

Research Data Centre (FDZ)  
of the German Federal  
Employment Agency (BA)  
at the Institute for  
Employment Research (IAB)

FDZ

# FDZ-Datenreport

Documentation of labour market data

01/2013

EN

## Sample of Integrated Labour Market Biographies (SIAB) 1975 - 2010

Philipp vom Berge,  
Marion König,  
Stefan Seth



Bundesagentur für Arbeit

# Sample of Integrated Labour Market Biographies (SIAB) 1975 - 2010

Philipp vom Berge, Marion König, Stefan Seth  
(Institute for Employment Research - Institut für Arbeitsmarkt- und Berufsforschung)

Documentation: Version 1 (SIAB 7510\_EN\_v1\_dok1)

Data: Version 1 (SIAB 7510 v1)

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

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

## Contents

<b>1</b>	<b>INTRODUCTION AND OUTLINE .....</b>	<b>8</b>
1.1	Introduction .....	8
1.2	Data use .....	9
1.2.1	Data access and data management .....	9
1.2.2	File names in the SIAB 7510 .....	11
1.3	Changes as compared to SIAB 7508 .....	11
1.4	Outline.....	12
1.5	List of variables .....	14
1.6	Volume structure .....	17
<b>2</b>	<b>DATA SOURCES .....</b>	<b>18</b>
2.1	Employee History (BeH).....	18
2.2	Benefit Recipient History (LeH).....	19
2.3	Unemployment Benefit II Recipient History (LHG / XLHG).....	20
2.4	Jobseeker History (ASU / XASU) .....	21
<b>3</b>	<b>DATA PREPARATION AND SAMPLING PROCEDURE.....</b>	<b>22</b>
3.1	Corrections and validation procedures.....	22
3.1.1	Employee History (BeH) .....	22
3.1.2	Benefit Recipient History (LeH) .....	22
3.1.3	Unemployment Benefit II Recipient History (LHG / XLHG) .....	23
3.1.4	Jobseeker History (ASU / XASU) .....	23
3.1.5	SGB II anonymisation .....	24
3.2	Episode splitting.....	24
3.3	Sampling procedure .....	25
3.4	Missing values .....	26
<b>4</b>	<b>DATA QUALITY AND PROBLEMS .....</b>	<b>26</b>
4.1	Entire IEB.....	26
4.2	Employee History (BeH).....	27
4.3	Benefit Recipient History (LeH).....	29
4.4	Unemployment Benefit II Recipient History (LHG / XLHG).....	29
4.5	Jobseeker History (ASU / XASU) .....	30
4.5.1	ASU.....	30

4.5.2	XASU .....	31
-------	------------	----

## **5 DESCRIPTION OF VARIABLES..... 31**

<b>5.1</b>	<b>Identifiers.....</b>	<b>31</b>
5.1.1	Artificial individual ID (persnr) .....	31
5.1.2	Artificial establishment number (betnr) .....	32
<b>5.2</b>	<b>Period of validity.....</b>	<b>33</b>
5.2.1	Original start date of observation (begorig) .....	33
5.2.2	Original end date of observation (endorig) .....	33
5.2.3	Start date of split episode (begepi) .....	33
5.2.4	End date of split episode (endeipi) .....	34
<b>5.3</b>	<b>Personal information.....</b>	<b>34</b>
5.3.1	Gender (frau) .....	34
5.3.2	Year of birth (gebjahr) .....	34
5.3.3	Nationality (nation) .....	35
5.3.4	Nationality, aggregated (nation_gr) .....	35
5.3.5	Marital status (famst) .....	35
5.3.6	Number of children (kind) .....	36
5.3.7	School education and vocational training (bild) .....	36
5.3.8	School-leaving qualification (schbild) .....	38
5.3.9	Skills level (quali) .....	38
5.3.10	Severe disability status (schweb) .....	39
5.3.11	Health problems (gesund_ein) .....	39
<b>5.4</b>	<b>Information on employment, benefit receipt and job search.....</b>	<b>40</b>
5.4.1	Reason for notification / reason for end of benefit receipt / reason for discontinuation of unemployment benefit II / reason for deregistration (grund) .....	40
5.4.2	Daily wage, daily benefit rate (tengtelt) .....	41
5.4.3	Transition zone (gleitz) .....	42
5.4.4	Occupation – activity performed / last activity (beruf) .....	43
5.4.5	Occupational status and working hours (stib) .....	43
5.4.6	Employment status (erwstat) .....	44
5.4.7	Employment status prior to job search (estatvor) .....	45
5.4.8	Type of benefit (lart) .....	46
5.4.9	BA client group (kundengr) .....	46
5.4.10	Type of termination of last job (art_kuend) .....	46
5.4.11	Desired working hours of the job sought (arbzeit) .....	47
5.4.12	Duration of remaining entitlement to unemployment benefit (restanspruch) .....	47
5.4.13	Type of institution (traeger) .....	47
<b>5.5</b>	<b>Establishment variables.....</b>	<b>48</b>
5.5.1	Economic activity 73 (w73_3) .....	48
5.5.2	Economic activity 93, 5-digit code (w93_5) .....	49
5.5.3	Economic activity 93, 3-digit code (w93_3) .....	49
5.5.4	Economic activity 03, 5-digit code (w03_5) .....	50
5.5.5	Economic activity 03, 3-digit code (w03_3) .....	50
5.5.6	Economic activity 08, 5-digit code (w08_5) .....	51
5.5.7	Economic activity 08, 3-digit code (w08_3) .....	51
5.5.8	Economic activity 73 generated – completed by extrapolation / imputation (w73_3_gen) .....	51
5.5.9	Economic activity 73 generated – type of completion (group_w73_3) .....	52
5.5.10	Economic activity 93 generated – completed by extrapolation / imputation (w93_3_gen) .....	52
5.5.11	Economic activity 93 generated – type of completion (group_w93_3) .....	53
5.5.12	Year of first appearance of establishment number (grd_jahr) .....	53
5.5.13	First appearance of establishment number (grd_dat) .....	53
5.5.14	Year of last appearance of establishment number (lzt_jahr) .....	54

5.5.15	Last appearance of establishment number (lzt_dat) .....	54
5.5.16	Total number of employees (az_ges) .....	55
5.5.17	Number of full-time employees (az_ges_vz) .....	55
5.5.18	Number of employees in marginal part-time employment (az_gf) .....	55
5.5.19	Median gross wage of all full-time employees (te_med) .....	56
<b>5.6</b>	<b>Location data .....</b>	<b>56</b>
5.6.1	Place of residence: district (Kreis) (wo_kreis) .....	56
5.6.2	Place of residence: employment agency (wo_aa).....	57
5.6.3	Place of residence: regional directorate (wo_rd) .....	58
5.6.4	Place of residence: federal state (Bundesland) (wo_bula).....	58
5.6.5	Place of work: district (Kreis) (ao_kreis) .....	59
5.6.6	Place of work: employment agency (ao_aa) .....	59
5.6.7	Place of work: regional directorate (ao_rd).....	60
5.6.8	Place of work: federal state (Bundesland) (ao_bula) .....	60
<b>5.7</b>	<b>Generated technical variables.....</b>	<b>60</b>
5.7.1	Source of the observation (quelle).....	60
5.7.2	Observation counter per person (spell) .....	61
5.7.3	Year (jahr).....	61
<b>6</b>	<b>REFERENCES .....</b>	<b>62</b>
<b>7</b>	<b>APPENDIX.....</b>	<b>63</b>
7.1	Frequency tables .....	63
<b>8</b>	<b>LIST OF ABBREVIATIONS.....</b>	<b>64</b>

## **Abstract**

This Datenreport describes the Sample of Integrated Labour Market Biographies (SIAB) 1975- 2010.

**Keywords:** German administrative micro data, labour market data, data manual

We would like to thank our colleagues in the Research Data Centre (FDZ) who were involved in the working group on the Sample of Integrated Labour Market Biographies as well as the ITM division of the Institute for Employment Research for their cooperation and support. This Datenreport is a revision of the FDZ-Datenreport 01/2010 (authors: Dorner, Heining, Jacobebbinghaus, Seth). Additionally, individual passages from IAB-internal data documentations by IAB ITM were adopted.

## **Data availability**

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

# 1 Introduction and outline

## 1.1 Introduction

The Sample of Integrated Labour Market Biographies (Stichprobe der integrierten Arbeitsmarktbiographien - SIAB) is a 2 % sample of the population of the Integrated Employment Biographies (IEB) of the Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung - IAB). The IEB comprises all individuals who showed one of the following statuses at least once during the observation period:

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

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

Whilst the Employee History (Beschäftigtenhistorik - BeH) is the origin of the information on employment subject to social security and marginal part-time employment, the receipt of benefits in accordance with Social Code Book III (SGB III) and Social Code Book II (SGB II) is recorded in the Benefit Recipient History (Leistungsempfängerhistorik - LeH) and the Unemployment Benefit II Recipient History (Leistungshistoriken Grundsicherung - LHG and XLHG). The Jobseeker Histories (Arbeitsuchenden-Historiken - ASU and XASU) are the data source for the periods of job search recorded by the BA, whilst participation in employment and training measures is recorded in the Participation-in-Measures History File (Maßnahmeteilnahmehistorik - MTH).

With the exception of participation in employment and training measures from the MTH, all of the data sources included in the IEB are also contained in the SIAB.

The SIAB is produced at the Research Data Centre (FDZ) of the Federal Employment Agency (BA) at the IAB.

The SIAB covers the employment histories of 1,639,325 individuals, and their employment biographies are documented in a total of 45,793,010 lines of data.

This Datenreport describes the variables of the weakly anonymous version of the SIAB, which largely comprises the original data, i.e. data which have not undergone anonymisation processes. In order to protect the anonymity of the data subjects, some variables are classi-



fied as particularly sensitive and are only disclosed on submission of a special application (see Chapter 1.2).

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

## **1.2 Data use**

### **1.2.1 Data access and data management**

The SIAB data in the weakly anonymous version may only be analysed in the context of a research visit at the FDZ and subsequent remote data access.

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

The SIAB data which include German as well as English labels<sup>1</sup> have a modular structure and are stored in several files. One module, which is henceforth called the Individual File, contains identifiers (artificial individual IDs and establishment numbers), the personal variables, the information on employment, benefit receipt and job search activity, the variables regarding place of residence, and technical variables. A second module, the Basis Establishment File, contains the establishment number, the year, and variables regarding the place of work and economic activities in aggregate form as well as other establishment information as of the reference date of 30 June.

Establishment variables and individual variables are therefore organised separately, which makes the structure of the data clear and saves storage space in the data management system (see Figure 1). The Basic Establishment File and the Individual File are linked via the establishment number and the year of the dataset. The variables which are marked with a “\*” in the List of variables (see p. 14 et seq.) are contained in the Basic Establishment File. Please note that the variable which is necessary for linking the two modules, “establishment number”, is available in both the Individual File and the Basic Establishment File.

---

<sup>1</sup> With the Stata command `label language en / label language de` labels can be switched to English or German.

