipt and job search
Origin	BeH
Data type	numerical
Hierarchy	none

Detailed description	<p>The occupational classification consists of a 3-digit code and comprises about 330 values. The three-digit occupational classification is aggregated to 120 occupational groups each consisting of at least 1,000 observations per year as of the reference date of June 30th. For further information on grouping see Table A 6.</p> <p>1) BeH</p> <p>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).</p> <p>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.</p> <p>Employment notifications with an end date later than 30.11.2011 are reported using the new occupation code 2010 (KldB2010) (see Section 1.3.2). These values are recoded to the KldB1988 by transferring the key area. This results in inaccuracies.</p> <p>2) ASU, XASU</p> <p>The variable is censored in the SUF (.z).</p>
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_gr)

Variable label	Occupational group - current/most recent (KldB 2010), grouped
Variable name	beruf2010_gr
Category	information on employment, benefit receipt and job search
Origin	BeH
Data type	numerical
Hierarchy	none
Detailed description	<p>The occupational group consists of a 3-digit code and comprises about 144 values. This classification was further grouped to 124 occupational groups, each consisting of at least 1,000 observations per year as of the reference date of June 30th. For further information on grouping see Table A 7.</p> <p>1) BeH</p> <p>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</p>

	<p>title that best defines the main activity performed (see Bundesagentur für Arbeit, 2005, p. V).</p> <p>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).</p> <p>Employment notifications with an end date earlier than 30.11.2011 are reported using the old occupation code 1988 (KldB 1988) (see Section 1.3.2). These values are recoded to the KldB2010 by transferring the key area. As the new occupation code is considerably more detailed than the old one, this results in substantial inaccuracies. This must be taken into account when analysing the data.</p> <p>2) LeH, ASU, XASU</p> <p>The variable is censored in the SUF (.z).</p>
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 Level of requirement – current/most recent (KldB 2010) (niveau)

Variable label	Level of requirement - current/most recent (KldB 2010)
Variable name	niveau
Category	information on employment, benefit receipt and job search
Origin	BeH, LeH, ASU, XASU
Data type	numerical
Hierarchy	none
Detailed description	<p>1) BeH</p> <p>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).</p> <p>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</p>

	<p>of the codes in the KIdB2010, is made available separately in the variable 'level of requirement' (niveau).</p> <p>Employment notifications with an end date earlier than 30.11.2011 are reported using the old occupation code 1988 (KIdB 1988) (see Section 1.3.2). These values are recoded to the KIdB2010 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.</p> <p>2) LeH, ASU, XASU</p> <p>The variable contains the level of requirement of the last job.</p>
Note on quality	There is a considerable increase in the number of missing values in 2011 due to the change in the reporting procedure.

5.5.5 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	<p>The employee's occupational status during the notification period is reported by the employer in the "employment details".</p> <p>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.</p> <p>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 1.3.2), 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.</p>
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).

5.5.6 Employment status (erwstat_gr)

Variable label	Employment status, grouped
Variable name	erwstat_gr
Category	information on employment, benefit receipt and job search
Origin	BeH, LeH, LHG, ASU, XASU
Data type	numerical
Hierarchy	none

Detailed description	<p>This variable assumes different values with different meanings for each data source. The variable is grouped in the SUF. For further information on aggregation see Table A 4.</p> <p>1) BeH</p> <p>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.</p> <p>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 1. Accordingly, it is possible that these employment relationships are slightly overestimated.</p> <p>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).</p> <p>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.</p> <p>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 value 3.</p> <p>2) LeH</p> <p>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.</p> <p>3) ASU/XASU</p> <p>The variable 'employment status' only contains the values 'not unemployed, but seeking work' as well as 'unemployed and simultaneously seeking work' in the SUF. Observations containing the value 'seeking advice' and 'without status' were deleted due to data confidentiality.</p> <p>Since the notification logic might in many cases permit re-identification of the exact day of birth, the original information on the date was changed by applying the anonymisation procedure described in Section 213.1.5.</p>
----------------------	--

	<p>4) LHG</p> <p>For LHG datasets, the 'employment status' variable shows whether the person is registered as an employable minor, an employable person of full age or not employable from the old-age pension threshold.</p> <p>Since the notification logic might in many cases permit re-identification of the exact day of birth, the original information on the date was changed by applying the anonymisation procedure described in Section 3.1.5.</p>
--	---

5.5.7 Transition zone (gleitz)

Variable label	Transition zone
Variable name	gleitz
Category	information on employment, benefit receipt and job search
Origin	BeH
Data type	numerical
Hierarchy	none
Detailed description	<p>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 (since 2013 from € 450.01 to € 850.00) (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.</p>

5.5.8 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	<p>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.</p>
Note on quality	<p>There is a considerable increase in the number of missing values in 2011 due to the change in the reporting procedure.</p>

5.5.9 Fixed-term contract (befrist)

Variable label	Fixed-term contract
Variable name	befrist
Category	information on employment, benefit receipt and job search

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

5.5.10 Reason of cancellation/notification/termination (grund_gr)

Variable label	Reason of cancellation/notification/termination, grouped
Variable name	grund_gr
Category	information on employment, benefit receipt and job search
Origin	BeH, LeH, LHG, ASU, XASU
Data type	numerical
Hierarchy	none
Detailed description	<p>The variable is grouped in the SUF. For further information on aggregation see Table A 5.</p> <p>1) BeH</p> <p>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.</p> <p>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).</p> <p>2) LeH</p> <p>In LeH observations, the variable indicates the reason for the end of receipt of unemployment benefit, unemployment assistance or maintenance allowance. There is no information about the reasons for the start of benefit receipt in the LeH, as the information comes from the notifications submitted by the employment agencies to the health insurance institutions regarding finished periods of benefit receipt.</p>

	<p>3) LHG</p> <p>The LHG observations contain the 'reason for discontinuation of Unemployment Benefit II' and indicate the reason why current benefits have been discontinued. The 'reason for discontinuation of Unemployment Benefit II' variable refers to the individual, not to the benefit community. If the Unemployment Benefit II receipt of a different member of the benefit community is discontinued, new observations for all members of the benefit community are started on this date, but the reason for discontinuation of Unemployment Benefit II is only available for the individual whose benefit is discontinued. This variable is valid exactly at the end of the original observation.</p> <p>4) ASU/ XASU</p> <p>In the case of (X)ASU observations, the variable contains the deregistration or exit reason. In the case of a change of legal sphere, the observation is split artificially and 'generated by data splitting' is entered as the reason for deregistration.</p>
Note on quality	The proportion of valid values (degree of completeness) for the reason for notification in the LHG data sources is very small (< 20%) across all years. The degree of completeness is also low in the XASU, but improves over time.

5.5.11 Employment before the age of 17 (tage_jung)

Variable label	days in employment before the lower age limit (year of 17th birthday)
Variable name	tage_jung
Category	information on employment
Origin	BeH
Data type	numerical
Hierarchy	none
Detailed description	This variable reports how many days an individual was employed before the year of his/her 17th birthday. Only the main employment is considered. (Stata syntax: quelle_gr==1 & level2==0).

5.5.12 Employment after the age of 62 (tage_alt)

Variable label	days in employment after the upper age limit (year of 62nd birthday)
Variable name	tage_alt
Category	information on employment
Origin	BeH
Data type	numerical
Hierarchy	none
Detailed description	This variable reports how many days an individual is employed after the year of his/her 62nd birthday. Only the main employment is considered. (Stata Syntax: quelle_gr==1 & level2==0).

