FDZ-DATENREPORT

Documentation of labour market data

Manfred Antoni, Andreas Ganzer, Philipp vom Berge

Bundesagentur für Arbeit

Sample of Integrated Labour Market Biographies Regional File (SIAB-R) 1975-2017

Manfred Antoni (IAB), Andreas Ganzer (IAB), Philipp vom Berge (IAB)

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

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

Contents

Zusammenfassung

Dieser Datenreport beschreibt das „Regionalfile" der Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB-R) 1975 - 2017. 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.

Abstract

This data report describes the "Regional File" of the Sample of Integrated Labour Market Biographies (SIAB-R) 1975 - 2017. This dataset represents the factually anonymous version of the Sample of Integrated Labour Market Biographies (SIAB) and may be transmitted to scientific research institutions after concluding a use agreement with the IAB.

Keywords

Acknowledgements

Data availability

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

1   Introduction and outline

1.1   Introduction

The Regional File of the Sample of Integrated Labour Market Biographies 1975-2017 (referred to as SIAB-R 7517 in the following) is a factually anonymous[1] scientific use file (SUF) of the Sample of Integrated Labour Market Biographies 1975-2017 (SIAB 7517)[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.

The starting point for the production of the SIAB-R 7517 is the weakly anonymous version of the SIAB 7517, a 2 % sample of the population of the Integrated Employment Biographies (IEB) of the Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung - IAB). The IEB comprises all individuals who showed one of the following statuses at least once during the observation period:

- employment subject to social security (recorded from 1975 onwards)

- marginal part-time employment (recorded from 1999 onwards)

- receipt of benefits in accordance with Social Code Book III (recorded from 1975 onwards) or Social Code Book II (recorded from 2005 onwards)

- registered with the Federal Employment Agency (Bundesagentur für Arbeit - BA) or at an institution responsible for implementing SGB II as a jobseeker (recorded from 1997 onwards)

- participation in an employment or training measure (recorded from 2000 onwards)

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

While the Employee History (Beschäftigtenhistorik - BeH) is the origin of the information on employment subject to social security and marginal part-time employment, the receipt of benefits in accordance with Social Code Book III (SGB III) and Social Code Book II (SGB II) is recorded in the Benefit Recipient History (Leistungsempfängerhistorik - LeH) and the Unemployment Benefit II Recipient History (Leistungshistorik Grundsicherung - LHG). The Jobseeker Histories (Arbeitsuchendenhistoriken - ASU and XASU) are the data source for the periods of job search recorded by the BA or by municipal institutions responsible for implementing SGB II, while participation in employment and training measures is recorded in the Participants-in-Measures History Files (Maßnahmeteilnahmehistoriken – MTH and XMTH). Unlike in previous versions of the SIAB, information about participation in employment and training measures from the XMTH is also contained in the current version of the SIAB. These participations were reported to the BA by approved municipal institutions via the transmission standard XSozial-BA-SGB II.

The SIAB-R 7517 is produced at the Research Data Centre (FDZ) of the Federal Employment Agency (BA) at the IAB. The SIAB covers the employment histories of 1,827,903 individuals, and their employment biographies are documented in a total of 62,340,521 lines of data.

---

[1] For further information on the concept of factual anonymisation see Section 3.4.

[2] For the weakly anonymous version of the SIAB 7517 see Antoni et al. (2019).

This Datenreport describes the preparation of the SIAB-R 7517 as well as the variables contained in the SUF. The Datenreport is structured as follows. In addition to the introduction, Section 1 contains information on data access as well as an outline of the data, the volume structure and a list of variables. A description of the individual data sources can be found in Section 2. Data preparation and the anonymisation concept of the SUF are discussed in Section 3. Section 4 discusses the data quality, and Section 5 describes the individual variables.

1.2 Data access

Upon request, the SIAB-R 7517 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, it is first necessary to submit a project application 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 website at http://fdz.iab.de/en.aspx.

1.3 Changes as compared to the SIAB-R 7514

1.3.1 Correction of incorrect entries in the municipality of Lahn

For the years 1975-1977 there was a misclassification of the establishments and employees in the municipality of Lahn (Lower Saxony) in the BeH. These misclassifications were corrected as far as possible. Further information is available in Section 4.2.

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

Unlike the previous version, the SIAB-R 7517 also contains information from the Participants-In-Measures History Files (MTH / XMTH, see section 2.5). Compared to the IEB, the categories of measures are only available in aggregated form.

1.3.3 Set of variables

The SIAB-R 7517 offers a number of additional variables compared with the SIAB-R 7514. The variable "Occupational status and working hours" (stib), which was removed in the previous version, is included again in the SIAB-R 7517. However, the restriction applies that the variable is only filled for notifications that were submitted before the introduction of the new occupation code (see Section 4.2). In addition, the variable "Vocational training, imputed" (ausbildung_imp, see Section 5.4.5) has been added. This variable contains imputed data for observations with originally missing information on the level of vocational training. In addition, a time-consistent version of the WZ08 industry classification was generated by extrapolation and imputation. Further information can be found in Section 5.7.1.

Table 1 gives 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
Vocational training, imputed (ausbildung_imp)	N	Newly included, but with fewer categories than the variable "Vocational training" (ausbildung)
Occupational status and working hours (stib)	N	Included again, but only filled for notifications submitted before the introduction of the new occupation code
w08_3 completed by extrapolation/imputation (w08_gen_gr)	N	Newly included; replaces the variable w93_gen_gr
w93_3 completed by extrapolation/imputation (w93_gen_gr)	D	Replaced by w08_gen_gr

N = new, D = dropped

1.3.4   Anonymisation

We tried to keep the anonymisation 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. Users of the old dataset versions should make themselves familiar with changes in the variables "ausbildung_gr", "erwstat_gr", "grund_gr", "ao_region" and "w08_gen_gr".

In addition, we changed the deletion/censoring of values of the variable "Nationality" (deutsch) to be able to censor fewer persons than in the SIAB-R 7514.

1.4   Outline

Table 2: Outline

Topics/ groups of variables	**Employee History (Beschäftigtenhistorik - BeH):**
	Annual notifications and end-of-employment notifications submitted to the social security agencies for employees covered by social security and employees in marginal part-time employment.
	Benefit Recipient History (Leistungsempfängerhistorik - LeH):
	Information on benefit receipt in accordance with Social Code Book III (SGB III) for recipients of unemployment benefit, unemployment assistance and maintenance allowance.
	Unemployment Benefit II Recipient History (Leistungshistorik Grundsicherung - LHG):
	Data on individuals in receipt of basic social security benefits in accordance with Social Code Book II (SGB II) (Types of institution: cooperation of employment agencies and municipalities/joint facilities, separated responsibilities/municipalities exercising their duties separately, authorised municipalities)

	Jobseeker History (Arbeitsuchendenhistorik - ASU) and from XSozial-BA-SGB II (Arbeitsuchendenhistorik aus XSozial-BA-SGB II - XASU):
	Information on job search activities including information that was transmitted to the BA by authorised municipalities via the transmission standard XSozial-BA-SGB II
	Participants-in-Measures History File (Maßnahmeteilnahmehistorik - MTH) and from XSozial-BA-SGB II (Maßnahmeteilnahmehistorik aus XSozial-BA-SGB II - XMTH):
	Information on participation in employment and training measures including information on participation in employment and training measures reported to the BA by authorised municipalities via the transmission standard XSozial-BA-SGB II
Data units	Employees covered by social security (including marginal part-time employees from 1999 onwards), benefit recipients, jobseekers, participants in measures, establishments
Number of cases	1,827,903 individuals 45,859,427 original observations 62,340,521 non-overlapping observations (after episode splitting)

Period covered	The period covered depends on the data source.
	BeH: 01.01.1975 - 31.12.2017
	                   (2014: 36-months-file
	                   2015 and 2016: 18-months-file,
	                   2017: 6- months-file)
	LeH: 01.01.1975 - 31.12.2017
	ASU: 01.01.1997 - 31.12.2017
	LHG: 01.01.2007 - 31.12.2017
	XASU: 01.01.2007 - 31.12.2017
	MTH: 01.01.2000 - 31.12.2017
	XMTH: 01.01.2007 - 31.12.2017

Time reference	exact to the day
Regional structure	328 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 31.12.2017
Data collection method	2% random sample
Institutions involved	Social security agencies, Federal Employment Agency (Bundesagentur für Arbeit), municipal institutions
Frequency of data collection	Continuous
File format and size	Stata: 3.1 GB; SPSS: 5.0 GB
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-R 7517 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	Data: "The data basis of this study is the factually anonymous Sample of Integrated Labour Market Biographies of the Institute for Employment Research (IAB) (version 1975 – 2017). Data access was provided via a Scientific Use File supplied by the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the IAB." DOI: 10.5164/IAB.SIAB-R7517.de.en.v1 Data documentation: Antoni, Manfred; Ganzer, Andreas; vom Berge, Philipp (2019): Sample of Integrated Labour Market Biographies Regional File (SIAB-R) 1975-2017. FDZ-Datenreport, 04/2019 (en), Nuremberg. DOI: 10.5164/IAB.FDZD.1904.en.v1
Dataset version	Sample of Integrated Labour Market Biographies Regional File (SIAB-R) – Version 7517 v1; DOI: 10.5164/IAB.SIAB-R7517.de.en.v1

1.5 List of variables

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

Table 3: Degrees of completeness of variables

	Variable is available for the data source. Degree of completeness > 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 < 0.05

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

Table 4: List of variables with degree of completeness

List of variables	Page	BeH	LeH	LHG	(X)ASU	(X)MTH
Identifiers	**35**					
Individual ID (persnr)	35	■	■	■	■	■
Establishment counter (bnn)	35	■				
Generated technical variables	**36**					
Spell counter per person (spell)	36	■	■	■	■	■
Source of spell, grouped (quelle_gr)	36	■	■	■	■	■
Period of validity	**36**					
Original start date (begorig)	36	■	■	■	■	■
Original end date (endorig)	37	■	■	■	■	■
Episode start date (begepi)	38	■	■	■	■	■
Episode end date (endepi)	38	■	■	■	■	■
Personal information	**38**					
Gender (frau)	38	■	■	■	■	■
Year of birth (gebjahr)	38	■	■	■	■	■
Nationality, grouped (deutsch)	39	■	■	■	■	■
Vocational training, grouped (ausbild-ung_gr)	39	▨			▨	▨
Vocational training, imputed (ausbild-ung_imp)	40	■				
School leaving qualification (schule)	40	▨			■	■
Information on employment, benefit receipt and job search	**41**					
Daily wage/daily benefit, rounded (tentgelt_gr)	41	■	■			
Occupation - current/most recent (KldB 1988) (beruf_gr)	43	■				

List of variables	Page	BeH	LeH	LHG	(X)ASU	(X)MTH
Occupational group - current/most recent (KldB 2010), grouped (beruf2010_gr)	43	■				
Level of requirement - current/most recent job (KldB 2010) (niveau)	44	■				
Part-time (teilzeit)	45	■				
Occupational status and working hours (stib)	45	■				
Employment status, grouped (erwstat_gr)	46	■	■	■	■	■
Transition zone (gleitz)	48	■				
Temporary agency work (leih)	48	■				
Fixed-term contract (befrist)	48	■				
Reason of cancellation/notification/termination, grouped (grund_gr)	49	■	■	▨	■	
Days in employment before age of 17 (tage_jung)	50	■	■	■	■	■
Days in employment after age of 62 (tage_alt)	50	■	■	■	■	■
Duration of unemployment (alo_dau)	50	■	■	■	■	■
Location data	**50**					
Place of work - region, grouped (ao_region)	50	■				
Commuter status (pendler)	51	■	▨	▨		
Establishment variables	**52**					
w08_3 completed by extrapolation/imputation, grouped (w08_gen_gr)	52	■				

1.6 Volume structure

Table 5: Volume structure

No. of cases	before splitting	after splitting
BeH	32,866,988	38,710,742
LeH	4,165,506	6,483,432
LHG	1,504,645	4,255,937
(X)MTH	1,022,395	1,785,856
(X)ASU	6,299,893	11,134,554
Total number of observations	45,859,427	62,340,521
Individuals	1,827,903	

2 Data sources

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

Figure 1: Data sources of the SIAB
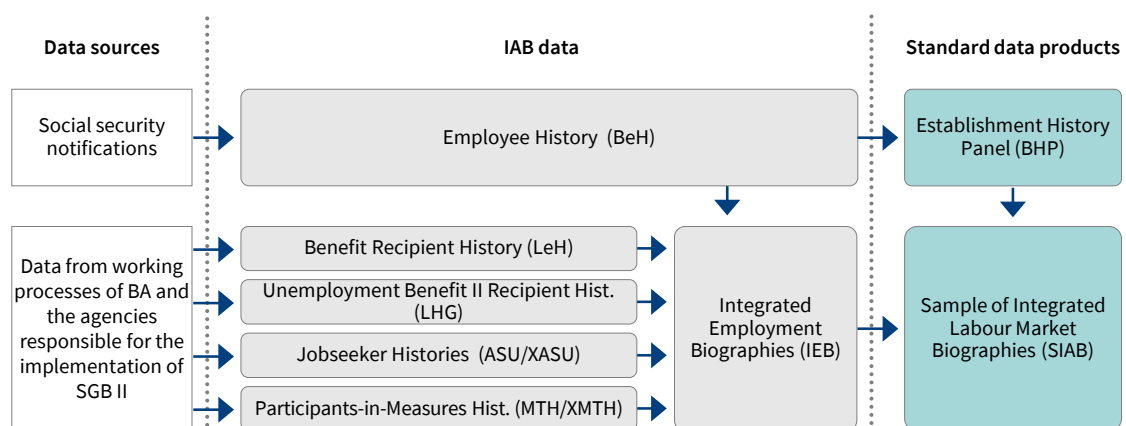
2.1 Employee History (BeH)

The source of data regarding employment is the Employee History (Beschäftigtenhistorik - BeH) of the IAB. The data basis is the integrated notification procedure for health, pension and unemployment insurance, which came into effect as of 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 et seq.; Wermter /Cramer 1988). Under this procedure,

employers are required to submit notifications to the responsible social security agencies concerning all of their employees covered by social security at least once a year. The BeH covers all white- and blue-collar workers as well as apprentices as long as they are not exempt from social security contributions. This means that civil servants, self-employed persons and regular students[3] (see Cramer 1985) are not recorded in the BeH in principle. Since the notification procedure was changed on 1 January 1999, employees in marginal part-time employment and unpaid family workers have also been recorded (not contained in the data until 1 April 1999). The data are recorded by the health insurance companies, collected and edited by the Federal Employment Agency (BA) and subsequently integrated into the History File by the IAB.