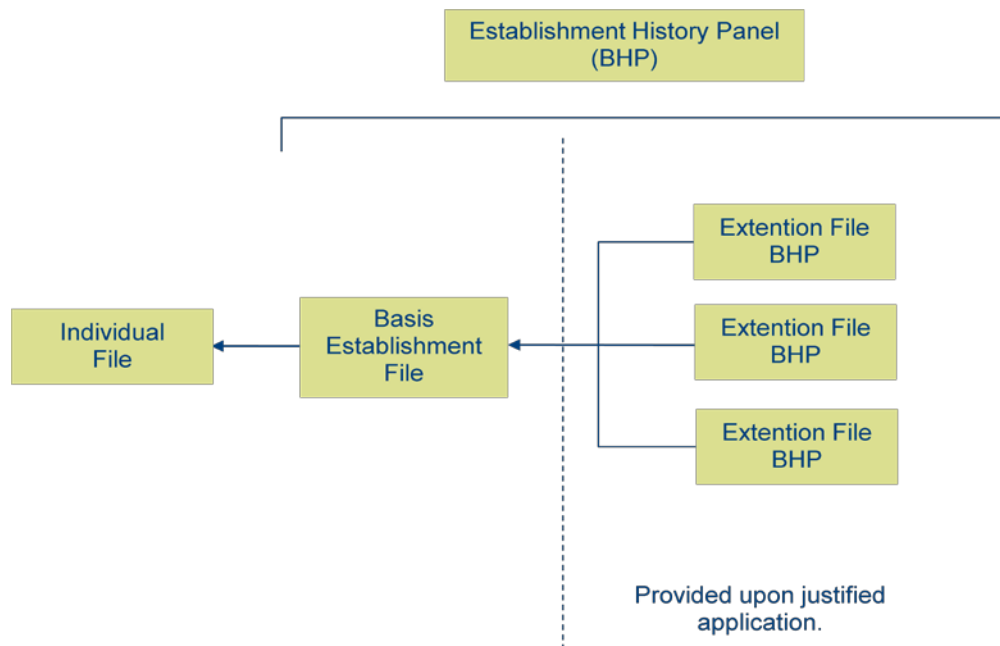


Figure 1: Data storage of SIAB data

Certain variables which make it possible to identify individuals or establishments are only disclosed in their original form if this is necessary for the analysis objective and is justified explicitly in the application for data access. The variables which are particularly sensitive from the viewpoint of data protection legislation are:

#### Individual File:

- nationality (nation)
- health problems (gesund\_ein)
- severe disability status (schweb)
- place of residence: employment agency area (wo\_aa)
- place of residence: district (Kreis) (wo\_kreis)
- place of work: employment agency (ao\_aa)

#### Basis Establishment File:

- place of work: district (Kreis) (ao\_kreis)
- economic activity 93 – sub-class of economic activity (five-digit code) (w93\_5)
- economic activity 03 - sub-class of economic activity (five-digit code) (w03\_5)
- economic activity 08 - sub-class of economic activity (five-digit code) (w08\_5)
- Date of first appearance (grd\_dat)
- Date of last appearance (lzt\_dat)

Following a justified application further establishment variables can be provided from the Establishment History Panel (Betriebs-Historik-Panel - BHP) (see figure 1). Further information on the BHP can be found under <http://fdz.iab.de/>.

### 1.2.2 File names in the SIAB 7510

#### Individual File

SIAB\_7510\_v1.dta

#### Basis Establishment File

SIAB\_7510\_v1\_bhp\_basis\_v1.dta

#### Extension Files BHP

Core dataset with variable blocks

SIAB\_7510\_v1\_bhp\_v1\_####.dta, #### = 1975 - 2010

Worker flows

SIAB\_7510\_v1\_bhp\_inflow\_v1.dta

SIAB\_7510\_v1\_bhp\_outflow\_v1.dta

Establishment dynamics

SIAB\_7510\_v1\_bhp\_ein\_v1.dta

SIAB\_7510\_v1\_bhp\_aus\_v1.dta

### 1.3 Changes as compared to SIAB 7508

#### *Corrected territorial allocation*

In contrast to SIAB 7508, in which only the BeH and LeH sources included the corrected territorial allocation, all locations stated in SIAB 7510 incorporate the territorial allocation as of 31 December 2010.

#### *Characteristics*

In SIAB 7508, the marital status (famst) and number of children (kind) characteristics included only information from the ASU source. These characteristics have now been complemented by information from the XASU, LHG, and XLHG sources.

The characteristics ao\_aa and ao\_rd are now also filled in for the LeH source.

The "willingness to seek work throughout Germany" (mobil) characteristic was no longer included in the SIAB 7510 variables, since it was only filled in until 2006 due to a change of the notification procedure when looking for a job.

The gesund\_ein variable was changed for the same reason. The "health problems - impact on placement" (gesund\_ein) characteristic now shows health problems if these affect the placement.

The level1 and level2 variables were no longer included in the variables of SIAB 7510 for data economy reasons. The variables can be generated by using the following (Stata) statements, if needed:

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

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

During the observation period, the economic sector classifications changed multiple times, so analyses over a longer period of time are more difficult. The FDZ therefore developed methods to create time-consistent characteristics from the different economic sector classifications. A detailed description can be found in Eberle et al. (2011). The economic sectors generated in this way have now been included in the Basic Establishment File by default.

#### *SGB II anonymisation*

A new section on SGB II anonymisation was included (see Chapter 3.1.5).

## 1.4 Outline

Current data version	Sample of Integrated Labour Market Biographies (SIAB) 1975 - 2010
Type and volume of the data	Weakly anonymous, 2 % random sample of individual accounts drawn from the Integrated Employment Biographies (IEB) of the IAB.
Outline of the contents	<p><b>Employee History (Beschäftigten-Historik - BeH):</b> 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><b>Benefit Recipient History (Leistungsempfänger-Historik - LeH):</b> 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><b>Unemployment Benefit II Recipient History drawn from A2LL (Leistungs-Historik-Grundsicherung aus A2LL - LHG):</b> Data on individuals in receipt of basic social security benefits in accordance with Social Code Book II (SGB II) (cooperation of employment agencies and municipalities (ARGE), separated responsibilities for the implementation of SGB II, joint facilities)</p> <p><b>Unemployment Benefit II Recipient History drawn from XSozial-BA-SGB II (XLHG):</b> Data on individuals in receipt of basic social security benefits in accordance with SGB II (authorised municipalities and separated responsibilities)</p> <p><b>Jobseeker History (Arbeitsuchenden-Historik - ASU):</b> Information on job search activity</p> <p><b>Jobseeker History drawn from XSozial-BA-SGB II (XASU):</b> Information on job search activity reported via the transmission</p>

	<p>standard XSozial-BA-SGBII to the BA by authorised municipalities.</p> <p><b>Sensitive variables:</b></p> <p>Nationality (nation)</p> <p>Health problems (gesund_ein)</p> <p>Severe disability status (schweb)</p> <p>Place of residence: employment agency (wo_aa)</p> <p>Place of residence: district (Kreis) (wo_kreis)</p> <p>Place of work: employment agency (ao_aa)</p> <p>Place of work: district (Kreis) (ao_kreis)</p> <p>Economic activity - sub-class of economic activity (five-digit code) (w93_5)</p> <p>Economic activity - sub-class of economic activity (five-digit code) (w03_5)</p> <p>Economic activity - sub-class of economic activity (five-digit code) (w08_5)</p> <p>First appearance of establishment (grd_dat)</p> <p>Last appearance of establishment (lzt_dat)</p>
Variables with regional relevance	<p>Place of work and place of residence, federal states, districts, regional directorates, employment agency areas.</p> <p>(Some of these variables are not available for all data sources and periods.)</p>
Date of territorial allocation	31 December 2010
Type of territorial allocation	<p>All data sources with corrected territorial allocation (date: 31 December 2010)</p> <p>Exception: original territorial allocation for the variable ao_aa in data sources LeH and ASU</p>
Period covered	<p>The period covered depends on the data source.</p> <p>BeH 01 January 1975 - 31 December 2010 (2008, 2009: 18-month file; 2010: 6-month file)</p> <p>LeH 01 January 1975 - 14 July 2011</p> <p>ASU 01 January 1997 - 14 July 2011</p> <p>LHG 01 January 2005 - 14 July 2011</p> <p>XASU 01 January 2005 - 14 July 2011</p> <p>XLHG 01 January 2005 - 14 July 2011</p>
Update frequency	Continuous
Data unit	Employees covered by social security (including marginal part-time employees from 1999 onwards), benefit recipients, jobseekers
Number of cases	<p>1,639,325 individuals</p> <p>45,793,010 non-overlapping observations (after episode splitting)</p> <p>35,683,562 original observations</p>
Corrections - outline	The data are largely taken from the IEB without corrections. Basically, only validation procedures are carried out, e.g. some variables are recoded.
Special features - outline	The data undergo an episode splitting procedure. This means that the data are prepared in such a way that only non-overlapping or completely parallel periods are created.

File organisation	The data are stored in two files. One contains individual-level information and the other establishment-related information. Further files with additional information on establishments can be provided following a justified application.
File format and size	Stata (individual file approx. 3 GB)
Data access	On-site usage at the FDZ of the BA at the IAB and subsequent remote data access
Degree of anonymisation	Weakly anonymous
Citation method	Sample of Integrated Labour Market Biographies (SIAB) 1975 - 2010, Nuremberg 2013

Table 1: Outline

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

List of variables	Module	Page	BeH	LeH	LHG	XLHG	ASU	XASU
<b>Identifiers</b>		<b>31</b>						
Artificial individual ID (persnr)		31						
Artificial establishment number (betnr)	*	32						
<b>Period of validity</b>		<b>33</b>						
Original start date of observation (begorig)		33						

List of variables	Module	Page	BeH	LeH	LHG	XLHG	ASU	XASU
Original end date of observation (endorig)		33						
Start date of split episode (begepi)		33						
End date of split episode (endepe)		34						
<b>Personal information</b>		34						
Gender (frau)		34						
Year of birth (gebjahr)		34						
Nationality (nation)		35						
Nationality, aggregated (nation_gr)		35						
Marital status (famst)		35						
Number of children (kind)		36						
School education and vocational training (bild)		36						
School-leaving qualification (schbild)		38						
Skills level (quali)		38						
Severe disability status (schweb)		38						
Health problems (gesund_ein)		39						
<b>Information on employment, benefit receipt and job search</b>		40						
Reason for notification / reason for end of benefit receipt / reason for discontinuation of unemployment benefit II / reason for deregistration (grund)		40						
Daily wage, daily benefit rate (tentgelt)		41						
Transition zone (gleitz)		42						
Occupation – activity performed / last activity (beruf)		43						
Occupational status and working hours (stib)		43						
Employment status (erwstat)		44						
Employment status prior to job search (estatvor)		45						
Type of benefit (lart)		46						
BA client group (kundengr)		46						
Type of termination of last job (art_kuend)		46						
Desired working hours of the job sought (arbeitszeit)		47						
Duration of remaining entitlement to unemployment benefit (restanspruch)		47						
Type of institution (traeger)		47						

List of variables	Module	Page	BeH	LeH	LHG	XLHG	ASU	XASU
<b>Establishment variables</b>		<b>48</b>						
Economic activity 73 – group of economic activity 73 (3-digit code) (w73_3)	*	48						
Economic activity 93 – sub-class of economic activity 93 (5-digit code) (w93_5)	*	49						
Economic activity 93 – group of economic activity 93 (3-digit code) (w93_3)	*	49						
Economic activity 03 – sub-class of economic activity 03 (5-digit code) (w03_5)	*	50						
Economic activity 03 – group of economic activity 03 (3-digit code) (w03_3)	*	50						
Economic activity 08 – sub-class of economic activity 08 (5-digit code) (w08_5)	*	51						
Economic activity 08 – group of economic activity 08 (3-digit code) (w08_3)	*	51						
Economic activity 73 generated – completed by extrapolation / imputation (w73_3_gen)								
Economic activity 73 generated – type of completion (group_w73_3)								
Economic activity 93 generated – completed by extrapolation / imputation (w93_3_gen)								
Economic activity 93 generated – type of completion (group_w93_3)								
Year of first appearance of establishment (grd_jahr)	*	51						
First appearance of establishment number (grd_dat)								
Year of last appearance of establishment (lzt_jahr)	*	53						
Last appearance of establishment number (lzt_dat)								
Total number of employees (az_ges)	*	54						
Number of full-time employees (az_ges_vz)	*	55						
Number of employees in marginal part-time employment (az_gf)	*	55						



List of variables	Module	Page	BeH	LeH	LHG	XLHG	ASU	XASU
Median gross daily wage of full-time employees (te_med)	*	56	■					
<b>Location data</b>		<b>56</b>						
Place of residence: district (Kreis) (wo_kreis)		56	■	■	■	■	■	■
Place of residence: employment agency (wo_aa)		56	■	■	■	■	■	■
Place of residence: regional directorate (wo_rd)		58	■	■	■	■	■	■
Place of residence: federal state (Bundesland) (wo_bula)		58	■	■	■	■	■	■
Place of work: district (Kreis) (ao_kreis)	*	59	■					
Place of work: employment agency (ao_aa)	*	59	■	■			■	
Place of work: regional directorate (ao_rd)	*	60	■	■			■	
Place of work: federal state (Bundesland) (ao_bula)	*	60	■					
<b>Generated technical variables</b>		<b>60</b>						
Source of the observation (quelle)		60	■	■	■	■	■	■
Observation counter per person (spell)		61	■	■	■	■	■	■
Year (jahr)		61	■	■	■	■	■	■

Table 2: List of variables with degree of completeness

## 1.6 Volume structure

No. of cases	before splitting	after splitting
BeH	26,312,013	29,741,469
LeH	3,749,512	5,336,121
LHG	791,719	2,180,759
XLHG	174,179	337,538
ASU	4,526,217	7,882,859
XASU	129,922	314,264
Total number of observations	35,683,562	45,793,010
Individuals		1,639,325

