

BASiD -

Biographical Data of Selected Social Security Agencies in Germany

Daniela Hochfellner,

(Institute for Employment Research (IAB) and Institute for Social Research at the University of Michigan (ISR))

Dana Müller, Anja Wurdack

(Institute for Employment Research (IAB))

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

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

Table of Contents

1	Introduction and Short Description	9
1.1	Introduction	9
1.2	Data Usage	10
1.3	Short Description	12
1.4	List of Variables	14
1.5	Content of the Variables	19
2	Data Sources	20
3	Data Preparation and Sampling Procedure	25
3.1	Correction and Validation Procedures	25
3.1.1	Employee History File – BeH	25
3.1.2	Benefit Recipient History – LeH	25
3.1.3	Unemployment Benefit II Recipient History (LHG/XLHG)	26
3.1.4	Jobseeker History (ASU/XASU)	26
3.1.5	Reassessment of the Overall File	26
3.2	Linking the Datasets	27
3.3	Structure of the Dataset	28
3.4	Sample Drawing and Projection	30
3.5	Missing Values	31
4	Time-Sensitive Personal Attributes (variable part)	32
4.1	Identifiers	32
4.1.1	Artificial Individual ID (persnr)	32
4.1.2	Artificial Establishment ID (betnr)	32
4.2	Period of Validity	33
4.2.1	Original Start Date of Observation (begorig)	33
4.2.2	Original End Date of Observation (endorig)	35
4.2.3	Original Start Date of Pension Insurance (BEGORIG)	36
4.2.4	Original End Date of Pension Insurance (ENDORIG)	36
4.2.5	Start Date of Split Episode (begepi)	37
4.2.6	End Date of Split Episode (endepe)	37
4.3	Personal Attributes	38
4.3.1	Gender (frau)	38
4.3.2	Date of Birth (gebdat)	38
4.3.3	Nationality (nation)	38
4.3.4	Nationality, Aggregated (nation_gr)	39
4.3.5	School Education and Vocational Training (bild)	39
4.3.6	School-Leaving Qualification (schbild)	41
4.3.7	Severe Disability Status (schweb)	41
4.3.8	State in employment history (zustand)	42
4.3.9	Type of Contribution (BYAT)	43
4.3.10	Type of Contribution, Aggregated (BYAT_GR)	44
4.3.11	Type of Insurance (VSGR)	44
4.3.12	Employment Status Prior to Job Search (estatvor)	45
4.4	Information on Employment, Benefit Receipt, and Job Search	45

4.4.1 Reason for Notification/ Reason for End of Benefit Receipt/ Reason for Discontinuation of Unemployment Benefit II / Reason for Deregistration (grund)	45
4.4.2 Daily Wage, Daily Benefit Rate (tentgelt)	46
4.4.3 Original Earnings points (EGPT)	47
4.4.4 Accredited Earnings points (EGPTAN)	48
4.4.5 Transition Zone (gleitz)	48
4.4.6 Occupation – Activity Performed/Most Recent Activity (beruf)	49
4.4.7 Occupational Status and Working Hours (stib)	50
4.4.8 Duration of Remaining Entitlement to Unemployment Benefit (restanspruch)	51
4.4.9 Type of Provider (traeger)	51
4.5 Location Data	52
4.5.1 Place of Residence: Employment Agency (wo_aa)	52
4.5.2 Place of Residence: District (wo_kreis)	54
4.5.3 Place of Residence: Federal State (Bundesland) (wo_bula)	55
4.5.4 Place of Work: Employment Agency (ao_aa)	55
4.5.5 Place of Work: District (ao_kreis)	56
4.5.6 Place of Work: Federal State (Bundesland) (ao_bula)	57
4.5.7 East-West Attributes (ow_knz)	57
4.5.8 Legal Basis to Determine Payment of Remuneration (RCEG)	58
4.6 Generated Technical Variables	59
4.6.1 Source of the Observation, grouped (quelle_gr)	59
4.6.2 Observation Counter per Person (spell)	59
4.6.3 Observation Counter per Episode (level2)	60
4.6.4 Observation Counter per Episode and Source (level1)	60
4.6.5 Type of Reassessment (berein)	60
4.7 Establishment Attributes	61
4.7.1 Economic Activity 73 (w73_3)	61
4.7.2 Economic Activity 93, 5-Digit Code (w93_5)	61
4.7.3 Economic Activity 93, 3-Digit Code (w93_3)	62
4.7.4 Economic Activity 03, 5-Digit Code (w03_5)	63
4.7.5 Economic Activity 03, 3-Digit Code (w03_3)	63
4.7.6 Economic Activity 08, 5-Digit Code (w08_5)	64
4.7.7 Economic Activity 08, 3-Digit Code 08 (w08_3)	64
4.7.8 Pension Receipt From Own Insurance (RTVS)	65
4.7.9 Month in the reference Period (ZRMO)	66
4.7.10 Parenting Period/Creditable Period (KI)	66
4.7.11 Application of special Regulations (KZSO)	67
4.7.12 Group of Employment for Application of FRG (BHBR)	68
4.7.13 FRG Group of Benefit/Group of Qualification (QLGR)	69
5 Personal Attributes, Due Date December 31 st 2007 (fixed part)	70
5.1 Artificial Individual ID (persnr)	70
5.2 Category of Persons (PSGR)	70
5.3 Category of Persons, Aggregated (PSGR_GR)	72
5.4 Type of Insurance (VSAT)	72
5.5 Account Balance Indicator (KTSD)	73
5.6 Year of Most Recent Notification on Adjustment of Account, with the Collaboration of the Insured Person (BXKLVS)	73

