

FDZ-Datenreport

Documentation of labour market data

07/2011

EN

Sample of Integrated Labour Market Biographies

Regional File 1975-2008 (SIAB-R 7508)

Matthias Dorner,
Marion König,
Stefan Seth



Sample of Integrated Labour Market Biographies – Regional File 1975-2008 (SIAB-R 7508)

Matthias Dorner, Marion König, Stefan Seth
(Institute for Employment Research, IAB)

Data: Version 1 (SIAB-R 7508 v1)
Documentation: Version 1 (SIAB-R_7508_EN_v1_dok1)

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 has two aims: first, users can ascertain whether the data are suitable for their research; second, the reports can be used to prepare analyses.

Abstract

This FDZ data report describes the Regional File of the Sample of Integrated Labour Market Biographies 1975-2008 (SIAB-R 7508). This dataset represents the factual anonymous version of the Sample of Integrated Labour Market Biographies (SIAB) and may be delivered to scientific research institutions after concluding a use agreement with the IAB.

Zusammenfassung

Dieser FDZ-Datenreport beschreibt das Regionalfile der Stichprobe der Integrierten Arbeitsmarktbiografien 1975-2008 (SIAB-R 7508). Der Datensatz stellt die faktisch anonymisierte Version der Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) dar und wird nach Abschluss eines Nutzungsvertrages mit dem IAB an Forschungseinrichtungen übermittelt.

Keywords: German administrative micro data, labour market data, data manual, factual anonymized data, scientific use file

Disclaimer

We would like to thank our FDZ colleagues for their helpful support and comments during the data production.

Parts of this data report are taken from the FDZ Datenreport 1/2010 on the weakly anonymous version of the SIAB 7508 and from internal data documentations on the raw data written by the IAB IT-Division, ITM. However, all remaining errors are our own.

Table of contents

1	Introduction and outline	6
1.1	Introduction	6
1.2	Data use	7
1.3	Differences between the SIAB and earlier versions	7
1.4	Outline	9
1.5	List of variables	10
1.6	Volume structure	12
2	Data sources	12
2.1	Employee History (BeH)	13
2.2	Benefit Recipient History (LeH)	14
2.3	Unemployment Benefit II Recipient Histories (LHG/XLHG)	14
2.4	Jobseeker Histories (ASU/XASU)	16
3	Data preparation, sampling procedure and anonymisation	16
3.1	Corrections of the basic data for the SIAB-R 7508	16
3.1.1	Employee History (BeH)	16
3.1.2	Benefit Recipient History (LeH)	17
3.1.3	Jobseeker Histories (X)ASU	18
3.1.4	Unemployment Benefit II Recipient Histories (X)LHG	18
3.1.5	Corrections of the aggregate IEB file	18
3.2	Sample design	19
3.3	Anonymisation of the SIAB-R 7508	19
3.3.1	Merging sources	20
3.3.2	Periods of validity of the sources	20
3.3.3	Censoring of data on age	23
3.3.4	Aggregation and deletion of variables	23
3.3.5	Consequences of the anonymisation	25
3.4	Definition of missing, invalid and anonymized values	26
3.5	Episode splitting	26
4	Data quality and problems	27
4.1	Employee History (BeH)	27
4.2	Benefit Recipient History (LeH)	28
4.3	Unemployment Benefit II Recipient Histories (LHG/XLHG)	29
4.4	Jobseeker Histories (ASU/XASU)	29
5	Description of variables	30
5.1	Identifiers	30
5.1.1	Artificial individual ID (persnr)	30
5.1.2	Establishment counter (bnn)	31
5.2	Period of validity	31
5.2.1	Original start date of observation (begorig)	31
5.2.2	Original end date of observation (endorig)	32
5.2.3	Start date of split episode (begepi)	32
5.2.4	End date of split episode (endepe)	32
5.3	Personal information	33
5.3.1	Gender (frau)	33

5.3.2	Year of birth (gebjahr)	33
5.3.3	Nationality german (deutsch)	33
5.3.4	School education and vocational training (bild)	34
5.3.5	School-leaving qualification (schbild)	35
5.4	Information on employment, benefit receipt and job-search	35
5.4.1	Reason for notification/ reason for end of benefit receipt/ reason for discontinuation of unemployment benefit II/ reason for deregistration, aggregated (grund_gr)	35
5.4.2	Daily wage, daily benefit rate, aggregated (tentgelt_gr)	37
5.4.3	Occupation – activity performed (beruf_gr)	39
5.4.4	Occupational status and working hours (stib)	39
5.4.5	Employment status (erwstat_gr)	40
5.4.6	Employment before the age of 17 (tage_jung)	41
5.4.7	Employment after the age of 62 (tage_alt)	41
5.5	Establishment variables	42
5.5.1	Economic activity WS 73 – group of economic activity 73 (w73_gen_gr)	42
5.5.2	Economic activity 03 – group of economic activity 03 (w03_gr)	42
5.6	Location data	43
5.6.1	Place of work: region (ao_region)	43
5.6.2	Commuter status (pendler)	44
5.7	Generated technical variables	44
5.7.1	Source of the observation (quelle_gr)	44
5.7.2	Observation counter per person (spell)	45
5.7.3	Observation counter per episode (level2)	45
5.7.4	Observation counter per episode and source (level1)	46
6	References	47
7	Appendix	49
7.1	Differences in the anonymisation of key variables in comparison to the IABS-R04	49
7.2	Additional correction steps in the IABS-Regional file (IABS-R04) in comparison with the SIAB-Regional file (SIAB-R 7508)	49
7.3	Overview of date adjustments and spell deletions	52
7.4	Comparison of the SIAB 7508 and SIAB-Regional File 7508	53
7.5	Recoding tables of aggregated variables in comparison with the weakly anonymous version of the SIAB 7508	54
7.6	Frequency tables	71
8	List of abbreviations	72

1 Introduction and outline

1.1 Introduction

The regional file of the Sample of Integrated Labour Market Biographies 1975-2008 (referred to as SIAB-R 7508 or SIAB-Regional File in the following) is a factually anonymous¹ scientific use file of the Sample of Integrated Labour Market Biographies 1975-2008 (SIAB 7508). The regional file is a specific type of scientific use file which distinguishes itself from other types of scientific use files by preserving a comparatively detailed regional data structure. However, it provides only a limited range of variables and values at the level of individuals and establishments for reasons of data confidentiality. Altogether the SIAB-Regional File in its current version covers the employment histories of 1,515,463 individuals. Their employment biographies are documented in a total of 34,862,777 lines of data. The starting point for the production of the SIAB-Regional File is the weakly anonymous version of the SIAB 7508, which is a 2 % sample of the Integrated Employment Biographies (Integrierte Erwerbsbiografien - IEB) of the Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung - IAB).

The IEB comprises all individuals who showed one of the following statuses at least once during the observation period from 1975 to 2008:

- 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 (SGB III) (recorded from 1975 onwards) or Social Code Book II (SGB II) (recorded from 2005 onwards)
- registered with the Federal Employment Agency (Bundesagentur für Arbeit - BA) as a jobseeker (recorded from 2000 onwards)
- planned or actual participation in an employment or training measure (recorded from 2000 onwards)

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

Whilst the Employee History (Beschäftigten-Historik - 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 SGB III and SGB II is recorded in the Benefit Recipient History (Leistungsempfänger-Historik - LeH) and the Unemployment Benefit II Recipient Histories (Leistungs-Historiken-Grundsicherung - LHG and XLHG). The Jobseeker Histories (Ar-

¹ For further information on the concept of factual anonymisation see Chapter 3.

beitsuchenden-Historiken - ASU and XASU) are the data source for the periods of job-search recorded by the BA, whilst participation and planned participation in employment and training measures is recorded in the Participation-in-Measures History File (Maßnahme-Teilnahme-Historik - MTH). With the exception of participation and planned participation in employment and training measures from the MTH, all of the data sources included in the IEB are also contained both in the SIAB-Regional File and the weakly anonymous version of the SIAB.

The SIAB replaces the IAB Employment Sample (IAB-Beschäftigtenstichprobe - IABS) and the Integrated Employment Biographies Sample (Stichprobe der Integrierten Erwerbsbiographien - IEBS). By merging the data sources of the IABS and the IEBS, it is now possible to provide a dataset that combines the advantages of the IABS (long observation period) with those of the IEBS (entire employment biographies). The SIAB is produced at the Research Data Centre (FDZ) of the Federal Employment Agency (BA) at the IAB.

This FDZ data report describes the preparation of the SIAB-Regional File and the variables of the scientific use file. The data report is structured as follows: besides an introduction Chapter 1 contains information on data access as well as an outline of the data, the volume structure and a list of variables. A description of the individual data sources can be found in Chapter 2. Data preparation and the anonymisation concept of the scientific use file are described in Chapter 3. Chapter 4 discusses the data quality and Chapter 5 closes with a description of the individual variables contained in the data set.

1.2 Data use

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

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

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

1.3 Differences between the SIAB and earlier versions

The Sample of Integrated Labour Market Biographies (SIAB) replaces the IAB Employment Sample (IAB-Beschäftigtenstichprobe - IABS) and the Integrated Employment Biographies Sample (Stichprobe der Integrierten Erwerbsbiographien - IEBS) as the major data product

for the analysis of individual labour market biographies. As the two predecessor data sets also form the data basis for their own scientific use files and therefore the predecessors of the SIAB-Regional File, the main differences between these data sets are outlined briefly below in Table 1.

Scientific use files of the FDZ (excerpt)			
	SIAB	IABS	IEBS
Type of scientific use file	regional file ² (SIAB-R 7508)	regional file (IABS-R04)	basis file ³
Data basis	Integrated Employment Biographies (IEB)	Employee and Benefit Recipient Histories (BLH)	Integrated Employment Biographies (IEB)
Sources	Employment, benefit receipt, job-search	Employment, benefit receipt	Employment, benefit receipt, job-search, participation in measures
Observation period Employment	1975-2008	1975-2004	1993-2003
Survey design	2% random sample of the individuals in the IEB	Stratified 2% random sample of the employees in the BLH	Birthday sample of the individuals in the IEB, 8/365 ≈ 2.19%
Annual notifications with person group 101 and remuneration == 0	Not contained	Contained	Not contained
Preparation of the basis data	Marginal corrections of the IEB	Comprehensive data corrections, completion procedure	Marginal corrections of the IEB

Table 1: Differences between scientific use files of the IABS, IEBS and SIAB.

Unlike with the SIAB-R 7508 and the IEBS-SUF, a number of additional correction procedures were conducted before anonymisation for the IABS-Regional File. A description of these correction procedures in the predecessor data of the IABS that were no longer carried out for the SIAB data can be found in the appendix (see Sections 7.1 and 7.2).

² We use the term *regional file* for a scientific use file in which the data have a comparatively detailed regional structure but for data protection reasons cover only a substantially reduced spectrum of information at individual and establishment level.

³ We use the term *basis file* for a scientific use file which shows detailed personal or establishment-specific information but for reasons of anonymity contains no or only very roughly structured spatial units.

1.4 Outline

Current data version	Sample of Integrated Labour Market Biographies - Regional File (SIAB-R) 1975-2008, Version 1 short form: SIAB-R_7508_EN_v1_dok1
Type and volume of the data	Factually anonymous data (scientific use file), 2% random sample of individual accounts drawn from the Integrated Employment Biographies (IEB) of the IAB.
Outline of the contents	<p>Employee History (Beschäftigten-Historik - BeH): Annual notifications and end-of-employment notifications submitted to the social security agencies for employees covered by social security and employees in marginal part-time employment.</p> <p>Benefit Recipient History (Leistungsempfänger-Historik - LeH): Information on benefit receipt in accordance with Social Code Book III (SGB III) for recipients of unemployment benefit, unemployment assistance and maintenance allowance.</p> <p>Unemployment Benefit II Recipient Histories drawn from A2LL (Leistungs-Historik-Grundsicherung aus A2LL - LHG) and XSozial-BA-SGB II (XLHG): Data on individuals in receipt of basic social security benefits in accordance with Social Code Book II (SGB II) (co-operations of employment agencies and municipalities – ARGE, separated responsibilities responsible for the implementation of SGB II – getrennte Trägerschaft, and authorised municipalities - zugelassene kommunale Träger)</p> <p>Jobseeker History (Arbeitsuchenden-Historik - ASU) and Jobseeker History drawn from XSozial-BA-SGB II (XASU): Information on job-search activity (applicant pool of the Federal Employment Agency) including information on job-search activity reported via the transmission standard XSozial-BA-SGBII to the BA by authorised municipalities or separated responsibilities.</p>
Variables with regional relevance	333 districts and aggregated regions based on the place of work, categorical variable for commuter status/place of residence
Date of territorial allocation	31.12.2008
Type of territorial allocation	BeH, LeH, (X)ASU, (X)LHG: corrected territorial allocation
Period covered	The period covered depends on the data source. BeH 1.1.1975 - 31.12.2008 (2007: 18-month file; 2008: 12-month file) LeH 1.1.1975 - 31.12.2008 (X)ASU 1.1.2000 - 31.12.2004 and 1.1.2007 - 31.12.2008 (X)LHG 1.1.2007 - 31.12.2008
Update frequency	Continuous

Data unit	Employees covered by social security (including marginal part-time employees from 1999 onwards), benefit recipients, job-seekers
Number of cases	1.515.463 individuals, 34.862.777 observations
Adjustments - outline	The data of the SIAB-Regional File undergo 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.
Special features - outline	The data in the SIAB-Regional File undergo an episode splitting procedure. For this, partly overlapping observations within an individual account are split in such a way that only non-overlapping or completely parallel periods are created.
File format and size	Stata (approx. 1.4 GB), SPSS (approx. 2.8 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)
Citation method	Sample of Integrated Labour Market Biographies – Regional File 1975-2008 (SIAB-R 7508), Nuremberg 2011 abbreviation: SIAB-R 7508 v1

1.5 List of variables

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

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

Example: the school-leaving qualification (schbild) is only available for (X)ASU observations; the observations of the other data sources contain missing “.n” in this variable. Another characteristic of the SIAB is that some variables have different contents depending on the data source. For instance, for BeH observations the employment status variable (erwstat_gr) contains the person group of the employment notification procedure, for LeH observations it contains the type of benefit, for (X)LHG observations it contains the SGB II status and for

(X)ASU observations the job-search status. These differences are not immediately obvious from the variable name for every variable.