5.5.13 Duration of unemployment (alo_dau)

Variable label	Duration of unemployment
Variable name	alo_dau

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

5.6 Location data

5.6.1 Place of work: region (ao_region)

Variable label	place of work: district / region
Variable name	ao_region
Category	location data
Origin	BeH
Data type	numerical
Hierarchy	federal state district region
Detailed description	<p>The variable indicates the district (urban district, rural district or aggregated if the population is lower than 100,000; see Section 3.4.4) in which the employee's establishment is located. The first two digits of the 5-digit district code (Kreisschlüssel) show the code for the federal state (Bundesland), positions 1-3 indicate the regional authority (Regierungsbezirk), and positions 1-5 show the district authority (Kreis). Federal states without a regional authority have a 0 in the third position. Due to the aggregation of the districts, a clear allocation to the regional author-</p>

	<p>ity is not possible in a few cases. However, district regions were generated in such a way that an unambiguous allocation to the spatial planning regions (Raumordnungsregionen) is possible.</p> <p>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.</p> <p>For reasons of data security, the place of work is only given when corresponding to the place of work according to the main spell (level2==0) of an episode. Further parallel places of work are censored (.z).</p> <p>Further information about grouped districts can be found in Table A 9.</p>
--	---

5.6.2 Commuter status (pendler)

Variable label	commuter status
Variable name	pendler
Category	location data
Origin	BeH, LeH, (X)ASU, (X)LHG
Data type	numerical
Hierarchy	none
Detailed description	<p>The variable 'commuter status' indicates whether an individual's current or most recent place of work corresponds to his or her current place of residence. As a reference for all sources, the current place of work or the place of work from the most recent BeH notification (main employment) is taken as a basis and compared with the current place of residence.</p> <p>The variable contains the following details: 0: the district of the place of residence corresponds to the district of the (last available) place of work (no commuter) 1: the district of the place of residence does not correspond to the district of the (last available) place of work, but both districts belong to the same district region (intra-regional commuter) 2: the district of the place of residence and the district of the (last available) place of work are located in different district regions (inter-regional commuter)</p> <p>This variable is available for: BeH: 1999-2014 LeH: 1999-2014 (X)ASU: 1999 – 2004 und 2007 – 2014 LHG: 2007 – 2014</p>

	<p>Values for the variable „commuter status“ are only given for main spells (level2==0).</p> <p>If the place of residence for one source is not available in the original data for an individual's episode, the variable 'commuter status' is re-coded with missing (.z). Missing (.z) is also coded if the individual's place of residence is available in one source but there is no previous BeH observation with a valid place of work that could be used as a reference.</p>
--	---

5.7 Establishment variables

5.7.1 Economic activity 93 generated – completed by extrapolation / imputation (w93_3_gen_gr)

Variable label	w93_3 completed by extrapolation/imputation
Variable name	W93_3_gen_gr
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	<p>The variable is grouped in the SUF. For further information on aggregation see Table A 8.</p> <p>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).</p> <p>WZ93 stands for “classification of economic activity, issue 1993” (Statistisches Bundesamt, 2002). The WZ93 is based on the Statistical Classification of Economic Activities in the European Community NACE Rev. 1 („Nomenclature générale 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“).</p> <p>Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected.</p>

6 References

- Antoni, Manfred; Ganzer, Andreas; vom Berge, Philipp** (2016): Stichprobe der Integrierten Arbeitsmarktbioografien (SIAB) 1975 - 2014. FDZ Datenreport, 04/2016 (de), Nürnberg
- Bender, Stefan; Hilzendegen, Jürgen; Rohwer, Götz; Rudolph, Helmut** (1996): Die IAB-Beschäftigtenstichprobe 1975-1990. Beiträge zur Arbeitsmarkt- und Berufsforschung 197, Nürnberg.
- Bertat, Thomas; Dundler, Agnes; Grimm, Christopher; Kiewitt, Jochen; Schomaker, Christine; Schridde, Henning; Zemann, Christian** (2013): Neue Erhebungsinhalte 'Arbeitszeit', 'ausgeübte Tätigkeit' sowie 'Schul- und Berufsabschluss' in der Beschäftigungsstatistik. Methodenbericht, Bundesagentur für Arbeit – Statistik, URL: <http://statistik.arbeitsagentur.de/Statistischer-Content/Grundlagen/Methodenberichte/Beschaeftigungsstatistik/Generische-Publikationen/Methodenbericht-Neue-Erhebungsinhalte-Arbeitszeit-ausgeuebte-Taetigkeit-sowie-Schul-und-Berufsabschluss-in-der-Beschaeftigungsstatistik.pdf> (Zugriff: 31.03.2016).
- Bundesagentur für Arbeit** (Hrsg.) (2005): Schlüsselverzeichnis für die Angaben zur Tätigkeit in den Meldungen zur Sozialversicherung. Ausgabe Januar 2005, Nürnberg.
- Bundesagentur für Arbeit** (2011): Klassifikation der Berufe 2010. Band 1: Systematischer und alphabetischer Teil mit Erläuterungen, Nürnberg. URL: <http://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen/Klassifikation-der-Berufe/KldB2010/Printausgabe-KldB2010/Printausgabe-KldB-2010-Nav.html> (Zugriff: 21.04.2016).
- Bundesanstalt für Arbeit** (1988): Klassifikation der Berufe – Systematisches und Alphabetisches Verzeichnis der Berufsbenennung, Nürnberg. URL: <http://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen/Klassifikation-der-Berufe/KldB1975-1992/KldB1975-1992-Nav.html> (Zugriff: 21.04.2016)
- Cramer, Ulrich** (1985): Probleme der Genauigkeit der Beschäftigtenstatistik. In: Allgemeines Statistisches Archiv 69: S. 56-68.
- Deutsche BKK** (2016): Ratgeber Sozialversicherung 2016, Wolfsburg, URL: https://www.deutschebkk.de/fileadmin/user_upload/microsites/arbeitgeber/medien/pdf/ratgeber-sozialversicherung-2016.pdf (Zugriff: 31.03.2016).
- Drews, Nils** (2006): Qualitätsverbesserung der Bildungsvariable in der IAB-Beschäftigtenstichprobe 1975-2001. FDZ Methodenreport, 05/2006 (de), Nürnberg.
- Eberle, Johanna; Jacobebbinghaus, Peter; Ludsteck, Johannes; Witter, Julia** (2011): Generation of time-consistent industry codes in the face of classification changes * Simple heuristic based on the Establishment History Panel (BHP). FDZ Methodenreport, 05/2011 (en), Nürnberg.
- Fitzenberger, Bernd; Osikominu, Aderonke; Völter, Robert** (2006): Imputation rules to improve the education variable in the IAB employment subsample. In: Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften, Jg. 126, H. 3, S. 405-436.
- Ludsteck, Johannes; Thomsen, Ulrich** (2016): Imputation of the Working Time Information for the Employment Register Data. FDZ Methodenreport 01/2016 (en), Nürnberg.

- Meinken, Holger; Koch, Iris** (2004): BA-Beschäftigtenpanel 1998-2002. Codebuch, Nürnberg.
- Paulus, Wiebke; Matthes, Britta** (2013): Klassifikation der Berufe * Struktur, Codierung und Umsteigeschlüssel. FDZ-Methodenreport, 08/2013 (de), Nürnberg.
- Schmucker, Alexandra; Seth, Stefan; Ludsteck, Johannes; Eberle, Johanna; Ganzer, Andreas** (2016): Betriebs-Historik-Panel 1975-2014. FDZ-Datenreport, 03/2016 (de), Nürnberg.
- Statistisches Bundesamt** (2002): Klassifikation der Wirtschaftszweige, Ausgabe 1993 (WZ 93), Wiesbaden. URL: <https://www.destatis.de/DE/Methoden/Klassifikationen/GueterWirtschaftsklassifikationen/Content75/KlassifikationWZ93.html> (Zugriff: 21.04.2016).
- Statistisches Bundesamt** (2003): Klassifikation der Wirtschaftszweige, Ausgabe 2003 (WZ 2003), Wiesbaden. URL: <https://www.destatis.de/DE/Methoden/Klassifikationen/GueterWirtschaftsklassifikationen/Content75/KlassifikationWZ2003.html> (Zugriff: 21.04.2016).
- Statistisches Bundesamt** (2008): Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008), Wiesbaden. URL: <https://www.destatis.de/DE/Methoden/Klassifikationen/GueterWirtschaftsklassifikationen/Content75/KlassifikationWZ08.html> (Zugriff: 21.04.2016).
- Wermter, Winfried; Cramer, Ulrich** (1988): Wie hoch war der Beschäftigtenanstieg seit 1983? – Ein Diskussionsbeitrag aus der Sicht der Beschäftigtenstatistik der Bundesanstalt für Arbeit. In: Mitteilungen aus der Arbeitsmarkt – und Berufsforschung 4/88, S. 468-482.
- Zimmermann, Ralf; Kaimer, Steffen; Oberschachtsiek, Dirk** (2007): Dokumentation des "Scientific Use Files der Integrierten Erwerbsbiographien" (IEBS-SUF V1) Version 1.0. (FDZ Datenreport, 01/2007 (de)), Nürnberg, 62 S.