5.7	Time of Current Commencement of Retirement (ZTPTRTBE)	73
5.8	Partial Pension Indicator (TLRT)	74
5.9	Partial Pension Indicator, Aggregated (TLRT_GR)	75
5.10	Information on Earnings points	75
5.10.1	Sum of Earnings points – West (SUEGPT_WEST)	75
5.10.2	Sum of Earnings points – East (SUEGPT_OST)	76
5.10.3	Personal earnings points – West (PSEGPT_WEST)	76
5.10.4	Personal Earnings points – East (PSEGPT_OST)	76
5.10.5	Sum of Earnings points for All Contribution Periods – West (BZEGPT_WEST)	77
5.10.6	Sum of Earnings points for All Contribution Periods – East (BZEGPT_OST)	77
5.10.7	Sum of Earnings points for Non-Contributory Periods – West (BYFHEGPT_WEST)	77
5.10.8	Sum of Earnings points for Non-Contributory Periods – East (BYFHEGPT_OST)	78
5.10.9	Sum of Earnings points from Full-Value Contribution Periods – West (BYVLEGPT_WEST)	78
5.10.10	Sum of Earnings points from Full-Value Contribution Periods – East (BYVLEGPT_OST)	78
5.10.11	Sum of Additional Earnings points for Periods of Reduced Contribution – West (BYGMEGPTZQ_WEST)	79
5.10.12	Sum of Additional Earnings points for Periods of Reduced Contribution – East (BYGMEGPTZQ_OST)	79
5.10.13	Additional/Accredited Earnings points for Consideration Periods Due to Parenting and/or Caregiving – West (ZQEGPTKIPE_WEST)	79
5.10.14	Additional/Accredited Earnings points for Consideration Periods Due to Parenting and/or Caregiving – East (ZQEGPTKIPE_OST)	80
5.10.15	Months with Additional/Accredited Earnings points for Consideration Periods Due to Parenting and/or Caregiving – West (ZQMOKIPE_WEST)	80
5.10.16	Months with Additional/Accredited Earnings points for Consideration Periods Due to Parenting and/or Caregiving – East (ZQMOKIPE_OST)	80
5.11	Full-Value Contribution Periods – West (BYVL_WEST)	81
5.12	Full-value Contribution Periods – East (BYVL_OST)	81
5.13	Share of Contribution in Earnings points East (OPXAZ)	81
5.14	Pension-Relevant Periods (AR/AV and KN) (RTZTMO)	82
5.15	Consideration Periods (BUEZT)	82
5.16	Additional EGPT Taken Into Consideration for Consideration Periods (BUEZTEGPT)	82
5.17	Extra Pay From Settlement of Pension Entitlements (Bonus) (VAZU)	82
5.18	Deduction From Settlements of Pension Entitlements (Malus) (VAAB)	83
5.19	Credited Periods	83
5.19.1	Total Credited Periods (AZ)	83
5.19.2	Sum of Credited Periods Due to School-Based Training (SCHULAZ)	84
5.19.3	Sum of Assessed Credited Periods of Technical Training or Participation in a Work Preparation Scheme (FASCHULAZ)	84
5.19.4	Non-contributory supplementary period(ZZ)	84
5.20	Start Date of Refund (VNZRES)	85
5.21	End Date of Refund (BSZRES)	85
5.22	Information on Children	85
5.22.1	Total Number of Children (KIND)	85
5.22.2	Children's Birth Dates (GBK11 – GBK110)	86
5.22.3	Consideration in Variable Part – (GBKIZ1 – GBKIZ10)	86

5.23	Country for Foreign Pensions (FRGLD)	87
5.24	Extrapolation Factor (HRF)	87
5.25	Persons Originally Insured Under Miners' Insurance (VSKN)	87
6	BHP Establishment Data	88
6.1	Artificial Establishment ID (betnr)	88
6.2	Year (jahr)	89
6.3	Total Number of Employees (az_ges)	89
6.4	Proportion of Women in an Establishment (at_f)	89
6.5	Proportion of Low-Qualified Employees in an Establishment (at_gq)	89
6.6	Proportion of Employees Under 30 in an Establishment (at_u30)	90
6.7	Proportion of Employees Over 50 in an Establishment (at_u30)	90
6.8	Gross Daily Rate of Pay of All Full-Time Employees (te_p25, te_med, te_p75)	90

Abstract

Das Forschungsdatenzentrum der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung (FDZ BA/IAB) sowie das Forschungsdatenzentrum der Deutschen Rentenversicherung (FDZ-RV) bieten Längsschnittdaten an, welche aus personenbezogenen Sozialdaten stammen. Diese Daten enthalten sowohl Angaben aus dem Meldeverfahren zur Sozialversicherung als auch Daten aus dem Verwaltungsverfahren beider Institutionen. Da beide Einrichtungen nur die Daten halten, welche für die jeweilige Institution relevant sind, bietet ein Zusammenführen der Daten neue Analysepotenziale für die Forschung. Das Ziel des Projekts ist daher, einen kombinierten Datensatz zu erstellen, der sowohl die Informationen seitens der Deutschen Rentenversicherung Bund als auch die der BA bzw. des IAB vereint. So können auf der einen Seite Lücken in den Datenquellen der jeweiligen Institutionen gefüllt werden. Auf der anderen Seite kann durch den Abgleich der unterschiedlichen Datenquellen die Qualität von administrativen Daten erhöht werden. Der fertig gestellte Datensatz wird der Wissenschaft als Scientific Use File über das FDZ der RV und als schwach anonymisierte Version über das FDZ der BA im IAB im Gastaufenthalt zur Verfügung gestellt. Dieser Datenreport beschreibt die Biografiedaten ausgewählter Sozialversicherungsträger in Deutschland (BASiD) 1951 – 2009 für die schwach anonymisierte Version.