Type of variables	Variables	Page	1	2	3	4
			BeH 1975- 2008	LeH 1975- 2008	(X)ASU 2000-2004 2007-2008	(X)LHG 2007- 2008
Identifiers	persnr	30				
	bnn	31				
Period of validity/ Dates	begorig	31				
	endorig	32				
	begepi	32				
	endepe	32				
Information on employment, benefit receipt and job- search	quelle_gr	44				
	erwstat_gr	40				
	grund_gr	35				
	tentgelt_gr	37				
Personal information	deutsch	33				
	gebjahr	33				
	frau	33				
	bild	34				
	beruf_gr	39				
	stib	39				
	schbild	35				
	tage_jung	41				
tage_alt	41					
(Generated) Establishment vari- ables	w73_gen_gr	42				
	w03_gr	42				
Regional identifiers	ao_region	43				
	pendler	44				
Generated technical variables	spell	45				
	level1	46				
	level2	45				

Table 2: List of variables and degree of completeness according to sources

1.6 Volume structure

Number of cases in the SIAB-R 7508	before splitting	after splitting
Employment (BeH)	23.642.269	25.825.849
Benefit receipt SGB III (LeH)	3.530.492	4.671.200
Job-search, also SGB II (ASU/XASU)	2.185.067	3.588.475
Benefit receipt SGB II (LHG/XLHG)	324.396	777.253
Total number of observations	29.682.224	34.862.777
Individuals		1.515.463

Table 3: Volume structure SIAB-Regional File 7508

2 Data sources

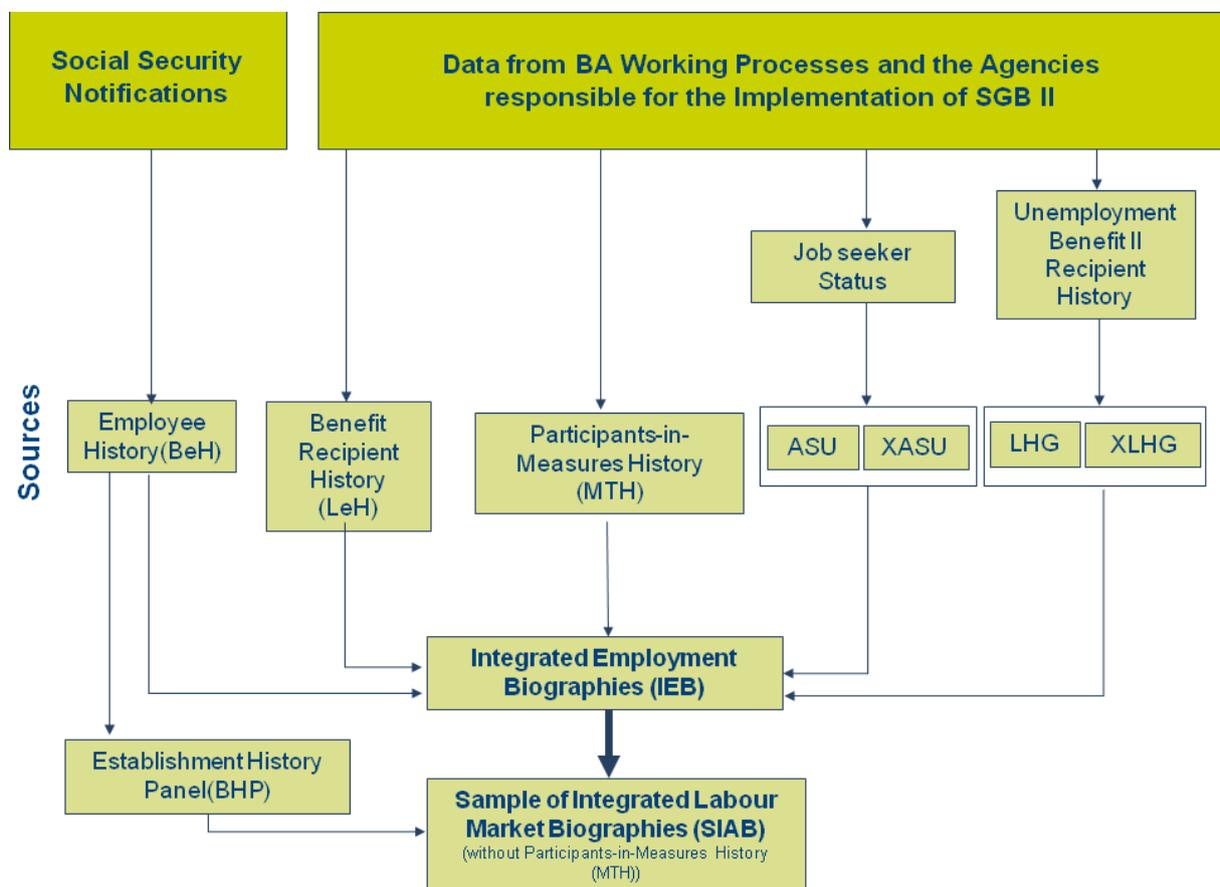


Figure 1: Data sources of the Integrated Employment Biographies (IEB)

The SIAB was drawn from the Integrated Employment Biographies (IEB) of the IAB as a 2% sample. The IEB in turn unite data from five different data sources, each of which may contain information from different administrative procedures.

Observations from all of the data sources shown, with the exception of participation (or planned participation) in employment and training measures (MTH), are incorporated into the SIAB from the IEB. Each data source of the IEB is specified below.

2.1 Employee History (BeH)

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

Every year in which an individual is in an employment relationship is depicted by at least one notification. The data are recorded by the health insurance companies, collected in a continuous file by the Federal Employment Agency (BA) and subsequently integrated into the History File of the IAB.

The variables of the SIAB-Regional file related to establishments are taken from the Establishment-History-Panel (Betriebs-Historik-Panel - BHP) which itself is also based on the BeH (see Hethey-Maier/Seth 2010). The following modified variables based on the BHP are included in the scientific use file:

- Place of Work (ao_region)
- Industry Classification WS 73 (extrapolated until 2008) (w73_gen_gr)
- Industry Classification WZ 03 (w03_gr)

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

1. Information on establishments is based on individual data and aggregated on the annual cut-off date of the BHP, June 30. Unlike the data on individuals, the establishment variables are not spell data but are only valid on 30 June precisely each year⁴.
2. Secondly, the BHP is created directly from the BeH, whereas the SIAB Individual File is generated from the IEB (see Figure 1). The IEB however does not incorporate all BeH observations (see Section 3.1).

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 (BA). The benefits comprise unemployment benefit, unemployment assistance and maintenance allowance, in other words exclusively benefits in the sphere of SGB III. The entitlement to benefits depends on legal requirements being fulfilled, e.g. not being classed as “in need” in the case of unemployment assistance, non-completion of the qualifying period for unemployment benefit or periods of unemployment during which the requirements were not met. Therefore, periods of unemployment without benefit receipt according to SGB III are not reported in the Benefit Recipient History.

2.3 Unemployment Benefit II Recipient Histories (LHG/XLHG)

Whereas the LeH contains the benefits in the sphere of SGB III, the Unemployment Benefit II Recipient Histories (Leistungs-Historik-Grundsicherung – LHG/XLHG) comprise receipts of benefits in accordance with SGB II. This covers both basic social security benefits (e.g. unemployment benefit II) and supplements to unemployment benefit or additional benefits. Unlike the benefits in the sphere of SGB III, the Bundesagentur für Arbeit (BA) is no longer necessarily the sole institution responsible for administering the benefits. The data therefore principally distinguish between the three possible types of institution responsible for implementing SGB II:

- Co-operations of employment agencies and municipalities (Arbeitsgemeinschaften - ARGE) in which the BA and the municipality deal with tasks jointly.
- In the case of separated responsibilities (getrennte Trägerschaft) the tasks continue to be divided between the BA and the municipality⁵.

⁴ An extreme example: there is an employment notification from 1.1.2006 until 30.5.2006; the establishment goes bankrupt in June 2006. There is then no information about this establishment in the BHP for 2006.

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

- In the case of authorised municipalities (zugelassener kommunaler Träger, optierende Kommune, Optionskommune), which are municipalities that have opted to be the sole agency administering benefits under SGB II, the municipality takes on the tasks in the field of SGB II independently.

The Unemployment Benefit II Recipient Histories cover two sources that differ in the way in which the data are transmitted: the data of the “Unemployment Benefit II Recipient History drawn from A2LL” (LHG) come from the BA administrative procedure A2LL, whilst the “Unemployment Benefit II Recipient History drawn from XLHG-BA-SGB II” (XLHG) records data which was reported in accordance with the transmission standard X-Sozial-BA-SBG II.

As a rule A2LL is used in all ARGE co-operations, whilst the standard X-Sozial-BA-SBG II is used by the authorised municipalities. Both of the procedures are used by municipalities with separated responsibilities.

The amount of benefits received is not determined at the level of the individual but at the level of the benefit community (Bedarfsgemeinschaft⁶). It is not possible, however, to link individuals with benefit receipt under SGB II at the level of benefit communities in this dataset. The LHG and the XLHG in principle contain all registered individuals who are entitled to receive benefits in accordance with § 7 SGB II, in other words needy individuals who are capable of work and, if applicable, also the members of their benefit community.

Each dataset depicts non-overlapping periods of entitlement to benefit of a person in a certain benefit community. However, the observations are split

- on transfer-relevant birthdays (14, 15, 18, 25 and 65) of the members of a benefit community,
- if there are changes in the composition of a benefit community (e.g. by members joining or leaving the benefit community),
- if there are changes in the variables of the benefit community client and

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

⁶ “A so-called ‘Bedarfsgemeinschaft’ (which is sometimes translated as ‘community of need’) includes all persons in a household receiving benefits jointly (i.e. as a joint payment). In the majority of cases, the ‘Bedarfsgemeinschaft’ and the household will be identical, which in particular applies in case of (married or unmarried) couples and parents with children below 25 years of age. However, under specific circumstances, the ‘Bedarfsgemeinschaft’ might not include all household members. Or, a household in which everybody receives benefit payments might be made up by more than one ‘Bedarfsgemeinschaft’. An example for the former would be a grown-up child living with its parents and earning just enough to make its own living but having insufficient means to support his mother and father – in this case the ‘Bedarfsgemeinschaft’ will only include the parents. An example for the latter would be a three- (or more) generation-household. Since a ‘Bedarfsgemeinschaft’ may only consist of two generations, this type of household would be made up of two such ‘Bedarfsgemeinschaften’, one consisting of the grandparent(s) and one of the parent(s) and child/children.” (Trappmann et al. 2009, p. 7). In this data report we also use the term benefit community.

- at the beginning and end of a sanction period for observations from 1 April 2006 onwards.

With regard to sanctions it must be taken into account that although the beginning and the end of a sanction period lead to the observations being split, it is not possible to determine from the data either the duration or the type of the sanction or the time when the sanction was imposed or when it began. The reason for this is the lack of a corresponding variable or value that indicates the start, the type or the duration of a sanction.

2.4 Jobseeker Histories (ASU/XASU)

Data about jobseekers are stored in the Jobseeker History (Arbeitsuchenden-Historik – ASU). The ASU data source contains information on jobseekers who are registered with employment agencies and from 2005 onwards also includes separated responsibilities responsible 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-SBG II standard.

3 Data preparation, sampling procedure and anonymisation

3.1 Corrections of the basic data for the SIAB-R 7508

The Integrated Employment Biographies of the IAB (IEB) with their different data sources are an important starting point for numerous datasets provided by the Research Data Centre (FDZ) of the Federal Employment Agency (BA) at the Institute for Employment Research (IAB). Before the IEB are ready for the preparation of the final FDZ data products they undergo some source-specific correction procedures.

The preparations of the IEB data which are relevant for the production of the SIAB 7508 and the scientific use file SIAB-R 7508 are described in the following sections.

3.1.1 Employee History (BeH)

- Observations with earnings amounting to zero or with no details on earnings and the value 101 for the person group variable and the value 50 for the reason for notification (annual notification) are not incorporated into the IEB.
- Some person groups whose data are not available for the full observation period are excluded in order to ensure that the population is as constant as possible over time. This means that from 1999 onwards only the person groups 101, 102, 103, 105, 106, 112,

118, 119, 120, 140, 141, 142, 143, 149, 201, 203 and 205 are supplied, as well as the two groups 109 and 209, which indicate people in marginal part-time employment. Groups that are not included from 1999 onwards include people in short-term employment, i.e. person groups 110, 202 and 210.

- Gender and date of birth are taken from the social security number. It must be taken into account that this individual information is harmonized across the different data sources. The territorial allocations for place of work and place of residence are updated to the status as of 31.12.2008.

3.1.2 Benefit Recipient History (LeH)

- For the IEB, only observations from the LeH for the period from 1.1.1975 (end date) until the latest available data are taken.
- It must be taken into account that in cases of ongoing receipt the end date corresponds to the end of the period for which the benefit is allowed and not the end of the benefit receipt period. There is no information on the end of the period for which the benefit is allowed in the SIAB-Regional File because benefit receipt notifications still valid on 31.12.2008 are right-censored (see Section 3.3.2). Observations with no end date or an invalid end date are excluded from further processing, since in these cases it cannot be assumed that a benefit payment was made at all.
- Observations without a valid start date are excluded.
- Observations whose end date precedes the start date are excluded.
- If the end date for unemployment assistance precedes the start date by one day, then the end date is deferred by one year.
- Overlapping LeH observations are corrected for the IEB in accordance with the following rules:
 - o All overlaps in which both of the observations involve the same type of benefit and the same amount of benefit are corrected as follows:
 - If one observation is completely embedded within the period of another one, the shorter observation is deleted. If the two observations are of the same length, one (randomly selected) observation is deleted.
 - In cases where observations partially overlap, the end date of the observation that began first is corrected.
 - o Overlaps of observations involving **different types of benefit** and/or different amounts of benefit paid are only corrected if the overlap concerns just one day and the original duration of both observations exceeds one day. In these cases the end date of the observation that began first is put back by one day.

3.1.3 Jobseeker Histories (X)ASU

- The data from the ASU that are used cover the period from January 2000 to December 2008. The registered periods of job-search activity of the individuals are regarded as complete.
- The ASU data basis only distinguishes between observations with the status “unemployed” and “jobseeker”. In the IEB, however, a distinction is made between “unemployed”, “incapacitated for work” and “jobseeker”. The employment status “incapacitated for work” is assigned to IEB observations when in the ASU data basis
 - o a preceding observation with the status “unemployed” exists which joins the next observation without a gap and has “incapacitated for work” as the reason for exit and
 - o a subsequent observation with the status “unemployed” exists which also follows without a gap and
 - o the observation itself does not have the status “unemployed” but “jobseeker”.
- Unlike the ASU, the XASU does not take into account periods of incapacity for work as there is no information about illnesses available in the XASU data.
- Observations whose end date precedes the start date are filtered out.

3.1.4 Unemployment Benefit II Recipient Histories (X)LHG

- Only observations of people who are capable of work and people under the age of 64 are read in.

Changes in the structure of benefit communities (Bedarfsgemeinschaften) as defined in SGB II, such as members of the benefit community reaching the retirement age of 65, lead to the splitting of spells in the data. The occurrence of splitting and the simultaneous availability of exact dates and individual information are sensitive with regard to data confidentiality. The spells concerned are therefore anonymised to ensure data confidentiality.

3.1.5 Corrections of the aggregate IEB file

After the raw data have been corrected and combined to form the population of the IEB, the aggregate file of the IEB is corrected across all of the sources. For this the following correction steps are carried out:

- Observations in which the age is under 13 or over 75 are deleted.
- Observations whose end date precedes the start date are deleted.
- Accounts in which the gender or date of birth change are corrected in accordance with priority rules.