7 Appendix

7.1 Overview of date adjustments and spell deletions

Table A 1: Overview of date adjustments and spell deletions within the preparations of the SIAB-Regional file 7514

	Start and End date observation	Date after censoring	BeH	LeH	(X)ASU	LHG
			1975-2014	1975-2014	1997-2004; 2007-2014	2007-2014
Deletions of spells	Start and end date between 1.1.2005 and 31.12.2006	-			X	
	Start and end date before 1.1.1997	-			X	
	Start and end date before 1.1.2007	-				X
Left-censoring of spells	Start date between 1.1.2005 and 31.12.2006, end date after 1.1.2007	Start date 1.1.2007			X	X
Right-censoring of spells	Start date before 31.12.2004, end date between 1.1.2005 and 31.12.2006	End date 31.12.2004			X	
Left- and right-censoring of spells	Start date before 31.12.2004, end date after 31.12.2006	End date spell 1 31.12.2004; Start date spell 2 1.1.2007			X	

7.2 Comparison of SIAB 7514 and SIAB-Regionalfile 7514

Table A 2: Reduction of the number of individuals in the course of the preparation of SIAB-Regional File 7514

Basic file: number of individuals SIAB 7514		1.757.925	100.0%	
1	Deletion of observations in all sources which end before an age of 17 or begin after the age of 62	1.716.446	97.6%	-2.4%
2	Deletion of observations in the source MTH	1.713.710	97,5 %	-0,2 %
3	Deletion of observations with the categories „seeking advice“ and „without status“ in the source ASU	1.711.649	97,4 %	-0,1 %
4	Deletion of ASU observations which end before 1997	1.710.286	97,3 %	-0,1 %
5	Deletion of LHG, XASU and XLHG observations which end before 2007	1.709.506	97,3 %	-0,0 %
6	Deletion of ASU observations which start after 2004 and end before 2007	1.707.228	97,1 %	-0,1 %
Number of individuals SIAB-Regional File 7514		1.707.228	97,1 %	- 2,9%

7.3 Recoding tables of aggregated variables in comparison with the weakly anonymous version of the SIAB 7510

Table A 3: Recoding of variable “Source of the observation, grouped“ (quelle_gr)

Data sources, grouped (quelle_gr)		
SIAB 7508	SIAB-R 7508 (SUF)	Categories
1	1	BeH Employee History File
2	2	LEH Benefit Recipient History File
7, 18	3	(X)ASU Combined Job-Search History File
16, 19	4	LHG Unemployment Benefit II Recipient History File

Table A 4: Recoding of variable “Source of the observation, grouped“ (erwstat_gr)

Employment status, grouped (erwstat_gr)		
SIAB 7514	SIAB-R 7514 (SUF)	Categories
101, 140, 143	1	Employees liable to social security without special characteristics
102, 141, 121, 122	2	Trainees
109, 209	3	Marginal part-time workers
103, 142	4	Employees in partial retirement
105, 106	5	Interns and student trainees

Employment status, grouped (erwstat_gr)		
SIAB 7514	SIAB-R 7514 (SUF)	Categories
118, 205	6	Other employment status
1	11	Unemployment benefit (ALG)
2	12	Unemployment assistance (ALHI)
3	13	Maintenance benefit (UHG)
5	14	Contribution §166b AFG/§207a SGB III
31, 41, 51	21	Unemployed (implying: registered as a job seeker with the BA)
33, 43, 53, 32	22	Not unemployed, but registered as a job seeker with the BA
10, 11	31	Minor and able to work
13, 14, 15	32	Adult and able to work

Table A 5: Recoding of variable “Reason for notification/ reason for end of benefit receipt/ reason for discontinuation of unemployment benefit II/ reason for deregistration, grouped” (grund_gr)

Reason for notification/ reason for end of benefit receipt/ reason for discontinuation of unemployment benefit II/ reason for deregistration, grouped (grund_gr)		
SIAB 7514	SIAB-R 7514 (SUF)	Categories
before 1999: 31, 32, 34, 35, 36, 37, 38, 39, 40	0	Deregistration due to end of employment
since 1999: 31	1	Deregistration due to change of health insurance company (since '99)
since 1999: 32	2	Deregistration due to change of contribution group (since '99)
since 1999: 34, 35	3	Deregistration due to interruption of employment for more than one month (also industrial conflict/dispute) (since '99)
since 1999: 36	4	Deregistration due to change of payroll accounting system (since '99)
since 1999: 40	5	Simultaneous registration and deregistration due to end of employment (since '99)
since 1999: 50, 70	6	Annual notification (all employment relationships that exist as of 31 December of the notification year must be declared)
51, 52, 53	7	Employment interruption notification
54	8	Notification of a lump sum payment
33, 55, 56, 57, 59, 71, 72	9	Deregistration for other reasons
2002, 2064	10	Begin of employment
2001, 2005, 2006, 2014, 2021, 2022, 2039, 2041, 2046, 2048, 2052, 2053, 2056, 2065, 2068	11	Entitlement to other benefit
2007, 2008, 2011, 2015, 2036, 2037, 2038, 2040, 2047, 2050, 2051, 2055, 2057, 2058, 2062	12	No availability for the labour market

Reason for notification/ reason for end of benefit receipt/ reason for discontinuation of unemployment benefit II/ reason for deregistration, grouped (grund_gr)		
SIAB 7514	SIAB-R 7514 (SUF)	Categories
2026, 2060	13	End/Drop out of measure
2003, 2009, 2017, 2020, 2024, 2028, 2031, 2042, 2049	15	Entitlement exhausted/Termination of period for which benefit is allowed
2004, 2012, 2016, 2019, 2027, 2029, 2032, 2033, 2035, 2044, 2054, 2061, 2066, 2067, 2069	16	Other reasons
5031, 5036, 5061, 5079, 5088, 5089, 6012	20	Measures of active labour market policy
5037, 5038, 5066, 6001	21	Employment - self-sought
5029, 5030, 5032, 5033, 5034, 5035, 5060, 5062, 5063, 5064, 5080, 5085, 5086, 5098, 6002, 6003	22	Employment - placement by third party
5082, 6005	23	Employment - no further details
5047, 5046, 5067, 6006	24	Self-employed
5053, 5068, 6007, 6015, 6025	25	Military/civilian service, military exercise, soldier, official
5041, 5052, 5076, 5084, 6008, 6009, 6010	26	Incapacity for work (illness, maternity protection), non-employment, retirement from working life
5042, 5056, 5065, 6004	27	Recall/re-employment at the same employer
5039, 5043, 5070, 5071, 5083, 6021, 6022, 6023, 6024, 6027	28	Vocational training, further training, other training
5040, 5049, 5069, 6016	29	School education, studies
5045, 5077, 5081, 6029	30	Change of provider
5044, 5051, 5074, 6018, 6026	31	Lack of co-operation or availability
5055, 6028	32	Discontinuation of fulfillments
5050, 5075, 6017, 6030, 6031	33	Special regulations
5027, 5048, 5054, 5073, 5078, 5087, 5090, 5091, 6011, 6013, 6014, 6019, 6020	34	Other reasons
5100, 6100	35	Technical split of the spell
3001	50	Begin of employment
3005	51	School education/studies
3006	52	Relocation
3007	53	Self-denotification
3008	54	End of benefit receipt/benefit continuance
3009, 3013	55	Retirement pension, basic financial security in old age
3015	57	Loss of ability to work
3016	58	End of need for financial support
3002, 3003, 3004, 3010, 3011, 3012	59	Other reasons

Table A 6: Recoding of variable “Occupation – current/most recent (KldB 1988), grouped” (beruf_gr)

Occupation – current/most recent (KldB 1988), grouped (beruf_gr)		
SIAB 7514	SIAB-R 7514 (SUF)	Categories
011, 012, 013, 014, 015, 016, 017, 018, 019, 020, 021, 022, 023, 024, 025, 026, 027, 028, 029, 030, 031, 032, 041, 042, 043, 044	1	Farmers until animal keepers and related occupations
051, 052, 053, 054, 055, 056, 057, 058, 059, 060, 061, 062	2	Gardeners, garden workers until forest workers, forest cultivators
071, 072, 073, 074, 075, 076, 077, 078, 079, 080, 081, 082, 083, 084, 085, 086, 087, 088, 089, 090, 091, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112	3	Miners until shaped brick/concrete block makers
121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135	4	Ceramics workers until glass processors, glass finishers
141	5	Chemical plant operatives
142, 143, 144	6	Chemical laboratory workers until vulcanisers
151	7	Plastics processors
161, 162, 163, 164	8	Paper, cellulose makers until other paper products makers
171, 172, 173, 174	9	Type setters, compositors until printers (flat, gravure)
175, 176, 177	10	Special printers, screeners until printer s assistants
181, 182, 183, 184	11	Wood preparers until basket and wicker products makers
191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203	12	Iron, metal producers, melters until semi-finished product fettlers and other mould casting occupations
211, 212, 213	13	Sheet metal pressers, drawers, stampers until other metal moulders (non-cutting deformation)
221	14	Turners
222, 223, 224	15	Drillers until borers
225, 226	16	Metal grinders until other metal-cutting occupations
231, 232, 233, 234, 235, 242, 243, 244	17	Metal polishers until metal bonders and other metal connectors
241	18	Welders, oxy-acetylene cutters
251, 252, 263	19	Steel smiths until pipe, tubing fitters
261	20	Sheet metal workers
262	21	Plumbers
270, 271, 272	22	Locksmiths, not specified until sheet metal, plastics fitters