Abstract

The Research Data Centre of the Federal Employment Agency in the Institute for Employment Research (FDZ BA/IAB) and the Research Data Centre of the German Pension Insurance (FDZ-RV - GRV) offer longitudinal individual-level datasets. These datasets contain on the one hand information of the social security notifications and on the other hand characteristics of the administrative procedures of both institutions. In each institution, only information for accomplishment of their own current tasks is kept. The ambition of this project is to compile a common dataset which contains data of the GRV and the BA and/or, respectively, the IAB. The amount of information on individuals will be increased by filling the gaps in the single data sources by using the information of the other data source, which will provide new potential for scientific research. The combination of different data sources also supports the improvement of the quality of administrative records. The data are provided to the scientific community as a scientific use file as well as a weakly anonymised dataset accessible by on-site use.

This data report describes the biographical data of selected social insurance agencies in Germany (BASiD) 1951 – 2009.

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

Disclaimer

We would like to thank our colleagues at both Research Data Centres, who provided valuable input and comments for the creation of the dataset and the data report.

The data report at hand incorporates parts of the data report of the weakly anonymised version of the Sample of Integrated Labour Market Biographies (SIAB) (FDZ data report 01/2010) (cf. Dorner et al., 2010). Moreover, individual passages are based on IAB internal data documentation of IAB-ITM raw data. The description of the pension-relevant features is based on the data documentation of the Research Data Centre of the German Pension Insurance (Deutsche Rentenversicherung Bund).

This report is based in part upon work supported by the National Science Foundation under Grant Number SES-1326365. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

The authors of the data report at hand are solely responsible for its content and any possible errors.

1 Introduction and Short Description

1.1 Introduction

The aim of the BASiD project – “Biographical Data of Selected Social Security Agencies in Germany” – is to merge individual-level data of the Federal Employment Agency (BA), the Institute for Employment Research (IAB) and the German Pension Insurance (GRV) into one dataset (cf. Hochfellner et al., 2010). An advantage in the integration of the new dataset is that the data are all based on the same legal grounds, namely the German Code Book (Sozialgesetzbuch, SGB), and were therefore mainly generated by the same data generation process. The newly generated dataset “BASiD 1951 – 2009” contains not only individual-level data but also information on establishments. Researchers can access the weakly anonymised version via the Research Data Centre of the BA at the IAB, and the scientific use file via the Research Data Centre of the GRV. The scientific use file was created at the FDZ-RV. It must not be confused with the weakly anonymised version, which we will discuss here.

The basic dataset is the Sample of Insured Persons and their Insurance Accounts (VSKT) of the GRV, due date December 31st 2007. The VSKT is a disproportionate stratified random sample which was first taken in 1983 and has since then persisted as a panel. The basic population includes all insured persons of up to 67 years of age whose insurance account contains an entry at the due date and is neither closed nor is the insured person reported dead.

Everyone in the VSKT is identified in the IAB database. The data sources used here are the Integrated Employment Biographies (IEB) and the Establishment History Panel (BHP). The IEB include people who, during the observation period of 1975 – 2009, were employed and covered by social security at least once (data collected since 1975) or hold marginal part-time employment (data collected since 1999), or who received benefits within the legal scope of Social Code Book III (SGB III) (data collected since 1975) or Social Code Book II (SGB II) (data collected since 2005) or were registered as seeking employment at the Federal Employment Agency (BA) (data collected since 2000), or participated in a labour market policy measure (data collected since 2000). The information listed here comes from various data sources, which are merged together in the IEB. While the information on employment covered by social security and marginal part-time employment stems from the Employee History (BeH), the reception of benefits within the legal scopes of SGB III and SGB II is recorded in the Benefit Recipient History (LeH) and/or the Unemployment Benefit II Recipient History

(LHG and XLHG). The Jobseeker Histories (ASU and XASU) serve as the data source for the periods of job-seeking as recorded by the BA. Any participation in labour market policy measures is recorded in the Participants-in-Measures History File (MTH). The BHP includes information on the employing establishments, due date June 30th, between 1975 and 2008.

1.2 Data Usage

The BASiD data in the weakly anonymous version may only be analysed in the context of a on-site use at the FDZ and subsequent remote execution. 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 website¹.

The BASiD dataset is divided in three parts. They include time-sensitive personal attributes (variable part of the data), personal attributes, due date December 31st 2007 (fixed part of the data), and the establishment data from the Establishment History Panel (BHP). Time-sensitive personal attributes include socio-demographic features as well as on-time information on employment, benefit receipt and job-seeking, regional and some technical features. Moreover, this part of the data also includes all periods of compulsory contribution, credited periods and periods of consideration, such as information on parenting, illness, and pension. The personal attributes, due date December 31st 2007, include, among others, the number of children and the children's birth dates as well as other aspects which factor into the assessment of standard benefits of pensions, such as personal earnings points which insured persons receive for their pension-relevant periods. The establishment file includes information on the size of the establishment as well as its employee structure, due date June 30th of every year.

These three parts of the data can be merged into one dataset. The time-sensitive personal attributes and the personal attributes, due date December 31st 2007, are linked through the artificial personal ID (persnr). The establishment file can be linked via the establishment number (betnr) and the year of the dataset. The establishment data already contain a year indicator (jahr). In the individual dataset, this indicator must still be created.