The weakly anonymous SIAB 7508 data based on this file do not undergo any further corrections and thus (virtually) correspond to the original data of the IEB. A few variables are simply recoded and for reasons of anonymity some of the data are aggregated.

These data represent the basis for producing the scientific use file of the Sample of the Integrated Employment Biographies (SIAB-R 7508).

3.2 Sample design

The basis for the scientific use file is the weakly anonymous version of the SIAB, which is a 2% sample of the individuals recorded in the IEB (see Dorner et al. 2010).

Due to the nature of the sampling, case numbers of the population can be estimated with a weighting factor of 50. This applies in principle to all the (person) groups across all of the data sources included in the SIAB-Regional File. However, depending on the research question and the chosen set of variables, the varying degrees of completeness of the data sources (see Table 2) and their specific validity periods have to be taken into account in order to evaluate the representativity of the research.

The sampling procedure for the SIAB observations is therefore comparable to that of the IEBS.⁷ However, there are two differences with regard to the IAB Employment Samples (IABS). First, the IABS is an employment sample, i.e. the individuals were drawn solely from the employee observations of the BeH, and the benefit observations of the LeH were simply added to the sample. Thus, there are no individuals with only LeH observations in the IABS, whereas the SIAB does contain individuals who do not have any employment spells at all in their biography (BeH observations). Second, unlike the IABS, the SIAB is not drawn as a stratified sample distinguishing between East and West Germany and between German/non-German. Owing to the large numbers of cases, however, these four groups are contained in the data with approximately the same selection probabilities even without explicit stratification.

3.3 Anonymisation of the SIAB-R 7508

The Sample of Integrated Labour Market Biographies, which serves as the data basis for the SIAB-Regional File, contains a large amount of personal information on the status of a person on the labour market. Additional establishment characteristics are available in the case of employment notifications from the social security notification procedure. The legislator stipulates that these social data (§ 67 SGB X) which are available at micro level at the FDZ

⁷ The sample for the IEBS is drawn using a random selection of eight birthdays (see Oberschachtsiek et al. 2009).

may only be transmitted for scientific purposes (§ 282 para. 7 SGB III, § 75 SGB X) if both the personal and the establishment-related data can be regarded as factually anonymous when tested. This means that a disproportionately large amount of time, expense and effort would be required in order to identify an individual in the data set. The costs of deanonymisation must therefore exceed its benefits.

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

3.3.1 Merging sources

When developing the SIAB-Regional File from the SIAB 7508, details on the source of certain observations were merged. For example, information from the original data sources of the Unemployment Benefit II Recipient Histories (LHG and XLHG) and information about job-search (ASU and XASU) were combined in each case to a joint data source, (X)ASU and (X)LHG, in the SIAB-Regional File. The reason for this is that the information about whether a person in a region receives his/her support from an authorised municipality (zugelassener kommunaler Träger - zkT) constitutes a risk of the data subject being re-identified. This risk results from the fact, that the information about the institution responsible for the person, in combination with regional information from other sources or from the employment history can make it possible to identify the location at local authority level.

3.3.2 Periods of validity of the sources

When developing the SIAB-Regional File, after aggregating the data sources (see Section 3.3.1) and in accordance with the data protection regulations mentioned above, information about the start and end dates of selected observations were adjusted. Figure 2 illustrates the periods for which the various sources are available in the SIAB-Regional File.

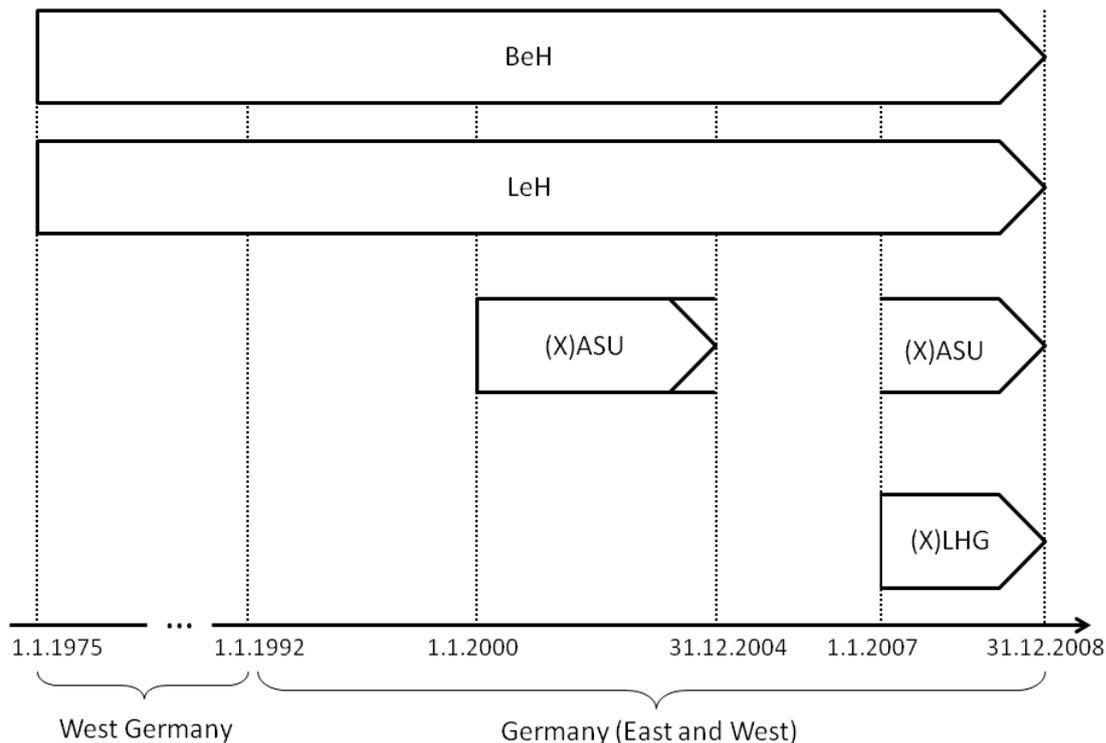


Figure 2: Periods of validity of the data sources

For the Employee History (BeH) and the Benefit Recipient History (LeH) the SIAB-R 7508 only contains observations for the period between 1.1.1975 and 31.12.2008. Notifications submitted in the context of the social security notification procedure in East Germany are only assumed to be complete from 1992 onwards. In employment notifications with valid places of work in East Germany before the year 1992 the information on the district region is set to missing (.z). It is therefore only possible to distinguish whether a person is employed in East or West Germany in the SIAB-Regional File from 1992 onwards.

The (X)ASU is available in the SIAB-Regional File for the period between 1.1.2000 and 31.12.2004 and the period between 1.1.2007 and 31.12.2008. As the ASU original source is only assumed to be complete from the year 2000 onwards, observations with an end date before 1.1.2000 are not reported in the SIAB-Regional File. The deletion of observations in the years 2005 and 2006 is due first to the quality-relevant under-representation of notifications from the authorised municipalities responsible for implementing SGB II (zugelassene kommunale Träger) in the IEB original data (see Dorner et al. 2010). Second, revealing these data is problematic in terms of data protection legislation because of the risk of individuals reported by authorised municipalities being identified in combination with regional information. The outlined problems associated with notifications from authorised municipalities also apply to the (X)LHG in 2005 and 2006. Accordingly the SIAB-Regional File only contains observations from the (X)LHG with episode start dates from 1.1.2007 onwards.

Observations whose start and end dates are outside these source-specific periods of validity are no longer contained in the data. In the case of observations from the LeH, (X)ASU and (X)LHG sources which span more than one year and in which either the start date or the end date falls in one of the source-specific periods of validity, the dates are changed. This means that the original information regarding either the start or end date is replaced by the respective start or end date of the source-specific period of validity.

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

The following fictitious examples in Figure 3 illustrate the procedure used for adjusting the dates. The white bar [1] always represents the original period and the grey bar [2] the period contained in the SIAB-Regional File after the date adjustment procedure.

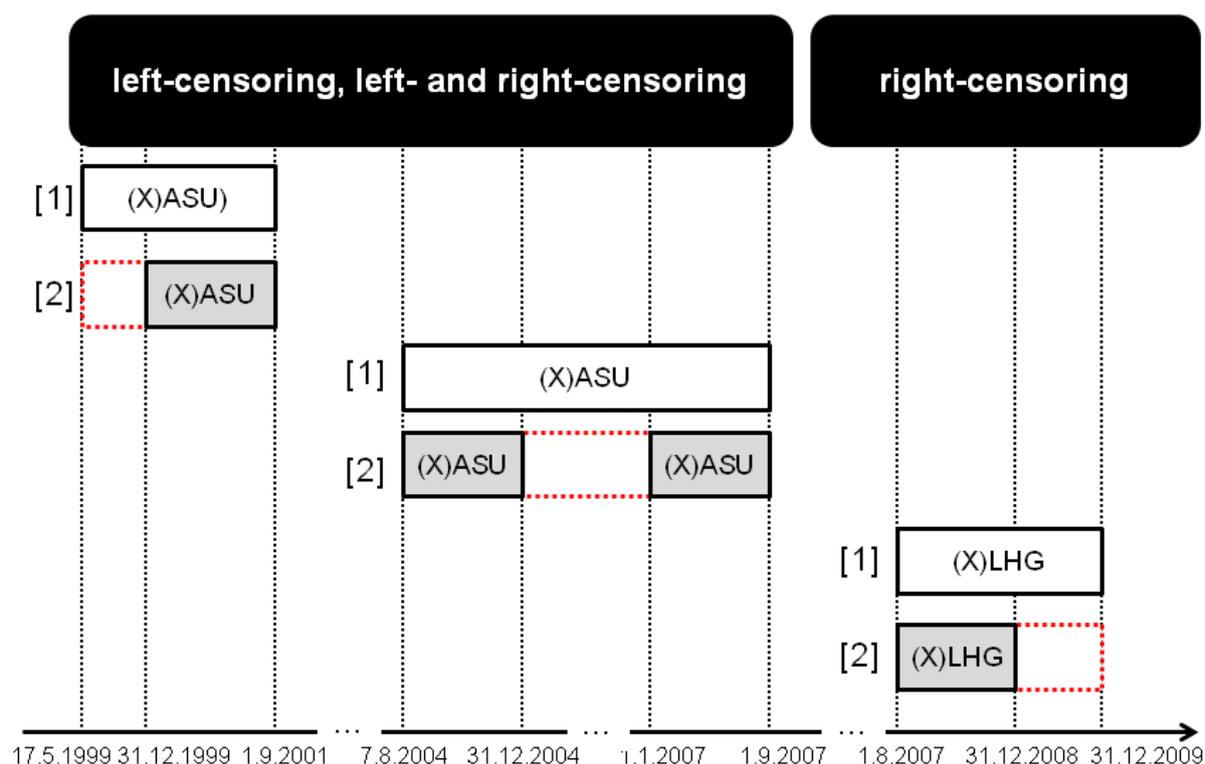


Figure 3: Examples of the date adjustment procedure for observations spanning more than one year

In the example showing the left-censoring of a job-search observation from the (X)ASU to 1.1.2000 and in the example depicting the right-censoring of an observation from the (X)LHG (also LeH or (X)ASU) to 31.12.2008, the number of data records in the data remains the

same. Only the parts of the multi-year observation which are outside the observation period covered by the SIAB-Regional File are deleted.

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

An overview of the date conversions carried out can be found in tabular form in the Appendix (see Section 7.3, Table A2).

3.3.3 Censoring of data on age

For data confidentiality reasons the scientific use file only contains those data records in which a person's age is between 17 and 62. Records that end before the age of 17 or begin after the age of 62 are excluded. In the case of records spanning more than one year which begin before or in the year of a person's 17th birthday but end after that date, the original start date ('begorig') was converted to 1st of Januar of the year in which the person reaches the age limit of 17. Similarly, the original end date ('endorig') was converted to 31 December of the year in which the person reaches the age limit of 62 if the record begins before or in the year of a person's 62nd birthday and end after that year. The variables 'tage_jung' and 'tage_alt' report the number of days in (regular) employment⁸ before the age of 17 and after the age of 62, respectively.

3.3.4 Aggregation and deletion of variables

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

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

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

⁸ The calculation of periods of employment takes the following restrictions into account: `quelle_gr==1 & level2==0` (Stata Syntax).

This demand was met by checking the number of inhabitants in the 413 autonomous municipal districts (kreisfreie Städte) and rural districts (Landkreise) as of 31.12.2008⁹. The regionalisation contained in the SIAB-Regional File 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. This regionalisation in the SIAB-Regional File is based on the procedure used to create the microcensus district regions (Mikrozensuskreisregionen - MZKR) (see Christians/Wirth 2009, p.22). The aggregation results in a total of 333 district regions which are shown in the variable 'place of work: region' (ao_region) (territorial allocation 31.12.2008). As was done when creating the microcensus district regions, superordinate spatial units such as standard statistical regions (Raumordnungsregionen) and federal states were taken into account when defining the district regions (see Section 7, Figure A1/Table A10).

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

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

On the basis of these criteria, further variables were aggregated when producing the SIAB-Regional File. Frequency counts of individuals as of 30th June of any year were taken as the basis. Attributes of variables with a high risk of deanonymisation were aggregated first if their univariate marginal distribution as of 30th June of the respective year did not comprise a minimum number of persons (2a). Second, they were aggregated if the cells of the contingency tables showing the relationship between the particular variable and the variable 'place of work: region' did not contain a minimum of three persons (2b). When checking the data on

⁹ Data basis: Population projection of the Federal Statistical Office (Statistisches Bundesamt) as of 31.12.2008 for the autonomous municipal authorities (kreisfreie Städte) and the administrative districts (Kreise) (NUTS 3-level) of the Federal Republic of Germany.

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 7508.

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

Aggregations were conducted for the following variables: nationality, occupation – activity performed, classification of economic activity (WS 73 and WZ 2003), reason for notification/ reason for end of benefit receipt/ reason for discontinuation of unemployment benefit II/ reason for deregistration, occupational status.

For these variables the years 1975 up to and including 1984 were not used to assess the aggregation level required for the factual anonymity of the variables. Due to the structural change on the 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.

3.3.5 Consequences of the anonymisation

The information on employment and benefit receipt in accordance with SGB III in the SIAB-Regional File are only affected to a small extent by deletions at the level of data records and individuals. The deletion of individual data records for anonymisation reasons in the process of adjusting the observation periods for the (X)ASU and (X)LHG sources resulted in some individuals being dropped from the SIAB-Regional File. As a result of the data preparation procedures mentioned above a total of 143,561 individual accounts were deleted with reference to the SIAB 7508. The SIAB-Regional File therefore still contains 91.3 percent of the persons included in the source file, SIAB 7508 (see Section 7.4, Table A3).

Some personal and establishment-related variables were either aggregated or deleted from the data during the anonymisation process across all sources. In the data sources with information on benefit receipt and job-search, the spectrum of variables was reduced considerably, while the employment-related variables were largely retained and were simply aggregated. The SIAB-Regional File thus provides the continuity for employment-related analyses to a large extent compared with the IAB Employment Samples (especially the IABS-Regional File 1975-2004).