Occupation – current/most recent (KldB 1988), grouped (beruf_gr)		
SIAB 7514	SIAB-R 7514 (SUF)	Categories
273	23	Engine fitters
274, 275	24	Plant fitters, maintenance fitters until steel structure fitters, metal shipbuilders
281	25	Motor vehicle repairers
282, 283, 284	26	Agricultural machinery repairers until precision mechanics
285, 286	27	Other mechanics until watch-, clockmakers
291, 301, 302	28	Toolmakers until precious metal smiths
303, 304, 305, 306	29	Dental technicians until doll makers, model makers, taxidermists
311	30	Electrical fitters, mechanics
312, 313, 315	31	Telecommunications mechanics, craftsmen until radio, sound equipment mechanics
314	32	Electrical appliance fitters
321	33	Electrical appliance, electrical parts assemblers
322	34	Other assemblers
323	35	Metal workers (no further specification)
331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 371, 372, 373, 374, 375, 376, 377, 378	36	Spinners, fibre preparers until skin processing operatives
351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362	37	Cutters until textile finishers
391, 392	38	Bakery goods makers until confectioners (pastry)
401, 402, 403	39	Butchers until fish processing operatives
411, 412	40	Cooks until ready-to-serve meals, fruit, vegetable preservers, preparers
421, 422, 423, 424, , 425, 426, 427, 428, 429, 430, 431, 432, 433	41	Wine coopers until sugar, sweets, ice-cream makers
441, 442	42	Bricklayers until concrete workers
451, 453	43	Carpenters until scaffolders
452	44	Roofers
461, 462	45	Paviors until road makers
463, 464, 465, 466	46	Tracklayers until other civil engineering workers
470, 471, 472	47	Building labourer, general until other building labourers, building assistants, n.e.c.
481, 482	48	Stucco workers, plasterers, rough casters until insulators, proofers
483, 484, 485, 486	49	Tile setters until screed, terrazzo layers
491, 492, 502, 503, 504	50	Room equippers until other wood and sports equipment makers
501	51	Carpenters
511	52	Painters, lacquerers (construction)
512, 513, 514	53	Goods painters, lacquerers until ceramics/glass painters

Occupation – current/most recent (KldB 1988), grouped (beruf_gr)		
SIAB 7514	SIAB-R 7514 (SUF)	Categories
521	54	Goods examiners, sorters, n.e.c.
522	55	Packagers, goods receivers, despatchers
531	56	Assistants (no further specification)
541, 542, 543, 544, 545, 546	57	Generator machinists until construction machine attendants
547, 548, 549	58	Machine attendants, machinists helpers until machine setters (no further specification)
601	59	Mechanical, motor engineers
602	60	Electrical engineers
603	61	Architects, civil engineers
604, 605, 606, 607	62	Survey engineers until other engineers
611, 612	63	Chemists, chemical engineers until physicists, physics engineers, mathematicians
621	64	Mechanical engineering technicians
622, 623	65	Electrical engineering technicians until building technicians
624, 625, 626, 627	66	Measurement technicians until remaining manufacturing technicians
628	67	Other technicians
629	68	Foremen, master mechanics
631, 632	69	Biological specialists until physical and mathematical specialists
633, 634	70	Chemical laboratory assistants until photo laboratory assistants
635	71	Technical draughtspersons
681	72	Wholesale and retail trade buyers, buyers
682	73	Salespersons
683, 684, 685, 686	74	Publishing house dealers, booksellers until service-station attendants
687, 688	75	Commercial agents, travellers until mobile traders
691, 692	76	Bank specialists until building society specialists
693, 694	77	Health insurance specialists (not social security) until life, property insurance specialists
701	78	Forwarding business dealers
702, 703, 704, 705, 706	79	Tourism specialists until cash collectors, cashiers, ticket sellers, inspectors
711, 712, 713, 715, 716	80	Railway engine drivers until street attendants
714	81	Motor vehicle drivers
721, 722, 723, 724, 725, 726	82	Navigating ships officers until air transport occupations
731, 732, 733, 734	83	Post masters until telephonists
741	84	Warehouse managers, warehousemen
742	85	Transportation equipment drivers

Occupation – current/most recent (KldB 1988), grouped (beruf_gr)		
SIAB 7514	SIAB-R 7514 (SUF)	Categories
743, 744	86	Stowers, furniture packers until stores/transport workers
751	87	Entrepreneurs, managing directors, divisional managers
752, 753	88	Management consultants, organisers until chartered accountants, tax advisers
761, 762, 763	89	Members of Parliament, Ministers, elected officials until association leaders, officials
771, 772	90	Cost accountants, valuers until accountants
773	91	Cashiers
774	92	Data processing specialists
781	93	Office specialists
782, 783	94	Stenographers, shorthand-typists, typists until data typists
784	95	Office auxiliary workers
791, 792	96	Factory guards, detectives until watchmen, custodians
793, 794	97	Doormen, caretakers until domestic and non-domestic servants
801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814	98	Soldiers, border guards, police officers until judicial enforcers
821, 822, 823	99	Journalists until librarians, archivists, museum specialists
831, 832, 833, 834	100	Musicians until scenery/sign painters
835, 836, 837, 838	101	Artistic and assisting occupations (stage, video and audio) until performers, professional sportsmen, auxiliary artistic occupations
841, 842, 843, 844	102	Physicians until Pharmacists
851, 852	103	Non-medical practitioners until masseurs, physiotherapists and related occupations
853	104	Nurses, midwives
854	105	Nursing assistants
855, 857	106	Dietary assistants, pharmaceutical assistants until medical laboratory assistants
856	107	Medical receptionists
861, 863, 891, 892, 893	108	Social workers, care workers until religious care helpers
862	109	Home wardens, social work teachers
864	110	Nursery teachers, child nurses
871, 872, 873, 874	111	University teachers, lecturers at higher technical schools and academies until technical, vocational, factory instructors
875, 876, 877	112	Music teachers, n.e.c. until other teachers
881, 882, 883	113	Economic and social scientists, statisticians until scientists n.e.c.

Occupation – current/most recent (KldB 1988), grouped (beruf_gr)		
SIAB 7514	SIAB-R 7514 (SUF)	Categories
901, 902	114	Hairdressers until other body care occupations
911, 912	115	Restaurant, inn, bar keepers, hotel proprietors, catering trade dealers until waiters, stewards
913	116	Others attending on guests
921, 922, 923, 924	117	Housekeeping managers until employees by household cheque procedure
931, 932	118	Laundry workers, pressers until textile cleaners, dyers and dry cleaners
933, 934	119	Household cleaners until glass, buildings cleaners
935, 936, 937	120	Street cleaners, refuse disposers until machinery, container cleaners and related occupations

Table A 7: Recoding of variable “Occupational group - current/most recent (KldB 2010), grouped” (beruf10_gr)

Occupational group - current/most recent (KldB 2010), grouped		
SIAB 7514	SIAB-R 7514 (SUF)	Categories
111, 112, 113, 114, 115, 116, 117	1	Occupations in agriculture, forestry, and farming
121	2	Occupations in gardening
122	3	Occupations in floristry
211	4	Occupations in underground and surface mining and blasting engineering
212	5	Conditioning and processing of natural stone and minerals, production of building materials
213	6	Occupations in industrial glass-making and -processing
214	7	Occupations in industrial ceramic-making and -processing
221	8	Occupations in plastic- and rubber-making and -processing
222	9	Occupations in colour coating and varnishing
223	10	Occupations in wood-working and -processing
231	11	Technical occupations in paper-making and -processing and packaging
232	12	Occupations in technical media design
233	13	Occupations in photography and photographic technology
234	14	Occupations in printing technology, print finishing, and book binding
241	15	Occupations in metal-making
242	16	Occupations in metalworking