Table 3: Volume structure

## 2 Data sources

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

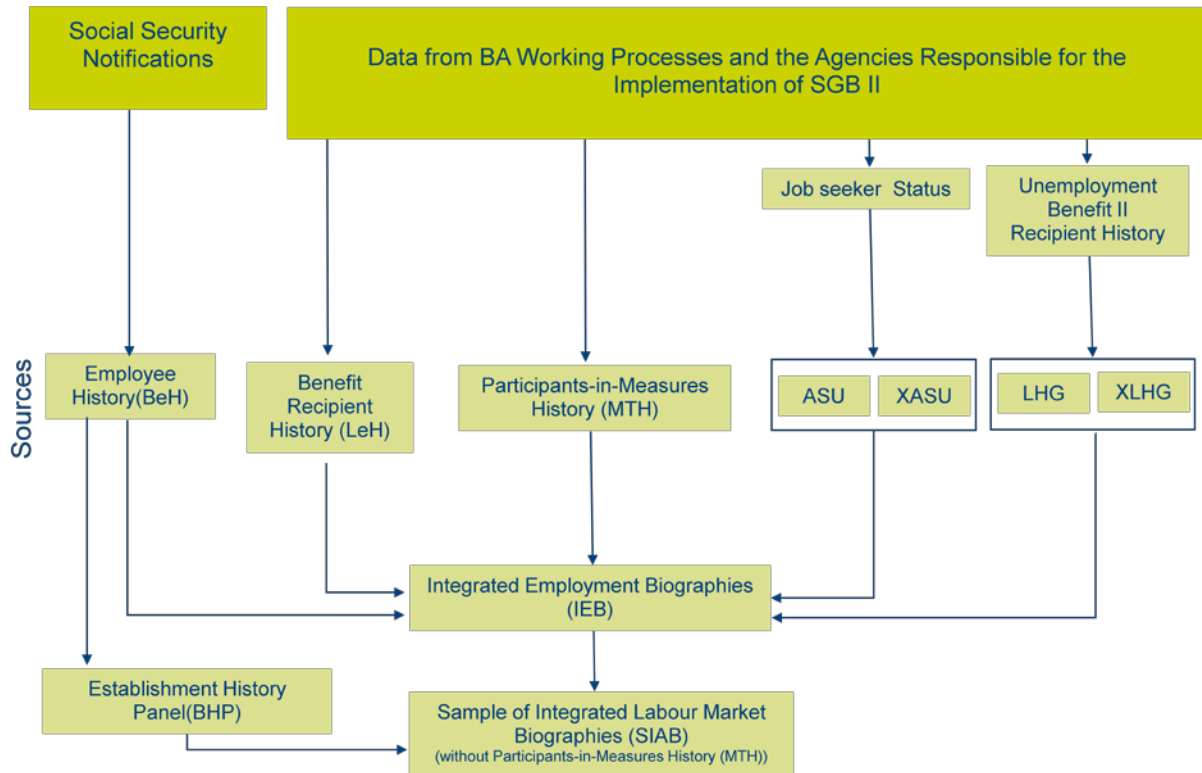


Figure 2: Data sources of the IEB

Observations from all of the data sources shown, with the exception of participation in employment and training measures (MTH), are incorporated from the IEB.

### 2.1 Employee History (BeH)

The source of data regarding employment is the Employment History (Beschäftigten-Historik - BeH) of the IAB. The data basis is the integrated notification procedure for health, pension and unemployment insurance, which came into effect as of 01 January 1973 (and was extended to cover Eastern Germany as of 01 January 1991) and is known by the abbreviation DEÜV (previously DEVO / DÜVO) (for further details see: Bender et al. 1996, p. 4 et seq.; Wermter /Cramer 1988). Under this procedure employers are required to submit notifications to the responsible social security agencies concerning all of their employees covered by social security at least once a year. The BeH covers all white- and blue-collar workers as well as apprentices as long as they are not exempt from social security contributions. This means

that civil servants, self-employed persons and regular students (see Cramer 1985) are in principle not recorded in the BeH. Since the notification procedure was changed on 01 January 1999, employees in marginal part-time employment and unpaid family workers have also been recorded (not contained in the data until 01 April 1999). The data are recorded by the health insurance companies, collected and edited by the Federal Employment Agency (BA) and subsequently integrated into the History File by the IAB.

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

When linking individual data with establishment data it has to be taken into account that the variables in the Basis Establishment File as well as in the BHP Extension Files are aggregated on 30 June of a year. Unlike the data on individuals, the establishment variables are therefore not spell data but are only valid on 30 June precisely each year.<sup>2</sup>

The Establishment File is linked with the Individual File via the programme-specific commands of the software packages used for preparing and analysing the data. In Stata, for instance, the two files can be linked using the “merge” command in connection with the relevant paths (see Box 1).

```
use siab_7510_v1.dta
gen jahr = year(begepi)
sort betnr jahr
merge m:1 betnr jahr using SIAB_7510_v1_bhp_basis_v1.dta
```

Box 1: Example code for Stata12

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

---

<sup>2</sup> An extreme example: an employment notification exists from 01 January 2006 to 30 May 2006; the establishment goes bankrupt in June 2006. There is then no information about this establishment in the BHP for 2006.

## 2.3 Unemployment Benefit II Recipient History (LHG / XLHG)

The Unemployment Benefit II Recipient Histories (Leistungshistoriken Grundsicherung LHG / XLHG) contain the receipt of benefits in accordance with Social Code Book II (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 Social Code Book III, the BA is no longer necessarily the sole institution responsible for administering the benefits. The data therefore distinguish between the three possible types of institution responsible for implementing SGB II:

- Joint facilities (Gemeinsame Einrichtungen) / Cooperation of employment agencies and municipalities (Arbeitsgemeinschaften – ARGE, until the end of 2010) in which the BA and the municipality deal with tasks jointly,
- Separated responsibilities (getrennte Trägerschaft) (until the end of 2011) – the tasks continue to be divided between the BA and the municipality<sup>3</sup> –,
- Authorised municipalities which are also called opting local authorities or opting municipalities according to the initial experimental clause of Section 6a - here the local authority is responsible for all tasks in the sphere of SGB II.

The data of the “Unemployment Benefit II Recipient History drawn from A2LL” (LHG) come from 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 until 2010 in all ARGE cooperation projects, and from 2011 onwards in joint facilities, 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.

An important difference compared with the LeH is that the amount of benefits received is not determined at the level of the individual but at the level of the benefit community (Bedarfsgemeinschaft<sup>4</sup>). It is not possible, however, to link individuals with benefit receipt under

---

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

<sup>4</sup> A so-called ‘Bedarfsgemeinschaft’, or ‘benefit community’, includes all individuals in a household who receive benefits jointly (i.e. as a joint payment). In the majority of cases, the ‘benefit community’ and the household will be the same, which in particular applies in the case of (married or unmarried) couples and parents with children under the age of 25. However, under specific circumstances the ‘benefit community’ might not include all household members, or a household in which everybody receives benefit payments might be made up of more than one ‘benefit community’. An example of the former is if a grown-up child lives with his/her parents and earns just enough to make his/her own living but has insufficient means to support his/her mother and father – in this case the ‘benefit community’ will only include the parents. An example of the latter is a three- (or more) generation house-

SGB II at the level of benefit communities in the SIAB. The LHG and the XLHG in principle contain all registered individuals who are entitled to receive benefits in accordance with Section 7 SGB II, in other words, needy individuals who are capable of work and, if applicable, also the members of their benefit community. However, only periods during which a person is capable of work or is over the age of 65 are incorporated.

Each dataset depicts non-overlapping periods of entitlement to benefit of a person in a certain benefit community. A new observation starts for the following administrative reasons:

- on relevant birthdays (14, 15, 18 and 65) of the members of a benefit community, which are statutory and relevant for structural changes of the benefit community (see Chapter 3.1.5),
- 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
- at the beginning and end of a sanction period for observations from 01 April 2006 onwards. However it must be taken into account that it is not possible to determine either the duration, the type of the sanction or the time when the sanction was imposed or when it began based on the data. 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.

The LHG data are available from 01 January 2005, the XLHG data from 01 November 2005. However, until the beginning of 2007 both data sources are incomplete (see Chapter 4).

## **2.4 Jobseeker History (ASU / XASU)**

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

---

hold: since a 'benefit community' may only consist of two generations, this type of household will be made up of two such 'benefit communities', one consisting of the grandparent(s) and one of the parent(s) and child(ren). (Trappmann et al. 2007)

### **3 Data preparation and sampling procedure**

#### **3.1 Corrections and validation procedures**

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

- Observations in which the age is under 13 or over 75 are deleted.
- Observations whose end date precedes the start date are deleted.

Inconsistent information on gender or date of birth within an account is corrected, datasets without information on the date of birth are deleted. Further corrections, like in the IABS (Completion procedure of presumably missing notifications, strike corrections) do not happen.

##### **3.1.1 Employee History (BeH)**

- To capture a person group that is as constant as possible over time, some person groups for which data are not available throughout the entire observation period are excluded. To be precise, from 1999 onwards only the person groups 101, 102, 103, 105, 106, 112, 118, 119, 120, 140, 141, 142, 143, 149, 201, 203 and 205 are supplied, as well as the two groups 109 and 209, which indicate people in marginal part-time employment. Groups that are not included from 1999 onwards include people in short-term employment, i.e. person groups 110, 202 and 210.
- Observations with earnings amounting to zero or with no details on earnings, and the value 101 for the person group variable, and the value 50 for the reason for notification (annual notification) are not incorporated into the IEB.
- Gender and date of birth are taken from the social security number. This information is harmonised across data sources.
- The territorial allocations for place of work and place of residence are updated to the status as of 31 December 2010.

##### **3.1.2 Benefit Recipient History (LeH)**

- It must be taken into account that in cases of ongoing receipt the end date corresponds to the end of the period for which the benefit is allowed and not the end of the benefit receipt period. LeH observations which go beyond 14 July 2011 thus depict the period for which the benefit is allowed. Observations with no end date or an invalid end date are excluded from further processing, since in these cases it cannot be assumed that a benefit payment was made at all.
- Observations without a valid start date are excluded.

- A very small proportion of the LeH observations begin before 1975. This can be explained by the fact that the IEB contains all LeH observations that end in 1975 or later.
- Observations whose end date precedes the start date are excluded.
- If the end date for the receipt of unemployment assistance precedes the start date by one day, then the end date is increased by one year.
- The territorial allocations are corrected analogous to the BeH.

### **3.1.3 Unemployment Benefit II Recipient History (LHG / XLHG)**

- It only contains observations of people who are capable of work and people over the age of 65.
- Individual-related variables that are available for the (X)LHG sources are valid for the entire observation period.
- The territorial allocations are corrected analogous to the BeH.

### **3.1.4 Jobseeker History (ASU / XASU)**

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

- In contrast to the ASU source, the XASU only distinguishes the status
  - “not unemployed, but seeking job” or
  - “unemployed and simultaneously seeking job”.
- Unlike the ASU, periods of illness are not taken into account when generating the “employment status”, since there is no information about illness available in the XASU data. When calculating the unemployment duration with XASU observations, gaps due to illness cannot be identified.
- The XASU contains non-overlapping time periods for individuals. If one of the following variables changes, in each case a new data spell is generated for the XASU:
  - Change of job search status
  - Change of availability
  - Change of SGB II institution (due to notification procedure)
  - Change of place of residence
- The territorial allocations are corrected analogous to the BeH.

### **3.1.5 SGB II anonymisation**

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

In order to prevent an indirect determination of the exact day of birth, the following procedure is applied. Observations divided on the 18<sup>th</sup> or 65<sup>th</sup> birthday will be merged into a single one. Besides the exact day of birth no other information will get lost in this case. Observations beginning on the 15<sup>th</sup> or 65<sup>th</sup> birthday (end on the 65<sup>th</sup> birthday), the start date (end date) of the observation will be set to the beginning (the end) of the respective quarter. This correction results in a bias of the duration of the observation.

Corrected observations are marked in the "employment status" variable.

## **3.2 Episode splitting**

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





Figure 3: Episode splitting

The original date variables for the beginning and the end of the original observation (begorig and endorig) are retained, the variables 'start date of the split episode' and 'end date of the split episode' (begepi and endepe) mark the beginning and the end of the split episodes. It is possible to establish whether observations have been split by comparing the original period (begorig and endorig) with the episode period (begepi and endepe).