More detailed information about deletions and aggregations at variable level due to anonymisation compared to the weakly anonymous version of the SIAB 7508 and compared to related FDZ data products can be found in the tables in the Appendix to the Datenreport.

3.4 Definition of missing, invalid and anonymized values

In the SIAB-R 7508 missing values or values that are invalid due to anonymisation are coded as follows:

Term	Value	Description
No (valid) details available, anonymised values	.z	Values of a variable which are not systematically missing, i.e. the variable is available in principle for the data source, but no details are available for the value considered or no allocation is possible. Otherwise: anonymised values due to data confidentiality obligations. Please note that it is not possible to distinguish between anonymised and invalid values originating from the weakly anonymous version of the SIAB.
Systematically not available	.n	A variable is not available in principle for a data source (dark grey cells in the overview of variables in Section 1.5) or is not available for a certain period.

Table 4: Coding of missing and anonymised values

3.5 Episode splitting

The data of the SIAB-Regional File are available with “split” episodes. If observations overlap within an individual 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.

Episode splitting results in observations being duplicated and the period of validity being changed. The original date variables for the beginning and the end of the original observation (begorig and endorig) are the starting point for episode splitting. 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 split episodes. Using the technical variables in the data which are generated additionally (level1, level2) it is feasible to check whether original observations have been multiplied. To restore the 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 (Stata syntax: `keep begorig == begepi`).

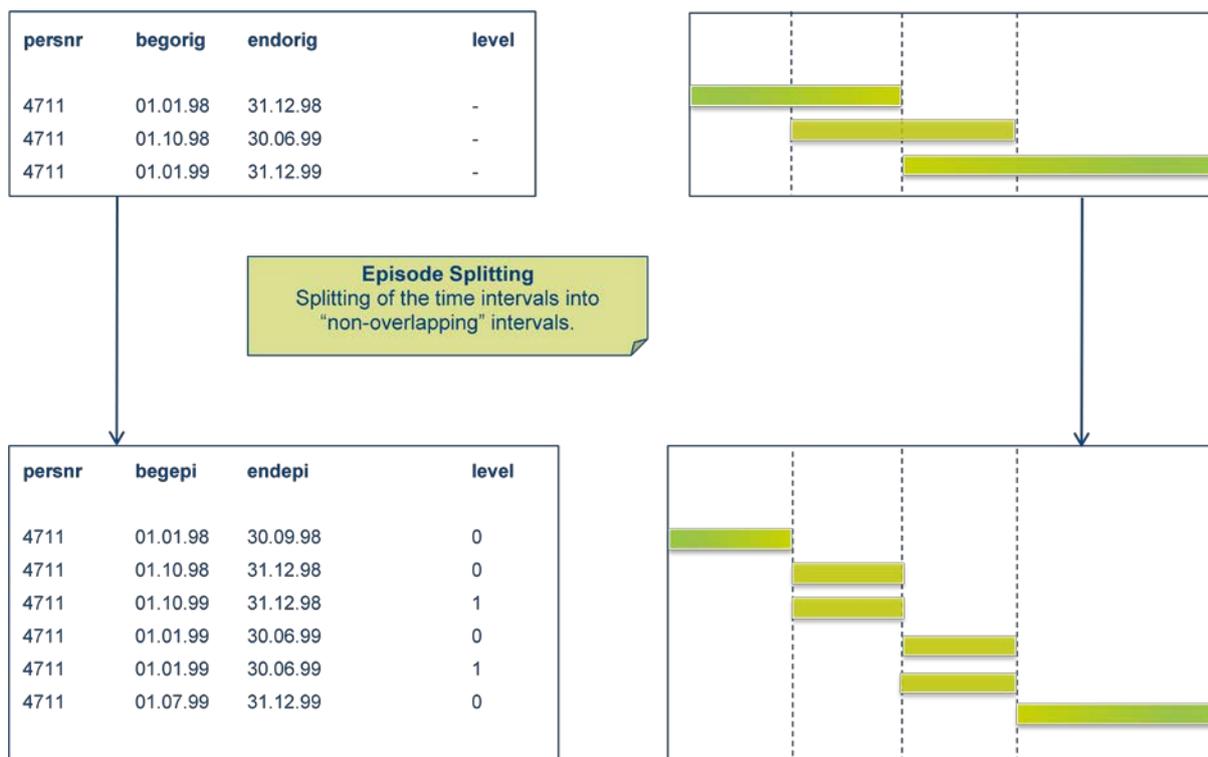


Figure 4: Episode splitting in the IEB and SIAB data

4 Data quality and problems

4.1 Employee History (BeH)

- Due to the introduction of the employment notification procedure in the federal states of East Germany, the notifications for BeH observations for East Germany can only be assumed to be sufficiently complete from 1993 onwards. 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.
- Within the employment notification procedure a certain time lag is unavoidable. Although changes in employment relationships have to be reported immediately and existing employment relationships have to be confirmed annually by April of the following year, some notifications actually arrive years later. The History File of the IAB is not updated continuously, however, but at certain intervals. This is done using files of employment notifications for one particular year which were submitted 36, 18, 12 or 6 months after the end of the reporting year (e.g. the 18-month file for 2007 can be created in July 2009 at the earliest). Notifications submitted more than three years late are not taken into ac-

count at the IAB, which means that a 36-month file shows a 100 % degree of completeness by definition.

- In the version of the IEB on which the SIAB data are based the degree of completeness of the BeH observations last stood at 100 % in 2006. 18-month files were used for 2007, and the observations for 2008 come from a 12-month file. It can therefore be assumed that the BeH observations for 2007 and in particular for 2008 are slightly under-reported in the SIAB. However, this should not reduce the ability to analyse the data at individual level. The missing notifications occur more frequently in a few establishments, however.
- 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 of the German Pension Insurance (Deutsche Rentenversicherung) increases considerably from that year onwards¹⁰.
- In 1975 a smaller number of cases than expected emerges for individual birth cohorts. 1975 is therefore a problematic year for analyses¹¹.
- Owing to the introduction of the new employment notification procedure in 1999 only the person groups 101, 102, 103, 105, 106, 112, 118, 119, 120, 140, 141, 142, 143, 201, 203 and 205 are supplied, as well as the two groups 109 and 209, which indicate people in marginal part-time employment. Groups that are not supplied include people in short-term employment, i.e. person groups 110, 202 and 210. In the SIAB-Regional file the person groups mentioned above are only contained in aggregated form. Exception: prior to 1999 observations with no (valid) details on the person group are also recorded, which is of importance in particular for 1991 because a great many employees were registered in this way in East Germany.

4.2 Benefit Recipient History (LeH)

- The benefit receipt data used to be saved on magnetic tapes. Owing to a fault in the magnetic tapes the benefit receipt data up to and including 1980 are only partially contained. Thus in the SIAB data, too, it can be assumed that information on benefit receipt in that period is not available in full. This is to be taken into account when conducting analyses with the LeH.

¹⁰ See Bender et al. 1996.

¹¹ See Bender et al. 1996.

4.3 Unemployment Benefit II Recipient Histories (LHG/XLHG)

- With regard to the completeness of case numbers or benefit histories from the LHG and XLHG data sources, there are only occasional gaps for individual institutions from 2007 onwards. It can be assumed that these gaps continue in the SIAB data. In addition, the quality of the notifications depends to a large extent on the institution submitting them, which must be taken into account in particular when conducting regional analyses.
- The degree of completeness of the variable “Reason for discontinuation of unemployment benefit II” (grund_gr, see Section 5.4.1) is comparatively low in the SIAB data. However, it must be taken into consideration here that the “Reason for discontinuation of unemployment benefit II” is a purely person-related variable. Thus if an observation of an individual in receipt of Unemployment Benefit II is split due to a change in the structure of his/her benefit community, the variable “Reason for discontinuation of unemployment benefit II” shows the value “missing” (.z).
- Although the beginning and the end of a benefit sanction lead to the observations being split, it is not possible to determine from the data either the duration or the type of the sanction or the time when the sanction was imposed or when it began.
- In the LHG as well as to a higher degree in the XLHG it can be observed that there is an above-average number of 15-year-olds and fewer 16- and 17-year-olds with references to a lack of capacity to work. 15-, 16- and 17-year-old individuals in need who are supported by the authorised municipalities may therefore be more poorly represented in the IEB, as “individuals not capable of work” are not included in the IEB.

4.4 Jobseeker Histories (ASU/XASU)

ASU

- The registered periods of individual job-search activities ending in the year of 2000 or later are regarded as complete.

XASU

- In the scientific use file the XASU notifications from the 69 authorised municipalities are assumed to be complete in the IEB as of the year 2007 onwards.
- As there is no central allocation of identification numbers, inconsistencies or implausible job-search histories are possible in the IEB and therefore also in the SIAB data.
- Several variables in the Scientific Use File sometimes have only a very low degree of completeness for the XASU. Variables which are affected by this include “nationality”, “reason for notification”, “employment status”. The degree of completeness of these

variables improves over time. In general it is to be taken into account that the degrees of completeness sometimes vary depending on the institution supplying the data.

- It remains unclear whether the institutions – as intended – only report periods in which they themselves are responsible for the jobseekers or whether preceding periods of support from the BA in accordance with SGB III are reported as continuous episodes if the job-search status does not change.
- For a number of institutions 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.

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 Artificial individual ID (persnr)

Variable label	artificial individual ID
Variable name	persnr
Category	identifiers
Origin	BeH, LeH, (X)LHG, (X)ASU
Data type	numerical
Hierarchy	none
Description	<p>The artificial individual ID indicates which observations belong to the same person. Artificial means that it is not possible to infer any of the person’s characteristics or any original identifiers from this individual ID.</p> <p>As there is no uniform individual indicator across the BA, the allocation of (BA) data from different data sources (e.g. employment and benefits) to individuals is not always clear. The formation of the individual identifier which spans all data sources is based on a heuristic developed by the BA.</p> <p>The artificial individual ID from the SIAB 7508 was replaced by an artificial, randomly assigned individual ID in the SIAB-Regional File. The values of this newly generated individual ID range from 1 to 1,515,463.</p>

5.1.2 Establishment counter (bnn)

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

5.2 Period of validity

5.2.1 Original start date of observation (begorig)

Variable label	original start date of observation
Variable name	begorig
Category	period of validity
Origin	BeH, LeH, (X)LHG, (X)ASU
Data type	date
Hierarchy	none
Description	<p>The original start date of the observation corresponds to the original start date of the notification. Due to date adjustments related to the anonymisation of the 'begorig' variable the information in the SIAB-Regional File can differ from the original information in the weakly anonymous version of the SIAB. 'begorig' can differ from the start date of the episodes (begepi) when episode splitting applies (see Section 3.5).</p> <p>Because of the rules of the notification procedure, in BeH observations the starting and ending year are always identical (obligation of the employer to submit annual employment notifications).</p>

5.2.2 Original end date of observation (endorig)

Variable label	original end date of observation
Variable name	endorig
Category	period of validity
Origin	BeH, LeH, (X)LHG, (X)ASU
Data type	date
Hierarchy	none
Description	<p>The original end date of the observation corresponds to the original end date of the notification. Due to date adjustments related to the anonymisation of the 'endorig' variable the information in the SIAB - Regional File can differ from the original information in the weakly anonymous version of the SIAB. 'endorig' can differ from the start date of the episodes (begepi) when episode splitting applies (see Section 3.5).</p> <p>Because of the rules of the notification procedure, in BeH observations the starting and ending year are always identical (obligation of the employer to submit annual employment notifications).</p>

5.2.3 Start date of split episode (begepi)

Variable label	start date of split episode
Variable name	begepi
Category	generated period of validity
Origin	BeH, LeH, (X)LHG,(X)ASU
Data type	date
Hierarchy	none
Description	<p>Start date of the split episode. The term "episode" means the period of an observation. This period can differ from the period of the original information when single episodes are right- or left-censored (see Section 3) and/or episode splitting applies (see Section 3.5).</p>

5.2.4 End date of split episode (endepi)

Variable label	end date of split episode
Variable name	endepi
Category	generated period of validity
Origin	BeH, LeH, (X)LHG,(X)ASU
Data type	date
Hierarchy	none
Description	<p>End date of the split episode. The term "episode" means the period of an observation. This period can differ from the period of the original information when single episodes are right- or left-censored (see Section 3) and/or episode splitting applies (see Section 3.5).</p>

5.3 Personal information

5.3.1 Gender (frau)

Variable label	gender
Variable name	frau
Category	personal variable
Origin	BeH, LeH, (X)LHG,(X)ASU
Data type	numerical
Hierarchy	None
Description	Gender dummy (0 - man, 1 - woman). The gender information is constant per account.
Notes on quality	In the original data it may happen that the gender changes. This is corrected during the data preparation process. Details from data sources regarding benefit payment (LeH, (X)LHG) are given highest priority here.

5.3.2 Year of birth (gebjahr)

Variable label	year of birth
Variable name	gebjahr
Category	personal variables
Origin	BeH, LeH, (X)LHG,(X)ASU
Data type	numerical
Hierarchy	none
Description	The year of birth is constant per account. Observations with a calculated age under 17 and over 62 years of age are not represented in the data (see Section 3.3.3).
Notes on quality	In the original data it may happen that the year of birth changes between the data sources. This is corrected during the data preparation process. The information from the social security number is given highest priority here.