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

¹ http://fdz.iab.de/de/FDZ_Data_Access.aspx

explicitly in the application for data access. The variables which are particularly sensitive for BASiD from the viewpoint of data protection legislation are:

- nationality (“nation”)
- severe disability status (“schweb”)
- place of residence: employment agency area (“wo_aa”)
- place of residence: district (“wo_kreis”)
- place of work: employment agency (“ao_aa”)
- place of work: district (“ao_kreis”)
- economic activity 93 – sub-class of economic activity (“w93_5”)
- economic activity 03 – sub-class of economic activity (“w03_5”)
- economic activity 08 – sub-class of economic activity (“w08_5”)
- type of contribution (“BYAT”)

1.3 Short Description

Current data version	Biographical Data of Selected Social Security Agencies in Germany (BASiD) 1951 – 2009
Type and volume of the data	Disproportionate stratified random sample by agency, gender, nationality and year of birth
Outline of the contents	<p>Insurance Account Sample (Versichertenkontenstichprobe – VSKT): Information of all relevant circumstances, i.e. all periods of compulsory pension insurance contributions included in the insured person's annuity computation. Apart from employment and benefit receipt, these include parenting and sick periods or caregiving periods for family members.</p> <p>Employee History (Beschäftigten-Historik – BeH): Annual notifications and end-of-employment notifications submitted to the social security agencies for employees covered by social security and employees in marginal part-time employment.</p> <p>Benefit Recipient History (Leistungsempfänger-Historik – LeH): Information on benefit receipt in accordance with Social Code Book III (SGB III) for recipients of unemployment benefit, unemployment assistance, and maintenance allowance.</p> <p>Unemployment Benefit II Recipient History drawn from (Leistungs-Historik-Grundsicherung – (X)LHG): 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) as well as separate responsibilities, and authorised municipalities and separate responsibilities)</p> <p>Jobseeker History (Arbeitsuchenden-Historik – (X)ASU): Information on job-seeking activity (applicant pool data) as well as job-seeking activity reported via the transmission standard XSozial-BA-SGBII to the BA by authorised municipalities and separate municipalities.</p> <p>Participants-in-Measures History File (Maßnahmenteilnahme-historik – MTH): Participation in measures to actively boost the labour market, such as job-creation measures and structural adjustment measures.</p> <p>Sensitive variables: Severe disability status (schweb) Place of residence: employment agency (wo_aa) Place of residence: district (wo_kreis) Nationality (nation) Place of work: employment agency (ao_aa) Place of work: district (ao_kreis) Economic activity – sub-class of economic activity (w93_5) Economic activity – sub-class of economic activity (w08_5) Economic activity – sub-class of economic activity (w08_5) Type of contribution (BYAT)</p>

Variables with regional relevance	Place of work and place of residence, federal states, districts, employment agency areas. (Some of these variables are not available for all data sources and periods.)
Type of territorial allocation	BeH, LeH, MTH, (X)ASU, (X)LHG, VSKT, with corrected territorial allocation (date: Dec 31 st 2009)
Period covered	The period covered depends on the data source. BeH Jan 1 st 1975 – Dec 31 st 2009 (2006, 2007: 18-month file; 2008: 12-month file, 2009: 6-month file) LeH Jan 1 st 1975 – Oct 14 th 2010 (Account by client number) LHG Jan 1 st 2005 – Oct 14 th 2010 XLHG Nov 1 st 2005 – Oct 14 th 2010 ASU Jan 1 st 2000 – Sep 14 th 2010 XASU Jan 1 st 2005 – Oct 14 th 2010 MTH Jan 1 st 2000 – Oct 14 th 2010 VSKT Jan 1 st 1951 – Dec 31 st 2007 The population of the individual attributes may differ from the depicted validity periods in individual cases. This is no mistake in the data. Before the official validity period, the respective attributes were ascertained on a voluntary basis from individual agencies.
Data unit	Individuals with at least one entry in their insurance account at pension insurance, due date December 31 st 2007.
Number of cases	568,468 individuals
Corrections – outline	The data were taken from VSKT, the IEB and the BHP, and edited some more. Some variables were recoded, for example.
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 BASiD data consist of three partial datasets which are filed individually. These include time-sensitive personal attributes, personal attributes (due date December 31 st 2007), and the establishment attributes from the BHP.
File format and size	STATA (time-sensitive personal attributes approx. 4 GB, personal attributes with due date December 31 st 2007 approx. 100 MB, establishment attributes from the BHP approx. 500 MB) SPSS and SAS available on request
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	Biographical Data of Selected Social Security Agencies in Germany (BASiD) 1951 – 2009, 2011.

1.4 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 no information on this variable. Therefore, the filling degree is only coloured for the source X(ASU).

List of variables	Page	School	Employment	AL O	Search	Measure	Household	Voluntary	Pension
Identifiers	Fehler! Textmarke nicht definiert.								
Artificial individual ID (persnr)	32								
Artificial establishment ID (betnr)	32								
Period of validity	33								
Original start date of observation (begorig)	333								
Original end date of observation (endorig)	35								
Original start date of pension insurance (BEGORIG)	36								
Original end date of pension insurance (ENDORIG)	36								
Start date of split episode (begepi)	37								

End date of split episode (endept)	37									
Personal attributes	38									
Gender (frau)	38									
Date of birth (gebdat)	38									
Nationality (nation)	38									
Nationality, aggregated (nation_gr)	39									
School education and vocational training (bild)	39									
School-leaving qualification (schbild)	41									
Severe disability status (schweb)	41									
State in working life (zustand)	42									
Type of contribution (BYAT)	43									
Type of contribution, aggregated (BYAT_GR)	44									
Type of insurance (VSGR)	44									
Employment status prior to job search (estatvor)	45									
Information on employment, benefit receipt, and job search	45									
Reason for notification / reason for end of benefit receipt / reason for discontinuation of Unemployment Benefit II / reason for deregistration (grund)	45									
Daily wage, daily benefit rate (tentgelt)	46									
Original earnings points (EGPT)	47									
Accredited earnings points (EGPTAN)	48									
Transition zone (gleitz)	49									
Occupation – activity performed/most recent activity (beruf)	49									
Occupational status and working hours (stib)	50									
Duration of remaining entitlement to unemployment benefit (restanspruch)	51									
Type of agency (traeger)	51									
Location data	52									