To restore the original data without the split episodes or to delete the episodes that were created artificially by means of episode splitting it is necessary to select all observations for which the start of the original observation is the same as the start of the split episode (begepi == begorig).

### 3.3 Sampling procedure

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

### 3.4 Missing values

In the SIAB, missing values are coded as follows:

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

## 4 Data quality and problems

### 4.1 Entire IEB

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

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

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

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

<b>Biographical gap</b>	<b>Information on gap, identifiable using the manifestation of the "grund" characteristic in the preceding dataset of the source, if necessary</b>
Civil servants, professional soldiers, judges, employees of bodies or foundations under public law, unsupported self-employed persons	LeH, ASU
Students, people in school-based further education	LeH, ASU

People who are ill / not able to work for more than 6 weeks (illness during unemployment, however, is represented in the ASU source under certain conditions, see Chapter 3.1.4)	BeH, LeH, ASU
People receiving old-age pension without employment if not member of a benefit community	LeH, ASU
Recipients of early retirement benefits	LeH, ASU
Trade professionals working from home	
Short-time employed persons	ASU
People in facilities of the youth welfare service organisation, in vocational training centres, approved workshops or similar facilities for disabled persons	ASU
Participants in benefit programmes for participation in working life (people in rehabilitation)	ASU
(Sideline) farmers	
Caregivers according to Section 19 SGB XI	
Conscripts	BeH, LeH, ASU
Persons in reserve duty training	BeH, LeH, ASU
Persons fulfilling community service	BeH, LeH, ASU
Persons fulfilling a voluntary social or ecological year instead of community service	
Other people not registered with the statutory pension insurance or the Federal Employment Agency (e.g. sabbatical, funding from personal assets or pensions, emigration, employment abroad, voluntary work, etc.)	LeH, ASU, BeH
Strikers in case of a duration of the strike of more than a month	LeH
Social benefit recipients (prior to the introduction of SGB II in 2005), social allowance recipients (according to SGB II)	
Compensation recipients according to FELEG (Gesetz zur Förderung der Einstellung der landwirtschaftlichen Erwerbstätigkeit, Act on the Support in Case of Termination of Farming Activities)	
People participating in measures	ASU

Table 4: Biographical gaps and identification possibilities

## 4.2 Employee History (BeH)

- Due to the introduction of the employment notification procedure in the federal states of Eastern Germany, the notifications for Eastern Germany can only be assumed to be sufficiently complete from 1993 onwards. For the same reason, a great deal of spells in 1991 show missing values for several variables (such as employment status).

- The increase in the number of BeH observations from 1999 onwards is due to the introduction of the obligation to submit employment notifications for people in marginal part-time employment from 01 April 1999 onwards.
- Especially in 1999, observations with part-time employment increase significantly. This is caused by the actually observed increase in part-time work as well as by the fact that since 1999 employment notifications have been completed more properly.
- Within the employment notification procedure, a certain time lag is unavoidable. Although changes in employment relationships have to be reported immediately, and existing employment relationships have to be confirmed annually by April of the following year, some notifications actually arrive years later. The History File of the IAB is not updated continuously, however, but at certain intervals. This is done using files of employment notifications for one particular year which were submitted 36, 18, 12 or 6 months after the end of the reporting year (e.g. the 18-month file for 2007 can be created in July 2009 at the earliest). Notifications submitted more than three years late are not taken into account at the IAB, which means that a 36-month file shows a 100 % degree of completeness by definition.
- In the version of the IEB on which the SIAB data are based the degree of completeness of the BeH observations last stood at 100 % in 2007. 18-month files were used for 2008 and 2009, and the observations for 2010 come from a 6-month file. It can therefore be assumed that the BeH observations for 2008, 2009 and in particular for 2010 are slightly underreported in the SIAB. However, this should not reduce the ability to analyse the data at individual level. The missing notifications occur more frequently in a few establishments, however. This means that in individual cases the establishment data, e.g. establishment size, are grossly incorrect and will change considerably in subsequent versions.
- In 1984, a change was made in the employment notification procedure. From that time onwards, one-off payments of gross earned income were reported as part of the annual earnings subject to social security contributions, which leads to an increase in the average daily wage. In particular, the proportion of wages and salaries above the upper earnings limit has increased considerably from that year onwards (see Bender et al. 1996).
- For the years 1992 until 2000 noticeable drops and rises in the number of notifications were noticed. Drops can be observed especially for the following 10 districts: Braunschweig (03101), Wolfsburg (03103), Emden (03402), Kassel (06633), Essen (05113), Neuss (05162), Erftkreis (05362), Hersfeld-Rotenburg (06632), Miltenberg (09676)

and Kempten (Allgäu) (09763). This is due to notification problems of single or several establishments in these regions.

- Considerable drops were also observed for the districts Salzgitter (03102) und Hoyer-swerda (14264).
- Concerning the notifications for full-time employment, especially the districts Main-Taunus (06436) und Alzey-Worms (07331) are noteworthy. They feature above-average rises. Also in this case, the reasons are notification problems at single or several establishments in these regions.
- In the years 1996 to 1998, the values 841-844 (doctors and pharmacies) within the "occupation – activity performed" variable are very rare compared to the neighbouring years.

#### **4.3 Benefit Recipient History (LeH)**

- For the states of Eastern Germany, the LeH observations were not fully recorded until 1992.
- The benefit receipt data used to be saved on magnetic tapes. Owing to a fault in one magnetic tape, the benefit receipt data up to and including 1980 are only partially contained. Thus, in the present data product, too, it can be assumed that information on benefit receipt in that period is not available in full.
- Caused by an internal change of systems, there is a break in the collection of periods of exclusion of benefits and of benefit suspension in 2004 due to failure to comply in 2004.

#### **4.4 Unemployment Benefit II Recipient History (LHG / XLHG)**

- With regard to the completeness of case numbers or benefit histories from the LHG and XLHG data sources, there are substantial gaps in the years 2005 and 2006. We therefore strongly advise against analysing the data for this time period based merely on the (X)LHG sources.
- Also as of 2007, underrecording of mostly one month depending on the institution has occurred, especially in the XLHG. The LHG shows underrecording and overrecording related to changes in the type of institution as well.
- Furthermore, for both of the data sources LHG and 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- and 16-year-old individuals in need who are supported by the authorised municipalities may therefore be more poorly represented, as "individuals not capable of work" are not included in the IEB.

- Some individuals for whom a (X)LHG spell exists are excluded completely or partly from the benefit receipt according to SGB II, for instance because they are in a subsidised training programme, receive old-age pensions, live in a stationary institution or receive insurance payments to avoid need for assistance. This affects on average 3 to 5 percent of all cases.

## **4.5 Jobseeker History (ASU / XASU)**

### **4.5.1 ASU**

- The registered periods of job search activity in the ASU source have been regarded as complete from the year 1997 onwards. Therefore, the analysis potential of the ASU spells before 1997 is limited.
- For the placement experts it is not possible to always immediately record the allocation to the legal sphere, since frequently it is only known after some time which institution is primarily responsible due to a possible entitlement to SGB II benefits. Therefore, we recommend comparing the manifestation of the "type of institution" characteristic from the ASU with the manifestation of the LHG and/or XLHG during the same period of time. Due to the recording gaps in the LHG and XLHG between 2005 and 2006 this is not always possible.
- For some people for whom an opting municipality has been responsible since 2005, "artificial" ASU datasets were created by the Federal Employment Agency in parallel. These can be identified via estatvor (transfer to an authorised municipality).
- From mid-2005 until mid-2006, the coArb IT procedure, from which the jobseeker and applicant pool data originate, was superseded by the VerBIS procedure at the Federal Employment Agency. In July 2005, coArb was first replaced by VerBIS in the employment agency in Wiesbaden as a pilot project. From December 2005 onwards, it was then gradually replaced by VerBIS in several stages in all employment agencies. The information for many of the variables recorded was gathered with different levels of differentiation and different qualitative weighting in the two systems. It is therefore very difficult to integrate these variables into the IEB, which is only possible using a special procedure (mapping). Unfortunately, a full conversion of the affected variables from coArb to VerBIS cannot be achieved by means of mapping, so for some variables there is an unusually large number of the values "no details available", "other" or "missing". Moreover, striking differences may occur in frequency counts, depending on whether the original source of the data was coArb or VerBIS. Important limitations of the analysis potential are mentioned in the corresponding description of variables.

## 4.5.2 XASU

- In contrast to the job search spells from the cooperation of employment agencies and municipalities (ARGE) and the separated responsibilities, systematic cases of under-recording have emerged for the 69 authorised municipalities since 01 January 2005. Thus, data from the XASU source should only be analysed from June 2006 onwards.
- A variety of variables sometimes have only a very low degree of completeness for the XASU. Variables which are affected by this include “school-leaving qualification”, “severe disability status”, “reason for notification” as well as “employment status prior to job search”. Although the degree of completeness of these variables improves over time, they are still partly unsatisfactory. The beruf “occupation – activity performed” variable is not available in the XASU for almost the entire period available.
- For a number of institutions (districts), the proportion of registered recipients of unemployment benefit II who are also registered jobseekers is implausibly large at times or continuously in the IEB. One possible reason for this could be an incorrect determination of the status “not unemployed but seeking work” by these institutions.
- The institution-related and period-related plausibility of the XASU data should be examined before use, taking the research question into account.

## 5 Description of variables

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

### 5.1 Identifiers

#### 5.1.1 Artificial individual ID (persnr)

Variable label	artificial individual ID
Variable name	persnr
Category	identifiers
Origin	BeH, LeH, LHG, XLHG, ASU, XASU
Data type	numerical
Hierarchy	none
Detailed 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 identifier in the different data sources, the allocation of the information from different data sources (e.g. employment and benefits) to individuals is not always unambiguous. In such cases, implausible employment histories may arise.</p> <p>The formation of the individual identifier which spans all data sources is based on a heuristic developed by the BA.</p>

### 5.1.2 Artificial establishment number (betnr)

Variable label	artificial establishment number
Variable name	betnr
Category	identifiers
Origin	BeH
Data type	numerical
Hierarchy	none
Detailed description	<p>The artificial establishment number indicates which observations belong to the same establishment. It is based on the establishment number allocated by the BA, which was replaced by an artificial number. (Further information on the allocation of establishment numbers by the BA can be found in Bender et al. 1996: p. 15 et seq. and pp. 27-30.) The establishment number and year specification can be used to merge individual and establishment information.</p> <p>For the establishment number, the following should be observed in general:</p> <ol style="list-style-type: none"> <li>a) If the company has one office only, or if the company has one office only in one municipality, this office is the establishment and receives an establishment number.</li> <li>b) If the company has <b>several</b> branch offices in <b>one</b> municipality, these establishment premises / workplaces must be merged into a single establishment under one establishment number, if they are part of the same economic class. Vice versa, if they are <b>not</b> part of the same economic class, each branch office is an establishment and receives an establishment number each.</li> <li>c) If the company has <b>several</b> branch offices in <b>several</b> municipalities, each of these branch offices is an establishment and receives an establishment number each.</li> </ol> <p>In this context, the following definitions with regards to the allocation of establishment numbers as part of the notification procedure for social security must be observed:</p> <ol style="list-style-type: none"> <li>a) An <b>establishment</b> is a regionally and economically delimited unit in which employees work and to which, according to the above-mentioned principles, an establishment number is to be allocated.</li> <li>b) A <b>workplace</b> is a unit in which employees work and to which, according to the above-mentioned principles, no establishment number is to be allocated.</li> <li>c) A <b>company</b> as a term combines establishment premises and workplaces of the same employer.</li> <li>d) An <b>employer</b> is any natural person or legal entity that employs at least one employee subject to social insurance contributions or in marginal part-time employment.</li> <li>e) Establishment and establishment premises are synonyms; branch office is a synonym for subsidiary, district office, out-sourced office, workplace, etc. if it is no establishment.</li> </ol>



Notes on quality	The establishment number is only missing in a very limited number of cases. These observations are notifications for the person group “205” (earnings notifications for casual workers). As establishment variables (place of work, economic activity, establishment size, etc.) are merged via the establishment number, they are missing in these observations.
------------------	---

## 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, LHG, XLHG, ASU, XASU
Data type	date
Hierarchy	none
Detailed description	<p>The original start date of the observation corresponds to the original start date of the notification. This can differ from the start date of the episodes (begepi) (see also the comments on episode splitting in Chapter 3.2)</p> <p>Since the notification logic might under certain circumstances allow for re-identification of the exact day of birth, the original information on the date were changed in these cases by applying the anonymisation procedure described in Chapter 3.1.5.</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, LHG, XLHG, ASU, XASU
Data type	date
Hierarchy	none
Detailed description	<p>The original end date of the observation corresponds to the original end date of the notification. This can differ from the end date of the relevant line of data, the so-called end date of the split episode (see also the comments on episode splitting in Chapter 3.2).</p> <p>Since the notification logic might under certain circumstances allow for re-identification of the exact day of birth, the original information on the date were changed in these cases by applying the anonymisation procedure described in Chapter 3.1.5.</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, LHG, XLHG, ASU, XASU
Data type	date
Hierarchy	none
Detailed description	Start date of the split episode It always equals or is greater than the start date of the original observation (see also the comments on episode splitting in Chapter 3.2).