The variables of the SIAB-R 7517 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. 2018). The following modified variables based on the BHP are included in the SUF:

- Place of work - region (ao_region)
- w08_3 completed by extrapolation/imputation, grouped (w08_gen_gr)

Two issues have to be taken into account when using the establishment variables:

1. Information on establishments is based on individual data and aggregated on the annual cut-off date of the BHP, 30 June. Unlike the data on individuals, the establishment variables are not spell data but are only valid on 30 June precisely each year.[4]
2. The BHP is created directly from the BeH, whereas the SIAB individual data are generated from the IEB (see Figure 2). 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, maintenance allowance, and contributions paid by the BA to private health or care insurance while benefits are being drawn. Benefits in the context of Social Code Book II (e.g., unemployment benefit II) are not included in the data. 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 benefits and capable of work, about the members of their benefit unit (Bedarfsgemeinschaft) in accordance with § 7 SGB II and about certain individuals associated with the benefit unit. In the SIAB, however, it is not possible to link individuals with benefit receipt in accordance with Social Code Book II (SGB II) within benefit units. The

---

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

[4] An extreme example: an employment notification exists from 1 January 2006 to 30 May 2006; the establishment goes bankrupt in June 2006. In this case, there would be no information about this establishment in the BHP for 2006.

receipt of benefits in accordance with SGB II covers both basic social security benefits (e.g., Unemployment Benefit II) and supplements to unemployment benefit or additional benefits. The LHG does not contain any information about the benefit rates, however. As the amount of benefit received is not determined at the level of the individual but at the level of the benefit unit in the case of Unemployment Benefit II, it is difficult to assign an individual benefit rate. Therefore, the SIAB also does not contain information about SGB-II-benefit rates.

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

- Cooperation of employment agencies and municipalities (Arbeitsgemeinschaften – ARGE) until the end of 2010 / joint facilities (gemeinsame Einrichtungen) since 2011), in which the BA and the municipality deal with tasks jointly,

- separated responsibilities (getrennte Trägerschaft - gt) / municipalities exercising their duties separately (until 2011) – here the tasks are divided between the BA and the municipality,[5]

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

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

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

2.4 Jobseeker Histories (ASU/XASU)

Data about jobseekers are stored in the Jobseeker Histories (Arbeitsuchendenhistoriken – ASU/XASU). The ASU data source contains information on jobseekers who are registered with employment agencies, and from 2005 onwards also includes ARGE cooperation projects and separated responsibilities for the implementation of SGB II. The XASU data source, on the other hand, contains the data of jobseekers in receipt of Unemployment Benefit II (ALG-II) from authorised municipalities from 2005 onwards. These data are reported in accordance with the X-Sozial-BA-SGB II standard. The earliest available data in the ASU are from 1 January 1997 and in the XASU from 1 January 2005, respectively.

---

[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.

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

The Participants-In-Measures History Files (Maßnahmeteilnahmehistoriken - MTH) contain information that can be assigned to different legal spheres. First, they contain active labour market policy measures in accordance with Social Code Book III and participation in such measures. Second, the MTH contain measures in the legal sphere of Social Code Book II if these are recorded in BA administrative procedures. Measures implemented by authorised municipalities or opting local authorities are recorded in the XMTH from 2005 onwards. The earliest available data in the MTH are from 1 January 2000, those in the XMTH are from 1 January 2005.

3 Data preparation and sampling procedure

3.1 Corrections and validation procedures

Before the data from the 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.

No further corrections (such as the addition of presumably missing notifications, strike corrections) are performed for the IEB. Please note that additional modifications are performed when creating the SIAB-R 7517 to ensure data protection. These additional measures are described in Section 3.4.

3.1.1 Employee History (BeH)

- To capture a person group that is as constant as possible over time, some person groups for which data are not available throughout the entire observation period are excluded. From the reporting year 2011 onwards, the BeH data originate from newly designed source data. As a result, a number of person groups have been introduced or reactivated as they are classified by the BA statistics as being subject to social security contributions. The person groups 101 - 107, 111 - 114, 118, 119, 120, 140, 141, 142, 143, 149, 201 and 203 - 205 are therefore contained from that time onwards as well as the two groups 109 and 209, which indicate people in marginal part-time employment. Groups that are not included are, for example, people in short-term employment, i.e. person groups 110, 202 and 210.

- Person groups 123, 124 and 127 have been newly introduced in 2011.

- For data protection reasons, the person groups 107, 111, 113, 114, 127 and 204 are combined to form the person group "other workers" (599).

- From the reporting year 2012 onwards apprentices were included as the new person groups 121 and 122.

- Observations with no details on earnings, 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 2017.

3.1.2 Benefit Recipient History (LeH)

- Observations without a valid start date are excluded.

- If the end date for the receipt of unemployment assistance precedes the start date by one day and the spell was not deleted, then the end date is increased by one year.

- Between 2004 and 2006 the notification procedure from which the data originate was changed. Overlaps occurring between the old and the new procedures were corrected.

- Observations with no end date or an invalid end date are excluded, since in these cases it cannot be assumed that a benefit payment was made at all.

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

3.1.3 Unemployment Benefit II Recipient History (LHG)

- Observations without a BA client number are deleted.

- Cancelled data records are not used.

- It only contains observations of people who are capable of work and people under the age of 65.

- In each case, non-overlapping periods of benefit entitlement of a person in a certain benefit unit (BG) are created. New observations are begun for the following administrative reasons:
  - on certain birthdays of members of the BG that are stipulated by law and relevant for structural changes in the benefit unit (14, 15, 18 and 25) and the individual retirement age of members of the BG (see Section 3.1.6),
  - when the structure of the benefit unit changes (e.g. due to entries/exits),
  - when there are changes in a variable of the BG client and
  - at the beginning and the end of a case of benefit sanctions for observations from 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 (see Section 4.1.2) were corrected as far as possible.

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

3.1.4 Jobseeker Histories (ASU/XASU)

- Observations with an end date before 1 January 1995 are not included.

- There is no consolidation of the ASU observations for individual persons. Therefore, overlaps between ASU observations might occur.

- Individual-related variables that are only available for the (X)ASU sources always refer to the beginning of the spell.

- A new ASU spell is generated as soon as a change of status occurs (e.g., from seeking work to unemployed). This also applies if the type of institution (employment agency, cooperation of employment agency and municipality, joint facility, authorised municipalities, separated responsibilities) changes.

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

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

3.1.5 Participants-In-Measures History File (MTH)

- Observations generated more than a year after the end of the measure are deleted if another observation exists that was generated within the year after completion of the measure.

- Only the most recent record of an individual case of participation in a measure is used.

- Only cases of participation in measures that are classified as "actually took place" are included in the IEB. Cases of participation that did not take place or have not yet taken place are deleted. Cases of participation are also classed as not having taken place when a deletion date is set during the participation in a measure.

- Certain types of measure are not included. These include services to support careers advice and job placement, mobility assistance and pure rehabilitation measures.

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

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

- For XMTH, the particular challenge is to identify and handle multiple notifications of participation that are caused by technical or organisational issues. The causes of these duplicate notifications are, on the one hand, a missing identification number for participation in measures until April 2009 or improper handling of the subsequently introduced promotion ID and, on the other hand, a change of provider numbers in connection with mergers or other reorganisations of institutions responsible for implementing SGB II due to the reorganisation of the employment agencies in 2012/2013. In addition, provider-specific problems arose with version or product changes of the municipal software or with version changes of the XSozial standard. At times, this results in a considerable number of reports of participations in measures of the

same type, which partly or completely overlap in time per participant. Real duplicates are identified and sorted out because they are highly likely to be technical duplicates. Hidden duplicates can also be identified and sorted out. In principle, the rule always applies that the most up-to-date information is retained.

- In addition, notifications with a starting date before 2005 are excluded.

- Overlapping and immediately adjacent notifications of the same type of measure (from XSozial) are combined to one observation. Measure-specific characteristics of the combined periods are set to system missings.

- Subsequently, the remaining characteristics are compiled and calculated. For this, consolidated individual data from XSozial-histories and LHG are used.

- Finally, the following measures are excluded:
  - one-off benefits (like UBV/Mobi-/Vermittlungsbudget /LES) and
  - specific rehabilitation measures

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

3.1.7 SGB II anonymisation

In order to reduce the risk of deanonymisation, only the year of birth is available in the SIAB by default. However, in the LHG and (X)ASU there is the risk that the exact date of birth may still be obvious due to the chronological structure of the observations. Observations might end systematically on certain birthdays and/or the day before, or start again on the birthday.

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

3.2 Episode splitting

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

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).

Figure 2: Episode splitting

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

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

persnr	begepi	endepi	level
4711	01.01.98	30.09.98	0
4711	01.10.98	31.12.98	0
4711	01.10.98	31.12.98	1
4711	01.01.99	30.06.99	0
4711	01.01.99	30.06.99	1
4711	01.07.99	31.12.99	0
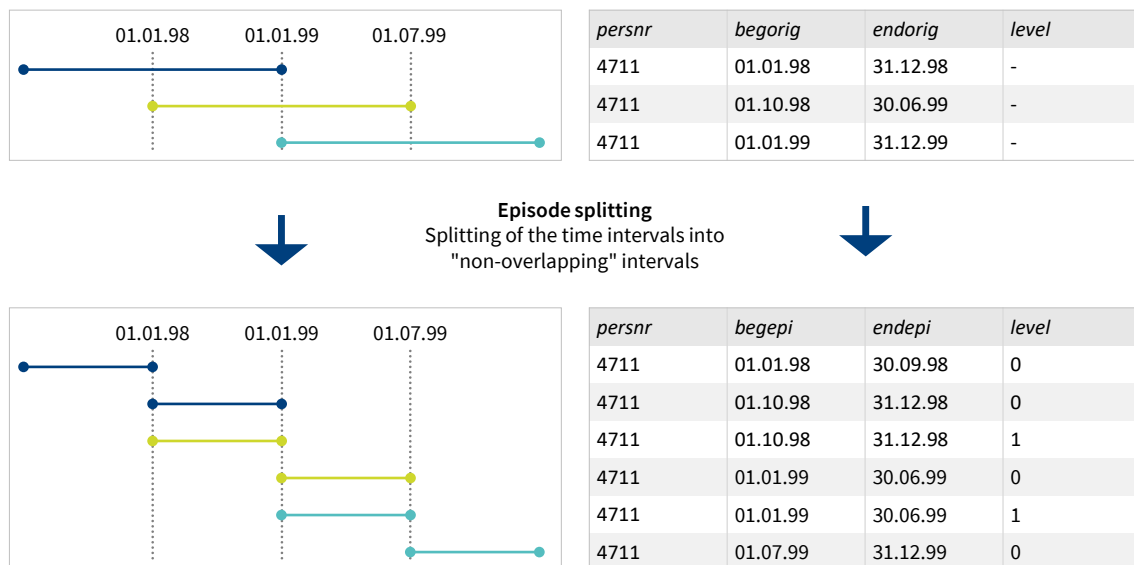
To restore the original data without the split episodes or to delete the episodes that were created artificially by means of episode splitting it is necessary to select all observations for which the start of the original observation is the same as the start of the split episode (`begepi == begorig`).

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 (see Box 1).

Box 1: Example code to create additional observation counters in Stata 14

```
bysort persnr begepi quelle (spell): gen byte level1 = _n-1
bysort persnr begepi (spell): gen byte level2 = _n-1
```

3.3   Sampling procedure

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

3.4   Anonymisation of the SIAB-R 7517

The Sample of Integrated Labour Market Biographies, which serves as the data basis for the SIAB-R 7517, 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 German Social Code (SGB) requires 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) 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 authoritative orientation framework for the preparation of scientific use files at the FDZ. The Sample of Integrated Labour Market Biographies SUF 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 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 of data sources

When developing the SIAB-R 7517 from the SIAB 7517, details on the source of certain observations were merged. For example, information from the original sources Jobseeker Histories (ASU and XASU) and the Participants-In-Measures History Files (MTH and XMTH) were combined to joint data sources (X)ASU and (X)MTH, respectively, in the SIAB-R 7517. The reason for this is that the information on whether a person receives their support from an authorised municipality (zugelassener kommunaler Träger - zkT) in a given region 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.

3.4.2 Periods of validity of data 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-R 7517. For the Employee History (BeH) and the Benefit Recipient History (LeH) the SIAB-R 7517 only contains observations for the period between 01.01.1975 and 31.12.2017. 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), as many observations have missing values during that period. It is therefore only possible to distinguish whether a person is employed in East or West Germany in the SIAB-R 7517 from 1992 onwards.

Figure 3: Periods of validity of the data sources
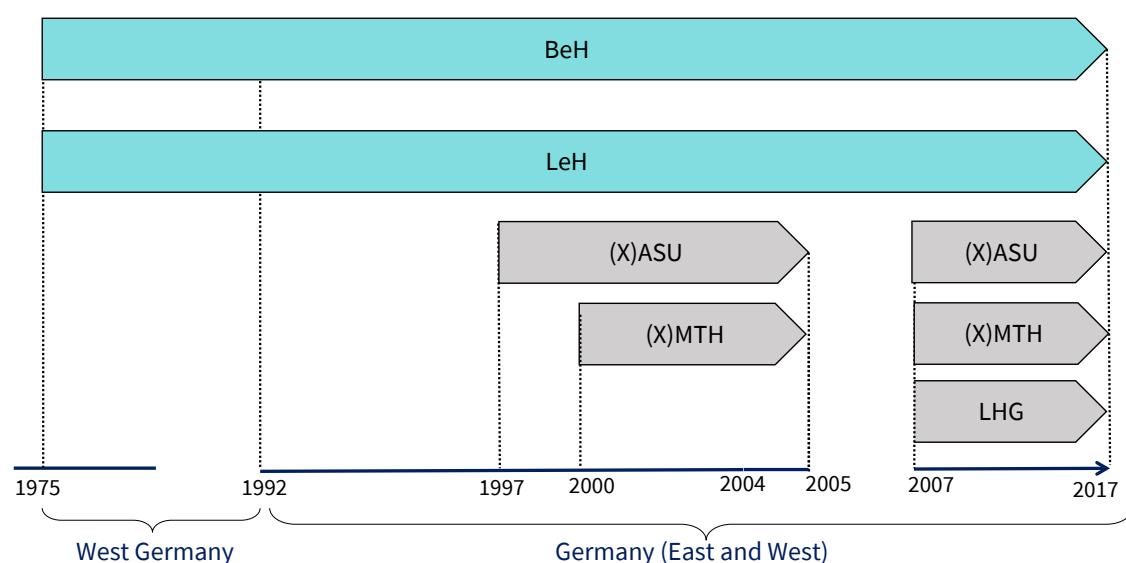
The (X)ASU is available in the SIAB-R 7517 for the period between 01.01.1997 and 31.12.2004 and the period between 01.01.2007 and 31.12.2017. As the ASU original source is only assumed to be complete from the year 1997 onwards, observations with an end date before 01.01.1997 are not reported in the SIAB-R 7517. The same applies to observations from the (X)MTH ending before 01.01.2000. The deletion of observations from the (X)ASU and the (X)MTH in the years 2005 and 2006 is, first, due to a quality-relevant under-representation of notifications from the authorised municipalities (zugelassene kommunale Träger) responsible for implementing SGB II in the IEB original data (see Antoni et al. 2019). Second, revealing these data is problematic for data privacy 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-R 7517 only contains observations from the LHG with episode start dates from 01.01.2007 onwards.