Place of residence: employment agency (wo_aa)	52								
Place of residence: district (wo_kreis)	54								
Place of residence: federal state (Bundesland) (wo_bula)	55								
Place of work: employment agency (ao_aa)	55								
Place of work: district (ao_kreis)	56								
Place of work: federal state (Bundesland) (ao_bula)	57								
East-West attributes (ow_knz)	57								
Legal basis to determine payment of remuneration (RCEG)	58								
Generated technical variables	59								
Source of the observation, grouped (quelle_gr)	59								
Observation counter per person (spell)	59								
Observation counter per episode (level2)	60								
Observation counter per episode and source (level1)	60								
Type of reassessment (berrein)	60								
Establishment attributes	61								
Economic activity 73 (w73_3)	61								
Economic activity 93, 5-digit code (w93_5)	61								
Economic activity 93, 3-digit code (w93_3)	62								
Economic activity 03, 5-digit code (w03_5)	63								
Economic activity 03, 3-digit code (w03_3)	63								
Economic activity 08, 5-digit code (w08_5)	64								
Economic activity 08, 3-digit code (w08_3)	64								
Pension receipt from own insurance (RTVS)	65								
Period of months (ZRMO)	66								
Parenting period/creditable	66								

period (KI)									
Attribute application of regulations (KZSO)	67								
Group of employment for application of FRG (BHBR)	68								
FRG group of benefit/group of qualification (QLGR)	69								

Table 1: Population degree of time-sensitive personal attributes

List of attributes	Page	12/31/2007
Artificial individual ID (persnr)	70	
Category of persons (PSGR)	70	
Category of persons, aggregated (PSGR_GR)	72	
Type of insurance (VSAT)	72	
Account balance indicator (KTSD)	73	
Year of most recent notification on adjustment of account, with the collaboration of the insured person (BXKLVS)	73	
Time of current commencement of retirement (ZTPTRTBE)	73	
Partial pension indicator (TLRT)	74	
Partial pension indicator, aggregated (TLRT_GR)	75	
Information on earnings points	75	
Sum of earnings points – West (SUEGPT_WEST)	75	
Sum of earnings points - East (SUEGPT_OST)	76	
Personal earnings points - West (PSEGPT_WEST)	76	
Personal earnings points - East (PSEGPT_OST)	76	
Sum of earnings points for all contribution periods - West (BZEGPT_WEST)	77	
Sum of earnings points for all contribution periods - East (BZEGPT_OST)	77	
Sum of earnings points for non-contributory periods – West (BYFHEGPT_WEST)	77	
Sum of earnings points for non-contributory periods – East (BYFHEGPT_OST)	78	
Sum of earnings points from full-value contribution periods – West (BYVLEGPT_WEST)	78	
Sum of earnings points from full-value contribution periods – East (BYVLEGPT_OST)	78	
Sum of additional earnings points for periods of reduced contribution – West (BYGMEGPTZQ_WEST)	79	
Sum of additional earnings points for periods of reduced contribution – East (BYGMEGPTZQ_OST)	79	
Additional/accredited earnings points for consideration periods due to parenting and/or caregiving – West (ZQEGPTKIPE_WEST)	79	
Additional/accredited earnings points for consideration periods due to parenting and/or caregiving – East (ZQEGPTKIPE_OST)	80	
Months with additional/accredited earnings points for consideration periods due to parenting and/or caregiving – West (ZQMOKIPE_WEST)	80	
Months with additional/accredited earnings points for consideration periods due to parenting and/or caregiving – East (ZQMOKIPE_OST)	80	
Full-value contribution period – West (BYVL_WEST)	81	
Full-value contribution period – East (BYVL_OST)	81	
Share of East German Earnings points (OPXAZ)	81	
Pension-relevant periods (AR/AV and KN) (RTZTMO)	82	
Consideration periods (BUEZT)	82	
Additional EGPT taken into consideration for consideration periods (BUEZTEGPT)	82	

Extra pay from statutory equalisation of pensions (VAZU)	82	
Deductions from statutory equalisation of pensions (VAAB)	83	
Credited periods	83	
Total credited periods (AZ)	83	
Sum of credited periods due to school-based training (SCHULAZ)	84	
Sum of assessed credited periods of technical training or participation in a work preparation scheme (FASCHULAZ)	84	
Creditable additional period of coverage (ZZ)	84	
Start date of refund (VNZRES)	85	
End date of refund (BSZRES)	85	
Information on children	85	
Total number of children (KIND)	85	
Birth of first child (GBK11)	86	
Birth of second child (GBK12)	86	
Birth of third child (GBK13)	86	
Birth of fourth child (GBK14)	86	
Birth of fifth child (GBK15)	86	
Birth of sixth child (GBK16)	86	
Birth of seventh child (GBK17)	86	
Birth of eighth child (GBK18)	86	
Birth of ninth child (GBK19)	86	
Birth of tenth child (GBK110)	86	
Consideration in variable part – first child (GBKIZ1)	86	
Consideration in variable part – second child (GBKIZ1)	86	
Consideration in variable part – third child (GBKIZ3)	86	
Consideration in variable part – fourth child (GBKIZ4)	86	
Consideration in variable part – fifth child (GBKIZ5)	86	
Consideration in variable part – sixth child (GBKIZ6)	86	
Consideration in variable part – seventh child (GBKIZ7)	86	
Consideration in variable part – eighth child (GBKIZ8)	86	
Consideration in variable part – ninth child (GBKIZ9)	86	
Consideration in variable part – tenth child (GBKIZ10)	86	
Country for foreign pensions (FRGLD)	87	
Extrapolation factor (HRF)	87	
Persons originally insured under miners' insurance (VSKN)	87	