#### 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, LHG, XLHG, ASU, XASU
Data type	date
Hierarchy	none
Detailed description	End date of the split episode. It always equals or is less than the end date of the original observation (see also the comments on episode splitting in Chapter 3.2).

### 5.3 Personal information

#### 5.3.1 Gender (frau)

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

#### 5.3.2 Year of birth (gebjahr)

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

### 5.3.3 Nationality (nation)

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

### 5.3.4 Nationality, aggregated (nation\_gr)

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

### 5.3.5 Marital status (famst)

Variable label	marital status
Variable name	famst
Category	personal variables
Origin	LeH, LHG, XLHG, ASU, XASU
Data type	numerical
Hierarchy	none
Detailed description	<p>This variable describes the marital status. The characteristic in the LeH source has only two manifestations (0 - not married, 1 - married), while in the ASU / XASU sources, a distinction is made between four manifestations, and in the LHG / XLHG sources between seven manifestations. The information from the sources was not compared.</p> <p><b>1) ASU / XASU</b>  Marital status includes categories of living in a joint household with children or a partner. However, it is not possible to distinguish clearly whether</p> <ul style="list-style-type: none"> <li>• a married person / person living in a joint household is bringing up children,</li> <li>• a married person / person living in a joint household / single person has any children at all,</li> </ul>

	<ul style="list-style-type: none"> <li>• a lone parent / person living in a joint household / single person is married</li> </ul> <p><b>2) XLHG</b> The quality of the variable is reduced due to a break in the notification logic in 2007.</p>
--	--

### 5.3.6 Number of children (kind)

Variable label	number of children
Variable name	kind
Category	personal variables
Origin	LeH, LHG, XLHG, ASU, XASU
Data type	numerical
Hierarchy	none
Detailed description	<p>This variable has a different meaning depending on the data source. Except for the LeH source, it indicates the number of children aged under 15 living in the household / in the benefit community (Bedarfsgemeinschaft) at the time when the application was made.</p> <p><b>1) LeH</b> The LeH only distinguishes between whether the number of children is zero or at least one. The following values are therefore available: 0 no children 100 One or more children The variable is not updated when there are changes in the type of benefit or the approval of benefits, but only when a new case of benefit receipt occurs after a period of employment. Children are generally included in the LeH up to their 18<sup>th</sup> year of age, under certain conditions also beyond that date if these conditions constitute a higher benefit rate for the beneficiary.</p> <p><b>2) ASU</b> Here, the value of the variable corresponds to the number of children. Until 30 June 2006, only up to nine children could be recorded. The value 0 does not exist. For datasets prior to 30 June 2006, the zero value was recoded to "missing", since it is not clear whether zero should be interpreted as "no children" or as "field not filled in". For datasets after 30 June 2006, the characteristic is only filled in if children exist.</p> <p><b>3) LHG, XLHG, XASU</b> The variable shows the number of children in the benefit community (Bedarfsgemeinschaft). In the LHG and XLHG sources, the value is valid for the entire original period.</p>

### 5.3.7 School education and vocational training (bild)

Variable label	school education and vocational training
----------------	--

Variable name	bild
Category	personal variables
Origin	BeH, ASU, XASU
Data type	numerical
Hierarchy	none
Detailed description	<p>It must be taken into account that this variable has a different meaning depending on the data source:</p> <p><b>1) BeH</b></p> <p>For BeH observations, the variable contains the school education / vocational training reported by the employers in the employment notification procedure. The 'school education and vocational training' indicator 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"> <li>1 Primary school / lower secondary school / intermediate school leaving certificate or equivalent school education, without a vocational qualification</li> <li>2 Primary school / lower secondary school / intermediate school leaving certificate or equivalent school education with a vocational qualification</li> <li>3 With upper secondary school leaving certificate (Abitur), without a vocational qualification</li> <li>4 With upper secondary school leaving certificate (Abitur) and a vocational qualification</li> <li>5 Degree from a university of applied sciences (Fachhochschule)</li> <li>6 University degree</li> </ol> <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 he/she joins a new firm. It can generally be assumed that when a person is employed in a firm for a longer period, the personal data that they reported when they joined the firm is simply continued." (Meinken / Koch 2004, p. 63).</p> <p>A method for correcting missing values or temporal inconsistencies in the education and training data in the predecessor sample, the IABS, can be found in Fitzenberger et al. (2006) and in Drews (2006). Only the BeH data source is used for this, however, as this was the only data source with information on education and training in the IABS.</p>

	<p><b>2) ASU / XASU</b></p> <p>ASU and XASU 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> <p>For the ASU source, it has not been possible to differentiate between in-firm and external vocational training since 01 July 2006 due to changes in the procedure. For the XASU source, this differentiation is not possible at all over the entire period of time. In both cases, the mixed category is coded to "in-firm training". For the ASU source, it has neither been possible to differentiate between "no completed vocational training" and "not specified" since 01 July 2006.</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> <p>We advise against a quantitative analysis of the variable for the ASU source from 2006 forward.</p>

### 5.3.8 School-leaving qualification (schbild)

Variable label	school-leaving qualification
Variable name	schbild
Category	personal variables
Origin	ASU, XASU
Data type	numerical
Hierarchy	none
Detailed 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 has continuously increased over the years, it remains comparatively low (partially under 50 %).

### 5.3.9 Skills level (quali)

Variable label	skills level
Variable name	quali
Category	personal variables
Origin	ASU
Data type	numerical
Hierarchy	none
Detailed description	This variable takes into account not only vocational training but also occupational experience. In the case of equivalent competence as a

	result of educational background, performance or experience, a skills level can also be entered when the applicant does not hold a corresponding qualification.
Notes on quality	Due to a change in the notification procedure, this value is only available until 2006.

### 5.3.10 Severe disability status (schweb)

Variable label	severe disability status
Variable name	schweb
Category	personal variables
Origin	ASU, XASU
Data type	numerical
Hierarchy	none
Detailed description	<p>A severe disability status is recognised from a level of disability of over 50 %. Equal status to severely disabled is granted on application in accordance with Section 2 of the Severely Disabled Persons Act (Schwerbehindertengesetz – SchwbG) when the level of disability is 30 or 40 %.</p> <p>The variable refers to the start date of the period of unemployment or the job search status.</p>
Notes	Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases.
Note on quality	The degree of completeness of this variable is poor in the XASU in 2005 and 2006.

### 5.3.11 Health problems (gesund\_ein)

Variable label	Placement-relevant health problems
Variable name	gesund_ein
Category	personal variables
Origin	ASU, XASU
Data type	numerical
Hierarchy	none
Detailed description	The variable provides information on whether placement-relevant, approved health problems exist for a person.
Notes	<p>The new <i>health problems - impact on placement</i> characteristic replaces the old <i>health problems</i> characteristic which was included in SIAB 7508 and filled in only until 2006 due to a change in the notification procedure.</p> <p>For this characteristic it is not known whether a customer refused to give details or the information was not collected by the placement agent.</p> <p>Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases.</p>

## 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 (grund)

Variable label	Reason for employment notification / end of benefit receipt
Variable name	grund
Category	information on employment, benefit receipt and job search
Origin	BeH, LeH, LHG, XLHG, ASU, XASU
Data type	numerical
Hierarchy	none
Detailed description	<p><b>1) BeH</b> In BeH observations, the 'reason for notification' variable indicates the reason why the employer submitted the employment notification in question to the social security agencies. However, not all of the possible reasons for submitting a notification that may occur in the context of the notification procedure are available in the IEB. For instance, the IEB only includes notifications that have information on earnings (i.e. annual, employment interruption and end of employment notifications), whilst initial registrations are not contained as they contain no information on earnings. However, this does not involve a loss of information, as the details from a registration are transmitted again with the following annual, employment interruption or end of employment notification. The reasons for submitting employment notifications are encoded according to the regulations of the notification procedure which has been in effect since 01 January 1999 (in accordance with DEÜV).</p> <p><b>2) LeH</b> In LeH observations, the variable indicates the reason for the end of receipt of unemployment benefit, unemployment assistance or maintenance allowance. In contrast, the LeH contains no information about the reasons for the start of benefit receipt, as the information for the LeH comes from the notifications submitted by the employment agencies to the health insurance institutions regarding finished periods of benefit receipt.</p> <p><b>3) LHG / XLHG</b> The LHG and XLHG 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 <b>refers to the individual</b>, not to the benefit community. 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. If the Unemployment Benefit II receipt of one member of the benefit community is discontinued, new observations for all members of the benefit community are started at this date, while the reason for end / discontinuation of benefit is only available for the person concerned. This variable is valid exactly at the end of the original observation.</p>



	<p><b>4) ASU / XASU</b></p> <p>In the case of 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 has been reduced from 26 April 2003 onwards. For analysis over long periods of time, the old values can be recoded to the currently valid ones using the table below:</p> <table border="1"> <thead> <tr> <th>old</th> <th>-&gt;</th> <th>new</th> <th>old</th> <th>-&gt;</th> <th>new</th> <th>old</th> <th>-&gt;</th> <th>new</th> <th>old</th> <th>-&gt;</th> <th>new</th> </tr> </thead> <tbody> <tr> <td>29</td> <td>-&gt;</td> <td>60</td> <td>36</td> <td>-&gt;</td> <td>61</td> <td>44</td> <td>-&gt;</td> <td>74</td> <td>51</td> <td>-&gt;</td> <td>74</td> </tr> <tr> <td>30</td> <td>-&gt;</td> <td>60</td> <td>37</td> <td>-&gt;</td> <td>66</td> <td>45</td> <td>-&gt;</td> <td>77</td> <td>52</td> <td>-&gt;</td> <td>76</td> </tr> <tr> <td>31</td> <td>-&gt;</td> <td>61</td> <td>38</td> <td>-&gt;</td> <td>66</td> <td>46</td> <td>-&gt;</td> <td>67</td> <td>53</td> <td>-&gt;</td> <td>68</td> </tr> <tr> <td>32</td> <td>-&gt;</td> <td>60</td> <td>39</td> <td>-&gt;</td> <td>71</td> <td>47</td> <td>-&gt;</td> <td>67</td> <td>54</td> <td>-&gt;</td> <td>78</td> </tr> <tr> <td>33</td> <td>-&gt;</td> <td>60</td> <td>40</td> <td>-&gt;</td> <td>69</td> <td>48</td> <td>-&gt;</td> <td>78</td> <td></td> <td></td> <td></td> </tr> <tr> <td>34</td> <td>-&gt;</td> <td>60</td> <td>42</td> <td>-&gt;</td> <td>65</td> <td>49</td> <td>-&gt;</td> <td>69</td> <td></td> <td></td> <td></td> </tr> <tr> <td>35</td> <td>-&gt;</td> <td>60</td> <td>43</td> <td>-&gt;</td> <td>70</td> <td>50</td> <td>-&gt;</td> <td>75</td> <td></td> <td></td> <td></td> </tr> </tbody> </table>	old	->	new	old	->	new	old	->	new	old	->	new	29	->	60	36	->	61	44	->	74	51	->	74	30	->	60	37	->	66	45	->	77	52	->	76	31	->	61	38	->	66	46	->	67	53	->	68	32	->	60	39	->	71	47	->	67	54	->	78	33	->	60	40	->	69	48	->	78				34	->	60	42	->	65	49	->	69				35	->	60	43	->	70	50	->	75			
old	->	new	old	->	new	old	->	new	old	->	new																																																																																						
29	->	60	36	->	61	44	->	74	51	->	74																																																																																						
30	->	60	37	->	66	45	->	77	52	->	76																																																																																						
31	->	61	38	->	66	46	->	67	53	->	68																																																																																						
32	->	60	39	->	71	47	->	67	54	->	78																																																																																						
33	->	60	40	->	69	48	->	78																																																																																									
34	->	60	42	->	65	49	->	69																																																																																									
35	->	60	43	->	70	50	->	75																																																																																									
Note on quality	<p>Proportion of valid values (degree of completeness) for the reason for notification in the LHG, XLHG, ASU and XASU data sources:</p> <table border="1"> <thead> <tr> <th></th> <th>2005</th> <th>2006</th> <th>2007</th> <th>2008</th> <th>2009</th> <th>2010</th> <th>2011</th> </tr> </thead> <tbody> <tr> <td>LHG</td> <td>0.17</td> <td>0.11</td> <td>0.19</td> <td>0.18</td> <td>0.15</td> <td>0.12</td> <td>0.05</td> </tr> <tr> <td>XLHG</td> <td>0.02</td> <td>0.06</td> <td>0.09</td> <td>0.12</td> <td>0.17</td> <td>0.20</td> <td>0.11</td> </tr> <tr> <td>ASU</td> <td>0.83</td> <td>0.95</td> <td>0.98</td> <td>0.97</td> <td>0.96</td> <td>0.89</td> <td>0.48</td> </tr> <tr> <td>XASU</td> <td>0.25</td> <td>0.53</td> <td>0.52</td> <td>0.53</td> <td>0.55</td> <td>0.49</td> <td>0.27</td> </tr> </tbody> </table>		2005	2006	2007	2008	2009	2010	2011	LHG	0.17	0.11	0.19	0.18	0.15	0.12	0.05	XLHG	0.02	0.06	0.09	0.12	0.17	0.20	0.11	ASU	0.83	0.95	0.98	0.97	0.96	0.89	0.48	XASU	0.25	0.53	0.52	0.53	0.55	0.49	0.27																																																								
	2005	2006	2007	2008	2009	2010	2011																																																																																										
LHG	0.17	0.11	0.19	0.18	0.15	0.12	0.05																																																																																										
XLHG	0.02	0.06	0.09	0.12	0.17	0.20	0.11																																																																																										
ASU	0.83	0.95	0.98	0.97	0.96	0.89	0.48																																																																																										
XASU	0.25	0.53	0.52	0.53	0.55	0.49	0.27																																																																																										