Occupational group - current/most recent (KldB 2010), grouped		
SIAB 7514	SIAB-R 7514 (SUF)	Categories
243	17	Occupations in treatment of metal surfaces
244	18	Occupations in metal constructing and welding
245	19	Occupations in precision mechanics and tool making
251	20	Occupations in machine-building and -operating
252	21	Technical occupations in the automotive, aeronautic, aerospace and ship building industries
261	22	Occupations in mechatronics, automation and control technology
261	23	Technical occupations in energy technologies
263	24	Occupations in electrical engineering
271	25	Occupations in technical research and development
272	26	Draftspersons, technical designers, and model makers
273	27	Technical occupations in production planning and scheduling
281	28	Occupations in textile making
282	29	Occupations in the production of clothing and other textile products
283	30	Occupations in leather- and fur-making and -processing
291	31	Occupations in beverage production
292	32	Occupations in the production of foodstuffs, confectionery and tobacco products
293	33	Cooking occupations
311	34	Occupations in construction scheduling and supervision, and architecture
312	35	Occupations in surveying and cartography
321	36	Occupations in building construction
322	37	Occupations in civil engineering
331	38	Floor layers
332	39	Painters and varnishers, plasterers, occupations in the waterproofing of buildings, preservation of structures and wooden building components
333	40	Occupations in the interior construction and dry walling, insulation, carpentry, glazing, roller shutter and jalousie installation
341	41	Occupations in building services engineering
342	42	Occupations in plumping, sanitation, heating, ventilating, and air conditioning
343	43	Occupations in building services and waste disposal
411	44	Occupations in mathematics and statistics
412	45	Occupations in biology

Occupational group - current/most recent (KldB 2010), grouped		
SIAB 7514	SIAB-R 7514 (SUF)	Categories
413	46	Occupations in chemistry
414	47	Occupations in physics
421	48	Occupations in geology, geography and meteorology
422	49	Occupations in environmental protection engineering
423	50	Occupations in environmental protection management and environmental protection consulting
431	51	Occupations in computer science
432	52	Occupations in IT-system-analysis, IT-application-consulting and IT-sales
433	53	Occupations in IT-network engineering, IT-coordination, IT-administration and IT-organisation
434	54	Occupations in software development and programming
511	55	Technical occupations in railway, aircraft and ship operation
512	56	Occupations in the inspection and maintenance of traffic infrastructure
513	57	Occupations in warehousing and logistics, in postal and other delivery services, and in cargo handling
514	58	Service occupations in passenger traffic
515	59	Occupations in traffic surveillance and control
516	60	Management assistants in transport and logistics
521	61	Driver of vehicles in road traffic
522	62	Drivers of vehicles in railway traffic
523	63	Aircraft pilots
524	64	Ship's officers and masters
525	65	Drivers and operators of construction and transportation vehicles and equipment
531, 532, 533	66	Occupations in physical security, personal protection, fire protection and workplace safety
541	67	Occupations in cleaning services
611	68	Occupations in purchasing and sales
612	69	Trading occupations
613	70	Occupations in real estate and facility management
621	71	Sales occupations in retail trade (without product specialisation)
622	72	Sales occupations (retail trade) selling clothing, electronic devices, furniture, motor vehicles and other durables
623	73	Sales occupations (retail) selling foodstuffs

Occupational group - current/most recent (KldB 2010), grouped		
SIAB 7514	SIAB-R 7514 (SUF)	Categories
624	74	Sales occupations (retail) selling drugstore products, pharmaceuticals, medical supplies and healthcare goods
625	75	Sales occupations (retail) selling books, art, antiques, musical instruments, recordings or sheet music
631	76	Occupations in tourism and the sports (and fitness) industry
632	77	Occupations in hotels
633	78	Gastronomy occupations
634	79	Occupations in event organisation and management
711	80	Managing directors and executive board members
712	81	Legislators and senior officials of special interest organisations
713	82	Occupations in business organisation and strategy
714	83	Office clerks and secretaries
715	84	Occupations in human resources management and personnel service
721	85	Occupations in insurance and financial services
722	86	Occupations in accounting, controlling and auditing
723	87	Occupations in tax consultancy
731	88	Occupations in legal services, jurisdiction, and other officers of the court
732	89	Occupations in public administration
733	90	Occupations in media, documentation and information services
811	91	Doctors' receptionists and assistants
812	92	Laboratory occupations in medicine
813	93	Occupations in nursing, emergency medical services and obstetrics
814	94	Occupations in human medicine and dentistry
816	95	Occupations in psychology and non-medical psychotherapy
817	96	Occupations in non-medical therapy and alternative medicine
818	97	Occupations in pharmacy
821	98	Occupations in geriatric care
822	99	Occupations providing nutritional advice or health counselling, and occupations in wellness
823	100	Occupations in body care
824	101	Occupations in funeral services

Occupational group - current/most recent (KldB 2010), grouped		
SIAB 7514	SIAB-R 7514 (SUF)	Categories
825	102	Technical occupations in medicine, orthopaedic and rehabilitation
831	103	Occupations in education and social work, and pedagogic specialists in social care work
832	104	Occupations in housekeeping and consumer counselling
833	105	Occupations in theology and church community work
841	106	Teachers in schools of general education
842	107	Teachers for occupation-specific subjects at vocational schools and in-company instructors in vocational training
843	108	Teachers and researcher at universities and colleges
844	109	Teachers at educational institutions other than schools (except driving, flying and sports instructors)
845	110	Driving, flying and sports instructors at educational institutions other than schools
912	111	Occupations in the humanities
913	112	Occupations in the social sciences
914	113	Occupations in economics
921	114	Occupations in advertising and marketing
922	115	Occupations in public relations
923	116	Occupations in publishing and media management
924	117	Occupations in editorial work and journalism
931, 932, 933, 934, 935, 936	118	Occupations in product design, artisan craftwork, fine arts and the making of musical instruments
941	119	Musicians, singers and conductors
942	120	Actors, dancers, athletes and related occupations
943	121	Presenters and entertainers
944	122	Occupations in theatre, film and television productions
945	123	Occupations in event technology, cinematography, and sound engineering
946	124	Occupations in stage, costume and prop design,
947	125	Technical and management occupations in museums and exhibitions
11, 12, 14, 815, 911	126	Other occupations

Table A 8: Recoding of variable “Economic activity 93 generated – completed by extrapolation / imputation, grouped” (w93_gen_gr)

Economic activity 93 generated – completed by extrapolation / imputation, grouped (w93_gen_gr)			
SIAB 7514	SIAB-R 7514 (SUF)	Industry	Anonymisation of industries in regions
011-015, 020, 050, 101-103, 111-112, 120, 131-132, 141-145	1	Agriculture, mining and quarrying	1002, 3101, 3257, 3356, 3403, 5120, 5122, 5316, 5915, 6413, 6434, 7111, 7132, 7337, 8128, 8212, 8221, 8231, 8311, 9171, 9174, 9177, 9362, 9372, 9373, 9562, 9563, 9572, 9573, 10043, 10045, 12054, 15002 16053
201-205, 231-233, 241-247, 251-252	2	Production of rubber and plastic products, processing of minerals, wood industry, chemical industry	1054, 1057, 3103, 3403, 7111, 7211, 7337, 9161, 9174, 9177, 9179, 9474, 10043, 12054, 12061 12064, 13003, 13073, 15087, 16053
271-275, 281-287, 291-297	3	Metal production and processing, mechanical engineering	3103, 12054
300-316, 321-323, 331-335, 341-343, 351-355	4	Automotive, production of data processing equipment, electrical and optical engineering	1051, 3157, 3256, 3356, 3451, 5366, 5512, 5558, 6634, 7140, 7235, 7340, 9171, 9172, 9176, 9771, 9774, 12062, 12066, 12073
151-160, 171-177, 181-183, 191-193, 211-212, 221-223, 261-268, 361-366, 371-372	5	Consumer goods	
551-555	6	Hospitality industry	3157, 3453, 5512
454-455	7	Building industry	
501-505, 511-517, 521-527	8	Sale, maintenance and repair of motor vehicles and household goods	
601-603, 611-612, 621-623, 631-634, 641-642	9	Transport and communication	9176
651-652, 660, 671-672, 701-703, 711-714, 721-726, 731-732, 741-748	10	Credit and insurance intermediation, land and housing, rentals	