Table 2: Population degree of personal attributes, due date December 31st 2007

List of attributes	Page	Employment
Artificial establishment ID (betnr)	88	
Year (jahr)	90	
Total number of employees (az_ges)	89	
Proportion of women in an establishment (at_f)	89	
Proportion of low-qualified people in an establishment (at_gq)	89	
Proportion of people under 30 in an establishment (at_u30)	90	
Proportion of people over 50 in an establishment (at_ü50)	90	
Gross daily rate of pay of all full-time employees (te_p25)	90	
Gross daily rate of pay of all full-time employees (te_med)	90	
Gross daily rate of pay of all full-time employees (te_p75)	90	

Table 3: Population degree of establishment attributes

1.5 Content of the Variables

Another characteristic of the BASiD data is that some variables differ in content, depending on their data source. Above all, the variables “state” (zustand) and “source” (quelle_gr) should be noted. During the merging of the data, identical or additional states of employment could be identified in the RV data and the data of the BA and/or the IAB; they are filed under the “state” variable. In order to find out whether the information on employment situation stems from the IAB or the RV database, the “source” variable may be used. If identical information is found in both IAB and RV data, this information is saved as a twin. If gaps were filled or parallel states were recorded in the data, the source will contain the dataset where the attribute in question comes from. Due to this structure, it is no longer possible to draw a direct conclusion as to the “state” by using only the “source” variable. While the BeH source only contains employment information, that is not true for information from the sources VSKT or twin. In order to find out whether, for example, there is employment present for a twin, both the state and the source must be considered. The following diagram depicts the mapping of states and the individual data sources.

Source of the observation, aggregated	BeH	LeH	(X)LHG	(X)ASU	MTH	VSKT	Twin
Employment state							
School-based training							
Vocational training							
Military/civilian service							
Employment subject to social security contributions							
Marginal part-time employment							
Self-employment							
Unemployment Benefit (ALG), Unemployment Benefit I (ALG I)							
Maintenance allowance (UHG) and other financial assistance pursuant to SGB III							

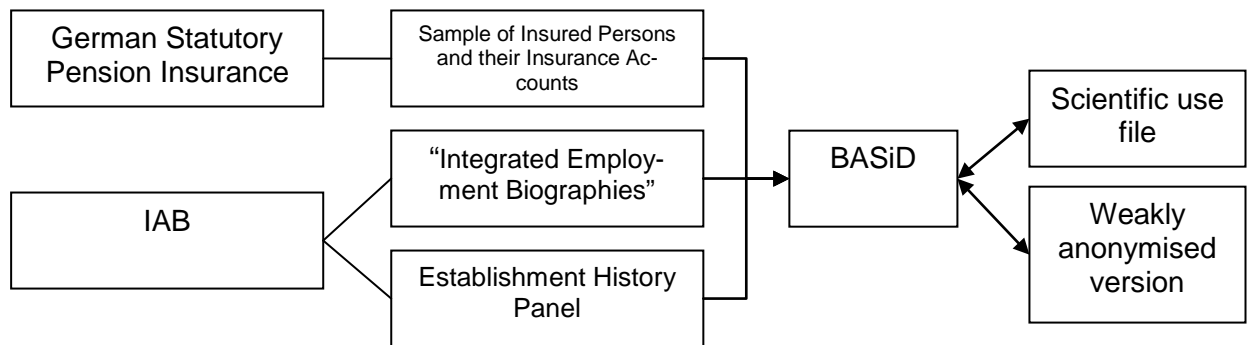
Unemployment Assistance (ALHI), Unemployment Benefit II (ALG II)							
ALG credited periods only							
Job-seeking (not registered as unemployed)							
Job-seeking (registered as unemployed)							
Improve prospects in the primary labour market (measure)							
Promotion of vocational training (measure)							
Measures accompanying employment (measure)							
Employment creation measures (measure)							
Discretionary support (measure)							
Other support (measure)							
Illness, inability to work							
Maternity, parenting, household							
Caregiving on a non-commercial basis, caregivers							
Voluntary insurance, subsequent insurance, refund on grounds of marriage, other periods							
Compulsory contribution periods of the EU							
Reduced earning capacity							
Part-time work for older employees and early retirement							
Pension, pension receipt							

Table 4: Mapping state – source

All data that this project is based on are taken from the social security notification procedure on the one hand and, on the other hand, from the respective institutions' administrative processes. The information is collected as part of the institutions' requirements to fulfil their lawful purpose, including the creation of statistics and information of the public. The data are process-produced, which means that they have a decisive advantage over survey data: they are not based on information provided by the respondents themselves, i.e. there is no danger of people having memory losses or no existence of panel attrition and non response.

2 Data Sources

The basic dataset für the project “Biographical Data of Selected Social Security Agencies in Germany – BASiD” is the Sample of Insured Persons and their Insurance Accounts (VSKT). In order to create the dataset, all persons and their existing information as contained in this sample are identified accordingly in the IAB database. The data sources of the IAB include the Integrated Employment Biographies (IEB) and the Establishment History Panel (BHP). The diagram below helps to visualise the setup of the BASiD dataset.



Source: proprietary

Figure 1: Data sources of the BASiD project

Below, we will further elaborate on the included data sources.