5.3.3 Nationality german (deutsch)

Variable label	nationality german
Variable name	deutsch
Category	personal variables
Origin	BeH, LeH, (X)LHG, (X)ASU
Data type	numerical
Hierarchy	none
Description	This variable distinguishes between German nationality and non-German nationality. Please note that the variable contains valid information only for individuals whose place of work is solely in West Germany throughout the entire employment biography. Valid information on nationality is not included for individuals who have

	<p>been employed in East Germany at least once.</p> <p>If place of work contains only missing values for an individual and hence a distinction between East and West Germany can not be made, nationality information is provided.</p>
--	--

5.3.4 School education and vocational training (bild)

Variable label	school education and vocational training
Variable name	bild
Category	personal variables
Origin	BeH, (X)ASU
Data type	numerical
Hierarchy	none
Description	<p>It must be taken onto account that this variable has a different meaning depending on the data source:</p> <p>1) BeH</p> <p>For BeH observations the variable contains the school education / vocational training reported by the employers in the employment notification procedure. The indicator 'school education and vocational training' is a combined variable which is used to gather information on both the employee's school education and his/her vocational qualifications (the highest qualification gained in each case). Here the school qualifications are to be established first and then the vocational qualifications. Only in the case of the values 5 (degree from a university of applied sciences) and 6 (university degree) are no other vocational qualifications determined (see BA 2005, p. IX.). The following values exist:</p> <ol style="list-style-type: none"> 1 Primary school/lower secondary school/intermediate school leaving certificate or equivalent school education, without a vocational qualification 2 Primary school/lower secondary school/intermediate school leaving certificate or equivalent school education with a vocational qualification 3 With upper secondary school leaving certificate (Abitur), without a vocational qualification 4 With upper secondary school leaving certificate (Abitur) and a vocational qualification 5 Degree from a university of applied sciences (Fachhochschule) 6 University degree <p>It cannot be inferred from the data whether the individual completed school education or not, both cases can occur in the values "1" and "2".</p> <p>Changes in the vocational training status frequently occur at the same time as a change of establishment. This is because the notification data are compiled anew in the new firm. If, for example, an employee has gained a higher qualification via a part-time further training course while still working then this change of status is probably not recorded until s/he joins a new firm. It can generally be assumed that when a person is employed in a firm for a longer period, the personal data that they reported when they joined the firm is simply continued." (Meinken/Koch 2004, p. 63).</p> <p>A method for correcting missing values or temporal inconsistencies in</p>

	<p>the education and training data in the predecessor sample, the IABS, can be found in Fitzenberger et al. (2006) and in Drews (2006). Only the BeH data source is used for this, however, as this was the only data source with information on education and training in the IABS.</p> <p>2) (X)ASU (X)ASU observations contain the completed vocational training at the beginning of the job-search activity. The following values exist:</p> <p>21 no completed vocational training 22 in-firm training 23 external training 24 full-time vocational school (Berufsfachschule) 25 technical college (Fachschule) 26 university of applied sciences (Fachhochschule) 27 university</p>
Notes on quality	<p>“For certain sub-groups there is a large proportion of missing data because this variable is not of importance as regards social security (this affects people in marginal part-time employment, for example).” (Meinken/Koch 2004, p. 63).</p>

5.3.5 School-leaving qualification (schbild)

Variable label	school-leaving qualification
Variable name	schbild
Category	personal variables
Origin	(X)ASU
Data type	numerical
Hierarchy	none
Description	This variable contains the school-leaving qualification held by the individual at the beginning of the period of unemployment or job-search.
Notes on quality	Although the degree of completeness in the XASU increases continuously until 2008, it remains comparatively low (under 50% in some cases). The sources ASU and XASU are aggregated and cannot be distinguished in the factually anonymous dataset.

5.4 Information on employment, benefit receipt and job-search

5.4.1 Reason for notification/ reason for end of benefit receipt/ reason for discontinuation of unemployment benefit II/ reason for deregistration, aggregated (grund_gr)

Variable label	reason for notification/ reason for end of benefit receipt/ reason for discontinuation of unemployment benefit II/ reason for deregistration, aggregated
Variable name	grund_gr
Category	information on employment, benefit receipt and job-search

Origin	BeH, LeH, (X)LHG, (X)ASU
Data type	numerical
Hierarchy	none
Description	<p>1) BeH In BeH observations the 'reason for notification' variable indicates the reason why the employer submitted the employment notification in question to the social security agencies. However, not all of the possible reasons for submitting a notification that may occur in the context of the notification procedure are available in the IEB or the SIAB. For instance, the IEB only includes notifications that have information on earnings (i.e. annual, employment interruption and end of employment notifications), whilst initial registrations are not included as they contain no information on earnings. However, this does not involve a loss of information as the details from a registration are transmitted again with the following annual, employment interruption or end of employment notification.</p> <p>In the SIAB the reasons for submitting employment notifications are encoded according to the regulations of the notification procedure which has been in effect since 01.01.1999 (in accordance with DEÜV).</p> <p>2) LeH In LeH observations the variable indicates the reason for the end of receipt of unemployment benefit, unemployment assistance or maintenance allowance. In contrast, neither the LeH nor the SIAB contains information about the reasons for the start of benefit receipt, as the information for the LeH comes from the notifications submitted by the employment agencies to the health insurance institutions regarding finished periods of benefit receipt.</p> <p>3) (X)LHG The (X)LHG datasets contain the 'reason for discontinuation of Unemployment Benefit II' and indicate the reason why current benefits have been discontinued. The 'reason for discontinuation of Unemployment Benefit II' variable refers to the individual, not to the benefit community. The variable is valid exactly at the end of the observation. If the value of the variable changes, the data record is split. It must be taken into account that the 'reason for discontinuation of Unemployment Benefit II' is a person-related variable. If the Unemployment Benefit II receipt of a different member of the benefit community is discontinued, all observations of the members of the benefit community are split on this date, but the reason for discontinuation of Unemployment Benefit II is only available for the individual whose benefit is discontinued.</p> <p>4) (X)ASU In the case of (X)ASU records the variable contains the deregistration or exit reason. It must be taken into account here that the number of values of the variable was reduced from 26.04.2003 onwards. The old values were recoded to the new values as far as this was possible. Due to data confidentiality, reasons for notification were aggregated within sources during the production of the SIAB-Regional File.</p>

Note on quality	Proportion of valid values (degree of completeness) for the reason for notification in the data sources (X)LHG und (X)ASU					

	quelle_gr	Year				
		2000	2001	2002	2003	2004

	BeH	1.000	1.000	1.000	1.000	1.000
	LeH	1.000	1.000	1.000	0.680	0.695
	(X)ASU	0.794	0.775	0.784	0.769	0.753

	quelle_gr	Year				
		2005	2006	2007	2008	

	BeH	1.000	1.000	1.000	1.000	
	LeH	0.969	1.000	1.000	1.000	
	(X)ASU			0.913	0.862	
	(X)LHG			0.156	0.131	

5.4.2 Daily wage, daily benefit rate, aggregated (tentgelt_gr)

Variable label	daily wage, daily benefit rate
Variable name	tentgelt_gr
Category	information on employment, benefit receipt and job-search
Origin	BeH, LeH
Data type	numerical
Hierarchy	none
Description	<p>1) BeH</p> <p>In BeH observations this variable shows the employee's gross daily wage. It was 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 for the entire period. In the SIAB-Regional File the daily wage is reported only in Euros and in whole numbers for reasons of data confidentiality. Wages below 0.5 Euros were rounded down to 0 Euros.</p> <p>Until 1998 employers in principle only reported the earnings which were subject to social security contributions. Earnings below the marginal part-time income threshold were not reported.</p> <p>Earnings exceeding the upper earnings limit for statutory pension insurance were only reported up to this limit. It must be taken into account here that there are two upper earnings limits in the statutory pension insurance scheme: the earnings limit of the pension insurance for wage and salary earners and that of the miners' pension insurance, which is generally higher. Please note that the SIAB-Regional File does not include information which makes it possible to distinguish which pension insurance institution is responsible for an individual.</p>

Since the inclusion of marginal part-time employees in the employment notification procedure as of 01.04.1999, 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 or the amount of wages paid, though this is likely to be seldom the case given the relevance of this information for social security contributions.

All wages (in whole numbers of Euros) above the upper earnings limit for statutory pension insurance are not reported with their actual value but are replaced by the value of the respective annual upper earnings limit for statutory pension insurance. The marginal part-time income threshold and the upper earnings limit for statutory pension insurance differ from year to year as well as between East and West 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/en.aspx>.

A daily wage reported as 0 Euros and the rounded down wages below 0.5 Euros can be put down to "employment interruption notifications". During these periods the employment relationship continues to exist in legal terms, but without pay. This is the case for periods of illness after the end of continued payment of wages, for periods of maternity leave and for sabbaticals.

2) LeH

For LeH observations the variable shows the daily benefit rate, rounded and converted into Euros in each case. Benefit rates below 0.5 Euros were rounded down to 0 Euros

It must be taken into account that for observations with an original start date prior to 01.01.1998 the daily benefit rate applies to working days, while for observations with an original start date from 01.01.1998 onwards it applies to calendar days.

A daily benefit rate of 0 Euros is reported either due to rounding down extremely low benefit payments below 0.5 Euros or, since 01.01.2005, also in the context of periods of benefit suspension / interruption. If a reason for deregistration / reason of end of benefit / reason for notification is reported for an observation with a daily benefit rate equal to 0, then it is a notification of interruption of benefit payments. In the case of observations that reflect a period of benefit suspension, the entitlement is the same as before the start of the benefit suspension period.

5.4.3 Occupation – activity performed (beruf_gr)

Variable label	occupation – activity performed, aggregated
Variable name	beruf_gr
Category	information on employment, benefit receipt and job-search
Origin	BeH, (X)ASU
Data type	numerical
Hierarchy	none
Description	<p>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.</p> <p>For this the employer encodes the employee's job in accordance with the "Classification of Occupations. Systematic and Alphabetical Directory of Job Titles" (published by the Federal Employment Agency, Nuremberg, 1988), which contains approx. 25,000 job titles. The occupational classification consists of a 3-digit code and comprises about 330 values. If more than one job title with different classification codes applies for one employee, the employer is required to select the job title that best defines the main activity performed (see BA 2005, p. V).</p> <p>For reasons of data confidentiality, the original occupation categories were aggregated into 120 categories.</p> <p>2) (X)ASU Information on the occupation was replaced with a missing value (.z) for (X)ASU data for reasons of data confidentiality.</p>

5.4.4 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
Hierarchy	none
Description	<p>The employee's occupational status during the notification period is reported by the employer in the 'employment details'.</p> <p>The variable "occupational status" first distinguishes between full-time and part-time employees; the decisive factor here is the ratio between the contracted hours and the usual working hours in the establishment. The variable only actually provides details about the occupational status for full-time employees, whilst for part-time employees it only records whether their working hours exceed a certain limit or not. This limit was 20 hours of work per week until 1978, between 1979 and 1987 it was 15 hours per week and since 1988 it has been 18 hours per week.</p> <p>The distinction between blue-collar employees in full-time employment (1,2) and white-collar employees centres solely on the type of pension-insurance institution (Federal Social Insurance Office for Salaried Em-</p>

	<p>ployees – BfA – for white-collar workers, and Land Social Insurance Office – LVA – for blue-collar workers). The category "employees in vocational training" (value 0) covers not only trainees/apprentices, placement workers and interns but also semi-skilled trainees, students at colleges for health occupations and participants in subsidized further vocational training, retraining and induction training.</p> <p>If more than one code is possible for an employee, the employer is required to classify the job according to the activity which is performed most. If this cannot be ascertained clearly, the code of the higher occupational status is to be entered (see BA (2005): p. VI).</p>
--	---

5.4.5 Employment status (erwstat_gr)

Variable label	employment status, aggregated
Variable name	erwstat_gr
Category	information on employment, benefit receipt and job-search
Origin	BeH, LeH, (X)ASU
Data type	numerical
Hierarchy	none
Description	<p>This variable assumes different values with different meanings for each data source.</p> <p>1) BeH</p> <p>For BeH observations the variable contains the person group of the employment notification, which was introduced along with the new notification procedure as of 01.01.1999. It indicates contribution- or benefit-related particularities of the employment relationship which concern at least one social security agency, differentiated according to the institutions submitting the notifications.</p> <p>Employees covered by social security whose employment relationship does not show any particularities are given the code 101. If more than one code applies then the smallest is to be indicated. Employees in marginal part-time employment can be distinguished in the SIAB as the person groups 109 and 209.</p> <p>The person group variable has only been a component of the social security notifications since the introduction of the new notification procedure on 01.01.1999. The person group can be contained in notifications that refer to the years prior to 1999 but were not received until 1999 or later. For notifications which were received before 1999 an attempt is made to allocate the notifications to the person groups on the basis of certain rules and with the aid of the variables 'school education and vocational training', 'occupational status and working hours' and 'occupation' as well as other information. In many cases, however, appropriate allocations are not possible. The obligation to submit employment notifications for people in marginal part-time employment has been in force since 01.04.1999. Before this date the variable contains no data about people in marginal part-time employment.</p> <p>2) LeH</p> <p>For LeH observations the grouped benefit type is reported, i.e. it states whether unemployment benefit, unemployment assistance or maintenance allowance is received or whether contributions to private long-term care insurance are paid by the BA. The values of this variable</p>

	<p>were generated by combining the values of the “type of benefit” variable.</p> <p>3) (X)ASU</p> <p>The (X)ASU contains the job-search status. As the data sources ASU and XASU have been aggregated, the two can no longer be distinguished from one other in the SIAB-Regional File. Only episodes with the status “not unemployed, but seeking work” and “seeking work and simultaneously unemployed” were integrated into the SIAB. Information regarding the job-search status was combined with the equivalent status from the ASU.</p> <p>Recipients of unemployment benefits (Unemployment Benefit I or II) over the age of 58 who receive benefits under the relaxed conditions according to § 428 of SGB III (or §65 para. 4 of SGB II) and individuals aged over 58 who are not benefit recipients and are not willing to be placed in employment in the sense of § 252 para. 8 SGB VI are recorded as individuals seeking advice.</p>
--	---

5.4.6 Employment before the age of 17 (tage_jung)

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

5.4.7 Employment after the age of 62 (tage_alt)

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

5.5 Establishment variables

5.5.1 Economic activity WS 73 – group of economic activity 73 (w73_gen_gr)

Variable label	economic activity WS 73, aggregated, generated from 2003 onwards
Variable name	w73_gen_gr
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	none
Description	<p>This variable is based on the classification of economic activity WS73. WS73 stands for the “Classification of Economic Activities for the Statistics of the Federal Employment Services, edition 1973” (“Klassifikation der Wirtschaftszweige für die Statistik der Bundesanstalt für Arbeit, Ausgabe 1973”). Using a 3-digit code the classification distinguishes between 269 classes of economic activity, whereby the first digit of the code defines the division of economic activity, of a total of 10, and the first two digits together define the particular group of economic activity, of a total of 95.</p> <p>Each establishment may only be assigned one code. If an establishment is active in different economic sectors, the main economic activity must be determined.</p> <p>Information on the WS 73 is only available from 1975 until 2002. The variable reports information on the economic activity of the notifying establishment according the classification of economic activities W73 for this period. From 2003 onwards all values are imputed on the basis of heuristics. For this, valid information from the WS 73 from earlier years is extrapolated into the future. If extrapolation is not possible due to missing values, information is replaced using correspondence tables. For further information on methodology and quality aspects please see Eberle et al. (2011).</p> <p>For data confidentiality reasons the SIAB-Regional File reports economic activity from the WS 73 only in 16 aggregated categories. In regions which did not reach the minimum number of observations required for factual anonymity, the economic activities concerned are anonymised and recoded to missing (.z) (see Appendix).</p>

5.5.2 Economic activity 03 – group of economic activity 03 (w03_gr)

Variable label	economic activity 03, aggregated
Variable name	w03_gr
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	none

Description	<p>This variable indicates the economic activity as a 5-digit code in accordance with the WZ03 classification. WZ03 stands for the "Classification of Economic Activities, Edition 2003" ("Klassifikation der Wirtschaftszweige Ausgabe 2003") of the Federal Statistical Office (eds.).</p> <p>The economic activity in accordance with the 2003 classification has only been available since 2003.</p> <p>Each establishment may only be assigned one code. If an establishment is active in different economic sectors, the main economic activity must be determined.</p> <p>For data confidentiality reasons the SIAB-Regional File reports economic activity from the WZ 03 only in 18 aggregated categories. In regions which did not reach the minimum number of observations required for factual anonymity, the economic activities concerned are anonymised and recoded to missing (.z) (see Appendix).</p>
-------------	--