#### 5.4.2 Daily wage, daily benefit rate (tentgelt)

Variable label	daily wage, daily benefit rate
Variable name	tentgelt
Category	information on employment, benefit receipt and job search
Origin	BeH, LeH
Data type	numerical
Hierarchy	none
Detailed description	<p><b>1) BeH</b></p> <p>In BeH observations, this variable shows the employee's gross daily wage. It is calculated from the fixed-period wages reported by the employer and the duration of the (unsplit) original notification period in calendar days. The daily wage is shown in euros.</p> <p>Until 1998, employers in principle only reported the earnings which were subject to social security contributions. Earnings below the marginal part-time income threshold were not reported. Earnings exceeding the upper earnings limit for statutory pension insurance are only reported up to this limit. There are two upper earnings limits in the statutory pension insurance scheme. The earnings limit of the miners' pension insurance is generally higher than the earnings limit of the pension insurance for wage and salary earners. However, in the data a differentiation between these two insurance carriers is not possible.</p> <p>Since the inclusion of marginal part-time employees in the employment notification procedure on 01 April 1999, earnings below the marginal part-time income threshold have also been recorded; the upper earn-</p>

	<p>ings limit still applies as the upper ceiling. In some cases, however, the reported earnings nonetheless exceed the upper earnings limit. Generally, this can probably be attributed to the payment of annual bonuses which the employer can add to the regular earnings in the annual, employment interruption or end of employment notifications. In this case, it is irrelevant whether the upper earnings limit in the statutory pension insurance which is decisive for the notification period is exceeded as a result of this addition. However, such earnings notifications could also be due to incorrect details in the employment period. (The earnings information, however, may be considered less error-prone due to their insurance relevance.) The marginal part-time income threshold and the upper earnings limit for statutory pension insurance differ from year to year as well as between Eastern and Western Germany (the decisive factor is the location of the establishment). An overview of these limits and thresholds can be found under <a href="http://fdz.iab.de">http://fdz.iab.de</a>.</p> <p>A daily wage reported as 0 euros can be put down to “employment interruption notifications”. During these periods, the employment relationship continues to exist in legal terms, but without pay. This is the case for periods of illness after the end of continued payment of wages, for periods of maternity leave and for sabbaticals.</p> <p>The daily wage is shown with two decimal places. All values greater than 0 and smaller than 0.01 were rounded up to 0.01. This makes it possible to identify the above-mentioned employment interruption notifications with the condition daily wage = 0.</p> <p><b>2) LeH</b></p> <p>For LeH observations, the variable shows the daily benefit rate, converted into euros in each case. It must be taken into account that for observations with an original start date prior to 01 January 1998 the daily benefit rate applies to working days, while for observations with an original start date from 01 January 1998 onwards it applies to calendar days.</p> <p>Since 01 January 2005, a daily benefit rate reported as 0 euros can be put down to benefit suspension periods or interruptions of benefit payments. If a reason for end of benefit is reported for an observation with a daily benefit rate equal to 0, then it is a notification of interruption of benefit payments. In the case of observations that reflect a period of benefit suspension, the entitlement is the same as before the start of the benefit suspension period.</p>
--	--

### 5.4.3 Transition zone (gleitz)

Variable label	transition zone
Variable name	gleitz
Category	information on employment, benefit receipt and job search
Origin	BeH
Data type	numerical
Hierarchy	none

Detailed description	This variable has only been available from 2003 onwards and only for BeH observations. It indicates whether the employment notification relates to employment in the low-wage sector, within the so-called transition zone. Jobs in the transition zone have a gross monthly wage of € 400.01 to € 800.00 (so-called midi jobs) for which the employee only has to pay a reduced overall social security contribution. As employees with earnings in the transition zone can voluntarily pay the “regular” social security contribution, not all employees with corresponding earnings are automatically classified as being in the transition zone. The corresponding legislation has been in force since 01 April 2003.
----------------------	---

#### 5.4.4 Occupation – activity performed / last activity (beruf)

Variable label	occupation – activity performed / last activity
Variable name	beruf
Category	employment / benefit receipt
Origin	BeH, ASU, XASU
Data type	numerical
Hierarchy	none
Detailed description	<p><b>1) BeH</b></p> <p>The occupational title of the job performed by the employee during the notification period is a component of the 'employment details' submitted by the employer.</p> <p>For this, the employer encodes the employee's job in accordance with the “Classification of Occupations. Systematic and Alphabetical Directory of Job Titles” (published by the Federal Employment Agency, Nuremberg, 1988), which contains approx. 25,000 job titles. The occupational classification consists of a 3-digit code and comprises about 330 values.</p> <p>If more than one job title with different classification codes applies to one employee, the employer is required to select the job title that best defines the main activity performed (see BA 2005, p. V).</p> <p><b>2) ASU / XASU</b></p> <p>For ASU observations, this variable contains the occupation of the last activity performed.</p> <p>For XASU observations, the variable contains the last successfully completed vocational training.</p>

#### 5.4.5 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

Detailed description	<p>The employee's occupational status during the notification period is reported by the employer in the "employment details".</p> <p>The "occupational status" variable 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 actually provides details about the occupational status for full-time employees only, 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 and white-collar employees centres solely on the type of pension insurance institution (Federal Social Insurance Office for Salaried Employees – BfA – for white-collar workers, and Land Social Insurance Office – LVA – for blue-collar workers). The "employees in vocational training" category covers not only trainees / apprentices, placement workers and interns but also semi-skilled trainees, students at colleges for health occupations and participants in subsidised 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>
----------------------	--

#### 5.4.6 Employment status (erwstat)

Variable label	employment status
Variable name	erwstat
Category	personal variables
Origin	BeH, LeH, LHG, XLHG, ASU, XASU
Data type	numerical
Hierarchy	none
Detailed description	<p>This variable assumes different values with different meanings for each data source.</p> <p><b>1) BeH</b></p> <p>For BeH observations, the variable contains the category of the employment notification, which was introduced along with the new notification procedure (DEÜV) as of 01 January 1999. It indicates contribution- or benefit-related particularities of the employment relationship.</p> <p>If multiple keys apply to an employment notification, the smallest must be indicated by the reporting employer. The majority of these cases are employment relationships subject to social security contributions without any exceptions which are kept under key number 101. Therefore, it is possible that these employees are slightly overestimated. The notification procedure states that changes in the employment status - e.g. when an apprentice is employed with his/her training company after completing his/her vocational training - must be indicated by a new notification (cf. Deutsche BKK 2012, p. 31).</p> <p>The status can be contained in employment notifications that refer to the years prior to 1999 but were not received until 1999 or later. For</p>

	<p>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 'school education and vocational training', 'occupational status and working hours' and 'occupation' variables as well as other information. In many cases, however, appropriate allocations are not possible. Since 01 April 1999, also employees in marginal part-time employment have been recorded in the DEÜV notification procedure. This group of people can be differentiated via the manifestations 109 and 209. For employees in marginal part-time employment, no data prior to the introduction of the notification obligation in 1999 could be collected.</p> <p><b>2) LeH</b></p> <p>For LeH observations, the variable employment status contains the grouped benefit type. Thus, it can be differentiated whether a person receives unemployment benefit, unemployment assistance or maintenance allowance or whether contributions to private long-term care insurance are paid by the BA.</p> <p><b>3) ASU / XASU</b></p> <p>For ASU observations, the "employment status" variable reports the job search status. Recipients of unemployment benefits (Unemployment Benefit I or II) over the age of 58 who receive benefits under the relaxed conditions according to Section 428 of Social Code Book III (or Section 65 Para. 4 of Social Code Book II) and individuals aged over 58 who are not benefit recipients and are not willing to be placed in employment in the sense of Section 252 Para. 8 Social Code Book VI are recorded as individuals seeking advice.</p> <p>In XASU observations, the variable employment status reports the "not unemployed, but registered as a job seeker" as well as "unemployed and registered as a job seeker" values.</p> <p>Since the notification logic might in many cases allow for re-identification of the exact day of birth, the original information on the date were changed by applying the anonymisation procedure described in Chapter 3.1.5.</p> <p><b>4) LHG / XLHG</b></p> <p>For LHG and XLHG datasets, the "employment status" characteristic shows whether the person is registered as underage employable, full-aged employable or not employable as of old-age pension threshold.</p> <p>Since the notification logic might in many cases allow for re-identification of the exact day of birth, the original information on the date were changed by applying the anonymisation procedure described in Chapter 3.1.5.</p>
--	---

#### 5.4.7 Employment status prior to job search (estatvor)

Variable label	employment status prior to job- search
Variable name	estatvor
Category	information on employment, benefit receipt and job search
Origin	ASU, XASU

Data type	numerical
Hierarchy	none
Detailed description	For ASU and XASU observations, this variable shows the employment status prior to the job search activity. From 12/2002 onwards the values of the variable have reduced substantially. The values of older observations were recoded to the currently valid values, which are thus valid for the entire period. This information does not refer to the start date of the episode, but to the start of the original time period.
Note on quality	The proportion of valid values (degree of completeness) is very small for XASU observations:

#### 5.4.8 Type of benefit (lart)

Variable label	type of benefit
Variable name	lart
Category	information on employment, benefit receipt and job search
Origin	LeH
Data type	numerical
Hierarchy	none
Detailed description	Type of earnings replacement benefit paid by the BA. In the "employment status" variable this variable is combined for the LeH data source.

#### 5.4.9 BA client group (kundengr)

Variable label	BA client group
Variable name	kundengr
Category	information on employment, benefit receipt and job search
Origin	ASU
Data type	numerical
Hierarchy	none
Detailed description	This variable contains the classification of the BA clients into different types, such as market client, client to be activated, client to be supported and counselling client. This differentiation takes into account the need for action with regard to the individual's potential obstacles to integration.

#### 5.4.10 Type of termination of last job (art\_kuend)

Variable label	type of termination of last job
Variable name	art_kuend
Category	information on employment, benefit receipt and job search
Origin	ASU
Data type	numerical
Hierarchy	none

Detailed description	This variable describes how the last employment or training relationship was terminated before a new period of job search. It can therefore be used to identify job-to-job placements.
----------------------	--

#### 5.4.11 Desired working hours of the job sought (arbzeit)

Variable label	desired working hours of the job sought
Variable name	arbzeit
Category	information on employment, benefit receipt and job search
Origin	ASU
Data type	numerical
Hierarchy	none
Detailed description	During the placement procedure, jobseekers indicate how many working hours the job they are seeking should have.