Sample of Insured Persons and their Insurance Accounts– (VSKT)

The VSKT is a special elicitation by the German Statutory Pension Insurance which contains all pension-relevant information on its insured persons (cf. Stegmann 2006). The VSKT includes living insured persons who, on December 31st of the respective reporting year (in the case of the BASiD project, 2007), are of the working age of up to 67 and have an insurance account with at least one entry. This means that the insurance account must neither be closed nor contain a registered date of death, and must either show a contribution or receipt of benefits. The VSKT is conceived as panel. The first sample was collected in 1983 in the form of a disproportionate stratified 1% random sample according to agency (blue collar workers, white collar workers, and miners and mine employees, sex, nationality (German/foreigner) and year of birth, taken from the RV insurance accounts. Ever since then, the sample has been continued as a panel. The panel form was chosen in order to be able to access clarified accounts on the one hand, which implies very good data quality, and on the other hand to avoid the effort of clarifying accounts, which would otherwise be incurred if the data were extracted annually from the general population of all insured individuals. For a more detailed description of the VSKT, please see Himmelreicher/Stegmann (2008) and Richter/Himmelreicher (2008).

Integrated Employment Biographies – IEB

The IAB uses the IEB (version V09.00.01 from March 25th 2011) to merge information to the VSKT data. The IEB contains five different sources: BeH, LeH, (X)LGH, MTH, and (X)ASU. Moreover, the IEB includes delayed acquisitions and subsequent corrections of the datasets. The time margin of the IEB differs from the one of the VSKT. With the IEB, it is possible to enlarge the time horizon of the VSKT by two years, because all information on employment is available until the end of 2009.²

Employee History File – BeH

The data source on employment is the Employment History File of the IAB. The data basis is the integrated notification procedure for health, pension and unemployment insurance, which came into effect as of January 1st 1973 (and was extended to cover Eastern Germany as of January 1st 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 individuals and regular students (see Cramer 1985) are generally not recorded in the BeH. Since the notification procedure was changed on January 1st 1999, employees in marginal part-time employment and unpaid family workers have also been recorded (not contained in the data until April 1st 1999). For every year that an individual is employed, at least one notification is available. 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 observed period for BeH is from January 1st 1975 to December 31st 2009.

² In fact, even longer. Parts of the data are follow-ups or planned training measures. A change of the information is possible for a changing employment status.

Benefit Recipient History – LeH

The Benefit Recipient History 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; i.e. no benefits within the scope 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.

The observed period for LeH spans from 1975 to 2010.

Unemployment Benefit II Recipient History – (X)LHG

The Unemployment Benefit II Recipient Histories (LHG/XLHG) contain the receipt of benefits in accordance with Social Code Book II (SGB II), other than LeH, which contains benefits in accordance with Social Code Book III (SGB III). Receipt of benefits within the scope of Unemployment Benefit II Recipient Histories cover both basic social security benefits (e.g. Unemployment Benefit II) and supplements to unemployment benefit or additional benefits. In contrast to benefits in accordance with SGB III, the BA is not necessarily the only responsible authority for administering the benefits. There are three different types of institutions responsible for implementing SGB II:

- Joint facilities/Cooperation of employment agencies and municipalities (ARGE), in which the BA and the municipality deal with tasks jointly.
- Separate responsibilities – the tasks continue to be divided between the BA and the municipality. 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) Sentence 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.
- Authorised municipalities also called opting local authorities or opting municipalities – the local authority is responsible for all tasks in the scope of SGB II.

An important difference compared with 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. It is not possible, however, to link individuals with benefit receipt under SGB II at the level of benefit

communities in BASiD. LHG and XLHG generally contain all registered individuals who are entitled to receive benefits in accordance with Section 7 SGB II, i.e. individuals in need 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 under the age of 64 are incorporated.

Each dataset contains non-overlapping periods of entitlement to benefit of a person in a certain benefit community. Regarding sanctions, it must be noted that the start and end dates of a sanction lead to a splitting of the records. 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 is that there is no appropriate value or variable which states the start, type, and duration of the sanctions. The LHG data source has been filled since establishing ALG II in the years from 2005 to 2010. Information from XLHG is available from November 2005 to 2010.

Participation-in-Measures History File – MTH

The MTH includes attendance of measures of active labour market programs. These include job-creation and structural adjustment measures as well as entry-level allowance (for recipients of Unemployment Benefit II) due to Section 29 SGB II, free facilitation, and measures of advanced vocational training. For the use of the IEB, it must be noted that information on participation in measures as well as job-seeking periods is fully available since January 1st 2000. There is also information for participation in measures before 2000, but these may very likely be incomplete. Generally, all observations with an entry date before 1995 are excluded.

Jobseeker History – (X)ASU

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

Establishment History Panel – BHP

The BHP includes all establishments in Germany with at least one employee liable to social security at June 30th of each year, or since January 1st 1999 with at least one marginal part time employee. For generating the BHP, yearly cross-sectional information is collected at the

reference date, for Western Germany since 1975 and for Eastern Germany since 1992, and linked to a panel (Spengler, 2009).

3 Data Preparation and Sampling Procedure

3.1 Correction and Validation Procedures

Before the data specified in Chapter 2 are merged, they undergo source-specific correction procedures (see the following chapters).

3.1.1 Employee History File – BeH

- Implausible dates are corrected.
- Registrations of employment covered by social security are filtered out, as the information on entering employment is also contained in the annual notifications.
- Cancellations are taken into consideration, and the corresponding observations are deleted. Notifications of modifications are implemented.
- Duplicates are removed according to the following criteria: insurance number, establishment ID, start and end date, reason for notification, payment, and date of data entry.
- Insurance numbers are updated effective April 8th 2008; closed numbers are deleted.