Observations whose start and end dates are outside these source-specific periods of validity are not contained in the data. In the case of observations from the LeH, LHG, (X)ASU and (X)MTH sources that span more than one calendar 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.

Records from the sources (X)ASU and (X)MTH 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-R 7517 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-R 7517 after the date adjustment procedure.

Figure 4: Examples of the date adjustment procedure for observations spanning more than one year
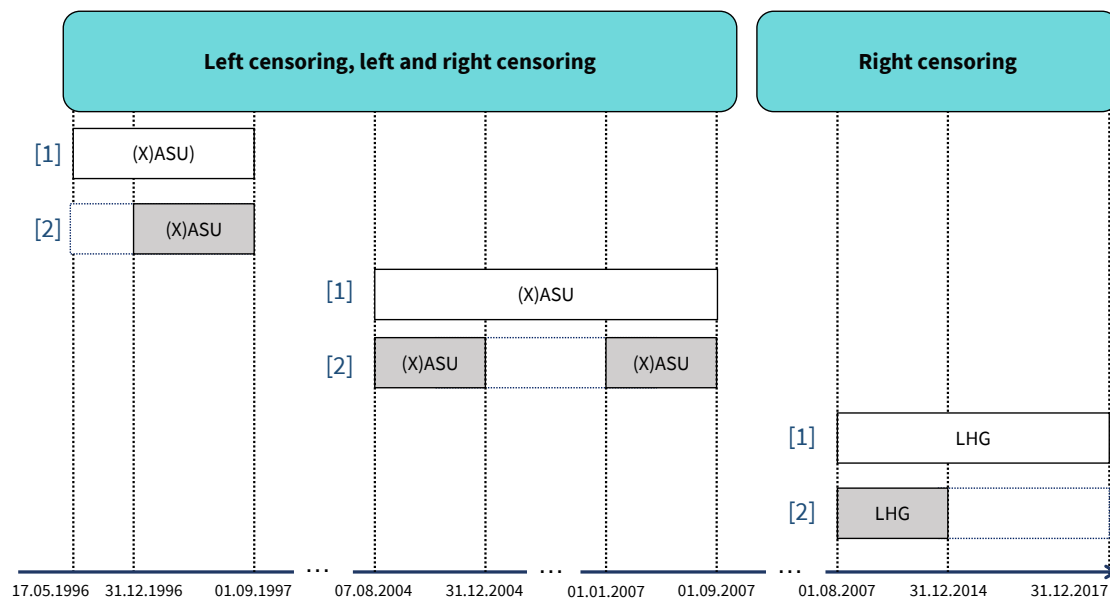
In the example showing the left-censoring of a job-search observation from the (X)ASU to 01.01.1997 and in the example depicting the right-censoring of an observation from the LHG (also applicable to LeH, (X)ASU or (X)MTH) to 31.12.2017, the number of data records in the data remains the same. Only the parts of the multi-year observation which are outside the observation period covered by the SIAB-R 7517 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-R 7517 increases compared with the original file. Table A 1 provides an overview of date adjustment in SIAB-R 7517.

3.4.3   Censoring of information on age

For data confidentiality reasons the SUF 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 January 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) employment[7] before the year of the 17th birthday and after the year of 62nd birthday, 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 set of variables was strongly reduced compared with the weakly anonymous version of the SIAB for

---

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

data protection reasons, the employment-related variables from the Employee History (BeH) have largely been retained and aggregated in the SIAB-R 7517.

The variables remaining in the SIAB-R 7517 constitute the basis for the application of further anonymisation procedures in the form of the aggregation of information. Müller et al. (1991, pp. 444 - 445) 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 requirement was met by checking the number of inhabitants in the 401 autonomous municipal districts (kreisfreie Städte) and rural districts (Landkreise) as of 31 December 2017[8]. The regionalisation contained in the SIAB-R 7517 combines the information on districts in the original data in such a way that each of the district regions created has no fewer than 100,000 inhabitants as of the reference date. The aggregation results in a total of 328 district regions which are shown in the variable "Place of work – region" (ao_region) (territorial allocation 31 December 2017). 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*
    2.1. *do not comprise at least 50,000 individuals in the population,*
    2.2. *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 data protection, 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-R 7517. Frequency counts of individuals as of 30 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 30 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 7517.

---

[8] Data basis: Population projection of the Federal Statistical Office (Statistisches Bundesamt) as of 31.12.2017 for the autonomous municipal authorities (kreisfreie Städte) and the administrative districts (Kreise) (NUTS 3-level) of the Federal Republic of Germany.

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", "vocational training", "occupation/activity performed", "classification of economic activity", "occupational status" and "reason of cancellation/notification/termination". For these variables the years 1975 up to and including 1986 were not used to assess the aggregation level required for the factual anonymity of the variables. Due to the structural change on the labour 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-R 7517 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 LHG, (X)ASU and (X)MTH sources resulted in some individuals being dropped from the SIAB-R 7517. As a result of the data preparation procedures mentioned above a total of 47,536 individual accounts were deleted compared with the SIAB 7517. The SIAB-R 7517 therefore still contains 97.5 percent of the persons included in the source file SIAB 7517 (see Table A 2 for more details).

Some individual- 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, job-search and participation in employment and training measures, 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 7517 and compared to related FDZ data products can be found in the tables in the Appendix to the data report (see Section A3).

3.5   Missing values

In the SIAB-R 7517, missing values and those no longer valid due to the anonymisation are coded as follows:

Term	Value	Description
No (valid) details available	.z	Values of a variable that are not systematically missing, i.e., the variable is available in principle for the data source, but no details are available for the value considered or cannot be interpreted reasonably. Additionally: anonymised values for data protection reasons. It is not possible to distinguish between anonymous data and invalid data in the source data.

Term	Value	Description
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.
Anonymised	.a	Indicates observations in which the variable "deutsch" has been anonymised in order to prevent the de-anonymisation of persons with foreign citizenship in certain district regions (see description in Section 5.4.3).

4 Data quality and problems

4.1 Entire IEB

4.1.1 Gaps in employment histories

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

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

- gaps with no information available at all, and

- gaps for which information may be available from the "reason of cancellation/notification/termination" variable of the observation immediately preceding the gap (if a corresponding observation exists).

These gaps were identified using the variables "Reason of cancellation/notification/termination" and "Employment status" in the various sources of the SIAB 7517. The list in Table 6 makes no claims to be exhaustive. The anonymisation of the variable "grund_gr" in the SIAB-R 7517 further reduces the identifiability of biographical gaps.

Table 6: Biographical gaps and possible ways of identifying them

Biographical gap	Information on gap, potentially identifiable using the details in the "grund" variable in the preceding observation of the source
Civil servants, professional soldiers, judges, employees of bodies or foundations under public law	XASU
Self-employed persons without support	LeH, ASU
Students, persons in school-based further education	LeH, LHG, ASU, XASU

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

4.1.2   Introduction of SGB II and subsequent institutional changes

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

- Usually, the Arbeitsgemeinschaften (ARGE) between BA and the district took over the tasks according to the SGB II. These were replaced in 2011 by the joint facilities (gE) - also known as "Job Centers".

- In addition, it was possible until the end of 2011 for the BA and the district to perform the tasks assigned to them in their own (separate) responsibility (gT and gAw).

- Since 2005, it has also been possible for authorised municipalities (zkT; also: opting munici-palities) to take over the tasks. Initially, 69 administrative districts or independent towns took sole responsibility for the basic provision for job seekers until 31.12.2010. With a constitutional amendment, the option was extended to 110 providers from 2012 onwards.

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

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

4.2 Employee History (BeH)

- Information on vocational training, the occupation/activity performed and the occupational status is transmitted by means of notifications made by the employer in accordance with the Data Collection and Transmission Regulation (DEÜV) (see Section 2.1) using a so-called occupation code. The new occupation code 2010 was adopted for notifications with an end date later than 30 November 2011 (for further details, see Bertat et al., 2013). The decision to switch to the new occupation code was made by the central organisations of the social security agencies as a number of facts could no longer be recorded in a way that was up-to-date and realistic using the occupation code 2003. As the notifications made by employers in accordance with DEÜV only enter the Employee History (BeH), the change of the occupation code only affects observations from this source. The measurement of the following characteristics previously reported using the occupation code 2003 is affected by this change: working hours, occupation, occupational status and school and vocational qualification 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 switch to a new occupational classification. Instead of the previous Classification of Occupations 1988 (Klassifikation

der Berufe 1988 (KldB 1988)), the more highly differentiated KldB 2010 is reported with the new occupation code.[9]

- In contrast to SIAB-R 7514, SIAB-R 7517 contains the variable "Occupational status and working hours" (stib) again. However, the variable is only filled for reports that were submitted before the introduction of the new occupation code, and the FDZ does not extrapolate or impute for later reports. The categories of the variable "stib" that can be consistently observed over the whole observation period (occupational status as a trainee, distinction between part-time and full-time) are to be reproduced in the variables "erwstat_gr" and "teilzeit".

- The introduction of the new occupation code in 2011 led to a number of problems. For example, during the transition period granted to employers in the social security notification procedure,[10] there was a temporary increase in the number of missing details. Analyses of the BA statistics (Bertat et al., 2013, p. 10) show that in 20 to 30 percent of cases no information was available in the new or converted variables "occupation/activity performed", "working time" and "vocational education and training" after the switch. This situation began to improve significantly in the first half of 2013. In order to improve the quality of the "working time" variable in the transition period, Ludsteck/Thomsen (2016) developed an imputation procedure to replace the missing values by imputed values. The imputed data are included in the SIAB-R 7517. No imputation is performed regarding the gaps in the other variables.

- Due to the introduction of the employment notification procedure in the federal states of eastern Germany, the notifications for eastern Germany can only be assumed to be sufficiently complete from 1993 onwards. For the same reason, a large number of spells for 1991 have missing values for several variables (such as "vocational training", "employment status", and "daily wage").

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

---

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

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

- In the version of the IEB on which the SIAB-R 7517 is based, the year 2014 is the last year with a degree of completeness of BeH observations of 100%.[11] For the years 2015 and 2016, the 18-month files were used, and the observations for 2017 originate from a 6-month file. It can therefore be assumed that employment notifications for 2015 and 2016 are slightly underreported in the SIAB, and that those for 2017 are underreported to a slightly higher degree. However, this should not reduce the ability to analyse the data at individual level. The missing notifications occur more frequently in a few establishments, however. This means that in individual cases the establishment data, e.g. establishment size, are grossly incorrect and will change considerably in subsequent versions.

- In 1984, a change was made in the employment notification procedure. From that time onwards, one-off payments of gross earned income were reported as part of the annual earnings subject to social security contributions, which leads to an increase in the average daily wage. In particular, the proportion of wages and salaries above the upper earnings limit increased considerably from that year onwards (see Bender et al. 1996).

- For the years 1992 until 2000, noticeable decreases and increases in the number of notifications were observed. Decreases can be observed especially for the following 10 districts: Braunschweig (03101), Wolfsburg (03103), Emden (03402), Kassel (06633), Essen (05113), Neuss (05162), Erftkreis (05362), Hersfeld-Rotenburg (06632), Miltenberg (09676) and Kempten (Allgäu) (09763). This is due to notification problems of one or more establishments in these regions.

- Considerable decreases were also observed for the districts Salzgitter (03102) and Hoyerswerda (14264).

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

- In the years 1996 to 1998, the value 102 (physicians up to pharmacists) within the variable "Occupation - current/most recent (KldB 1988)" (beruf_gr) are very rare compared to the neighbouring years. The reasons for this are not known.

- In the years 1975 and 1977, there were so far considerably too many employees with a place of work municipality of Lahn (district of Emsland, Lower Saxony). Instead of the expected double-digit figure, there were up to 90,000 employee registrations per year with this place of work in the population of the BeH. The reason for this is a historical misclassification of employment reports from the city of Lahn. This was a merger of the Hessian municipalities Wetzlar, Gießen, Heuchelheim, Wettenberg and Lahnau (districts of Gießen and Lahn-Dill-Kreis) which was dissolved after a short time. On the basis of these findings and assuming that there were no real company relocations between these regions, the following adjustment rule was implemented at the level of the company number: As soon as the company location changes from the mu-

---

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

nicipality of Lahn to one of the listed Hessian municipalities in the years 1975 to 1978, the former specification of the municipality of Lahn is overwritten with the later correct specification. This rule significantly reduces the overhang and the municipality of Lahn in Lower Saxony then only has less than 2,000 employees in the population of the BeH in the critical years.

4.3 Benefit Recipient History (LeH)

- For the states of eastern Germany, the LeH observations were not fully recorded until 1992.

- The benefit receipt data used to be saved on magnetic tapes. Owing to a fault in one magnetic tape, the benefit receipt data up to and including 1980 are only partially contained. Thus, in the present data product, too, it can be assumed that information on benefit receipt in that period is not available in full.

- Due to an internal change of systems, there is a break in the recording of periods of exclusion from benefits and of benefit suspension in 2004. Until 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. The SIAB-R 7517 therefore does not contain any observations from the LHG for that period.

- 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 and of 1 January 2013 are already corrected as far as possible in the IEB. Nonetheless double notifications may still occur.