Economic activity 93 generated – completed by extrapolation / imputation, grouped (w93_gen_gr)			
SIAB 7514	SIAB-R 7514 (SUF)	Industry	Anonymisation of industries in regions
401-403, 410, 900, 911-913, 921-927, 930, 950	11	Public and personal services, household services	
801-804, 851-853	12	Education, social and health-care facilities	
751-753, 990	13	Public administration, social security	

Table A 9: Regionalisation of SIAB-Regional file 7514 (ao_region)

Regionalisation of districts (NUTS 3), territorial allocation 31.12.2014 (ao_region)		
SIAB 7514	SIAB-R 7514 (SUF)	District region
1002	1002	Kiel, City
1003	1003	Luebeck, City
1051	1051	Dithmarschen
1053	1053	Herzogtum Lauenburg
1054	1054	Nordfriesland
1055	1055	Ostholstein
1056	1056	Pinneberg
1057	1057	Ploen
1058, 1004	1058	Rendsburg-Eckernfoerde/Neumuenster, City
1059, 1001	1059	Schleswig-Flensburg/Flensburg, City
1060	1060	Segeberg
1061	1061	Steinburg
1062	1062	Stormarn
2000	2000	Hamburg, City
3101	3101	Braunschweig, City
3103	3103	Wolfsburg, City
3151	3151	Gifhorn
3152	3152	Goettingen
3153	3153	Goslar
3155, 3255, 3156	3155	Northeim/Holzminden/Osterode am Harz
3157	3157	Peine
3158, 3102, 3154	3158	Wolfenbuettel/Salzgitter, City/Helmsted
3241	3241	Region Hannover
3251, 3401	3251	Diepholzau
3252	3252	Hameln-Pyrmont
3254	3254	Hildesheim
3256	3256	Nienburg (Weser)
3257	3257	Schaumburg District

Regionalisation of districts (NUTS 3), territorial allocation 31.12.2014 (ao_region)		
SIAB 7514	SIAB-R 7514 (SUF)	District region
3351	3351	Celle
3352, 3461	3352	Cuxhaven/Wesermarsch
3353	3353	Harburg
3355	3355	Lueneburg
3356	3356	Osterholz
3357	3357	Rotenburg (Wuemme)
3358	3358	Heide District
3359	3359	Stade
3360, 3354	3360	Uelzen/Luechow-Dannenberg
3361	3361	Verden
3403	3403	Oldenburg, City
3404	3404	Osnabrueck, City
3451	3451	Ammerland
3452, 3402	3452	Aurich/Emden, City
3453	3453	Cloppenburg
3454	3454	Emsland
3455, 3405, 3462	3455	Friesland/Wilhelmshaven, City/Wittmund
3456	3456	Grafschaft Bentheim
3457	3457	Leer
3458	3458	Oldenburg
3459	3459	Osnabrueck
3460	3460	Vechtachen
4011	4011	Bremen, City
4012	4012	Bremerhaven, City
5111	5111	Duesseldorf, City
5112	5112	Duisburg, City
5113	5113	Essen, City
5114	5114	Krefeld, City
5116	5116	Moenchengladbach, City
5117	5117	Muelheim an der Ruhr, City
5119	5119	Oberhausen, City
5120	5120	Remscheid, City
5122	5122	Solingen, City
5124	5124	Wuppertal, City
5154	5154	Kleve
5158	5158	Mettmann
5162	5162	Rhein-District Neuss
5166	5166	Viersen
5170	5170	Wesel
5314	5314	Bonn, City
5315	5315	Koeln, City
5316	5316	Leverkusen, City

Regionalisation of districts (NUTS 3), territorial allocation 31.12.2014 (ao_region)		
SIAB 7514	SIAB-R 7514 (SUF)	District region
5334	5334	Staedtregion Aachen
5358	5358	Dueren
5362	5362	Rhein-Erft-District
5366	5366	Euskirchen
5370	5370	Heinsberg
5374	5374	Oberbergischer District
5378	5378	Rheinisch-Bergischer District
5382	5382	Rhein-Sieg-District
5512	5512	Bottrop, City
5513	5513	Gelsenkirchen, City
5515	5515	Muenster, City
5554	5554	Borken
5558	5558	Coesfeld
5562	5562	Recklinghausen
5566	5566	Steinfurt
5570	5570	Warendorf
5711	5711	Bielefeld, City
5754	5754	Guetersloh
5758	5758	Herford
5762	5762	Hoexter
5766	5766	Lippe
5770	5770	Minden-Luebbecke
5774	5774	Paderbornlgebirge
5911	5911	Bochum, City
5913	5913	Dortmund, City
5914	5914	Hagen, City
5915	5915	Hamm, City
5916	5916	Herne, City
5954	5954	Ennepe-Ruhr-District
5958	5958	Hochsauerland District
5962	5962	Maerkischer District
5966	5966	Olpe
5970	5970	Siegen-Wittgenstein
5974	5974	Soest
5978	5978	Unna
6411	6411	Darmstadt, City
6412	6412	Frankfurt am Main, City
6413	6413	Offenbach am Main, City
6414	6414	Wiesbaden, City
6431, 6437	6431	Bergstrasse/Odenwald District
6432	6432	Darmstadt-Dieburg
6433	6433	Gross-Gerau

Regionalisation of districts (NUTS 3), territorial allocation 31.12.2014 (ao_region)		
SIAB 7514	SIAB-R 7514 (SUF)	District region
6434	6434	Hochtaunus District
6435	6435	Main-Kinzig-District
6436	6436	Main-Taunus-District
6438	6438	Offenbach
6439	6439	Rheingau-Taunus-District
6440	6440	Wetterau District
6531	6531	Giessen, City
6532	6532	Lahn-Dill-District
6533	6533	Limburg-Weilburg
6534	6534	Marburg-Biedenkopf
6535	6535	VogelsbergDistrict
6611	6611	Kassel, City
6631	6631	Fulda
6632	6632	Hersfeld-Rotenburg
6633	6633	Kassel
6634	6634	Schwalm-Eder-District
6635	6635	Waldeck-Frankenberg
6636	6636	Werra-Meissner-District
7111	7111	Koblenz, City
7131	7131	Ahrweiler
7132	7132	Altenkirchen (Westerwald)
7133, 7134	7133	Bad Kreuznach/Birkenfeld
7137	7137	Mayen-Koblenz
7138	7138	Neuwiedr Havel, City
7140, 7135	7140	Rhein-Hunsrueck-District/Cochem-Zell
7141	7141	Rhein-Lahn-District City
7143	7143	WesterwaldDistrict
7211	7211	Trier, City
7231	7231	Bernkastel-Wittlich
7232, 7233	7232	Eifel District Bitburg-Pruem/Vulkaneifel
7235	7235	Trier-Saarburg
7314	7314	Ludwigshafen am Rhein, City
7315	7315	Mainz, City
7331, 7319	7331	Alzey-Worms/Worms, City
7332, 7316	7332	Bad Duerkheim/Neustadt an der Weinstrasse, City
7333, 7336	7333	Donnersberg District/Kuselwald
7334	7334	Germersheim
7335, 7312	7335	Kaiserslautern/Kaiserslautern, City
7337, 7313	7337	Suedliche Weinstrasse/Landau in der Pfalz, City
7338, 7311, 7318	7338	Rhein-Pfalz-District/Frankenthal (Pfalz), City/Speyer, City
7339	7339	Mainz-Bingen

Regionalisation of districts (NUTS 3), territorial allocation 31.12.2014 (ao_region)		
SIAB 7514	SIAB-R 7514 (SUF)	District region
7340, 7317, 7320	7340	Suedwestpfalz/Pirmasens, City/Zweibruecken, City
8111	8111	Stuttgart
8115	8115	Boeblingen
8116	8116	Esslingen
8117	8117	Goeppingen
8118	8118	Ludwigsburg
8119	8119	Rems-Murr-District
8121	8121	Heilbronnz-Osterzgebirge
8125	8125	Heilbronn
8126	8126	Hohenlohe District
8127	8127	Schwaebisch Hall
8128	8128	Main-Tauber-District
8135	8135	Heidenheimadt
8136	8136	Ostalb District
8212	8212	Karlsruhe
8215	8215	Karlsruhe
8216, 8211	8216	Rastatt/Baden-Baden
8221	8221	Heidelberg
8222	8222	Mannheim
8225	8225	Neckar-Odenwald-District
8226	8226	Rhein-Neckar-District
8231	8231	Pforzheim
8235	8235	Calw
8236	8236	Enz District
8237	8237	Freudenstadt
8311	8311	Freiburg im Breisgau
8315	8315	Breisgau-Hochschwarzwald
8316	8316	Emmendingen
8317	8317	Ortenau District
8325	8325	Rottweil
8326	8326	Schwarzwald-Baar-District
8327	8327	Tuttlingen
8335	8335	Konstanz
8336	8336	Loerrach District
8337	8337	Waldshut
8415	8415	Reutlingen
8416	8416	Tuebingen
8417	8417	Zollernalb District
8421	8421	Ulm
8425	8425	Alb-Donau-District
8426	8426	Biberach
8435	8435	BodenseeDistrict