5.6 Location data

5.6.1 Place of work: region (ao_region)

Variable label	place of work: district / region
Variable name	ao_region
Category	location data
Origin	BeH
Data type	numerical
Hierarchy	district region federal state
Description	<p>German districts were aggregated to 333 district regions (see Section 7.5) with a minimum of at least 100,000 inhabitants on December 31 2008.</p> <p>The weakly anonymous data contain the original district identifier where the employee's establishment is located. The first two digits of the 5-digit district code (Kreisschlüssel) show the code for the federal state (Bundesland) (NUTS 1), positions 1-3 indicate the regional authority (Regierungsbezirk) (NUTS 2), and positions 1-5 show the district authority (Kreis) (NUTS 3). Federal states without a regional authority have a 0 in the third position.</p> <p>In order to guarantee consistent regional allocations across the entire observation period, the information on the district was recoded to the territorial allocation of 31.12.2008, 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 31.12.2008. As the district boundaries have changed over time, cases would occur in which the district code of the location of the establishment would change without the establishment concerned having relocated, if the territorial allocations were not updated.</p>

5.6.2 Commuter status (pendler)

Variable label	commuter status
Variable name	pendler
Category	location data
Origin	BeH, LeH, (X)ASU, (X)LHG
Data type	numerical
Hierarchy	none
Description	<p>The variable 'commuter status' indicates whether an individual's current or most recent place of work corresponds to his or her current place of residence. As a reference for all sources, the current place of work or the place of work from the most recent BeH notification (main employment) is taken as a basis and compared with the current place of residence.</p> <p>The variable contains the following details: 0: the district of the place of residence corresponds to the district of the (last available) place of work (no commuter) 1: the district of the place of residence does not correspond to the district of the (last available) place of work, but both districts belong to the same district region (intra-regional commuter) 2: the district of the place of residence and the district of the (last available) place of work are located in different district regions (inter-regional commuter)</p> <p>The variable is available for: BEH: 1999-2008 LEH: 2003-2008 (X)ASU: 2000-2004 and 2007-2008 (X)LHG: 2007-2008</p> <p>If the place of residence for one source is not available in the original data for an individual's episode, the variable 'commuter status' is recoded with .z. .z is also coded if the individual's place of residence is available in one source but there is no previous BeH observation with a valid place of work that could be used as a reference.</p>

5.7 Generated technical variables

5.7.1 Source of the observation (quelle_gr)

Variable label	source of the observation, aggregated
Variable name	quelle_gr
Category	generated technical variables
Origin	BeH, LeH, (X)ASU, (X)LHG
Data type	numerical
Hierarchy	none
Description	The variable indicates the data source. The original sources ASU and

	<p>XASU as well as LHG and XLHG were each aggregated to one source for data confidentiality reasons.</p> <p>Hence, it is not possible to distinguish between the types of institution. Because of the adjustments made, the designated time periods of the sources differ from those provided in the SIAB 7508. For further information on data preparation see Chapter 3.</p>
--	--

5.7.2 Observation counter per person (spell)

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

5.7.3 Observation counter per episode (level2)

Variable label	observation counter per episode
Variable name	level2
Category	generated technical variables
Origin	BeH, LeH, (X)ASU, (X)LHG
Data type	numerical
Hierarchy	none
Anonymisation	none
Description	<p>A counter that counts parallel observations. The variable is generated during the episode splitting procedure and refers to the split observations.</p> <p>This level counter indicates whether observations overlap temporally within one account. This is the case, for example, when a person has more than one employment relationship at one time or when benefit receipt and marginal part-time employment coincide. The "level2" variable numbers the observations of such an episode, starting with 0. Within the episode the observations are first sorted by data source (BeH, LeH, (X)LHG, (X)ASU). Within the BeH data source non-marginal observations come before marginal ones, then the observations are sorted by daily wages, starting with the highest wage. Within the LeH data source the observations are sorted by type of benefit; within the other data sources the sorting process is random.</p>

5.7.4 Observation counter per episode and source (level1)

Variable label	observation counter per episode and source
Variable name	level1
Category	generated technical variables
Origin	BeH, LeH, (X)ASU, (X)LHG
Data type	numerical
Hierarchy	none
Anonymisation	none
Description	<p>With the aid of this level variable (simultaneous) observations per episode and source are counted. The variable is generated during the episode splitting procedure and refers to the split observations.</p> <p>This level counter indicates whether observations overlap temporally within one account. Level1 differentiates in a different way to Level2, as it counts parallel observations additionally according to the data source from which the observations come. In each case the "level1" variable thus assumes the value 0 for the first observation available per episode and data source. Further parallel observations for the same episode and source are then counted upwards from there.</p> <p>Within the BeH data source non-marginal observations come before marginal ones, then the observations are sorted by daily wages, starting with the highest wage. Within the LeH data source the observations are sorted by type of benefit; within the other data sources the sorting process is random. (For each data source the following applies: $level2 - level1 \equiv const$; for the BeH data source, $const = 0$.)</p>

6 References

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

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

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

Christians, Helga; Wirth, Heike (2009): Mikrozensus-Regionalfile: Hochrechnung, Stichprobenfehler und Anonymisierung. Düsseldorf, Mannheim. URL: http://www.forschungsdatenzentrum.de/bestand/mikrozensus_regionalfile/fdz_1_mz-regionalfile_dokumentation.pdf (Accessed: 01.07.2011).

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

Dorner, Matthias; Heining, Jörg; Jacobebbinghaus, Peter; Seth, Stefan (2010): Sample of Integrated Labour Market Biographies (SIAB) 1975-2008. (FDZ Datenreport, 01/2010 (en)), Nuremberg, 63 p.

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

Drews, Nils (2008): Das Regionalfile der IAB-Beschäftigtenstichprobe 1975-2004 - Handbuch-Version 1.0.2. (FDZ Datenreport, 02/2008 (de)), Nuremberg, 87 p.

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, 21 p.

Fitzenberger, Bernd; Osikominu, Aderonke; Völter, Robert (2006): Imputation rules to improve the education variable in the IAB employment subsample. In: Schmollers Jahrbuch. Journal of Applied Social Science Studies, Vol. 126 (3), pp. 405-436.

Hethey-Maier, Tanja; Seth, Stefan (2010): The Establishment History Panel (BHP) 1975-2008 * Handbook Version 1.0.2. (FDZ Datenreport, 04/2010 (en)), Nuremberg, 80 p.

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

Müller, Walter; Blien, Uwe; Knoche, Peter; Wirth, Heike (1991): Die faktische Anonymität von Mikrodaten; Band 19 der Schriftenreihe Forum der Bundesstatistik herausgegeben vom Statistischen Bundesamt; Stuttgart.

Oberschachtsiek, Dirk; Scioch, Patrycja; Seysen, Christian; Heining, Jörg (2009): Stichprobe der Integrierten Erwerbsbiografien IEBS - Handbuch für die IEBS in der Fassung 2008. (FDZ Datenreport, 03/2009 (de)), Nuremberg, 59 p.

Trappmann, Mark; Christoph, Bernhard; Achatz, Juliane; Wenzig, Claudia; Müller, Gerrit; Gebhardt, Daniel (2009): Design and stratification of PASS - A New Panel Study for Research on Long Term Unemployment. IAB-Discussion Paper 5/2009, 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.

Zimmermann, Ralf; Kaimer, Steffen; Oberschachtsiek, Dirk (2007): Dokumentation des "Scientific Use Files der Integrierten Erwerbsbiographien" (IEBS-SUF V1) Version 1.0. (FDZ Datenreport, 01/2007 (de)), Nuremberg, 62 p.

7 Appendix

7.1 Differences in the anonymisation of key variables in comparison to the IABS-R04

	SIAB-Regional File 7508	IABS-Regional File 04
Information on place of residence	Implied by commuter status (pendler)	Information on place of residence (region) omitted
Years of birth/information on age	Notifications with age from 17 to 62 (via gebjahr) only. Additional variables indicate days of employment before 17 and after 62 years (tage_alt, tage_jung).	Notifications with age from 16 to 62, observations below or above the age limit without valid information on age (gebjahr)
Daily wage/daily benefit rates	Rounded wages and benefit rates; wages above the upper earnings limit for statutory pension insurance replaced by the value of the upper earnings limit for statutory pension insurance (tentgelt_gr)	Truncated daily wages in whole numbers. No valid information on daily benefit rates, values above the upper earnings limit for statutory pension insurance replaced by 999 and those below the marginal part-time income threshold replaced by 998 (tentgelt)

Table A 1: Differences in the anonymisation of key variables in comparison with the IABS-R04

7.2 Additional correction steps in the IABS-Regional file (IABS-R04) in comparison with the SIAB-Regional file (SIAB-R 7508)

Section 1.3 already went into differences between the SIAB, the IABS and the IEBS. This section shows which correction procedures that were conducted for the IABS are no longer carried out for the SIAB. The number of observations affected in the IABS 2004 gives some idea of the significance of the individual steps.

Corrections in the IABS 1975-2004
Observations deleted if reason for notification == 35 (industrial action)
Observations deleted if reason for notification 33 and remuneration == 0
Observations deleted if reason for notification missing (LeH only)

Correction in the case of receipt of multiple benefits

If one observation is completely embedded within another longer one, the longer observation is retained and the shorter one is deleted.

If two observations are equally long, the one with the highest priority is retained.

If both observations have the same level of priority, the one with the highest benefit rate is retained.

In cases where observations partially overlap, the observation that began first is shortened.

Correction of overlaps of one day

All cases which overlap by exactly one day are corrected. The following procedure is used for the correction:

- First observation has earnings greater than zero:
First observation is retained in full, all other observations in the episode are shortened by one day.

- First observation has earnings equal to zero:
If there are only BeH observations in the episode with earnings equal to zero than the first observation is retained. Subsequent observations are shortened by one day. If there are also observations regarding marginal part-time employment or LeH observations in the episode, then the first of these observations is retained in full. All other observations are corrected.

The correction is not applied in the case of observations with an original duration of one day.

Corrections of overlaps of BeH and LeH observations

Here it applies that the BeH observation generally has priority over the LeH observation.

BeH observations with earnings equal to zero and BeH observations which are completely embedded in the period of an LeH observation are exceptions. In these cases only the BeH observation is deleted or corrected.

Episodes in which only BeH observations of marginal part-time employment and LeH observations overlap are not corrected.

If an observation is deleted, however, all other cases of the same episode and the same source, in other words also observations of marginal part-time employment, are also deleted.

Strike correction

Due to a lockout during the strike involving establishments in the automotive industry in Baden-Württemberg and Hesse in June/July 1984 there are employment interruptions. A strike correction is therefore conducted for the individuals concerned. All observations that fall in the gap are deleted.

Completion procedure

In the IABS, information is added to fill gaps if the following conditions are fulfilled:

- a gap between two observations is greater than seven days
- the observation preceding the gap or the last observation in a person's account:
 - describes employment covered by social security (not marginal part-time employment),
 - shows a valid entry for the variable blh_ao_kreis and
 - the reason for submitting the notification indicates employment (not a deregistration)
- the observation after the gap is a BeH observation
- no information is added to annual employment notifications at the end of an account.

The length of the amended period depends on whether the period falls in the old or the new employment notification procedure. If the gap is smaller than the maximum period that can be amended, the period is amended completely.

For all amended observations the earnings, the daily wage and the reason for submitting the notification are set to 0, start and end dates of the split episode are set in accordance with the periods to be filled. The remaining variables are taken over from the observation preceding the gap.

7.3 Overview of date adjustments and spell deletions

	Start and End date observation	Date after censoring	BeH	LeH	(X)ASU	(X)LHG
			1975-2008	1975-2008	2000-2004; 2007-2008	2007-2008
Deletions of spells	Start and end date before 1.1.1975	-	X	X		
	Start and end date between 1.1.2005 and 31.12.2006	-			X	
	Start and end date before 1.1.2000	-			X	
	Start and end date before 1.1.2007	-				X
	Start and end date after 31.12.2008	-		X	X	X
Left-censoring of spells	Start date before 1.1.1975, end date after 1.1.1975	Start date 1.1.1975		X		
	Start date before 1.1.2000, end date after 1.1.2000	Start date 1.1.2000			X	
	Start date between 1.1.2005 and 31.12.2006, end date after 1.1.2007	Start date 1.1.2007			X	X
Right-censoring of spells	Start date before 31.12.2008, end date after 31.12.2008	end date 31.12.2008		X	X	X
	Start date before 31.12.2004, end date between 1.1.2005 and 31.12.2006	end date 31.12.2004			X	

Left- and right-censoring of spells	Start date before 31.12.2004, end date after 31.12.2006	end date spell 1 31.12.2004; end date spell 2 1.1.2007			X	
-------------------------------------	---	--	--	--	---	--

Table A 2: Outline of date corrections and deletions within the preparations of the SIAB-Regional file 7508

7.4 Comparison of the SIAB 7508 and SIAB-Regional File 7508

Basic file: number of individuals SIAB 7508		1659024	100.0%	
1	Deletion of observations in the sources BeH and LeH which end before 1975	1659024	100.0%	0
2	Deletion of observations in the sources LeH, (X)ASU and (X)LHG which start after 2008	1652618	99,6 %	-0,4 %
3	Deletion of ASU observations which end before 2000	1557912	93,9 %	-5,7 %
4	Deletion of LHG, XASU and XLHG observations which end before 2007	1555976	93,7 %	-0,2 %
5	Deletion of ASU observations which start after 2004 and end before 2007	1552083	93,5 %	-0,2 %
6	Deletion of observations in all sources which end before the age of 17 or start after the age of 62	1515463	91,3 %	-2,2 %
Number of individuals SIAB-Regional File 7508		1515463	91,3 %	- 8,7%

Table A 3: Reduction of the number of individuals within the preparations of the SIAB-Regional File

7.5 Recoding tables of aggregated variables in comparison with the weakly anonymous version of the SIAB 7508

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

Table A 4: Recoding of variable "Source of the observation, aggregated" (quelle_gr)

Employment status (erwstat_gr)		
SIAB 7508	SIAB-R 7508 (SUF)	Categories
101, 140, 143	1	Employees liable to social security without special characteristics
102, 141	2	Trainees
109, 209	3	Marginal part-time workers
103, 142	4	Employees in partial retirement
105, 106	5	Interns and student trainees
118, 205	6	Casual workers
112, 119, 120, 201, 203	7	Other employment status
1	11	Unemployment benefit (ALG)
2	12	Unemployment assistance (ALHI)
3	13	Maintenance benefit (UHG)
5	14	Contribution §166b AFG/§207a SGB III
31	21	Unemployed (implying: registered as a job seeker with the BA)
32	22	Illness up to six weeks/incapacitated during unemployment
33	23	Not unemployed, but registered as a job seeker with the BA
34	24	Without status
35	25	Seeking advice

Table A 5: Recoding of variable "Employment status, aggregated" (erwstat_gr)