- In the following job centres there are inaccuracies with regard to the allocation of benefit cases:
  – between Emden and Norden between September and December 2009
  – between Döbeln and Mittelsachsen from October to December 2012
  – between Tirschenreuth and Wunsiedel from November 2012 to March 2013

- Some individuals for whom a (X)LHG spell exists are excluded entirely or partly from benefit receipt according to SGB II, for instance because they take part in a subsidised training programme, receive an old-age pension, live in an in-patient facility or a residential institution or receive insurance payments aimed at avoiding need. This affects on average 3 to 5 percent of all cases. In XSozial this person group is sometimes underrecorded by some institutions. Exclusion from benefits cannot be identified in the SIAB-R 7517.

- Due to the reporting logic, information from the XSozial transmission standard can only be updated monthly.

- In the official performance statistics of the BA, reporting gaps are supplemented by a statistical estimation procedure at an aggregated regional level. No supplementary data sets are provided in the LHG. A comparison of the IEB with the performance statistics is therefore only possible to a limited extent.

4.5 Jobseeker Histories (ASU/XASU)

Following Antoni et al. (2019), the remarks listed below refer to the two data sources of the jobseeker histories (ASU and XASU) separately. These sources are combined in the SIAB-R 7517 for data protection reasons, which limits the relevance of some of the following remarks.

4.5.1 ASU

- The registered periods of job search activity in the ASU source are regarded as complete from the year 1997 onwards. In the SIAB-R 7517, the whole period before 1997 is not filled for observations from the ASU.

- With the introduction of SGB II on 1 January 2005, jobseekers are no longer fully covered by BA procedures. From this date, the ASU only covers persons who are supported by the BA in the sphere of the SGB III (employment promotion) or by ARGEn, gE or gT in the sphere of the SGB II (basic security).

- 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". The SIAB-R 7517 therefore does not contain any observations from the ASU from that period.

- The coArb procedure, which was used until June 2006, supported only the placement of unemployed persons and jobseekers. Some data were also collected about individuals who were only seeking advice, but these data are incomplete. The careers advice data were collected in a separate system. In VerBIS the attributes of the job-search status were extended to include

"seeking advice" and individuals "without status" The latter group includes individuals eligible for Unemployment Benefit II who are only available for job placement to a limited degree. The recording of this group in VerBIS is only regarded as largely complete since January 2008.

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

- The job search status is hardly corrected afterwards, which is why the case numbers largely coincide with the BA statistics until 2005. Since spring 2011, the jobseeker data from BA procedures and XSozial have been consolidated in the integrated unemployment statistics. This may lead to larger deviations. At the current margin, however, the stock data are identical.

4.5.2 XASU

- Observations from the XASU reported by the authorised municipalities are reasonably complete in the SIAB-R 7517 from 2007 onwards. The SIAB-R 7517 does not contain any observations from the XASU from before that year.

- From 2007 onwards, individual months are also repeatedly affected by delivery failures, analogous to the benefit receipt data (LHG).

- Over-reporting may also occur at certain points. A large number of technical problems can be responsible for this, including systematically missing deregistrations, incorrect reversals of cancellations and parallel job search notifications in several districts.

- A variety of variables sometimes have only a very low degree of completeness for the XASU. Variables which are affected by this include "School leaving qualification" and "Reason of notification". Although the degree of completeness of these variables improves over time, some of them are still unsatisfactory. The variable "Occupation - current/most recent (KldB 1988)" is not available in the XASU for almost the entire period available. This variable is therefore set to missing in the (X)ASU.

- For a number of institutions (districts), the proportion of registered recipients of unemployment benefit II who are also registered jobseekers is implausibly large at times or continuously in the IEB. One possible reason for this could be an incorrect determination of the status 'not unemployed but seeking work' by these institutions.

- The institution-related and period-related plausibility of the XASU data should be examined before use, taking the research question into account.

- Due to the reporting logic, information from the XSozial reporting procedure can only be updated monthly.

- Differences in consolidation rules, time references and regional assignments may result in differences to the published BA statistics.

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

Following Antoni et al. (2019), the remarks listed below refer to the two data sources of the participants-in-measures history files (MTH and XMTH) separately. These sources are combined in the

SIAB-R 7517 for data protection reasons, which limits the relevance of some of the following remarks.

4.6.1 MTH

- The MTH is incomplete for measures with a start date before 1 January 2000. The SIAB-R 7517 does not contain any observations from the MTH from that period.

- As of 1 January 2005, there is an inconsistency in the data as participants in measures were allocated to different institutions with the introduction of Social Code Book II (see Sections 2.5 and 4.1.2).

- The MTH only contains notifications that are recorded in BA procedures. The use of these procedures in cooperations of employment agencies and municipalities/separated responsibilities/municipalities exercising their duties separately increases continuously between 2005 and 2007. The notifications for these institutions are complete from March 2007 onwards. Measures that are reported by authorised municipalities via the XSozial standard are contained in the XMTH. The SIAB-R 7517 does not contain any observations from the participants-in-measures history files for the period 2005 through 2007.

- Because of the reorganisation of the institutions responsible for implementing SGB-II in 2011 to 2014, a split of the documentation of participations in measures in the MTH and in the XMTH may occur when there is a change in the reporting procedure.[12] This might result in a split or a duplication of the spell of the measure (see Section 3.1.6).

- In the case of notifications regarding the bridging allowance (Überbrückungsgeld) the maximum permissible duration of six months is sometimes exceeded. In most cases this can be explained by a default setting in the input mask of the data recording system.

- The MTH is supplemented by applicant characteristics (e.g., vocational training) from other BA procedures. For these variables the administrative procedure was switched from coArb to VerBIS in 2006. The same quality limitations as for the ASU apply here.

4.6.2 XMTH

- In the years 2005 to 2007 the notifications of participation in measures are incomplete. The degree of under-reporting is unclear due to a lack of or inaccurate comparative figures. The SIAB-R 7517 does not contain any observations from the XMTH from that period.

- Between 2005 and the beginning of 2017, about 13% of all institutions responsible for implementing SGB II report almost no municipal integration benefits (formerly accompanying benefits), which are included in the summarised category "other support". Many other job centers report only temporarily and/or only selected types. This leads to an under-recording of the "other support" category.

- Over-registration of participations: Total stocks are considered to be stable from the end of 2008. However, the majority of the providers still register notification profiles which are temporarily conspicuous and which raise doubts about the quality. Only the introduction of the

---

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

XSozial promotion-ID in 2009 and the meanwhile several years of experience of all participants stabilise the reporting process and the subsequent data processing. Also the new zkT, introduced in2012, report inconspicuously for the most part, so that the scope and duration of presumed over-recording decrease noticeably.

- The reorganisation of SGB II providers in 2011-2014 (see Section 4.1.2) lead to a split in the documentation of participation in measures in MTH and XMTH if the reporting procedure was changed. This could result in a split of the measure spell, but also in duplications (see Section 3.1.6).

- The figures or persons received in the XMTH from monthly key date counts differ in several respects from the published BA statistics.

5 Description of variables

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

5.1 Identifiers

5.1.1 Individual ID (persnr)

Variable label	Individual ID
Variable name	persnr
Category	identifiers
Origin	BeH, LeH, LHG, ASU, XASU, MTH, XMTH
Data type	numerical
Detailed description	The individual ID indicates which observations belong to the same person. Artificial means that it is not possible to infer any of the person's characteristics or any original identifiers from this individual ID. As there is no uniform individual identifier in the different data sources, the allocation of the information from different data sources (e.g., employment and benefits) to individuals is not always unambiguous. In such cases, implausible employment histories may arise. The formation of the individual identifier which spans all data sources is based on a heuristic developed by the BA.

5.1.2 Establishment counter (bnn)

Variable label	Counter of different establishment IDs
Variable name	bnn
Category	Identifiers
Origin	BeH
Data type	numerical

Detailed description	Numbers the establishments in a person's working life in ascending order.
	Example: The first establishment in which a person was employed receives the value 1. If the person moves to a different establishment, this establishment receives the value 2, etc. If the person returns to an establishment in which they were 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. Establishment numbers that are missing in the original data are set to missing (.z) in the SIAB-R 7517.
	Missing establishment numbers in the original data were also set to missing (.z) in the SUF.

5.2 Generated technical variables

5.2.1 Spell counter per person (spell)

Variable label	Spell counter per person
Variable name	spell
Category	generated technical variables
Origin	BeH, LeH, LHG, ASU, XASU, MTH, XMTH
Data type	numerical
Detailed description	The observation counter per person counts a person's observations, beginning with 1. The variable is generated during the episode splitting procedure and refers to the split observations. Using the "observation counter per person" variable, it is easy to restore the original sorting order. The observations are sorted first by the start date of the split episode and then by the data source. Within employment notifications, persons subject to social insurance contributions are sorted before marginal employment notifications and higher daily wages before lower ones. One exception is one-time payments, which are sorted backwards.

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, MTH, XMTH
Data type	numerical
Detailed description	The variable indicates the data source. The original sources ASU and XASU as well as MTH and XMTH, respectively, were aggregated to one source. Hence, for data confidentiality reasons, 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 7517. For further information on data preparation see Section 3.4. For further information on aggregation see Table A 3.

5.3 Period of validity

5.3.1 Original start date (begorig)

Variable label	Original start date
Variable name	begorig

Category	period of validity
Origin	BeH, LeH, LHG, ASU, XASU, MTH, XMTH
Data type	date
Detailed description	The original start date of the observation corresponds to the original start date of the notification. This can differ from the start date of the episodes (begepi) (see also the comments on episode splitting in Section 3.2). 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 3.1.7. **1) BeH** 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. **2) LHG, ASU, XASU** Certain changes lead to the creation of a new observation (see Sections 3.1.3 and 3.1.4). The variable "begorig" indicates the start date of the new period. **3) MTH, XMTH** In MTH and XMTH, new data records are created when a change of provider takes place during participation in a labour market policy measure.

5.3.2   Original end date (endorig)

Variable label	Original end date
Variable name	endorig
Category	period of validity
Origin	BeH, LeH, LHG, ASU, XASU, MTH, XMTH
Data type	date
Detailed description	The original end date of the observation corresponds to the original end date of the notification. This can differ from the end date of the relevant line of data, the so-called end date of the split episode (see also the comments on episode splitting in Section 3.2). Since the notification logic might under certain circumstances per-mit 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 3.1.7. **1) BeH** 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. **2) LHG, ASU, XASU** Certain changes lead to the creation of a new observation (see Sections 3.1.3 and 3.1.4). The variable "endorig" indicates the end date of the new period. **3) MTH, XMTH**

	In MTH and XMTH, new data records are created when a change of provider takes place during participation in a labour market policy measure.

5.3.3   Episode start date (begepi)

Variable label	Episode start date
Variable name	begepi
Category	generated period of validity
Origin	BeH, LeH, LHG, ASU, XASU, MTH, XMTH
Data type	date
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   Episode end date (endepi)

Variable label	Episode end date
Variable name	endepi
Category	generated period of validity
Origin	BeH, LeH, LHG, ASU, XASU, MTH, XMTH
Data type	date
Detailed description	The end date of the split episode is always equal to or smaller than the end date of the original observation (see also the comments on episode splitting in Section 3.2.)

5.4   Personal information

5.4.1   Gender (frau)

Variable label	Gender
Variable name	frau
Category	personal variable
Origin	BeH, LeH, LHG, ASU, XASU, MTH, XMTH
Data type	numerical
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, XMTH
Data type	numerical
Detailed description	The year of birth is constant within one individual account.

Notes on quality	In the original data, it may happen that the date of birth changes between the data sources. This is corrected during the data preparation process. The information from the social security number is given highest priority here. See Section 3.4.3 for further information on the anonymisation of this variable.

5.4.3   Nationality (deutsch)

Variable label	Nationality, grouped
Variable name	deutsch
Category	personal variables
Origin	BeH, LeH, LHG, ASU, XASU, MTH, XMTH
Data type	numerical
Detailed description	The variable reports whether the person is a German citizen (1) or whether they are not (0). The value for a person may change over time.
Notes on quality	The variable is not filled well in the LeH before 1983.
Anonymisation	In some district regions, persons with foreign citizenship are so sparsely represented in the data that the criteria described in Section 3.4.4 are not sufficiently met. This means that in the district region less than 3 persons have foreign citizenship on the reference date of one year. In these cases, the procedure is as follows: All persons who work in the respective district region this year are selected and their nationality is set to Anonymised (.a) for the entire biography. The majority of the persons anonymised in that way are thus German (>99%), but not all. It is up to the data users whether they wish to exclude these persons from the analyses or treat them as persons with German citizenship.

5.4.4   Vocational training (ausbildung_gr)

Variable label	Vocational training, grouped
Variable name	ausbildung_gr
Category	personal variables
Origin	BeH, ASU, XASU, MTH, XMTH
Data type	numerical
Detailed description	It must be taken into account that this variable has a different meaning depending on the data source: **1) BeH** For observations obtained from the BeH, the variable contains the vocational education reported by the employers as part of the employment notification procedure. **2) ASU, XASU, MTH, XMTH** For these observations the vocational education completed most recently is reported.
Notes on quality	**1) BeH** "Changes in the vocational training status frequently occur at the same time as a change of establishment. This is because the notification data are compiled anew in the new firm. If, for example, an employee obtained a higher qualification via a part-time further training course while still working then this change of status is probably not recorded until he/she joins a new 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 carried forward" (own translation of Meinken / Koch 2004, p. 63).
	The share of missing values increases almost continuously over time. Due to the introduction of the new occupation code in 2011, the share even temporarily strongly increased to around 51%. Since 2014, however, the proportion of missing values in the BeH has levelled off at around 40%.
	Missing values occur particularly frequently in the following groups: marginal part-time employees, part-time workers, foreign employees and workers from Eastern German. The reason for this is that the variable is not of particular importance as regards social security contributions (see Meinken/Koch, 2004, p. 63).
	For the variable "Vocational training, imputed" (ausbildung_imp, see Section 5.4.9), a method was applied to correct missing values or inconsistent changes of the training variable. However, this variable is only filled in the source BeH and has different categories than the variable "ausbildung_gr".
	2) ASU, MTH
	As a result of the switchover from coArb to VerBIS it is not possible to distinguish correctly between "no completed vocational training" and "no information available" in the ASU and MTH data sources between 2006 and 2008. A missing value in this period therefore does not necessarily mean that the person has no vocational training or that there is no information available on vocational education and training, but may also mean that it was not possible to apply the relevant data generation procedure. In the source MTH, older categories partially still occur even after 1 July 2006.
	2) XASU, XMTH
	The degree of completeness in the XASU and the XMTH is generally low.