#### 5.4.12 Duration of remaining entitlement to unemployment benefit (restanspruch)

Variable label	duration of remaining entitlement to unemployment benefit
Variable name	restanspruch
Category	information on employment, benefit receipt and job search
Origin	LeH
Data type	numerical
Hierarchy	none
Detailed description	<p>Remaining entitlement is the entitlement that remains after the end of the current period of benefit receipt. The remaining entitlement used to remain in place for seven years, since February 2006, this has been four years. If the end date of the benefit receipt is before 01 January 1998, the remaining entitlement is reported in working days, after this date it is reported in calendar days. This entitlement counts up to a maximum of the longest possible duration for the individual's respective age.</p> <p>This information does not refer to the start date of the episode, but to the start of the original time period.</p>

#### 5.4.13 Type of institution (traeger)

Variable label	type of institution
Variable name	traeger
Category	information on employment, benefit receipt and job search
Origin	LHG, XLHG, ASU, XASU
Data type	numerical
Hierarchy	none

Detailed description	<p>Until 2011, this variable shows the three possible types of institution responsible for the sphere of Social Code Book II (cooperation of employment agencies and municipalities (ARGE), authorised municipalities (zkT), separated responsibilities (gT)) and the type of institution for the sphere of Social Code Book III (Federal Employment Agency). From 2011 onwards, ARGEs have generally been organised as joint facilities (gE). By 01 January 2012 at the latest, separated responsibilities formed joint facilities or authorised municipalities. The characteristics manifestation 2 indicates the cooperation of employment agencies and municipalities (ARGE) type of institution until 2010, since 2011 it has been joint facilities (gE).</p> <p>The cooperation of employment agencies and municipalities (ARGE) is formed by the Federal Employment Agency and municipalities dealing with tasks jointly. In the case of authorised municipalities (zkT) (at first also called "opting local authorities" according to the experimental clause of Section 6a), the relevant municipality takes on the tasks in the field of SGB II independently. In the case of separated responsibilities (gAw), also called municipalities exercising their duties separately (gAw), the federal and/or municipal SGB II benefits are granted on their own responsibility.</p> <p>It must be taken into account that the district territories covered by the institutions are not always clear and thus may not necessarily correspond to the boundaries of the districts. The same applies to the employment agencies.</p> <p>ASU and XASU observations contain the type of institution that holds the records of the Applicant Pool Data.</p>
----------------------	---

## 5.5 Establishment variables

### 5.5.1 Economic activity 73 (w73\_3)

Variable label	WZ73 3-digit code
Variable name	w73_3
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	division (1-digit code)   group (2-digit code)   class (3-digit code) of economic activity
Detailed description	<p>This variable indicates the economic activity as a 3-digit code in accordance with the WS73 classification and is available from 1975 up to and including 2002.</p> <p>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 is only assigned one code. The assignment to the relevant class of economic activity is carried out under consideration of the institutional orientation of the establishment.</p>



### 5.5.2 Economic activity 93, 5-digit code (w93\_5)

Variable label	WZ93 5-digit code
Variable name	w93_5
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	section (1-digit code)   division (2-digit code)   group (3-digit code)   class (4-digit code)   sub-class (5 digit code) of economic activity
Detailed description	<p>This variable indicates the economic activity as a 5-digit code in accordance with the WZ93 classification and is available from 1998 up to and including 2003. WZ93 stands for the “Classification of Economic Activities for the Statistics of the Federal Employment Services, edition 1993” (“Klassifikation der Wirtschaftszweige für die Statistik der Bundesanstalt für Arbeit, Ausgabe 1993”). The WZ93 is based on the Statistical Classification of Economic Activities in the European Community NACE Rev. 1 („Nomenclature générale des activités économiques dans les communautés européennes“) which has four levels the first two of which are based on the international standard ISIC Rev. 3 („International Standard Industrial Classification of All Economic Activities“).</p> <p>Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected.</p>
Notes	Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, the economic activity is only shown as the 3-digit code (w93_3).

### 5.5.3 Economic activity 93, 3-digit code (w93\_3)

Variable label	WZ93 3-digit code
Variable name	w93_3
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	section (1-digit code)   division (2-digit code)   group (3-digit code)   class (4-digit code)   sub-class (5 digit code) of economic activity
Detailed description	<p>This variable indicates the economic activity as a 3-digit code in accordance with the WZ93 classification and is available from 1998 up to and including 2003. WZ93 stands for the “Classification of Economic Activities for the Statistics of the Federal Employment Services, edition 1993” (“Klassifikation der Wirtschaftszweige für die Statistik der Bundesanstalt für Arbeit, Ausgabe 1993”). The WZ93 is based on the Statistical Classification of Economic Activities in the European Community NACE Rev. 1 („Nomenclature générale des activités économiques dans les communautés européennes“) which has four levels the first two of which are based on the international standard ISIC Rev. 3 („International Standard Industrial Classification of All Economic Activities“).</p>

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

#### 5.5.4 Economic activity 03, 5-digit code (w03\_5)

Variable label	WZ03 5-digit code
Variable name	w03_5
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	section (1-digit code)   division (2-digit code)   group (3-digit code)   class (4-digit code)   sub-class (5-digit code) of economic activity
Detailed description	<p>This variable indicates the economic activity as a 5-digit code in accordance with the WZ03 classification and is available from 2003 up to 2008. WZ03 stands for the “Classification of Economic Activities, Edition 2003” (“Klassifikation der Wirtschaftszweige Ausgabe 2003”) of the Federal Statistical Office (eds.). Like the WZ93, the WZ03 is based on the Statistical Classification of Economic Activities in the European Community NACE Rev. 1 (see description of variables w93_3, w93_5). The classification of the economic activity has been updated, whereas the structure of the WZ93 has been largely maintained.</p> <p>Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected.</p>
Notes	Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, the economic activity is only shown as the 3-digit code (w03_3).

#### 5.5.5 Economic activity 03, 3-digit code (w03\_3)

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

	be reflected.
--	---------------

### 5.5.6 Economic activity 08, 5-digit code (w08\_5)

Variable label	WZ08 5-digit code
Variable name	w08_5
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	section (1-digit code)   division (2-digit code)   group (3-digit code)   class (4-digit code)   sub-class (5-digit code) of economic activity
Detailed description	<p>This variable indicates the economic activity as a 5-digit code in accordance with the WZ08 classification and is available from 2008 onwards. WZ08 stands for the "Classification of Economic Activities, Edition 2008" ("Klassifikation der Wirtschaftszweige Ausgabe 2008") of the Federal Statistical Office (eds.). The WZ08 is based on the Statistical Classification of Economic Activities in the European Community NACE Rev. 2.</p> <p>Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected.</p>
Notes	Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise the economic activity is only shown as the 3-digit code (w08_3).

### 5.5.7 Economic activity 08, 3-digit code (w08\_3)

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

### 5.5.8 Economic activity 73 generated – completed by extrapolation / imputation (w73\_3\_gen)

Variable label	w73_3 Completed by extrapolation / imputation
----------------	---

Variable name	w73_3_gen
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	division (1-digit code)   group (2-digit code)   class (3-digit code)
Detailed description	This variable indicates the economic activity as a 3-digit code in accordance with the WZ73 classification. From 1975 up to and including 2002, the variable contains the original values from w73_3. From 2003 onwards, the information is either continued or replaced with the help of recoding tables. Thus the variable provides time-consistent information on the economic activity based on the economic activity classification WS73. A detailed description can be found in Eberle et al. (2011). Further information on the WS73 classification can be found in the description of variables w73_3.

### 5.5.9 Economic activity 73 generated – type of completion (group\_w73\_3)

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

### 5.5.10 Economic activity 93 generated – completed by extrapolation / imputation (w93\_3\_gen)

Variable label	W93_3 Completed by extrapolation / imputation
Variable name	W93_3_gen
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	section (1-digit code)   division (2-digit code)   group (3-digit code)   class (4-digit code)   sub-class (5 digit code) of economic activity
Detailed description	This variable indicates the economic activity as a 3-digit code in accordance with the WZ93 classification. From 1998 up to and including 2003, the variable contains the original values from w93_3. Before 1998 and after 2003, the information is either written back / continued or replaced with the help of recoding tables, so that the variable contains time-consistent information on the economic activity based on the economic activity classification WS93. A detailed description can be found in Eberle et al. (2011).

	Further information on the WS93 classification can be found in the description of variables w93_3.
--	--

#### 5.5.11 Economic activity 93 generated – type of completion (group\_w93\_3)

Variable label	Type of completion w93_3
Variable name	group_w93_3
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	None
Detailed description	<p>This variable indicates the type of completion for the w93_3_gen variable. It is denoted whether the relevant value in w93_3_gen is consistent with the original value from w93_3, still missing / extrapolated or imputed based on recording tables.</p> <p>A detailed description of the procedure can be found in Eberle et al. (2011).</p>

#### 5.5.12 Year of first appearance of establishment number (grd\_jahr)

Variable label	Year of first appearance of establishment number
Variable name	grd_jahr
Category	establishment variables
Origin	BeH
Data type	numerical
Hierarchy	none
Detailed description	<p>This variable indicates the first appearance of the establishment number in the dataset.</p> <p>If an establishment number is only determined for the first time after 1975 (or after 1992 in Eastern Germany), this variable could indicate the date when the respective establishment was founded. However, it could also be an establishment that has been in existence for a longer time but has been allocated a new establishment number following a change of owner or a change in the legal form of the establishment. (For the allocation of establishment numbers see Bundesagentur für Arbeit 2007, p. 9-11). It might also be that the establishment already existed before, but did not have any employees liable to social security, or from 1999 onwards, neither marginal part-time workers.</p>

#### 5.5.13 First appearance of establishment number (grd\_dat)

Variable label	First appearance of establishment number
Variable name	grd_dat
Category	establishment variables
Origin	BeH
Data type	date
Hierarchy	none

Detailed description	This variable indicates the first appearance of the establishment number in the BeH to the day. If an establishment number in Western Germany is determined for the first time only after 1975, or after 1992 in Eastern Germany, this variable could indicate the date when the respective establishment was founded. However, it could also be an establishment that has been in existence for a longer time but has been allocated a new establishment number following a change of ownership or a change in the legal form of the establishment. (For the allocation of establishment numbers see Bundesagentur für Arbeit 2007, p. 9-11). It might also be that the establishment already existed before, but did not have any employees liable to social security, or from 1999 onwards, neither marginal part-time workers.
Notes	Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, the economic activity is only shown as the 3-digit code (grd_jahr).

#### 5.5.14 Year of last appearance of establishment number (lzt\_jahr)

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

#### 5.5.15 Last appearance of establishment number (lzt\_dat)

Variable label	Last appearance of establishment number
Variable name	lzt_dat
Category	establishment variables
Origin	BeH
Data type	date
Hierarchy	none

Detailed description	This variable indicates the last appearance of the establishment number in the dataset to the date (see Bender et. al. 1996). If the existence of an establishment number in the BHP already ends before 2008, it could indicate the closure of the establishment. However, other possible reasons for this are an “arbitrary change of the establishment number following a change of ownership or a change in the legal form of the establishment”, the “outsourcing of parts of the firm under a new establishment number” or other administrative changes (see Bender et al. 1996 or Bundesagentur für Arbeit 2007, p. 9-11).
Notes	Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, the economic activity is only shown as the 3-digit code (lzt_jahr).

#### 5.5.16 Total number of employees (az\_ges)

Variable label	total number of employees
Variable name	az_ges
Category	generated establishment variables
Origin	BeH
Detailed description	This variable contains the total number of an establishment’s employees reported to the social security agencies as of 30 June of a year. Since the introduction of the new notification regulations in 1999, people in marginal part-time employment have also been recorded. Part-time employees and dormant employment relationships (daily wage of zero) are also included.

#### 5.5.17 Number of full-time employees (az\_ges\_vz)

Variable label	total number of full-time employees
Variable name	az_ges_vz
Category	generated establishment variables
Origin	BeH
Hierarchy	none
Detailed description	This variable contains the number of an establishment’s full-time employees on 30 June of a year. People in marginal part-time employment and dormant employment relationships (daily wage of zero) are not included here.

#### 5.5.18 Number of employees in marginal part-time employment (az\_gf)

Variable label	number of employees in marginal part-time employment
Variable name	az_gf
Category	generated establishment variables
Origin	BeH
Hierarchy	none
Detailed description	The number of employees in marginal part-time employment is generated using the person group code – values 109 and 209. This variable has only been contained in the dataset since 1999 as it has only been included in the social security notification procedure since that year.