Reason for notification/ reason for end of benefit receipt/ reason for discontinuation of unemployment benefit II/ reason for deregistration, aggregated (grund_gr)		
SIAB 7508	SIAB-R 7508 (SUF)	Categories
30, 49 prior 1999 additionally: 31,32, 34, 35, 36, 40	1	Deregistration due to end of employment (also to death)
after 1999: 31	2	Deregistration due to change of health insurance company
after 1999: 32	3	Deregistration due to change of contribution group
after 1999: 34, 35	4	Deregistration due to interruption of employment for more than one month (also industrial conflict/dispute)
after 1999: 36	5	Deregistration due to change of payroll accounting system
after 1999: 40	6	Simultaneous registration and deregistration due to end of employment
33, 54, 55, 56, 57, 59	7	Deregistration for other reasons
50	8	Annual notification (all employment relationships that exist as of 31 December of the notification year must be declared)
51, 52, 53	9	Employment interruption notification
2002	10	Begin of employment
2001, 2021, 2039, 2041, 2046, 2048, 2051, 2052, 2053, 2056	11	Entitlement to other benefit
2026, 2034, 2038	12	Other reasons
2008, 2015, 2036, 2040, 2042, 2055, 2057, 2058	13	No availability for the labour market
2060	14	End of measure
2023, 2043, 2045, 2059	15	Period of exclusion/shortfall from benefit
2028	16	Entitlement exhausted
2020	17	End of period for maintenance allowance
2003, 2004, 2006, 2007, 2010, 2011, 2012, 2014, 2016, 2017, 2019, 2022, 2024, 2027, 2029, 2031, 2032, 2033, 2035, 2037, 2044, 2047, 2049, 2050, 2054, 2061, 2062, 2063, 2064	18	Other reasons
5031, 5061	20	Placement in ABM, SAM, BSI
5032, 5033, 5034, 5040	21	Placement with integration contract (also disabled), with settling-in allowance and recruitment subsidy for new businesses
5035, 5036, 5079, 5088, 6012, 6027	22	Other measures of active labour market policy
5039, 5071	23	FbW or German language courses
5029, 5037, 5038, 5066, 6001	24	Employment - self sought
5060, 5062, 5063, 5064, 5081,	25	Employment (>15 hours) - placed by private job

5086, 5098, 6002, 6003, 6022, 6023, 6024		agent or labour administration - found by agency/applicant in job exchange
5030, 5048, 5082, 6005, 6021	26	Placement in short-term employment, Employment (non-honorary) – unknown
5046, 5047, 5067, 6006	27	Self-employed
5053, 5068, 6007, 6015, 6025	28	Military/civilian service, military exercise, soldier
5041, 6008	29	Incapacity, Incapacity to work
5042, 5065, 6004	30	Recall / re-employment at the same employer
5056	31	Continuance of the current employment
5043, 5070	32	Vocational training
5049, 5069, 6016	33	School education, studies
5045, 5077	34	Relocation
5044, 5074, 6018, 6026	35	Lack of co-operation or availability
5052, 5076	36	Withdrawal from working life
5051, 5055, 6028	37	End of need for financial support, temporary discontinuation of fulfilments
5073	38	Non-renewal of notification
5050, 5075, 5087, 6017, 6030, 6031	39	Special regulations
5054, 5078, 5083, 6009, 6010, 6011, 6013, 6014, 6019, 6020, 6029	40	Other reasons
5027	41	Other benefits
3003, 3001, 4004, 4017, 4018	50	Begin of employment
3011, 3012, 3013, 4028, 4007, 4008, 4009	51	Military/civilian service, military exercise
4010, 3005	52	Transitional benefit
4013, 3006	53	School education/studies
4022, 4023, 3007	54	Relocation
4014, 3008	55	Self-denotification/local absence without benefit receipt
4029, 3009, 4006	56	End of benefit receipt/benefit continuance
4011, 4025, 3010, 3017, 4001	57	Retirement pension, basic financial security in old age
4002, 4012, 3016, 3004	58	(temp.) loss of ability to work
4016, 4019, 4020, 4021, 4027, 3015	59	End of need for financial support
4003, 4024, 4026, 4015, 3014, 4005, 3002	60	Other reasons

Table A 6: Recoding of variable “Reason for notification/ reason for end of benefit receipt/ reason for discontinuation of unemployment benefit II/ reason for deregistration, aggregated” (grund_gr)

Occupation – activity performed (beruf_gr)		
SIAB 7508	SIAB-R 7508 (SUF)	Categories
011, 012, 013, 014, 015, 016, 017, 018, 019, 020, 021, 022, 023, 024, 025, 026, 027, 028, 029, 030, 031, 032, 041, 042, 043, 044	1	Farmers until animal keepers and related occupations
051, 052, 053, 054, 055, 056, 057, 058, 059, 060, 061, 062	2	Gardeners, garden workers until forest workers, forest cultivators
071, 072, 073, 074, 075, 076, 077, 078, 079, 080, 081, 082, 083, 084, 085, 086, 087, 088, 089, 090, 091, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112	3	Miners until shaped brick/concrete block makers
121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135	4	Ceramics workers until glass processors, glass finishers
141	5	Chemical plant operatives
142, 143, 144	6	Chemical laboratory workers until vulcanisers
151	7	Plastics processors
161, 162, 163, 164	8	Paper, cellulose makers until other paper products makers
171, 172, 173, 174	9	Type setters, compositors until printers (flat, gravure)
175, 176, 177	10	Special printers, screeners until printers assistants
181, 182, 183, 184	11	Wood preparers until basket and wicker products makers
191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203	12	Iron, metal producers, melters until semi-finished product fettlers and other mould casting occupations
211, 212, 213	13	Sheet metal pressers, drawers, stampers until other metal moulders (non-cutting deformation)
221	14	Turners
222, 223, 224	15	Drillers until borers
225, 226	16	Metal grinders until other metal-cutting occupations
231, 232, 233, 234, 235, 242, 243, 244	17	Metal polishers until metal bonders and other metal connectors
241	18	Welders, oxy-acetylene cutters
251, 252, 263	19	Steel smiths until pipe, tubing fitters
261	20	Sheet metal workers
262	21	Plumbers
270, 271, 272	22	Locksmiths, not specified until sheet metal, plastics fitters
273	23	Engine fitters
274, 275	24	Plant fitters, maintenance fitters until steel struc-

		ture fitters, metal shipbuilders
281	25	Motor vehicle repairers
282, 283, 284	26	Agricultural machinery repairers until precision mechanics
285, 286	27	Other mechanics until watch-, clockmakers
291, 301, 302	28	Toolmakers until precious metal smiths
303, 304, 305, 306	29	Dental technicians until doll makers, model makers, taxidermists
311	30	Electrical fitters, mechanics
312, 313, 315	31	Telecommunications mechanics, craftsmen until radio, sound equipment mechanics
314	32	Electrical appliance fitters
321	33	Electrical appliance, electrical parts assemblers
322	34	Other assemblers
323	35	Metal workers (no further specification)
331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 371, 372, 373, 374, 375, 376, 377, 378	36	Spinners, fibre preparers until skin processing operatives
351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362	37	Cutters until textile finishers
391, 392	38	Bakery goods makers until confectioners (pastry)
401, 402, 403	39	Butchers until fish processing operatives
411, 412	40	Cooks until ready-to-serve meals, fruit, vegetable preservers, preparers
421, 422, 423, 424, , 425, 426, 427, 428, 429, 430, 431, 432, 433	41	Wine coopers until sugar, sweets, ice-cream makers
441, 442	42	Bricklayers until concrete workers
451, 453	43	Carpenters until scaffolders
452	44	Roofers
461, 462	45	Paviors until road makers
463, 464, 465, 466	46	Tracklayers until other civil engineering workers
470, 471, 472	47	Building labourer, general until other building labourers, building assistants, n.e.c.
481, 482	48	Stucco workers, plasterers, rough casters until insulators, proofers
483, 484, 485, 486	49	Tile setters until screed, terrazzo layers
491, 492, 502, 503, 504	50	Room equippers until other wood and sports equipment makers
501	51	Carpenters
511	52	Painters, lacquerers (construction)
512, 513, 514	53	Goods painters, lacquerers until ceramics/glass painters
521	54	Goods examiners, sorters, n.e.c.
522	55	Packagers, goods receivers, despatchers
531	56	Assistants (no further specification)
541, 542, 543, 544, 545, 546	57	Generator machinists until construction machine attendants

547, 548, 549	58	Machine attendants, machinists helpers until machine setters (no further specification)
601	59	Mechanical, motor engineers
602	60	Electrical engineers
603	61	Architects, civil engineers
604, 605, 606, 607	62	Survey engineers until other engineers
611, 612	63	Chemists, chemical engineers until physicists, physics engineers, mathematicians
621	64	Mechanical engineering technicians
622, 623	65	Electrical engineering technicians until building technicians
624, 625, 626, 627	66	Measurement technicians until remaining manufacturing technicians
628	67	Other technicians
629	68	Foremen, master mechanics
631, 632	69	Biological specialists until physical and mathematical specialists
633, 634	70	Chemical laboratory assistants until photo laboratory assistants
635	71	Technical draughtspersons
681	72	Wholesale and retail trade buyers, buyers
682	73	Salespersons
683, 684, 685, 686	74	Publishing house dealers, booksellers until service-station attendants
687, 688	75	Commercial agents, travellers until mobile traders
691, 692	76	Bank specialists until building society specialists
693, 694	77	Health insurance specialists (not social security) until life, property insurance specialists
701	78	Forwarding business dealers
702, 703, 704, 705, 706	79	Tourism specialists until cash collectors, cashiers, ticket sellers, inspectors
711, 712, 713, 715, 716	80	Railway engine drivers until street attendants
714	81	Motor vehicle drivers
721, 722, 723, 724, 725, 726	82	Navigating ships officers until air transport occupations
731, 732, 733, 734	83	Post masters until telephonists
741	84	Warehouse managers, warehousemen
742	85	Transportation equipment drivers
743, 744	86	Stowers, furniture packers until stores/transport workers
751	87	Entrepreneurs, managing directors, divisional managers
752, 753	88	Management consultants, organisers until chartered accountants, tax advisers
761, 762, 763	89	Members of Parliament, Ministers, elected officials until association leaders, officials
771, 772	90	Cost accountants, valuers until accountants
773	91	Cashiers
774	92	Data processing specialists

781	93	Office specialists
782, 783	94	Stenographers, shorthand-typists, typists until data typists
784	95	Office auxiliary workers
791, 792	96	Factory guards, detectives until watchmen, custodians
793, 794	97	Doormen, caretakers until domestic and non-domestic servants
801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814	98	Soldiers, border guards, police officers until judicial enforcers
821, 822, 823	99	Journalists until librarians, archivists, museum specialists
831, 832, 833, 834	100	Musicians until scenery/sign painters
835, 836, 837, 838	101	Artistic and assisting occupations (stage, video and audio) until performers, professional sportsmen, auxiliary artistic occupations
841, 842, 843, 844	102	Physicians until Pharmacists
851, 852	103	Non-medical practitioners until masseurs, physiotherapists and related occupations
853	104	Nurses, midwives
854	105	Nursing assistants
855, 857	106	Dietary assistants, pharmaceutical assistants until medical laboratory assistants
856	107	Medical receptionists
861, 863, 891, 892, 893	108	Social workers, care workers until religious care helpers
862	109	Home wardens, social work teachers
864	110	Nursery teachers, child nurses
871, 872, 873, 874	111	University teachers, lecturers at higher technical schools and academies until technical, vocational, factory instructors
875, 876, 877	112	Music teachers, n.e.c. until other teachers
881, 882, 883	113	Economic and social scientists, statisticians until scientists n.e.c.
901, 902	114	Hairdressers until other body care occupations
911, 912	115	Restaurant, inn, bar keepers, hotel proprietors, catering trade dealers until waiters, stewards
913	116	Others attending on guests
921, 922, 923, 924	117	Housekeeping managers until employees by household cheque procedure
931, 932	118	Laundry workers, pressers until textile cleaners, dyers and dry cleaners
933, 934	119	Household cleaners until glass, buildings cleaners
935, 936, 937	120	Street cleaners, refuse disposers until machinery, container cleaners and related occupations

Table A 7: Recoding of variable "Occupation – activity performed" (beruf_gr)

Economic activity 73 – group of economic activity 73 (w73_gen_gr)			
SIAB 7508	SIAB-R 7508 (SUF)	Industry	Anonymisation of industries in regions
001-031, 040, 050-080	1	Agriculture, energy, mining	3101, 3102, 3356, 6413, 7132, 8221, 8231, 9161, 9175, 9176, 9572, 9573, 9771, 9775
090-110, 130-146, 170-200, 220,221, 400,401, 430-433	2	Basic production, production of goods	1054, 12052
230-240, 260-300	3	Facture of structural metal products, mechanical engineering	
210, 211, 301-379	4	Steel deformation, vehicle construction, equipment construction	3151, 9177
120, 150-162, 380-390, 410-421, 440-530	5	Consumer goods industry	3151
540-581	6	Food and luxury food industry	3102, 5512, 6413, 12054
590-601	7	Main construction industry	
250, 610-616	8	Expansion industry	
620, 621	9	Wholesale trade	9171, 9672
622-625, 850	10	Retail industry	
630-683	11	Transport and communication	
690, 691, 721, 774, 790-830, 851, 861, 862, 863, 865	12	Economic services	
700, 703, 720, 730, 731, 760-773	13	Household services	
860, 864, 900	14	Homes, hospitals, education	
701, 702, 710, 711, 712, 740-758, 781-785	15	(Street)Cleaning, organisations	
722, 780, 840-845, 870-890	16	Public administration, social security	5512

Table A 8: Recoding of variable “Economic activity 73 – group of economic activity 73” (w73_gen_gr)

Economic activity 03 – group of economic activity 03 (w03_gr)			
SIAB 7508	SIAB-R 7508 (SUF)	Industry	Anonymisation of industries in regions
001-014, 040, 041	1	Agriculture, energy, mining	6413
015, 016	2	Food and luxury food industry	3102, 5512, 5916, 12054
017-022, 037	3	Textile and clothing, wood, paper, publishing industry	3158, 13053
023-028, 036	4	Primary production, recycling, other consumer goods	
029-035	5	Mechanical engineering, vehicle construction, precision mechanics	
045	6	Construction industry	
050	7	Automobile trade	
051	8	Wholesale trade, trade negotiations	
052	9	Retail industry	
055	10	Hospitality industry	
056-064	11	Transport and communication	
065-071	12	Credit and insurance industry, estate service, letting	
072-074	13	Economic services, R&D, data processing	
075, 090, 096-098	14	Public administration, public services	
080	15	Education	
085	16	Health, veterinary and social services	
091, 092	17	Religious association, culture, sports	
093-095	18	Other services, household services	

Table A 9: Recoding of variable “Economic activity 03 – group of economic activity 03” (w03_gr)



Legend

- Bundesländer (NUTS1)
- Regions, SIAB-Regional File 7508
- Districts, SIAB 7508 (NUTS 3)

Territorial Allocation 31.12.2008; Cartography: Matthias Dörner

Figure A 1: Regionalisation SIAB-Regional File 7508 (ao_region)