5.4.5   Vocational training, imputed (ausbildung_imp)

Variable label	Vocational training, imputed
Variable name	ausbildung_imp
Category	personal variables
Origin	BeH
Data type	numerical
Detailed description	The variable "Vocational training, imputed", is a supplement to the variable "Vocational training" (ausbildung_gr) and contains additional and harmonised information on the vocational training of employees for BeH spells. The variable thus offers a solution to the problems concerning the variable "Vocational training" (ausbildung_gr) described in Section 5.4.5. The imputation procedure is described in Thomsen et al (2018), which is based on the work of Fitzenberger et al. (2006).
	As the variable only uses the training information from BeH notifications, and because the educational categories of the old and the new occupation codes had to be harmonised for the variable "Vocational training" (ausbildung_gr), the variable "ausbildung_imp" has other categories than the variable "ausbildung_gr".

5.4.6   School leaving qualification (schule)

Variable label	School leaving qualification
Variable name	schule
Category	personal variables

Origin	BeH, ASU, XASU, MTH, XMTH
Data type	numerical
Detailed description	This variable contains the school leaving qualification. Different values are possible depending on the source. **1) BeH** With the switch to the new occupation code (see Section 4.2) the possible values of the variable change. The values from the old occupation code are: 5 Grade-/lower school certificate, intermediate school or equivalent qualification 8 Completion of education at a specialised upper secondary school/completion of higher education at a specialised college or upper secondary school leaving certificate, A-level equivalent, qualification for university; 13 years of schooling 9 Upper secondary school leaving certificate, A-level equivalent, qualification for university; 13 years of schooling With the new occupation code the values are: 1 No school leaving certificate 4 Lower secondary school certificate/ grade school certificate 6 Intermediate school leaving certificate 8 Completion of education at a specialised upper secondary school/completion of higher education at a specialised college or upper secondary school leaving certificate, A-level equivalent, qualification for university; 13 years of schooling **2) ASU, XASU, MTH, XMTH** The following values are possible for these data sources: 1 No school leaving certificate 4 Lower secondary school certificate/ grade school certificate 6 Intermediate school leaving certificate 7 Completion of education at a specialised upper secondary school/completion of higher education at a specialised college 9 Upper secondary school leaving certificate, A-level equivalent, qualification for university; 13 years of schooling They are valid at the beginning of the period of job-search or participation in a measure. In the case of people seeking an apprenticeship position, the variable may also contain the school qualification they are working towards in the XASU data source.
Notes on quality	The degree of completeness in the BeH has been decreasing continuously over time and seems to have levelled off at under 2/3 in recent years. In the XASU and the XMTH it has been increasing continuously and has levelled off at over 2/3 since 2013 (XASU) and 2012 (XMTH), respectively. In the ASU and the MTH the degree of completeness is generally high.

5.5   Information on employment, benefit receipt and job search

5.5.1   Daily wage/daily benefit (tentgelt_gr)

Variable label	Daily wage/daily benefit, rounded
Variable name	tentgelt
Category	information on employment, benefit receipt and job search
Origin	BeH, LeH

Data type	numerical
Detailed description	**1) BeH** In BeH observations, this variable shows the employee's gross daily wage. It is calculated from the fixed-period wages reported by the employer and the duration of the (unsplit) original notification period in calendar days. The daily wage is shown in euros. Until 1998, employers in principle only reported the earnings which were subject to social security contributions. Earnings below the marginal part-time income threshold were not reported. Earnings exceeding the upper earnings limit for statutory pension insurance are only reported up to this limit. There are two upper earnings limits in the statutory pension insurance scheme. The earnings limit of the miners' pension insurance is generally higher than the earnings limit of the pension insurance for wage and salary earners. However, it is not possible to differentiate between these two insurance providers in the data. Since the inclusion of marginal part-time employees in the employment notification procedure on 1 April 1999, earnings below the marginal part-time income threshold have also been recorded; the upper earnings limit still applies as the upper ceiling. In some cases, however, the reported earnings nonetheless exceed the upper earnings limit. Generally, this can probably be attributed to the payment of annual bonuses which the employer can add to the regular earnings in the annual, employment interruption or end of employment notifications. In this case, it is irrelevant whether the upper earnings limit in the statutory pension insurance which is decisive for the notification period is exceeded as a result of this addition. However, such earnings notifications could also be due to incorrect details in the employment period. (The earnings in-formation, however, may be considered less error-prone due to its insurance relevance.)The marginal part-time income threshold and the upper earnings limit for statutory pension insurance differ from year to year as well as between eastern and western Germany (the decisive factor is the location of the establishment). An overview of these limits and thresholds can be found under http://fdz.iab.de. A daily wage reported as 0 euros can be put down to "employment interruption notifications". During these periods, the employment relation-ship continues to exist in legal terms, but without pay. This is the case for periods of illness after the end of continued payment of wages, for periods of maternity leave and for sabbaticals. From 2013 onwards, the number of notifications with a reason for deregistration of 54 (notification of a one-off wage) increases sharply (see Section 5.5.12). It is likely that special payments which were reported with the annual declarations before 2013 are now reported separately. It is therefore advisable to add these variable one-time payments to the corresponding wages for simultaneous employment episodes within the same establishment when analysing wages over time. The daily wage is rounded to integer Euro values. All values greater than 0 and smaller than 0.5 were set 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 truncated. **2) LeH** For LeH observations, the variable shows the daily benefit rate, converted into euros in each case. It must be taken into account that for observations with an original end date prior to 1 January 1998 the daily benefit rate applies to working days (i.e., including Saturdays but excluding Sundays and public holidays), while for observations with an original end date from 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.
	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.

5.5.2   Occupation - current/most recent (KldB 1988) (beruf_gr)

Variable label	Occupation – current/most recent (KldB 1988), grouped
Variable name	beruf_gr
Category	information on employment, benefit receipt and job search
Origin	BeH
Data type	numerical
Detailed description	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 30 June. For further information on grouping see Table A 6. **1) BeH** The occupational title of the job performed by the employee during the notification period is a component of the 'employment details' submitted by the employer. If more than one job title with different classification codes applies to one employee, the employer is required to select the job title that best defines the main activity performed (see BA 2005, p. V). For this, the employer encodes the employee's job in accordance with the "Classification of Occupations. Systematic and Alphabetical Directory of Job Titles" (published by the Federal Employment Agency, Nuremberg, 1988), which contains approx. 25,000 job titles. The occupational classification consists of a 3-digit code and comprises about 330 values. Employment notifications with an end date later than 30 November 2011 are reported using the new occupation code 2010 (KldB2010) (see Section 4.2). These values are transcoded to the KldB1988 via a priority switch. This results in inaccuracies. **2) ASU, XASU, MTH, XMTH** The variable is censored in the SUF (.z).
Notes on quality	**BeH** There is a considerable increase in the number of missing values in 2011 due to the change in the occupation code. A similar accumulation of missing values occurs in 1991 due to reunification.

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, LeH, ASU, XASU, MTH, XMTH
Data type	numerical

	The occupational group consists of a 3-digit code and comprises about 144 values. This classification was further grouped to 126 occupational groups, each consisting of at least 1,000 observations per year as of the reference date of 30 June. For further information on grouping see Table A 7.
Detailed description	**1) BeH** The occupational title of the job performed by the employee during the notification period is a component of the 'employment details' submitted by the employer. If more than one job title with different classification codes apply for one employee, the employer is required to select the job title that best defines the main activity performed (see Bundesagentur für Arbeit, 2005, p. V). For this the employer encodes the employee's job in accordance with the "Classification of Occupations 2010" (Klassifikation der Berufe 2010, KldB2010, Bundesagentur für Arbeit, 2011). The occupational class consists of a 5-digit code and comprises about 1300 values. The less detailed occupational group is recorded by the first three digits of the code. The skill level required for a job, which is recorded in the fifth digit of the codes in the KldB2010, is made available separately in the variable "level of requirement" (niveau). Employment notifications with an end date earlier than 30 November 2011 are reported using the old occupation code 1988 (KldB 1988) (see Section 4.2). These values are recoded to the KldB2010 by transferring the key area. As the new occupation code is considerably more detailed than the old one, this results in substantial inaccuracies. This must be taken into account when analysing the data. **2) LeH, ASU, XASU, MTH, XMTH** The variable is censored in the SUF (.z).
Notes on quality	**BeH** There is a considerable increase in the number of missing values in 2011 due to the change in the occupation code. A similar accumulation of missing values occurs in 1991 due to reunification.

5.5.4 Level of requirement - current/most recent job (KldB 2010) (niveau)

Variable label	Level of requirement - current/most recent (KldB 2010)
Variable name	niveau
Category	information on employment, benefit receipt and job search
Origin	BeH, LeH, ASU, XASU, MTH, XMTH
Data type	numerical
Detailbeschreibung	**1) BeH** The occupational title of the job performed by the employee during the notification period is a component of the 'employment details' submitted by the employer. If more than one job title with different classification codes apply for one employee, the employer is required to select the job title that best defines the main activity performed (see Bundesagentur für Arbeit, 2005, p. V). For this the employer encodes the employee's job in accordance with the "Classification of Occupations 2010" (Klassifikation der Berufe 2010, KldB2010, Bundesagentur für Arbeit, 2011). The occupational class consists of a 5-digit code and comprises about 1300 values. The less detailed occupational sub-group is recorded by the first four digits of the code. The skill level required for a job, which is recorded in the fifth digit of the codes in the KldB2010, is made available separately in the variable "level of requirement" (niveau).

	Employment notifications with an end date earlier than 30 November 2011 are reported using the old occupation code 1988 (KldB 1988) (see Section 4.2). These values are recoded to the KldB2010 by transferring the key area. As the new occupation code is considerably more detailed than the old one, this results in substantial inaccuracies. This must be taken into account when analysing the data. **2) LeH, ASU, XASU, MTH, XMTH** The variable contains the occupation of the last job. See 1) with regard to the occupation code.
Notes on quality	**1) BeH** There is a considerable increase in the number of missing values in 2011 due to the change in the occupation code. A similar accumulation of missing values occurs in 1991 due to the reunification. **2) LeH** This variable was previously not filled in the LeH, since the occupational data in that source was of poor quality. By now, the occupation in the LeH is adopted from the job seeker history, leading to an increase in quality. **3) XASU, XMTH** The variable is not filled for almost the entire period available.

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
Detailed description	The variable "Part-time" (teilzeit) distinguishes between full-time and part-time employees. The decisive factor is the ratio between the contracted hours and the usual working hours in the establishment.
Notes on quality	There is a considerable increase in the number of missing values in 2011 due to the change in the reporting procedure. In order to reduce this problem, the working hours were imputed at the IAB for the period in question. Further information about the procedure can be found in Ludsteck/Thomsen (2016). A similar accumulation of missing values occurs in 1991 due to the reunification. No imputation is performed here.

5.5.6 Occupational status and working hours (stib)

Variable label	Occupational status and working hours
Variable name	stib
Category	information on employment, benefit receipt and job search
Origin	BeH
Data type	numerical
Detailed description	The employee's occupational status during the notification period is reported by the employer as part of the "employment details". The variable "occupational status and working hours" distinguishes between full-time and part-time employees. The decisive factor for this differentiation is the ratio between the contracted hours and the usual working hours in the establishment. For part-time employees

the variable only indicates whether their working hours exceed a certain threshold. Until 1978, this threshold was 20 hours of work per week, between 1979 and 1987 it was 15 hours per week and since 1988 it is 18 hours per week.

The variable only provides information regarding the occupational status for full-time employees, distinguishing among other things between blue-collar and white-collar employees in full-time employment and apprentices. The distinction between (full-time) blue-collar and white-collar employees is solely based on the type of compulsory pension insurance (Federal Social Insurance Office for Salaried Employees – BfA – for white collar workers, and State Social Insurance Office – LVA – for blue-collar workers). Master craftsmen and foremen are only included in a separate category if they are compulsorily insured in the workers' pension insurance. The assignment of a master craftsman or foreman to the blue-collar or white-collar employees can only be made via the respective pension provider. The "employees in vocational training" category covers not only apprentices, volunteers and interns but also semi-skilled trainees, students at colleges for health occupations and participants in subsidised further vocational training, retraining and induction training.

If more than one code is eligible for an employee, the employer is required to classify the job according to the activity which is predominantly performed. If this cannot be determined clearly, the code of the higher occupational status is to be entered (see BA 2005, p. VI).

Owing to the introduction of the new occupation code (see Section 4.2), however, this distinction is no longer possible. The variable "stib" is therefore only filled for notifications which date back to before the introduction of the new occupation code.

| Notes on quality | There is a considerable number of missing values in 1991 due to the German reunification. |

5.5.7   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, MTH, XMTH
Data type	numerical
Detailed description	This variable takes on different values with different meanings for each data source. The variable is grouped in the SUF. For further information on the aggregation see Table A 4. **1) BeH** For BeH observations, the variable "employment status" corresponds to the person group recorded in the new notification procedure (DEÜV) from 1 January 1999 onwards. It indicates contribution- or benefit-related particularities of the employment relationship. If multiple codes apply to an employment notification, the smallest must be indicated by the reporting employer. The majority of these cases are employment relationships subject to social security contributions without any distinctive characteristics, which are recorded under code number 1. Accordingly, it is possible that these employment relationships are slightly overestimated. The notification procedure stipulates that changes in the employment status - e.g., when an apprentice is taken on by his/her training company after completing his/her vocational training - must be indicated by a new notification.

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

Since 1 April 1999, employees in marginal part-time employment have also been recorded in the DEÜV notification procedure. This person group can be distinguished via the value 3. For employees in marginal part-time employment, no data prior to the introduction of the notification obligation in 1999 could be collected.

2) LeH

For LeH observations, the variable "employment status" contains the grouped benefit type. Thus, it is possible to distinguish whether a person receives unemployment benefit, unemployment assistance or maintenance allowance or whether contributions to private long-term care insurance are paid by the BA.

3) LHG

For LHG spells, the variable shows whether the person is underage and able to work or adult and able to work.

Since the reporting logic would make it possible to re-identify the exact date of birth in many cases, the original dates were changed by means of the anonymization procedure described in Section 3.1.7.

4) ASU, XASU

For ASU/XASU observations, the variable "employment status" reports the job search status.

A distinction is made between those who are "unemployed, seeking work", "unemployed, unable to work", "not unemployed, seeking work" and "not seeking work; seeking advice". Applicants who only want advice from the BA are considered as "seeking advice". In addition, there are rehabilitants and, before 2008, persons aged 58 or older who are not fully available for placement. As of 1 August 2016, this may also include persons who no longer require support, but who are nevertheless still under the job center's care.