### 5.5.19 Median gross wage of all full-time employees (te\_med)

Variable label	median gross wage of all full-time employees
Variable name	te_med
Category	generated establishment variables
Origin	BeH
Data type	numerical
Hierarchy	none
Detailed description	<p>This variable contains the median gross daily wage of the full-time employees of a company. Employees in marginal part-time employment and dormant employment relationships (daily wage of zero) are not included. The values are in euros for all years.</p> <p>According to the social insurance notification regulations, employers must indicate the amount of gross income subject to contributions of the employer for a certain period of time (income for a period of time). Until the end of 1998, employers had to report the gross wage subject to social insurance contributions only. So only income above the marginal threshold and below the contribution assessment ceiling were recorded. Since 1999, also incomes below the contribution assessment ceiling have been recorded as part of the new notification procedure. Gross wages above the contribution assessment ceiling, however, are still cut. In order to determine the gross daily wage, the income for a period of time is divided by the number of calendar days of that period, and this value is rounded to two decimal places. These data are then aggregated on the establishment level.</p>

## 5.6 Location data

### 5.6.1 Place of residence: district (Kreis) (wo\_kreis)

Variable label	place of residence: district (Kreis)
Variable name	wo_kreis
Category	location data
Origin	BeH, LeH, LHG, XLHG, ASU, XASU
Data type	numerical
Hierarchy	federal state   district
Detailed description	<p>In BeH and LeH observations, the place of residence at district level is only available for the years from 1999 onwards. The variable indicates the district (urban district or rural district) in which the social security contributor lives. The first two digits of the 5-digit district code (Kreisschlüssel) show the code for the federal state (Bundesland), positions 1-3 indicate the regional authority (Regierungsbezirk), and positions 1-5 show the district authority (Kreis). Federal states without a regional authority have a 0 in the third position.</p> <p>In the BeH, the place of residence is determined at the end of each year and added consistently to all datasets of a year. For the LHG, XLHG and XASU sources, the place of residence applies to the period of the original observation. For the ASU and LeH, the characteristic contains the place of residence at the beginning of the original period</p>



	<p>of time.</p> <p>In order to guarantee consistent regional allocations across the entire observation period, the information on the district was recoded with reference to the territorial allocation of 31 December 2010 for all sources, i.e. in all calendar years, a place of residence is assigned to a district in accordance with the boundaries that the district had on 31 December 2010. As the district boundaries have changed over time, cases would occur in which the district code changes without the individual concerned having relocated if the territorial allocations of the districts were not updated.</p> <p>For observations with an original start date after 2010, the location data was set to “missing”, because changes in the territorial allocation implemented from 2011 onwards cannot be recoded precisely to the allocation of 31 December 2010. For LHG, XLHG and XASU observations, the variable is available from 2005 onwards.</p>
Notes	<p>Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, only the federal state (wo_bula) is shown as regional information.</p>

### 5.6.2 Place of residence: employment agency (wo\_aa)

Variable label	place of residence: employment agency
Variable name	wo_aa
Category	location data
Origin	BeH, LeH, LHG, XLHG, ASU, XASU
Data type	numerical
Hierarchy	regional directorate   employment agency
Detailed description	<p>This variable contains the employment agency of the employee's / BA client's place of residence. This information is extracted from the residence address. For the LHG, XLHG and XASU data sources, the place of residence is valid for the period of the original observation. In the case of the LeH and ASU, the variable contains the place of residence at the start of the period of unemployment or the job search. For the BeH and LeH, it is available from 1999 onwards, for the LHG, XLHG and XASU from 2005 onwards, and for ASU observations, the variable is already available before 1999.</p> <p>In order to guarantee consistent regional allocations across the entire observation period, the information on the employment agency was recoded to the territorial allocation of 31 December 2010 for all data sources observations, i.e. in all calendar years, a place of residence is assigned to an employment agency area in accordance with the boundaries that the employment agency area had on 31 December 2010. As the boundaries of the employment agency areas have changed over time, cases would occur in which a person's employment agency area changes without him/her having relocated if the territorial allocations were not updated. Berlin constitutes a problematic case with regard to updating territorial allocations, however: The boundaries of Berlin's employment agency areas have been changed repeatedly over the years, which could not be fully corrected even by recoding the territorial allocations. For observations with an original</p>

	start date after 2010, the location data was set to “missing”, because changes in the territorial allocation implemented from 2011 onwards cannot be recoded precisely to the allocation of 31 December 2010.
Notes	Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, only the area of the regional directorate in which the social security contributor’s place of residence is located is shown.

### 5.6.3 Place of residence: regional directorate (wo\_rd)

Variable label	place of residence: regional directorate
Variable name	wo_rd
Category	location data
Origin	BeH, LeH, LHG, XLHG, ASU, XASU
Data type	numerical
Hierarchy	regional directorate   employment agency
Detailed description	<p>This variable contains the regional directorate of the place of residence. For the LHG, XLHG and XASU data sources, the place of residence is valid for the period of the observation. A change of place of residence (already at the level of local employment office (Geschäftsstelle) or local authority district (Gemeinde)) causes a new observation. In the case of the LeH and ASU, the variable contains the place of residence at the start of the period of unemployment or the jobseeker status. The variable is available for the BeH and the LeH from 1999 onwards, and for the LHG, XLHG and XASU from 2005 onwards. For ASU observations, the variable is already available before 1999.</p> <p>For all data sources, the territorial allocations were corrected to the status of 31 December 2010. For observations with an original start date after 2010, the location data was set to “missing”, because changes in the territorial allocation implemented from 2011 onwards cannot be recoded precisely to the allocation of 31 December 2010.</p>

### 5.6.4 Place of residence: federal state (Bundesland) (wo\_bula)

Variable label	place of residence: federal state (Bundesland)
Variable name	wo_bula
Category	location data
Origin	BeH, LeH, LHG, XLHG, ASU, XASU
Data type	numerical
Hierarchy	federal state   district
Detailed description	<p>This variable is an aggregation of the “district” variable to the 16 German federal states.</p> <p>In BeH observations, the federal state of the place of residence is only available for the years from 1999 onwards. Further information on the district of the place of residence can be found under Place of residence: district (Kreis) (wo_kreis).</p>

### 5.6.5 Place of work: district (Kreis) (ao\_kreis)

Variable label	place of work: district (Kreis)
Variable name	ao_kreis
Category	location data
Origin	BeH
Data type	numerical
Hierarchy	federal state   district
Detailed description	<p>The variable indicates the district (urban district or rural district) in which the employee's establishment is located. The first two digits of the 5-digit district code (Kreisschlüssel) show the code for the federal state (Bundesland), positions 1-3 indicate the regional authority (Regierungsbezirk), and positions 1-5 show the district authority (Kreis). 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 December 2010, i.e. in all calendar years, a place of work is assigned to a district in accordance with the boundaries that the district had on 31 December 2010. 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>
Notes	Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, only the federal state (ao_bula) is shown as regional information.

### 5.6.6 Place of work: employment agency (ao\_aa)

Variable label	place of work: employment agency
Variable name	ao_aa
Category	location data
Origin	BeH, LeH, ASU
Data type	numerical
Hierarchy	regional directorate   employment agency
Detailed description	<p>For employment notifications from the BeH, the characteristic contains the employment agency of the job location. For the LeH, this characteristic indicates the employment agency responsible for the payment. The employment agency taking care for the customer is included in the ASU. In contrast to the "place of residence: employment agency" characteristic, this characteristic has been available for the BeH and LeH datasets since 1975.</p> <p>The characteristic includes the corrected territorial allocation as of 31 December 2010. However, this does not apply to the LeH and ASU. For observations with an original start date after 2010, the location data was set to "missing", because changes in the territorial allocation implemented from 2011 onwards cannot be recoded precisely to the</p>

	allocation of 31 December 2010.
Notes	Owing to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form on application and only in well-founded cases. Otherwise, only the area of the regional directorate (ao_rd) in which the social security contributor's place of residence is located is shown.

### 5.6.7 Place of work: regional directorate (ao\_rd)

Variable label	place of work: regional directorate
Variable name	ao_rd
Category	location data
Origin	BeH, LeH, ASU
Data type	numerical
Hierarchy	regional directorate   employment agency
Detailed description	<p>The variable indicates the regional directorate of the BA client's place of work. In contrast to the "place of residence: regional directorate" variable, the variable is available for BeH and LeH observations from 1975 onwards and for ASU observations from 2000 onwards.</p> <p>The characteristic includes the corrected territorial allocation as of 31 December 2010. For observations in 2011 and 2012, the location data was set to "missing", because changes in the territorial allocation implemented from 2011 onwards cannot be recoded precisely to the allocation of 31 December 2010.</p>

### 5.6.8 Place of work: federal state (Bundesland) (ao\_bula)

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

## 5.7 Generated technical variables

### 5.7.1 Source of the observation (quelle)

Variable label	source of the observation
Variable name	quelle
Category	generated technical variables
Origin	BeH, LeH, LHG, XLHG, ASU, XASU
Data type	numerical
Hierarchy	none
Detailed description	The variable indicates the data source.

### 5.7.2 Observation counter per person (spell)

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

### 5.7.3 Year (jahr)

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

## 6 References

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

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

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

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

**Deutsche BKK** (2012): Ratgeber Sozialversicherung 2012, Wolfsburg, 56 S., URL: [http://www.deutschebkk.de/fileadmin/user\\_upload/Service/Download-Center/Ratgeber\\_Sozialversicherung\\_2012\\_01.pdf](http://www.deutschebkk.de/fileadmin/user_upload/Service/Download-Center/Ratgeber_Sozialversicherung_2012_01.pdf), (28 March 2012).

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

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

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

**Meinken, Holger; Koch, Iris** 2004: BA-Beschäftigtenpanel 1998-2002. Codebuch, 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, 468-482.

## **7 Appendix**

### **7.1 Frequency tables**

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

## 8 List of abbreviations

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



IABS	IAB-Beschäftigtenstichprobe	IAB Employment Samples
IEB	Integrierte Erwerbsbiographien	Integrated Employment Biographies
IEBS	Stichprobe der Integrierten Erwerbsbiographien	Integrated Employment Biographies Sample
ISIC	International Standard Industrial Classification of All Economic Activities	International Standard Industrial Classification of All Economic Activities
ITM	IT- und Informationsmanagement des Instituts für Arbeitsmarkt- und Berufsforschung	IT and information management of the Institute for Employment Research
LIAB	Linked-Employer-Employee-Daten des IAB	Linked employer-employee data of the IAB
LeH	Leistungsempfänger-Historik	Benefit Recipient History
LHG	Leistungs-Historik Grundsicherung	Unemployment Benefit II Recipient History
LVA	Landesversicherungsanstalt	Land Social Insurance Office
MTH	Maßnahmeteilnehmer-Historik	Participants-in-Measures History File
NACE	Nomenclature générale des activités économiques dans les communautés européennes	Nomenclature générale des activités économiques dans les communautés européennes
Schwbg	Gesetz zur Sicherung der Eingliederung Schwerbehinderter in Arbeit, Beruf und Gesellschaft – Schwerbehindertengesetz	law to guarantee the integration of persons with severe disabilities into employment and society – Severely Disabled Persons Act
SGB	Sozialgesetzbuch	German Social Code
SIAB	Stichprobe der Integrierten Arbeitsmarktbiographien	Sample of Integrated Labour Market Biographies
VerBIS	Vermittlungs- und Beratungsinformationssysteme	Information System for Placement and Counselling
XASU	Arbeitsuchenden-Historik aus XSozial-BA-SGB II	Jobseeker History from XSozial-BA-SGB II
XLHG	Leistungsempfänger-Historik Grundsicherung aus XSozial-BA-SGB II	Unemployment Benefit II Recipient History from XSozial-BA-SGB II
zkT	Zugelassener kommunaler Träger	Authorised municipalities

## Imprint

FDZ-Datenreport 01/2013\_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, Heiner Frank

### Technical production

Heiner Frank

### All rights reserved

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

### Download

[http://doku.iab.de/fdz/reporte/2013/DR\\_01-13\\_EN.pdf](http://doku.iab.de/fdz/reporte/2013/DR_01-13_EN.pdf)

### Internet

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

### Corresponding author:

Philipp vom Berge  
Institute for Employment Research,  
Regensburger Str. 104,  
D-90478 Nuremberg  
<mailto:philipp.berge@iab.de>