Regionalisation of districts (NUTS 3), territorial allocation: 2008 (ao_region)		
SIAB 7508	SIAB-R 7508 (SUF)	District region
1002	1002	KS Kiel
1003	1003	KS Luebeck
1051	1051	Dithmarschen
1053	1053	Herzogtum Lauenburg
1054	1054	Nordfriesland
1055	1055	Ostholstein
1056	1056	Pinneberg
1057	1057	Ploen
1058, 1004	1058	Rendsburg-Eckernfoerde, KS Neumuenster
1059, 1001	1059	Schleswig-Flensburg, KS Flensburg
1060	1060	Segeberg
1061	1061	Steinburg
1062	1062	Stormarn
2000	2000	KS Hamburg
3101	3101	KS Braunschweig
3102	3102	KS Salzgitter
3151	3151	Gifhorn
3152	3152	Goettingen
3153	3153	Goslar
3154, 3103	3154	Helmstedt, KS Wolfsburg
3155, 3156, 3255	3155	Northeim, Holzminden, Osterode am Harz
3157	3157	Peine
3158	3158	Wolfenbuettel
3241	3241	Region Hannover
3251, 3401	3251	Diepholz, KS Delmenhorst
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
3354, 3360	3354	Luechow-Dannenberg, Uelzen
3355	3355	Lueneburg
3356	3356	Osterholz
3357	3357	Rotenburg (Wuemme)
3358	3358	Soltau-Fallingbostel
3359	3359	Stade
3361	3361	Verden
3403	3403	KS Oldenburg

3404	3404	KS Osnabrueck
3451	3451	Ammerland
3452, 3402	3452	Aurich, KS Emden
3453	3453	Cloppenburg
3454	3454	Emsland
3455, 3405, 3462	3455	Friesland, Wittmund, KS Wilhelmshaven
3456	3456	Grafschaft Bentheim
3457	3457	Leer
3458	3458	Oldenburg
3459	3459	Osnabrueck
3460	3460	Vechta
4011	4011	KS Bremen
4012	4012	KS Bremerhaven
5111	5111	KS Duesseldorf
5112	5112	KS Duisburg
5113	5113	KS Essen
5114	5114	KS Krefeld
5116	5116	KS Moenchengladbach
5117	5117	KS Muelheim an der Ruhr
5119	5119	Ks Oberhausen
5120	5120	Ks Remscheid
5122	5122	KS Solingen
5124	5124	KS Wuppertal
5154	5154	Kleve
5158	5158	Mettmann
5162	5162	Rhein-Kreis Neuss
5166	5166	Viersen
5170	5170	Wesel
5313	5313	KS Aachen
5314	5314	KS Bonn
5315	5315	KS Koeln
5316	5316	KS Leverkusen
5354	5354	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	KS Bottrop
5513	5513	KS Gelsenkirchen
5515	5515	KS Muenster
5554	5554	Borken

5558	5558	Coesfeld
5562	5562	Recklinghausen
5566	5566	Steinfurt
5570	5570	Warendorf
5711	5711	KS Bielefeld
5754	5754	Guetersloh
5758	5758	Herford
5762	5762	Hoexter
5766	5766	Lippe
5770	5770	Minden-Luebbecke
5774	5774	Paderborn
5911	5911	KS Bochum
5913	5913	KS Dortmund
5914	5914	KS Hagen
5915	5915	KS Hamm
5916	5916	KS Herne
5954	5954	Ennepe-Ruhr-Kreis
5958	5958	Hochsauerlandkreis
5962	5962	Maerkischer Kreis
5966	5966	Olpe
5970	5970	Siegen-Wittgenstein
5974	5974	Soest
5978	5978	Unna
6411	6411	KS Darmstadt
6412	6412	KS Frankfurt am Main
6413	6413	KS Offenbach am Main
6414	6414	KS Wiesbaden
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	KS Kassel
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	KS Koblenz
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
7141	7141	Rhein-Lahn-Kreis
7143	7143	Westerwaldkreis
7211	7211	KS Trier
7231	7231	Bernkastel-Wittlich
7233, 7232	7233	Daun, Bitburg-Pruem
7235	7235	Trier-Saarburg
7314	7314	KS Ludwigshafen am Rhein
7315	7315	Ks Mainz
7331, 7319	7331	Alzey-Worms, KS Worms
7332, 7316	7332	Bad Duerkheim, KS Neustadt an der Weinstraße
7333, 7336	7333	Donnersbergkreis, Kusel
7334	7334	Germersheim
7335, 7312	7335	Kaiserslautern, KS Kaiserslautern
7337, 7313	7337	Suedliche Weinstrasse, KS Landau in der Pfalz
7338, 7311, 7318	7338	Rhein-Pfalz-Kreis, KS Speyer, KS Frankenthal (Pfalz)
7339	7339	Mainz-Bingen
7340, 7320, 7317	7340	Suedwestpfalz, KS Zweibruecken, KS Pirmasens
8111	8111	KS Stuttgart
8115	8115	Boeblingen
8116	8116	Esslingen
8117	8117	Goeppingen
8118	8118	Ludwigsburg
8119	8119	Rems-Murr-Kreis
8121	8121	KS Heilbronn
8125	8125	Heilbronn
8126	8126	Hohenlohekreis
8127	8127	Schwaebisch Hall
8128	8128	Main-Tauber-Kreis
8135	8135	Heidenheim
8136	8136	Ostalbkreis
8212	8212	KS Karlsruhe
8215	8215	Karlsruhe

8216, 8211	8216	Rastatt, KS Baden-Baden
8221	8221	KS Heidelberg
8222	8222	KS Mannheim
8225	8225	Neckar-Odenwald-Kreis
8226	8226	Rhein-Neckar-Kreis
8231	8231	KS Pforzheim
8235	8235	Calw
8236	8236	Enzkreis
8237	8237	Freudenstadt
8311	8311	KS Freiburg im Breisgau
8315	8315	Breisgau-Hochschwarzwald
8316	8316	Emmendingen
8317	8317	Ortenaukreis
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	KS Ulm
8425	8425	Alb-Donau-Kreis
8426	8426	Biberach
8435	8435	Bodenseekreis
8436	8436	Ravensburg
8437	8437	Sigmaringen
9161	9161	KS Ingolstadt
9162	9162	KS Muenchen
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, KS Rosenheim

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, KS Landshut
9275, 9262	9275	Passau, KS Passau
9277, 9279	9277	Rottal-Inn, Dingolfing-Landau
9278, 9263	9278	Straubing-Bogen, KS Straubing
9362	9362	KS Regensburg
9371, 9361	9371	Amberg-Sulzbach, KS Amberg
9372	9372	Cham
9373	9373	Neumarkt i.d.Opf.
9374, 9363, 9377	9374	Neustadt a.d. Waldnaab, Tirschenreuth, KS Weiden i.d.Opf.
9375	9375	Regensburg
9376	9376	Schwandorf
9471, 9461	9471	Bamberg, KS Bamberg
9472, 9462, 9477	9472	Bayreuth, KS Bayreuth, Kulmbach
9473, 9463	9473	Coburg, KS Coburg
9474	9474	Forchheim
9475, 9464, 9479	9475	Hof, KS Hof, Wunsiedel i.Fichtelgebirge
9478, 9476	9478	Lichtenfels, Kronach
9562	9562	KS Erlangen
9563	9563	KS Fuerth
9564	9564	KS Nuernberg
9571, 9561, 9577, 9575	9571	Ansbach, KS Ansbach, Neustadt a.d.Aisch-Bad Windsheim, Weißenburg-Gunzenhausen
9572	9572	Erlangen-Hoechstadt
9573	9573	Fuerth
9574	9574	Nuernberger Land
9576, 9565	9576	Roth, KS Schwabach
9663	9663	KS Wuerzburg
9671, 9661	9671	Aschaffenburg, KS Aschaffenburg
9672	9672	Bad Kissingen
9674, 9673	9674	Haßberge, Rhoen-Grabfeld
9676	9676	Miltenberg
9677	9677	Main-Spessart
9678, 9662	9678	Schweinfurt, KS Schweinfurt
9679, 9675	9679	Wuerzburg, Kitzingen
9761	9761	KS Augsburg
9771	9771	Aichach-Friedberg
9772	9772	Augsburg
9774	9774	Guenzburg

9775	9775	Neu-Ulm
9777, 9762	9777	Ostallgaeu, KS Kaufbeuren
9778, 9764	9778	Unteralldgaeu, KS Memmingen
9779, 9773	9779	Donau-Ries, Dillingen a.d.Donau
9780, 9763, 9776	9780	Oberallgaeu, KS Kempten (Allgaeu), Lindau (Bodensee)
10041	10041	Stadtverband Saarbruecken
10042, 10046	10042	Merzig-Wadern, St. Wendel
10043	10043	Neunkirchen
10044	10044	Saarlouis
10045	10045	Saarpfalz-Kreis
11000, 11100, 11200	11000	KS Berlin (Ost und West)
12052	12052	KS Cottbus
12054	12054	KS Potsdam
12060	12060	Barnim
12061	12061	Dahme-Spreewald
12062	12062	Elbe-Elster
12063	12063	Havelland
12064	12064	Maerkisch-Oderland
12065	12065	Oberhavel
12066	12066	Oberspreewald-Lausitz
12067, 12053	12067	Oder-Spree, KS Frankfurt (Oder)
12068, 12070	12068	Ostprignitz-Ruppin, Prignitz
12069, 12051	12069	Potsdam-Mittelmark, KS Brandenburg an der Havel
12071	12071	Spree-Neisse
12072	12072	Teltow-Flaeming
12073	12073	Uckermark
13003	13003	KS Rostock
13051	13051	Bad Doberan
13053	13053	Guestrow
13054	13054	Ludwigslust
13055, 13002	13055	Mecklenburg-Strelitz, KS Neubrandenburg
13056, 13052	13056	Mueritz, Demmin
13057, 13005, 13061	13057	Nordvorpommern, KS Stralsund, Ruegen
13058, 13006	13058	Nordwestmecklenburg, KS Wismar
13059, 13001, 13062	13059	Ostvorpommern, KS Greifswald, Uecker-Randow
13060, 13004	13060	Parchim, KS Schwerin
14511	14511	KS Chemnitz
14521	14521	Erzgebirgskreis
14522	14522	Mittelsachsen
14523	14523	Vogtlandkreis
14524	14524	Zwickau
14612	14612	KS Dresden
14625	14625	Bautzen

14626	14626	Goerlitz
14627	14627	Meissen
14628	14628	Saechs. Schweiz-Osterzgebirge
14713	14713	KS Leipzig
14729	14729	Leipzig
14730	14730	Nordsachsen
15002	15002	KS Halle/Saale
15003	15003	KS Magdeburg
15082, 15001	15082	Anhalt-Bitterfeld, Dessau-Rosslau
15083, 15086	15083	Boerde, Jerichower Land
15084	15084	Burgenland
15085	15085	Harz
15087	15087	Mansfeld-Suedharz
15088	15088	Saalekreis
15089	15089	Salzland
15090, 15081	15090	Altmarkkreis Salzwedel, Stendal
15091	15091	Wittenberg
16051	16051	KS Erfurt
16052	16052	KS Gera
16061	16061	Eichsfeld
16062, 16065	16062	Nordhausen, Kyffhaeuserkreis
16063, 16056	16063	Wartburgkreis, KS Eisenach
16064	16064	Unstrut-Hainich-Kreis
16066, 16054	16066	Schmalkalden-Meiningen, KS Suhl
16067, 16068	16067	Gotha, Soemmerda
16070	16070	Ilm-Kreis
16071, 16055	16071	Weimarer Land, KS Weimar
16072, 16069	16072	Sonneberg, Hildburghausen
16073	16073	Saalfeld-Rudolstadt
16074, 16053	16074	Saale-Holzland-Kreis, KS Jena
16075, 16076	16075	Saale-Orla-Kreis, Greiz
16077	16077	Altenburger Land

Table A 10: Regionalisation SIAB-Regional file 7508 (ao_region)

7.6 Frequency tables

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

8 List of abbreviations

ALG	Arbeitslosengeld	unemployment benefit
ARGE	Arbeitsgemeinschaft	co-operation of employment agencies and municipalities
ASU	Arbeitsuchenden-Historik	Jobseeker History
BA	Bundesagentur für Arbeit	Federal Employment Agency
BeH	Beschäftigten-Historik	Employee History
BfA	Bundesversicherungsanstalt für Angestellte	Federal Social Insurance Office for Salaried Employees
BHP	Betriebs-Historik-Panel	Establishment History Panel
DEÜV	Verordnung über die Erfassung und Übermittlung von Daten für die Träger der Sozialversicherung – Datenerfassungs- und – übermittlungsverordnung	Data Collection and Transmission Regulation - regulation on the collection and transmission of data for the social security agencies
DEVO	Zweite VO über die Erfassung von Daten für die Träger der Sozialversicherung und für die BA – Datenerfassungs-Verordnung	Data Collection Regulation - second regulation on the collection of data for the social security agencies and for the Federal Employment Agency
DÜVO	Zweite VO über die Datenübermittlung auf maschinell verwertbaren Datenträgern im Bereich der Sozialversicherung und der BA – Datenübermittlungs-Verordnung	Data Transmission Regulation - second regulation on the transfer of data on machine-readable data media in the field of social security and the BA
FDZ	Forschungsdatenzentrum	Research Data Centre
IAB	Institut für Arbeitsmarkt- und Berufsforschung	Institute for Employment Research
IABS	IAB-Beschäftigtenstichprobe	IAB Employment Samples
IEB	Integrierte Erwerbsbiographien	Integrated Employment Biographies
IEBS	Stichprobe der Integrierten Erwerbsbiographien	Integrated Employment Biographies Sample
LeH	Leistungsempfänger-Historik	Benefit Recipient History
LHG	Leistungs-Historik Grundsicherung	Unemployment Benefit II Recipient History
LVA	Landesversicherungsanstalt	Land Social Insurance Office
MTH	Maßnahmeteilnehmer-Historik	Participants-in-Measures History File
NUTS	Systematik der Gebietseinheiten für die Statistik der Europäischen Union	Nomenclature of Territorial Units for Statistics of the European Union
SGB	Sozialgesetzbuch	(German) Social Code Book
SIAB	Stichprobe der Integrierten Arbeitsmarktbiographien	Sample of Integrated Labour Market Biographies
SUF	Scientific Use File	scientific use file
XASU	Arbeitsuchenden-Historik aus XSozial-BA-SGB II	Jobseeker History from XSozial-BA-SGB II
XLHG	Leistungsempfänger-Historik Grundsicherung aus XSozial-BA-SGB II	Unemployment Benefit II Recipient History from XSozial-BA-SGB II

Imprint

FDZ-Datenreport 07/2011 (EN)

Publisher

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

Editorial staff

Stefan Bender, Iris Dieterich

Technical production

Iris Dieterich

All rights reserved

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

Download

http://doku.iab.de/fdz/reporte/2011/DR_07-11_EN.pdf

Internet

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

Corresponding author:

Matthias Dorner
Research Data Centre (FDZ) of the Federal
Employment Agency at the Institute for
Employment Research (IAB),
Regensburger Str. 104,
D - 90478 Nuremberg
Telefon: +49 (0) 911 / 179-8895
E-Mail: matthias.dorner@iab.de