"Not seeking work" mainly subsumes persons of whom activation or placement cannot be reasonably expected according to § 10 SGB II. Similarly, persons with an incapacity to work of more than 42 days who continue to receive ALG II are listed in the system under this status.

Since the reporting logic would make it possible to re-identify the exact date of birth in many cases, the original dates were changed by the anonymization procedure described in Section 3.1.7.

5) MTH, XMTH

For observations in the MTH and the XMTH, the variable "employment status" indicates the measure-type category. This is the highest level in the hierarchy of the measure-type classifications of the BA.

Notes on quality	**1) LHG** In some cases, there may be conflicting information on a person's ability to work if he or she is a member of different benefit units (BGs) at the same time. A possible reason for this is the determination of the ability to work by the job center specialist supervising the BG or by the respective responsible medical services. They can

	come to different results or document these results at different times. A further reason may be different birthday records for simultaneous BG customers, which may affect the status of their ability to work.
	2) ASU, XASU
	The categories "seeking advice" and "not seeking work" have only existed since the introduction of Verbis (see Section 4.5). Due to late notifications, however, entries before 2006 can also be found. The category "not seeking work" is considered to be under-recorded before 2008. In the XASU, the quality of status determination is partly limited for some institutions, especially in the first years.

5.5.8   Transition zone (gleitz)

Variable label	Transition zone
Variable name	gleitz
Category	information on employment, benefit receipt and job search
Origin	BeH
Data type	numerical
Detailed description	This variable is only available from 2003 onwards and only for BeH observations. It indicates whether the employment notification relates to employment in the low-wage sector, within the so-called transition zone. Jobs in the transition zone, so-called midi jobs, have a gross monthly wage of € 400.01 to € 800.00 (or € 450.01 to € 850.00 from 2003 onwards) for which the employee only has to pay a reduced overall social security contribution. As employees with earnings in the transition zone can voluntarily pay the "regular" social security contribution, not all employees with corresponding earnings are automatically classified as being in the transition zone. The corresponding legislation has been in force since 1 April 2003.

5.5.9   Temporary agency work (leih)

Variable label	Temporary agency work
Variable name	leih
Category	information on employment, benefit receipt and job search
Origin	BeH
Data type	numerical
Detailed description	The variable reports whether the person's employment is a temporary job via an employment agency. The variable is derived from the occupation code 2010 and is only available for notifications with an end date later than 30 November 2011.
Notes on quality	There is a considerable increase in the number of missing values in 2011 due to the change in the reporting procedure. By 2012, the share of missing values is down to about 3%.

5.5.10   Fixed-term contract (befrist)

Variable label	Fixed-term contract
Variable name	befrist
Category	information on employment, benefit receipt and job search
Origin	BeH
Data type	numerical
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 November 2011.

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

5.5.11 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
Detailed description	The variable is grouped in the SUF. For further information on the aggregation see Table A 5.

1) BeH
In BeH observations, the variable indicates the reason why the employer submitted the employment notification in question to the social security agencies. However, not all of the possible reasons for submitting a notification that may occur in the context of the notification procedure are available in the IEB. For instance, the IEB only includes notifications that have information on earnings (i.e., annual, employment interruption and end of employment notifications), while initial registrations are not contained as they contain no information on earnings. However, this does not involve a loss of information, as the details from a registration are transmitted again with the following annual, employment interruption or end of employment notification.
The reasons for submitting employment notifications are encoded according to the regulations of the notification procedure, which has been in effect since 1 January 1999 (in accordance with DEÜV).

2) LeH
In the LeH, this variable specifies the reason for the end of the receipt of unemployment benefits, unemployment assistance, or maintenance allowance. There is no information in the LeH about the reasons for the start of the benefit receipt, since the LeH is filled with the notifications from the employment agencies to the health insurance about completed benefit receipt durations.

3) LHG
The LHG observations contain the reason for discontinuation of Unemployment Benefit II and indicate the reason why current benefits have been discontinued. The reason for discontinuation of Unemployment Benefit II refers to the individual, not to the benefit unit. If the Unemployment Benefit II receipt of a different member of the benefit unit is discontinued, new observations for all members of the benefit unit are started on this date, but the reason for discontinuation of Unemployment Benefit II is only available for the individual whose benefit is discontinued. This variable is valid exactly at the end of the original observation.

4) ASU, XASU
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 "Technical split of the spell" is entered as the reason for deregistration. |

Notes on quality	**LHG, XASU**
	The proportion of valid values (degree of completeness) for the reason for notification in the LHG data sources is very small (< 20%) across all years. The degree of completeness is also low in the XASU, but improves over time.

5.5.12 Days in employment before age of 17 (tage_jung)

Variable label	Days in employment before age of 17
Variable name	tage_jung
Category	information on employment, benefit receipt and job search
Origin	BeH
Data type	numerical
Detailed description	This variable reports how many days an individual was employed before the year of their 17th birthday. Only the main employment is considered. (Stata syntax: quelle_gr==1 & level2==0).

5.5.13 Days in employment after age of 62 (tage_alt)

Variable label	Days in employment after age of 62
Variable name	tage_alt
Category	information on employment, benefit receipt and job search
Origin	BeH
Data type	numerical
Detailed description	This variable reports how many days an individual is employed after the year of their 62nd birthday. Only the main employment is considered. (Stata Syntax: quelle_gr==1 & level2==0).

5.5.14 Duration of unemployment (alo_dau)

Variable label	Duration of unemployment
Variable name	alo_dau
Category	information on employment, benefit receipt and job search
Origin	LeH, LHG, ASU, XASU, MTH, XMTH
Data type	numerical
Detailed description	The variable reports the duration (in days) of an uninterrupted sequence of periods of unemployment and is valid at the beginning of the observation.
	The following gaps do not result in an interruption of the period of unemployment:
	• any gap lasting seven days or less
	• periods of illness lasting up to 42 days
	However, when calculating the duration these gaps are not considered.
	Prior to 1997, the value "0" does not mean that the individual was not unemployed, as the ASU/XASU sources are not available here.

5.6 Location data

5.6.1 Place of work - region (ao_region)

Variable label	Place of work - region, grouped
Variable name	ao_region

Category	location data
Origin	BeH
Data type	numerical
Hierarchy	federal state | district region
Detailed description	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 at the third position. Due to the aggregation of the districts, a clear allocation to the regional authority 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.

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 2017, i.e. in all calendar years, a place of work is assigned to a district in accordance with the boundaries that the district had on 31 December 2017. As the district boundaries have changed over time, cases would occur in which the district code of the location of the establishment would change without the establishment concerned having relocated, if the territorial allocations were not updated.

For reasons of data privacy, 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).

Further information about grouped districts can be found in Table A 9. |
| Notes on quality | There are inaccuracies in the information provided for some employees with regard to where they live. The reporting requirement does not clarify which residence - main or secondary residence with predominant residence - is to be reported by the employer. In the employment history, this can lead to the impression of "long-distance commuters" between the registered main residence and the place of work, even though the person is in fact employed at the secondary residence, i.e. does not actually commute.

In the year 2015, the data show a reduction in the change of residence of approx. 10-15% compared to the usual level of the surrounding years, with regional differences occurring. It has not yet been possible to determine the exact reason for this deviation. |

5.6.2   Commuter status (pendler)

Variable label	Commuter status
Variable name	pendler
Category	location data
Origin	BeH, LeH, LHG, ASU, XASU, MTH, XMTH
Data type	numerical
Detailed description	The variable 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.

The variable contains the following categories: |

	0: the municipality of the place of residence corresponds to the municipality of the (last available) place of work (no commuter)
	1: the municipality of the place of residence does not correspond to the municipality of the (last available) place of work, but both municipalities belong to the same district region (intra-regional commuter)
	2: the municipality of the place of residence and the municipality of the (last available) place of work are located in different district regions (inter-regional commuter)
	This variable is available for: BeH: 1999-2017 LeH: 1999-2017 LHG: 2007 – 2017 (X)ASU: 1997 – 2004 and 2007 – 2017 (X)MTH: 2000 – 2004 and 2007 - 2017
	Values for the variable "Commuter status" are only given for main spells (level2==0). 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.

5.7 Establishment variables

5.7.1 Economic activity w08_3 completed by extrapolation/imputation (w08_gen_gr)

Variable label	w08_3 completed by extrapolation/imputation, grouped
Variable name	w08_gen_gr
Category	establishment variables
Origin	BHP
Data type	numerical
Hierarchy	section (1-digit code) | division (2-digit code) | group (3-digit code) | class (4-digit code) | sub-class (5-digit code) of economic activity
Detailed description	The variable is grouped in the SUF. For further information on aggregation see Table A 8.
	This variable indicates the economic activity in accordance with the WZ08 classification. From 2008 up to and including 2017, the variable contains the original values from the 3-digit code that have been grouped to protect data privacy. Before 2008, the information is either written back or replaced with the help of recoding tables, so that the variable contains time-consistent information on the economic activity based on the economic activity classification WZ08. A detailed description can be found in Eberle et al. (2011). The grouping has been implemented in such a way that it is compatible, as far as possible, with classifications by economic section or, in manufacturing, with the aggregate category A*38.
	WZ08 stands for the "Classification of Economic Activities, Edition 2008" ("Klassifikation der Wirtschaftszweige Ausgabe 2008") of the Federal Statistical Office (Statistisches Bundesamt, 2008). The WZ08 is based on the Statistical Classification of Economic Activities in the European Community NACE Rev. 2.
	Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected.

Anonymisation	In some district regions, establishments with a certain category of economic activity are so sparsely represented in the data that the criteria described in Section 3.4.4 are not sufficiently met. This means that less than three establishments in that category of economic activity exist in the district region on the reference date of one year. In these cases, the procedure is as follows: All establishments that operated in the respective district region this year are selected and their economic activity is set to missing (.z) for the entire observation period of the SIAB-R 7517. These establishments can therefore no longer be distinguished from other establishments for which there is no information on their economic activity in the data.

6 References

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

Bertat, Thomas / Dundler, Agnes / Grimm, Christopher / Kiewitt, Jochen / Schomaker, Christine / Schridde, Dr. Henning / Zemann, Dr. Christian (2013): Neue Erhebungsinhalte 'Arbeitszeit', 'ausgeübte Tätigkeit' sowie 'Schul- und Berufsabschluss' in der Beschäftigungsstatistik. Methodenbericht, Bundesagentur für Arbeit – Statistik, URL: https://statistik.arbeitsagentur.de/Statischer-Content/Grundlagen/Methodenberichte/Beschaeftigungsstatistik/Generische-Publikationen/Methodenbericht-Neue-Erbebungsinhalte-Arbeitszeita-ausgeuebte-Taetigkeit-sowie-Schul-und-Berufsabschluss-in-der-Beschaeftigungsstatistik.pdf (accessed: 12 June 2019).

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

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

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

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

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

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

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

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

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

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

Müller, Walter; Blien, Uwe; Knoche, Peter; Wirth, Heike, et al. (1991): Die faktische Anonymität von Mikrodaten. In: Forum der Bundesstatistik 19; Stuttgart.

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

Schmucker, Alexandra/Eberle, Johanna/ Ganzer, Andreas/Stegmaier, Jens/ Umkehrer, Matthias (2018): Establishment History Panel1975-2016. FDZ-Datenreport, 01/2018 (en), Nuremberg.

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

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

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

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

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

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

A Appendix

A1 Overview of date adjustments and spell deletions

Table A 1: Overview of date adjustments and spell deletions within the preparations of the SIAB-R 7517

	Start and End date observa-tion	Date after censoring	BeH	LeH	LHG	(X)ASU	(X)MTH
			1975-2017	1975-2017	2007-2017	1997-2004; 2007-2017	2000-2004; 2007-2017
Deletions of spells	Start and end date between 01.01.2005 und 31.12.2006	-				X	X
	Start and end date before 01.01.1997	-				X	
	Start and end date before 01.01.2000						X
	Start and end date before 01.01.2007	-			X		
Left-censoring of spells	Start date be-tween 01.01.2005 and 31.12.2006, end date after 01.01.2007	Start date 01.01.2007			X	X	X
Right-censoring of spells	Start date be-fore 31.12.2004, end date be-tween 01.01.2005 and 31.12.2006	End date 31.12.2004				X	X
Left- and right-censoring of spells	Start date be-fore 31.12.2004, end date after 31.12.2006	Start date spell 1 31.12.2004; end date spell 2 01.01.2007				X	X

A2 Comparison of SIAB 7517 und SIAB-R 7517

Table A 2: Reduction of the number of individuals in the course of the preparation of SIAB-R 7517

		Basic file: number of individuals SIAB 7517	1,875,439	100.0%	
	1	Deletion of observations in all sources which end before the age of 17 or begin after the age of 62	1,833,313	97.8%	-2.2%
	2	Deletion of ASU spells that end before 1997 and of MTH spells that end before 2000	1,831,716	97.7%	-0.1%
	3	Deletion of LHG, XASU and XMTH spells that end before 2007	1,830,776	97.6%	-0.1%
	4	Deletion of ASU- and MTH spells that begin after 2004 and end before 2007	1,827,903	97.5%	-0.1%
		Number of individuals SIAB-R 7517	**1,827,903**	**97.5%**	**-2.5%**

A3 Recoding tables of aggregated variables in comparison with the weakly anonymous version of the SIAB 7517

Table A 3: Recoding of variable "Source of spell" (quelle_gr)

SIAB 7517	SIAB-R 7517 (SUF)	Categories
1	1	BeH Employee History
2	2	LeH Benefit Recipient History
3	3	LHG Unemployment Benefit II Recipient History
4, 5	4	(X)MTH Participants-in-Measures Histories
6, 7	5	(X)ASU Job-Search Histories

Table A 4: Recoding of variable "Employment status" (erwstat_gr)

SIAB 7517	SIAB-R 7517 (SUF)	Categories
101, 140,143	1	Employees subject to social security
102, 141, 121, 122	2	Trainees
since 1999: 109, 209	3	Marginal part-time workers
since 1999: 103, 142	4	Employees in partial retirement

SIAB 7517	SIAB-R 7517 (SUF)	Categories
since 1999: 105, 106	5	Interns and student trainees
104, 112, 118 - 120, 123, 124, 149, 201, 203, 205, 599 before 1999: 103, 105, 106, 109, 118, 142, 205, 209	6	Other employment status
1	11	Unemployment benefit (ALG)
2	12	Unemployment assistance (ALHI)
3	13	Maintenance benefit (UHG)
5	14	Contributions to private long-term care insurance
31, 41, 51	21	Unemployed (ALO), registered as a job seeker)
32	22	Unemloyed - unable to work up to 42 days (ALO-AU)
33, 43, 53	23	Not unemployed, registered as a job seeker (NALO)
34, 35, 44, 54, 55	24	Not seeking work - without status (NASU); Seeking advice (RASU)
10, 11	31	Minor and able to work
13, 14, 15	32	Adult and able to work
10001	41	Activation and vocational integration
10002	42	Start an employment
10003	43	Career choice and vocational training
10004	44	Employment-generating measures
10005	45	Free promotion
10006	46	Other promotion
10007	47	Vocational retraining and further education