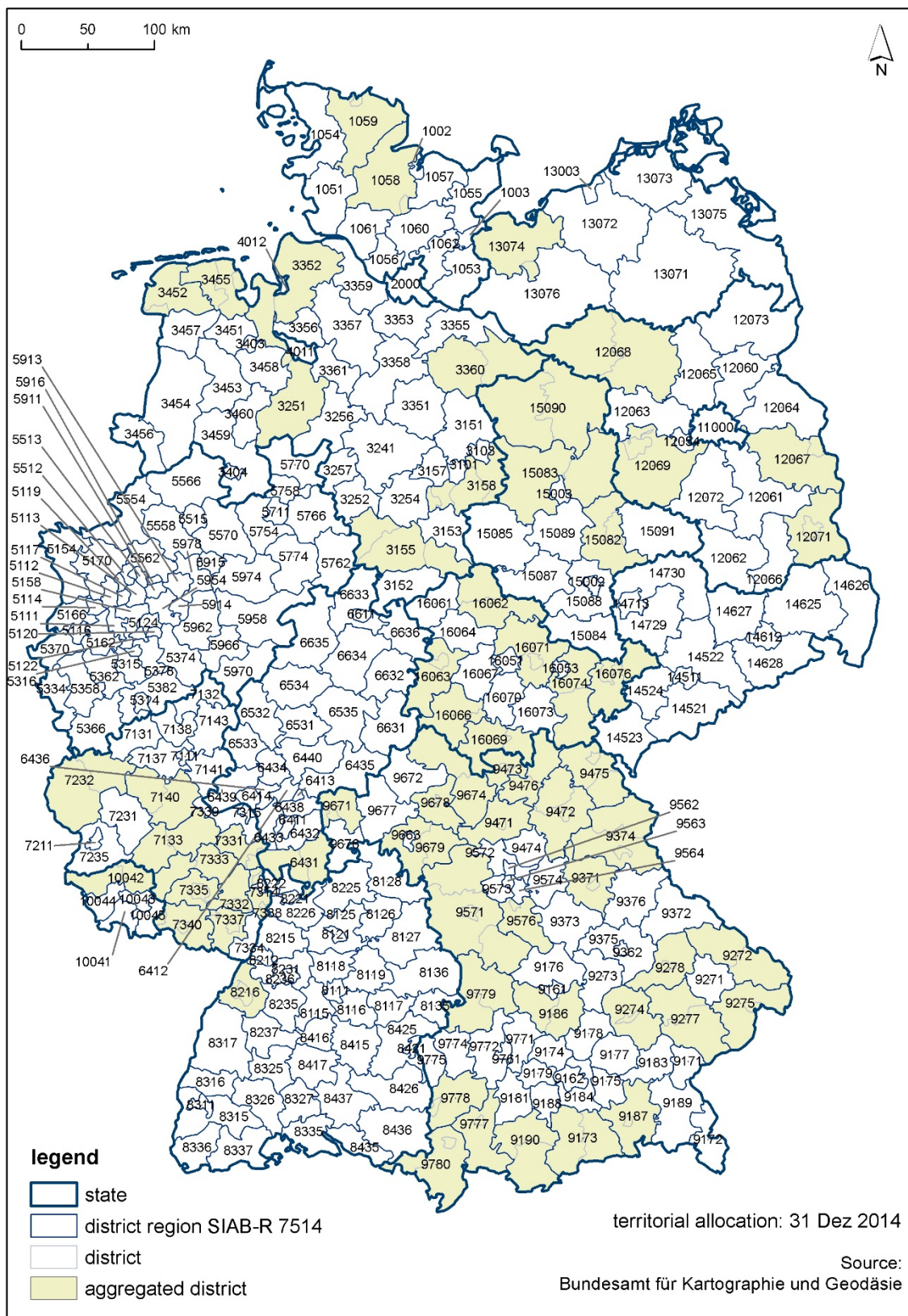
Regionalisation of districts (NUTS 3), territorial allocation 31.12.2014 (ao_region)		
SIAB 7514	SIAB-R 7514 (SUF)	District region
8436	8436	Ravensburg
8437	8437	Sigmaringeneis
9161	9161	Ingolstadt, City
9162	9162	Muenchen, City
9171	9171	Altoetting
9172	9172	Berchtesgadener Land
9173, 9182	9173	Bad Toelz-Wolfratshausen/Miesbach
9174	9174	Dachau
9175	9175	Ebersberg
9176	9176	Eichstaett
9177	9177	Erding
9178	9178	Freising
9179	9179	Fuerstenfeldbruck
9181	9181	Landsberg am Lech
9183	9183	Muehldorf a. Inn
9184	9184	Muenchen
9186, 9185	9186	Pfaffenhofen a.d. Ilm/Neuburg-Schrobenhausen
9187, 9163	9187	Rosenheim/Rosenheim, City
9188	9188	Starnberg
9189	9189	Traunstein
9190, 9180	9190	Weilheim-Schongau/Garmisch-Partenkirchen
9271	9271	Deggendorf
9272, 9276	9272	Freyung-Grafenau/Regen
9273	9273	Kelheim
9274, 9261	9274	Landshut/Landshut, City
9275, 9262	9275	Passau/Passau, City
9277, 9279	9277	Rottal-Inn/Dingolfing-Landau
9278, 9263	9278	Straubing-Bogen/Straubing, City
9362	9362	Regensburg, City
9371, 9361	9371	Amberg-Sulzbach/Amberg, City
9372	9372	Cham
9373	9373	Neumarkt i.d. OPf
9374, 9363, 9377	9374	Neustadt a.d. Waldnaab/Weiden i.d. OPf., City/Tirschenreuth
9375	9375	Regensburg
9376	9376	Schwandorf
9471, 9461	9471	Bamberg/Bamberg, City
9472, 9462, 9477	9472	Bayreuth/ Bayreuth, City/Kulmbach
9473, 9463	9473	Coburg/Coburg, City
9474	9474	Forchheim
9475, 9464, 9479	9475	Hof/Hof, City/Wunsiedel i. Fichtelgebirge
9476, 9478	9476	Kronach/Lichtenfels

Regionalisation of districts (NUTS 3), territorial allocation 31.12.2014 (ao_region)		
SIAB 7514	SIAB-R 7514 (SUF)	District region
9562	9562	Erlangen, City
9563	9563	Fuerth, City
9564	9564	Nuernberg, City
9571, 9561, 9575, 9577	9571	Ansbach/Ansbach, City/Neustadt a.d.Aisch-Bad Windsheim/Weissenburg-Gunzenhausen
9572	9572	Erlangen-Hoechstadt
9573	9573	Fuerth
9574	9574	Nuernberger Land
9576, 9565	9576	Roth/Schwabach, City
9663	9663	Wuerzburg, City
9671, 9661	9671	Aschaffenburg/Aschaffenburg, City
9672	9672	Bad Kissingen
9673, 9674	9674	Hassberge/Rhoen-Grabfeld
9676	9676	Miltenberg
9677	9677	Main-Spessart
9678, 9662	9678	Schweinfurt/Schweinfurt, City
9679, 9675	9679	Wuerzburg/Kitzingen
9761	9761	Augsburg, City
9771	9771	Aichach-Friedberg
9772	9772	Augsburg
9774	9774	Guenzburg
9775	9775	Neu-Ulm
9777, 9762	9777	Ostallgaeu/Kaufbeuren, City
9778, 9764	9778	Unterallgaeu/Memmingen, City
9779, 9773	9779	Donau-Ries/Dillingen a.d. Donau
9780, 9763, 9776	9780	Oberallgaeu/Kempten (Allgaeu), City/Lindau (Bodensee)
10041	10041	Regionalverband Saarbruecken
10042, 10046	10042	Merzig-Wadern/St. Wendel
10043	10043	Neunkirchen
10044	10044	Saarlouis
10045	10045	Saarpfalz-District
11000	11000	Berlin, City
12054	12054	Potsdam, City
12060	12060	Barnim
12061	12061	Dahme-Spreewald
12062	12062	Elbe-Elster
12063	12063	Havelland
12064	12064	Maerkisch-Oderland
12065	12065	Oberhavel
12066	12066	Oberspreewald-Lausitz
12067, 12053	12067	Oder-Spree/Frankfurt (Oder), City