Table A 5: Recoding of variable "Reason of cancellation/notification/termination" (grund_gr)

SIAB 7517	SIAB-R 7517 (SUF)	Categories
30, 49 before 1999: 31, 32, 34 - 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)

SIAB 7517	SIAB-R 7517 (SUF)	Categories
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)
50, 70	6	Annual notification (all employment relationships that exist as of 31 December of the notification year must be declared)
51	7	Employment interruption notification due to entitlement to other compensation
since 1999: 52, 53	8	Employment interruption notification due to parental leave/statutory official duties
54	9	Notification of a lump sum payment
33, 55 - 57, 59, 71, 72; before 1999: 52, 53	10	Deregistration for other reasons
2002, 2064	20	Begin of employment
2001, 2005, 2006, 2014, 2021, 2022, 2039, 2041, 2046, 2048, 2052, 2053, 2056, 2065, 2068	21	Entitlement to other benefit
2007, 2008, 2011, 2036 - 2038, 2040, 2047, 2050, 2051, 2055, 2057, 2058, 2062	22	No availability for the labour market
2015	23	Maternity benefit
2026, 2060	24	End/Drop out of measure
2003, 2009, 2017, 2020, 2024, 2028, 2031, 2042, 2049	26	Entitlement exhausted/Termination of period for which benefit is allowed
2004, 2010,2012, 2016, 2019, 2023, 2027, 2029, 2032, 2033, 2035, 2043 - 2045, 2054, 2059, 2061, 2063, 2066, 2067, 2069	27	Other reasons
5031, 5036, 5061, 5079, 5088, 5089, 6012	30	Measures of active labour market policy
5037, 5038, 5066, 6001	31	Employment - self-sought
5029, 5030, 5032 - 5035, 5060, 5062 - 5064, 5080, 5085, 5086, 5098, 6002, 6003	32	Employment - placement by third party
5082, 6005	33	Employment - no further details
5046, 5047, 5067, 6006	34	Self-employed

SIAB 7517	SIAB-R 7517 (SUF)	Categories
5053, 5068, 6007, 6015, 6025	35	Military/civilian service, military exercise, soldier, official
5041, 5052, 5076, 5084, 6008 - 6010	36	Incapacity for work (illness, maternity protection), non-employment, retirement from working life
5042, 5056, 5065, 6004	37	Recall/re-employment at the same employer
5039, 5043, 5070, 5071, 5083, 6021, 6022 - 6024, 6027	38	Vocational training, further training, other training
5040, 5049, 5069, 6016	39	School education, studies
5045, 5077, 5081, 6029	40	Change of provider
5044, 5051, 5074, 6018, 6026	41	Lack of co-operation or availability
5055, 6028	42	Discontinuation of fulfillments
5050, 5075, 6017, 6030, 6031	43	Special regulations
5027, 5048, 5054, 5073, 5078, 5087, 5090, 5091, 6011, 6013, 6014, 6019, 6020	44	Other reasons
5100, 6100	45	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
3014	56	Expiration of the authorization period
3015	57	Loss of ability to work
3016	58	End of need for financial support
3002 - 3004, 3010 - 3012	59	Other reasons

Table A 6: Recoding of variable "Occupation - current/most recent (KldB 1988)" (beruf_gr)

SIAB 7517	SIAB-R 7517 (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 up to animal keepers and related occupations
051, 052, 053, 054, 055, 056, 057, 058, 059, 060, 061, 062	2	Gardeners, garden workers up to forest workers, forest cultivators

SIAB 7517	SIAB-R 7517 (SUF)	Categories
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 up to shaped brick/concrete block makers
121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135	4	Ceramics workers up to glass processors, glass finishers
141	5	Chemical plant operatives
142, 143, 144	6	Chemical laboratory workers up to vulcanisers
151	7	Plastics processors
161, 162, 163, 164	8	Paper, cellulose makers up to other paper products makers
171, 172, 173, 174	9	Type setters, compositors up to printers (flat, gravure)
175, 176, 177	10	Special printers, screeners up to printer's assistants
181, 182, 183, 184	11	Wood preparers up to basket and wicker products makers
191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203	12	Iron, metal producers, melters up to semi-finished product fettlers and other mould casting occupations
211, 212, 213	13	Sheet metal pressers, drawers, stampers up to other metal moulders (non-cutting deformation)
221	14	Turners
222, 223, 224	15	Drillers up to borers
225, 226	16	Metal grinders up to other metal-cutting occupations
231, 232, 233, 234, 235, 242, 243, 244	17	Metal polishers up to metal bonders and other metal connectors
241	18	Welders, oxy-acetylene cutters
251, 252, 263	19	Steel smiths up to pipe, tubing fitters
261	20	Sheet metal workers
262	21	Plumbers
270, 271, 272	22	Locksmiths, not specified up to sheet metal, plastics fitters
273	23	Engine fitters
274, 275	24	Plant fitters, maintenance fitters up to steel structure fitters, metal shipbuilders

SIAB 7517	SIAB-R 7517 (SUF)	Categories
281	25	Motor vehicle repairers
282, 283, 284	26	Agricultural machinery repairers up to precision mechanics
285, 286	27	Other mechanics up to watch-, clockmakers
291, 301, 302	28	Toolmakers up to precious metal smiths
303, 304, 305, 306	29	Dental technicians up to doll makers, model makers, taxidermists
311	30	Electrical fitters, mechanics
312, 313, 315	31	Telecommunications mechanics, craftsmen up to 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 up to skin processing operatives
351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362	37	Cutters up to textile finishers
391, 392	38	Bakery goods makers up to confectioners (pastry)
401, 402, 403	39	Butchers up to fish processing operatives
411, 412	40	Cooks up to ready-to-serve meals, fruit, vegetable preservers, preparers
421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433	41	Wine coopers up to sugar, sweets, ice-cream makers
441, 442	42	Bricklayers up to concrete workers
451, 453	43	Carpenters up to scaffolders
452	44	Roofers
461, 462	45	Paviors up to road makers
463, 464, 465, 466	46	Tracklayers up to other civil engineering workers
470, 471, 472	47	Building labourer, general up to other building labourers, building assistants, n.e.c.
481, 482	48	Stucco workers, plasterers, rough casters up to insulators, proofers
483, 484, 485, 486	49	Tile setters up to screed, terrazzo layers
491, 492, 502, 503, 504	50	Room equippers up to other wood and sports equipment makers
501	51	Carpenters
511	52	Painters, lacquerers (construction)

SIAB 7517	SIAB-R 7517 (SUF)	Categories
512, 513, 514	53	Goods painters, lacquerers up to ceramics/glass painters
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 up to construction machine attendants
547, 548, 549	58	Machine attendants, machinists' helpers up to 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 up to other engineers
611, 612	63	Chemists, chemical engineers up to physicists, physics engineers, mathematicians
621	64	Mechanical engineering technicians
622, 623	65	Electrical engineering technicians up to building technicians
624, 625, 626, 627	66	Measurement technicians up to remaining manufacturing technicians
628	67	Other technicians
629	68	Foremen, master mechanics
631, 632	69	Biological specialists up to physical and mathematical specialists
633, 634	70	Chemical laboratory assistants up to 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 up to service-station attendants
687, 688	75	Commercial agents, travellers up to mobile traders
691, 692	76	Bank specialists up to building society specialists
693, 694	77	Health insurance specialists (not social security) up to life, property insurance specialists
701	78	Forwarding business dealers
702, 703, 704, 705, 706	79	Tourism specialists up to cash collectors, cashiers, ticket sellers, inspectors
711, 712, 713, 715, 716	80	Railway engine drivers up to street attendants
714	81	Motor vehicle drivers

SIAB 7517	SIAB-R 7517 (SUF)	Categories
721, 722, 723, 724, 725, 726	82	Navigating ships officers up to air transport occupations
731, 732, 733, 734	83	Post masters up to telephonists
741	84	Warehouse managers, warehousemen
742	85	Transportation equipment drivers
743, 744	86	Stowers, furniture packers up to stores/transport workers
751	87	Entrepreneurs, managing directors, divisional managers
752, 753	88	Management consultants, organisors up to chartered accountants, tax advisers
761, 762, 763	89	Members of Parliament, Ministers, elected officials up to association leaders, officials
771, 772	90	Cost accountants, valuers up to accountants
773	91	Cashiers
774	92	Data processing specialists
781	93	Office specialists
782, 783	94	Stenographers, shorthand-typists, typists up to data typists
784	95	Office auxiliary workers
791, 792	96	Factory guards, detectives up to watchmen, custodians
793, 794	97	Doormen, caretakers up to 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 up to judicial enforcers
821, 822, 823	99	Journalists up to librarians, archivists, museum specialists
831, 832, 833, 834	100	Musicians up to scenery/sign painters
835, 836, 837, 838	101	Artistic and assisting occupations (stage, video and audio) up to performers, professional sportsmen, auxiliary artistic occupations
841, 842, 843, 844	102	Physicians up to Pharmacists
851, 852	103	Non-medical practitioners up to masseurs, physiotherapists and related occupations
853	104	Nurses, midwives
854	105	Nursing assistants
855, 857	106	Dietary assistants, pharmaceutical assistants up to medical laboratory assistants

SIAB 7517	SIAB-R 7517 (SUF)	Categories
856	107	Medical receptionists
861, 863, 891, 892, 893	108	Social workers, care workers up to 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 up to technical, vocational, factory instructors
875, 876, 877	112	Music teachers, n.e.c. up to other teachers
881, 882, 883	113	Economic and social scientists, statisticians up to scientists n.e.c.
901, 902	114	Hairdressers up to other body care occupations
911, 912	115	Restaurant, inn, bar keepers, hotel proprietors, catering trade dealers up to waiters, stewards
913	116	Others attending on guests
921, 922, 923, 924	117	Housekeeping managers up to employees by household cheque procedure
931, 932	118	Laundry workers, pressers up to textile cleaners, dyers and dry cleaners
933, 934	119	Household cleaners up to glass, buildings cleaners
935, 936, 937	120	Street cleaners, refuse disposers up to machinery, container cleaners and related occupations

Table A 7: Recoding of variable "Occupational group - current/most recent (KldB 2010)" (beruf10_gr)

SIAB 7517	SIAB-R 7517 (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

SIAB 7517	SIAB-R 7517 (SUF)	Categories
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
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
262	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

SIAB 7517	SIAB-R 7517 (SUF)	Categories
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
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

SIAB 7517	SIAB-R 7517 (SUF)	Categories
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
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

SIAB 7517	SIAB-R 7517 (SUF)	Categories
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
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 "w08_3 completed by extrapolation/imputation" (w08_gen_gr)

SIAB 7517	SIAB-R 7517 (SUF)	Industry	Anonymised in regions
011 - 032, 051 - 099, 351 - 353, 360 - 390	1	Agriculture, forestry and fishing; mining and quarrying and other industry (A,B,D,E)	3103, 5120, 6413, 6434, 7132, 8221, 8421, 9474, 8562, 8563, 8573, 16053
101 - 120, 131 - 152, 161 - 182, 310 - 332	2	Manufacturing 1 (CA,CB,CC,CM)	
191, 192,201 - 206, 211, 212, 221 - 239, 241 - 259	3	Manufacturing 2 (CD,CE,CF,CG,CH)	12052
261 - 268, 271 - 279, 281 - 289, 291 - 309	4	Manufacturing 3 (CI,CJ,CK,CL)	3151, 3256, 12052, 12054, 12071
411 - 439	5	Construction (F)	
451 - 479	6	Wholesale and retail trades (G)	
491 - 532	7	Transportation and storage (H)	9562
551 - 563	8	Accommodation and food service activities (I)	3453, 9183, 9573, 9771
581 - 639, 681 - 829	9	Information and communication; Real estate activities; Business services (J,L,M,N)	
641 - 663	10	Financial and insurance activities (K)	5512, 7235, 8236, 12071, 13072, 16061
841 - 843	11	Public administration and defence (O)	5512
851 - 856	12	Education (P)	3361, 7231, 9375
861 - 889	13	Human health services; Residential care and social work activities (Q)	
900 - 990	14	Other services (R,S,T,U)	9183

Table A 9: Regionalisation of districts (NUTS 3) in SIAB-R 7517, territorial allocation 31 December 2017 (ao_region)

SIAB 7517	SIAB-R 7517 (SUF)	District region
1002	1002	Kiel, Capital of the Federal State
1003	1003	Luebeck, Hanseatic City
1051	1051	Dithmarschen

SIAB 7517	SIAB-R 7517 (SUF)	District region
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, Free and Hanseatic City
3101	3101	Braunschweig, City
3102	3102	Salzgitter, City
3103	3103	Wolfsburg, City
3151	3151	Gifhorn
3153	3153	Goslar
3155, 3255	3155	Northeim/Holzminden
3157	3157	Peine
3158, 3154	3158	Wolfenbuettel/Helmstedt
3159	3159	Goettingen
3241	3241	Region Hannover
3251, 3401	3251	Diepholz/Delmenhorst City
3252	3252	Hameln-Pyrmont
3254	3254	Hildesheim
3256	3256	Nienburg (Weser)
3257	3257	Schaumburg
3351	3351	Celle
3352, 3461	3352	Cuxhaven/Wesermarsch
3353	3353	Harburg
3355	3355	Lueneburg
3356	3356	Osterholz
3357	3357	Rotenburg (Wuemme)
3358	3358	Heidekreis
3359	3359	Stade
3360, 3354	3360	Uelzen/Luechow-Dannenberg
3361	3361	Verden
3403	3403	Oldenburg (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

SIAB 7517	SIAB-R 7517 (SUF)	District region
3458	3458	Oldenburg
3459	3459	Osnabrueck
3460	3460	Vechta
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-Kreis Neuss
5166	5166	Viersen
5170	5170	Wesel
5314	5314	Bonn, City
5315	5315	Koeln, City
5316	5316	Leverkusen, City
5334	5334	Staedteregion Aachen
5358	5358	Dueren
5362	5362	Rhein-Erft-Kreis
5366	5366	Euskirchen
5370	5370	Heinsberg
5374	5374	Oberbergischer Kreis
5378	5378	Rheinisch-Bergischer Kreis
5382	5382	Rhein-Sieg-Kreis
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