Regionalisation of districts (NUTS 3), territorial allocation 31.12.2014 (ao_region)		
SIAB 7514	SIAB-R 7514 (SUF)	District region
12068, 12070	12068	Ostprignitz-Ruppin/Prignitz
12069, 12051	12069	Potsdam-Mittelmark/Brandenburg an der Havel, City
12071, 12052	12071	Spree-Neisse/Cottbus, City
12072	12072	Teltow-Flaeming
12073	12073	Uckermark
13003	13003	Rostock, City
13071	13071	Mecklenburgische Seenplatte
13072	13072	Landkreis Rostock
13073	13073	Vorpommern-Ruegen
13074, 13004	13074	Nordwestmecklenburg/Schwerin, City
13075	13075	Vorpommern-Greifswald
13076	13076	Ludwigslust-Parchim
14511	14511	Chemnitz, City
14521	14521	Erzgebirgs District
14522	14522	Mittelsachsen
14523	14523	Vogtland District
14524	14524	Zwickau
14612	14612	Dresden, City
14625	14625	Bautzen
14626	14626	Goerlitz
14627	14627	Meissen
14628	14628	Saechsische Schweiz-Osterzgebirge
14713	14713	Leipzig, City
14729	14729	Leipzig
14730	14730	Nordsachsen
15002	15002	Halle (Saale), City
15003	15003	Magdeburg, City
15082, 15001	15082	Anhalt-Bitterfeld/Dessau-Rosslau, City
15083, 15086	15083	Boerde/Jerichower Land
15084	15084	Burgenland District
15085	15085	Harz
15087	15087	Mansfeld-Suedharz
15088	15088	Saale District
15089	15089	Salzland District
15090, 15081	15090	Stendal/Altmark District Salzwedel
15091	15091	Wittenberg
16051	16051	Erfurt, City
16053	16053	Jena, City
16061	16061	Eichsfeld
16063, 16056	16062	Nordhausen/Kyffhaeuser District
16064	16063	Wartburg District/Eisenach, City
16065, 16062	16064	Unstrut-Hainich-District

Regionalisation of districts (NUTS 3), territorial allocation 31.12.2014 (ao_region)		
SIAB 7514	SIAB-R 7514 (SUF)	District region
16066, 16054	16066	Schmalkalden-Meiningen/Suhl, City
16067	16067	Gotha
16069, 16072	16069	Hildburghausen/Sonneberg
16070	16070	Ilm-District
16071, 16055, 16068	16071	Weimarer Land/Weimar, City/Soemmerda
16073	16073	Saalfeld-Rudolstadt
16074, 16075	16074	Saale-Holzland-District/Saale-Orla-District
16076, 16052, 16077	16076	Greiz/Gera, City/Altenburger Land

Figure A 1: Regionalisation SIAB-Regional file 7514 (ao_region)



7.4 Frequency tables

Frequency tables and overviews of the individual values and labels of the variables in English and German can be found in separate files at:

http://doku.iab.de/fdz/reporte/2017/DR_01-17_frequencies_labels_de.zip

http://doku.iab.de/fdz/reporte/2017/DR_01-17_frequencies_labels_en.zip

8 List of abbreviations

AA	Agentur für Arbeit / Arbeitsamt	employment agency / employment office
ALG	Arbeitslosengeld	unemployment benefit
ARGE	Arbeitsgemeinschaft	cooperation of employment agencies and municipalities
ASU	Arbeitsuchende-Historik	Jobseeker History
A2LL	Arbeitslosengeld II – Leistungen zum Lebensunterhalt	unemployment benefit II - benefits to secure a livelihood
BA	Bundesagentur für Arbeit	Federal Employment Agency
BeH	Beschäftigten-Historik	Employee History
BfA	Bundesversicherungsanstalt für Angestellte	Federal Social Insurance Office for Salaried Employees
BHP	Betriebs-Historik-Panel	Establishment History Panel
BMAS	Bundesministerium für Arbeit und Soziales	Federal Ministry of Labour and Social Affairs
coArb	Computerunterstützte Arbeitsvermittlung (operatives Verfahren zur Verwaltung der Vermittlung (Altverfahren))	computer-aided job placement (procedure for the administration of job placements – old procedure)
DEÜV	Verordnung über die Erfassung und Übermittlung von Daten für die Träger der Sozialversicherung – Datenerfassungs- und – übermittlungsverordnung	Data Collection and Transmission Regulation - regulation on the collection and transmission of data for the social security agencies
DEVO	Zweite VO über die Erfassung von Daten für die Träger der Sozialversicherung und für die BA – Datenerfassungs-Verordnung –	Data Collection Regulation - second regulation on the collection of data for the social security agencies and for the Federal Employment Agency
DÜVO	Zweite VO über die Datenübermittlung auf maschinell verwertbaren Datenträgern im Bereich der Sozialversicherung und der BA – Datenübermittlungs-Verordnung –	Data Transmission Regulation - second regulation on the transfer of data on machine-readable data media in the field of social security and the BA

EDV	Elektronische Datenverarbeitung	Electronic data processing
FDZ	Forschungsdatenzentrum	Research Data Centre
FELEG	Gesetz zur Förderung der Einstellung der landwirtschaftlichen Erwerbstätigkeit	Act on the Support in Case of Termination of Farming Activities
gAw	Träger mit getrennter Aufgabenwahrnehmung	Municipalities exercising their duties separately
gE	Gemeinsame Einrichtung	Joint facility
gT	Getrennte Trägerschaft	Separated responsibilities
IAB	Institut für Arbeitsmarkt- und Berufsforschung	Institute for Employment Research
IABS	IAB-Beschäftigtenstichprobe	IAB Employment Samples
IEB	Integrierte Erwerbsbiographien	Integrated Employment Biographies
IEBS	Stichprobe der Integrierten Erwerbsbiographien	Integrated Employment Biographies Sample
ISIC	International Standard Industrial Classification of All Economic Activities	International Standard Industrial Classification of All Economic Activities
ITM	IT- und Informationsmanagement des Instituts für Arbeitsmarkt- und Berufsforschung	IT and information management of the Institute for Employment Research
LIAB	Linked-Employer-Employee-Daten des IAB	Linked employer-employee data of the IAB
LeH	Leistungsempfänger-Historik	Benefit Recipient History
LHG	Leistungs-Historik Grundsicherung	Unemployment Benefit II Recipient History
LVA	Landesversicherungsanstalt	Land Social Insurance Office
MTH	Maßnahmeteilnehmer-Historik	Participants-in-Measures History File
NACE	Nomenclature générale des activités économiques dans les communautés européennes	Nomenclature générale des activités économiques dans les communautés européennes
Schwbg	Gesetz zur Sicherung der Eingliederung Schwerbehinderter in Arbeit, Beruf und Gesellschaft – Schwerbehindertengesetz-	law to guarantee the integration of persons with severe disabilities into employment and society – Severely Disabled Persons Act
SGB	Sozialgesetzbuch	German Social Code
SIAB	Stichprobe der Integrierten Arbeitsmarktbio graphien	Sample of Integrated Labour Market Biographies
SUF	Scientific Use File	Scientific Use File
VerBIS	Vermittlungs- und Beratungsinformationssysteme	Information System for Placement and Counselling
XASU	Arbeitsuchenden-Historik aus XSozial-BA-SGB II	Jobseeker History from XSozial-BA-SGB II
zkT	Zugelassener kommunaler Träger	Authorised municipalities

Imprint

FDZ–Datenreport 1/2017 (EN)

Publisher

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

Editorial staff

Dana Müller, Dagmar Theune

Technical production

Dagmar Theune

All rights reserved

Reproduction and distribution in any form, also in parts,
requires the permission of FDZ

Download

http://doku.iab.de/fdz/reporte/2017/DR_01-17_EN.pdf

Internet

<http://fdz.iab.de/>

Corresponding author:

Philipp vom Berge
Institute for Employment Research (IAB)
Research Data Centre (FDZ)
Regensburger Str. 104
D-90478 Nürnberg
Phone: +49-911-179-5020
Email: Philipp.Berge@iab.de