SIAB 7517	SIAB-R 7517 (SUF)	District region
5774	5774	Paderborn
5911	5911	Bochum, City
5913	5913	Dortmund, City
5914	5914	Hagen, City
5915	5915	Hamm, City
5916	5916	Herne, City
5954	5954	Ennepe-Ruhr-Kreis
5958	5958	Hochsauerland Kreis
5962	5962	Maerkischer Kreis
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, Capital of the Federal State
6431, 6437	6431	Bergstrasse/Odenwaldkreis
6432	6432	Darmstadt-Dieburg
6433	6433	Gross-Gerau
6434	6434	Hochtaunuskreis
6435	6435	Main-Kinzig-Kreis
6436	6436	Main-Taunus-Kreis
6438	6438	Offenbach
6439	6439	Rheingau-Taunus-Kreis
6440	6440	Wetteraukreis
6531	6531	Giessen
6532	6532	Lahn-Dill-Kreis
6533	6533	Limburg-Weilburg
6534	6534	Marburg-Biedenkopf
6535	6535	Vogelsbergkreis
6611	6611	Kassel, City
6631	6631	Fulda
6632	6632	Hersfeld-Rotenburg
6633	6633	Kassel
6634	6634	Schwalm-Eder-Kreis
6635	6635	Waldeck-Frankenberg
6636	6636	Werra-Meissner-Kreis
7111	7111	Koblenz, City
7131	7131	Ahrweiler
7132	7132	Altenkirchen (Westerwald)
7133, 7134	7133	Bad Kreuznach/Birkenfeld
7137	7137	Mayen-Koblenz
7138	7138	Neuwied
7140, 7135	7140	Rhein-Hunsrueck-Kreis/Cochem-Zell

SIAB 7517	SIAB-R 7517 (SUF)	District region
7141	7141	Rhein-Lahn-Kreis
7143	7143	Westerwaldkreis
7211	7211	Trier, City
7231	7231	Bernkastel-Wittlich
7232, 7233	7232	Eifelkreis 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	Donnersbergkreis/Kusel
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-Kreis/Frankenthal (Pfalz), City/Speyer, City
7339	7339	Mainz-Bingen
7340, 7317, 7320	7340	Suedwestpfalz/Pirmasens, City/Zweibruecken, City
8111	8111	Stuttgart, Capital of the Federal State
8115	8115	Boeblingen
8116	8116	Esslingen
8117	8117	Goeppingen
8118	8118	Ludwigsburg
8119	8119	Rems-Murr-Kreis
8121	8121	Heilbronn, City
8125	8125	Heilbronn
8126	8126	Hohenlohekreis
8127	8127	Schwaebisch Hall
8128	8128	Main-Tauber-Kreis
8135	8135	Heidenheim
8136	8136	Ostalbkreis
8212	8212	Karlsruhe, City
8215	8215	Karlsruhe
8216, 8211	8216	Rastatt/Baden-Baden
8221	8221	Heidelberg, City
8222	8222	Mannheim, City
8225	8225	Neckar-Odenwald-Kreis
8226	8226	Rhein-Neckar-Kreis
8231	8231	Pforzheim
8235	8235	Calw
8236	8236	Enzkreis
8237	8237	Freudenstadt
8311	8311	Freiburg im Breisgau, City
8315	8315	Breisgau-Hochschwarzwald
8316	8316	Emmendingen
8317	8317	Ortenaukreis

SIAB 7517	SIAB-R 7517 (SUF)	District region
8325	8325	Rottweil
8326	8326	Schwarzwald-Baar-Kreis
8327	8327	Tuttlingen
8335	8335	Konstanz
8336	8336	Loerrach
8337	8337	Waldshut
8415	8415	Reutlingen
8416	8416	Tuebingen
8417	8417	Zollernalbkreis
8421	8421	Ulm
8425	8425	Alb-Donau-Kreis
8426	8426	Biberach
8435	8435	Bodenseekreis
8436	8436	Ravensburg
8437	8437	Sigmaringen
9161	9161	Ingolstadt, City
9162	9162	Muenchen, Capital of the Federal State
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

SIAB 7517	SIAB-R 7517 (SUF)	District region
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
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 Windsh./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-Kreis
11000	11000	Berlin, City
12052	12052	Cottbus, City
12054	12054	Potsdam, City
12060	12060	Barnim

SIAB 7517	SIAB-R 7517 (SUF)	District region
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
12068, 12070	12068	Ostprignitz-Ruppin/Prignitz
12069, 12051	12069	Potsdam-Mittelmark/Brandenburg an der Havel, City
12071	12071	Spree-Neisse
12072	12072	Teltow-Flaeming
12073	12073	Uckermark
13003	13003	Rostock, Hanseatic 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	Erzgebirgskreis
14522	14522	Mittelsachsen
14523	14523	Vogtlandkreis
14524	14524	Zwickau
14612	14612	Dresden, City
14625	14625	Bautzen
14626	14626	Goerlitz
14627	14627	Meissen
14628	14628	Saechs. Schweiz-Osterzgebirge
14713	14713	Leipzig, City
14729	14729	Leipzig
14730	14730	Nordsachsen
15002	15002	Halle (Saale), City
15003	15003	Magdeburg, Capital of the Federal State
15082, 15001	15082	Anhalt-Bitterfeld/Dessau-Rosslau, City
15083, 15086	15083	Boerde/Jerichower Land
15084	15084	Burgenlandkreis
15085	15085	Harz
15087	15087	Mansfeld-Suedharz
15088	15088	Saalekreis
15089	15089	Salzlandkreis
15090, 15081	15090	Stendal/Altmarkkreis Salzwedel
15091	15091	Wittenberg
16051	16051	Erfurt, City
16053	16053	Jena, City

SIAB 7517	SIAB-R 7517 (SUF)	District region
16061	16061	Eichsfeld
16062, 16065	16062	Nordhausen/Kyffhaeuser Kreis
16063, 16056	16063	Wartburgkreis/Eisenach, City
16064	16064	Unstrut-Hainich-Kreis
16066, 16054	16066	Schmalkalden-Meiningen/Suhl, City
16067	16067	Gotha
16069, 16072	16069	Hildburghausen/Sonneberg
16070	16070	Ilm-Kreis
16071, 16055, 16068	16071	Weimarer Land/Weimar, City/Soemmerda
16073	16073	Saalfeld-Rudolstadt
16074, 16075	16074	Saale-Holzland-Kreis/Saale-Orla-Kreis
16076, 16052, 16077	16076	Greiz/Gera, City/Altenburger Land

Figure A 1: Regionalisation SIAB-R 7517, territorial allocation 31.12.2017 (ao_region)

state border

district border

district

aggregated district region border (SIAB-R 7517)

aggregated district region (SIAB-R 7517)
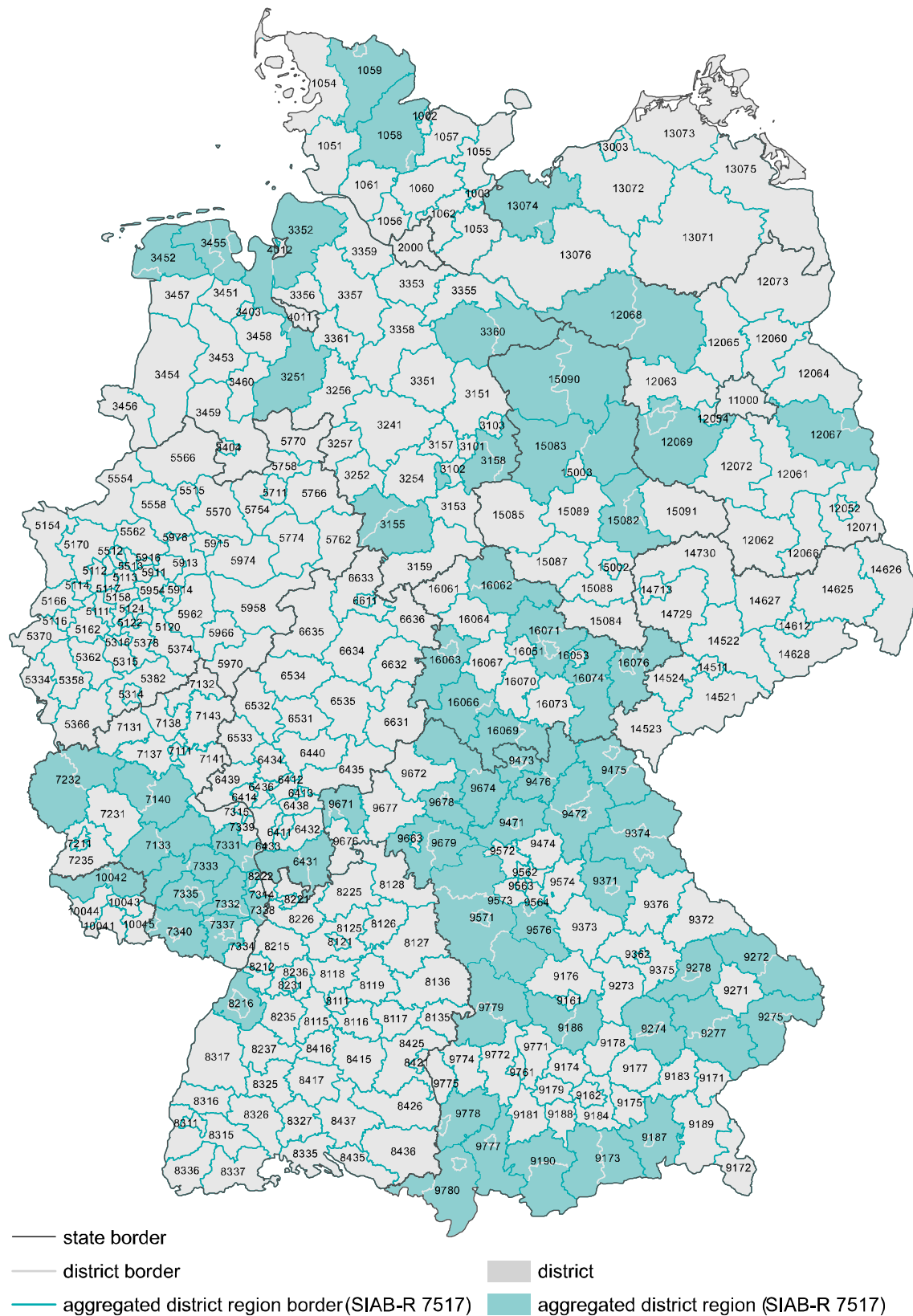
A4   Frequency tables

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

A5   List of abbreviations

ALG	Arbeitslosengeld	unemployment benefit
ARGE	Arbeitsgemeinschaft	cooperation of employment agencies and municipalities
ASU	Arbeitsuchendenhistorik	Jobseeker History
A2LL	Arbeitslosengeld II – Leistungen zum Lebensunterhalt	unemployment benefit II - benefits to secure a livelihood
BA	Bundesagentur für Arbeit	Federal Employment Agency
BeH	Beschäftigtenhistorik	Employee History
BfA	Bundesversicherungsanstalt für Angestellte	Federal Social Insurance Office for Salaried Employees
BG	Bedarfsgemeinschaft	Benefit unit
BHP	Betriebs-Historik-Panel	Establishment History Panel
coArb	Computerunterstützte Arbeitsvermittlung (operatives Verfahren zur Verwaltung der Vermittlung (Altverfahren))	computer-aided job placement (procedure for the administration of job placements – old procedure)
DEÜV	Verordnung über die Erfassung und Übermittlung von Daten für die Träger der Sozialversicherung – Datenerfassungs- und –übermittlungsverordnung	Data Collection and Transmission Regulation - regulation on the collection and transmission of data for the social security agencies
DEVO	Zweite VO über die Erfassung von Daten für die Träger der Sozialversicherung und für die BA – Datenerfassungs-Verordnung	Data Collection Regulation - second regulation on the collection of data for the social security agencies and for the Federal Employment Agency
DIM	Daten- und IT-Management	Data and IT Management
DÜVO	Zweite VO über die Datenübermittlung auf maschinell verwertbaren Datenträgern im Bereich der Sozialversicherung und der BA – Datenübermittlungs-Verordnung	Data Transmission Regulation - second regulation on the transfer of data on machine-readable data media in the field of social security and the BA
DWH	Data Warehouse	Data Warehouse
EDV	Elektronische Datenverarbeitung	Electronic data processing
FDZ	Forschungsdatenzentrum	Research Data Centre
FELEG	Gesetz zur Förderung der Einstellung der landwirtschaftlichen Erwerbstätigkeit	Act on the Support in Case of Termination of Farming Activities
gAw	Träger mit getrennter Aufgabenwahrnehmung	Municipalities exercising their duties separately
gE	Gemeinsame Einrichtung	Joint facility
gT	Getrennte Trägerschaft	Separated responsibilities
IAB	Institut für Arbeitsmarkt- und Berufsforschung	Institute for Employment Research
IEB	Integrierte Erwerbsbiographien	Integrated Employment Biographies
LeH	Leistungsempfängerhistorik	Benefit Recipient History
LHG	Leistungshistorik Grundsicherung	Unemployment Benefit II Recipient History

LVA	Landesversicherungsanstalt	Land Social Insurance Office
MTH	Maßnahmeteilnehmerhistorik	Participants-in-Measures History File
NACE	Nomenclature génerale des activités économiques dans les communautés européennes	Nomenclature génerale des activités économiques dans les communautés européennes
NUTS	Nomenclature des unités territoriales statistiques	Nomenclature des unités territoriales statistiques
SGB	Sozialgesetzbuch	German Social Code
SIAB	Stichprobe der Integrierten Arbeitsmarktbiographien	Sample of Integrated Labour Market Biographies
SUF	Scientific Use File	Scientific Use File
VerBIS	Vermittlungs- und Beratungsinformationssystems	Information System for Placement and Counselling
XASU	Arbeitsuchendenhistorik aus XSozial-BA-SGB II	Jobseeker History from XSozial-BA-SGB II
XMTH	Maßnahmeteilnahmehistorik aus XSozial-BA-SGB II	Participants-in-Measures History File from XSozial-BA-SGB II
zkT	Zugelassene kommunale Träger	Authorised municipalities

Imprint

Corresponding author
Philipp vom Berge
Phone: +49 911 179-5020
Email: Philipp.vom-Berge@iab.de