3.1.2 Benefit Recipient History – LeH

- For constructing the IEB, only observations from LeH for the period from January 1st 1975 (end date) until the latest available data are used.
- 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. Observations with no end date or an invalid end date are excluded from further processing, since in these cases it cannot be assumed that a benefit payment was made at all.
- Observations without a valid start date are excluded.
- Observations whose end date precedes the start date are excluded.
- If the end date for the receipt of unemployment assistance precedes the start date by one day, then the end date is increased by one year.
- Overlapping LeH observations are corrected for IEB in accordance with the following rules:
 - All overlaps in which both of the observations involve the same type of benefit and the same amount of benefit are corrected as follows: If one observation is completely embedded

within the period of another observation, the shorter observation will be deleted. If the two have the same length, one (randomly selected) observation will be deleted.

- In cases where observations partially overlap, the end date of the observation that began first is corrected.
- Overlaps of observations involving different types of benefit and/or different amounts of benefit paid are only corrected if the overlap concerns just one day and the original duration of both observations exceeds one day. In these cases, the end date of the observation that began first is put back by one day.

3.1.3 Unemployment Benefit II Recipient History (LHG/XLHG)

- Only observations of individuals who are capable of work and individuals under the age of 64 are used.

3.1.4 Jobseeker History (ASU/XASU)

- The used data from ASU cover the period from January 2000 to December 2010. The registered periods of job-search activity of the individuals are regarded as complete.
- The ASU data basis only distinguishes between observations with the status “unemployed” and “jobseeker,” if:
 - o a preceding observation with the status “unemployed” exists which joins the next observation without a gap and has “illness” as the reason for exit, and
 - o a subsequent observation with the status “unemployed” exists which also follows without a gap, and
 - o the observation itself does not have the status “unemployed” but “jobseeker”.
- Unlike ASU, XASU does not take into account periods of illness, as there is no information about illness available in the XASU data.
- Observations whose end date precedes the start date are filtered out.

3.1.5 Reassessment of the Overall File

After correcting the raw data and merging it into the basic IEB data, the overall IEB file is corrected and reassessed across all sources. The following reassessment steps are taken:

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

- Accounts in which the gender or date of birth change are corrected in accordance with priority rules.
- A territorial reassessment is carried out as of December 31st 2009.

3.2 Linking the Datasets

The data of both institutions are combined by using the social security number. Thanks to this unique identifier, identical persons can be identified in both institutions' databases. In addition, the start and end date of the notification, the reported wage or benefit receipt, and the state of employment is included in the process of linking employment biographies. However, only very few identical biographies (“twins”) are found when merging the single data on this basis. The reason for that are the different methods of data management within the two social security agencies. For identical information to be found, the identifiers have to be standardised first. The following modifications are made to the data material for the purposes of standardisation:

- Episode splitting: While the IAB data mainly contains annual notifications, the RV data material often contains shorter reporting periods. After episode splitting, the periods no longer overlap.
- All information was adapted to the periods of time modified through episode splitting.
- In the RV data, payment information was adjusted. This includes currency conversion into euros as well as evaluation of East-German payment according to Appendix 10 SGB VI.
- In the IAB data, all payments were converted to calendar days.

Identical information may now be found using the conditions “parallel observation time” “identical employment state” and “identical daily wage” of the respective biographical section. The diagram below shall serve to clarify this.

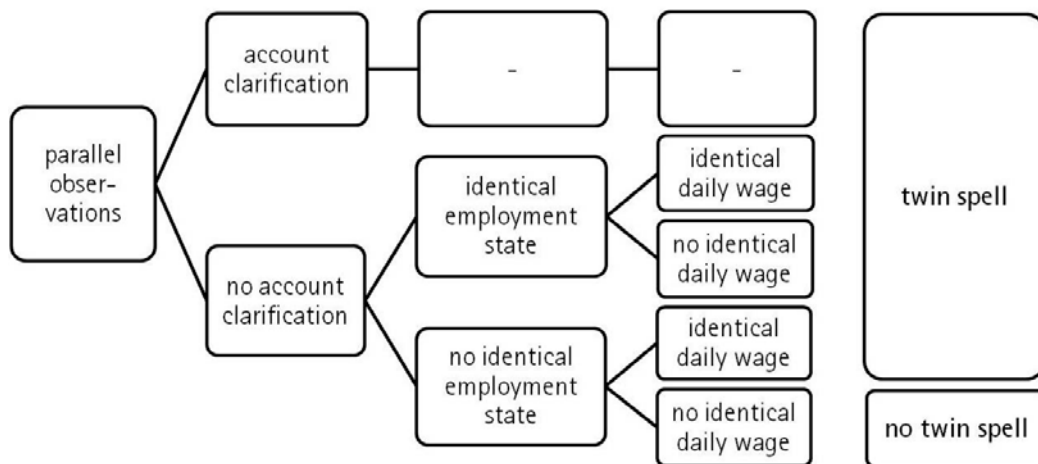


Figure 2: Linking the datasets

Pension insurance accounts are clarified on a regular basis, which means they are examined as to their plausibility and completeness. For biographical sections which have already been clarified, the data sources of IAB and RV were not reconciled. The biography was taken over from the RV database and was merely extended by including characteristics from the IAB data sources. Moreover, individual sections were marked as twins where no 100 % match could be determined. In these cases, reassessment was performed whenever the inconsistencies could be identified from the context or by reconciliation with the German Social Code of Law. This has been the way to correct wrong information. The type of correction that has been carried out is recorded in the “berein” variable (reassessment). All remaining observations not marked as twins have been retained as parallel observations in the data, if plausible.

3.3 Structure of the Dataset

To merge the data sources an episode splitting has been executed. Notifications within individuals' accounts in each of the data sources have to be rearranged according to their lengths. An episode splitting has been conducted to break up observations that overlap in time. The splitting process involves duplicating segments of observed periods of time and aligning start and end dates, so they are identical in all data sources for identical notifications. The variables referring to the original observed periods are adjusted to reflect the newly split time intervals. This procedure is shown in the following figure.

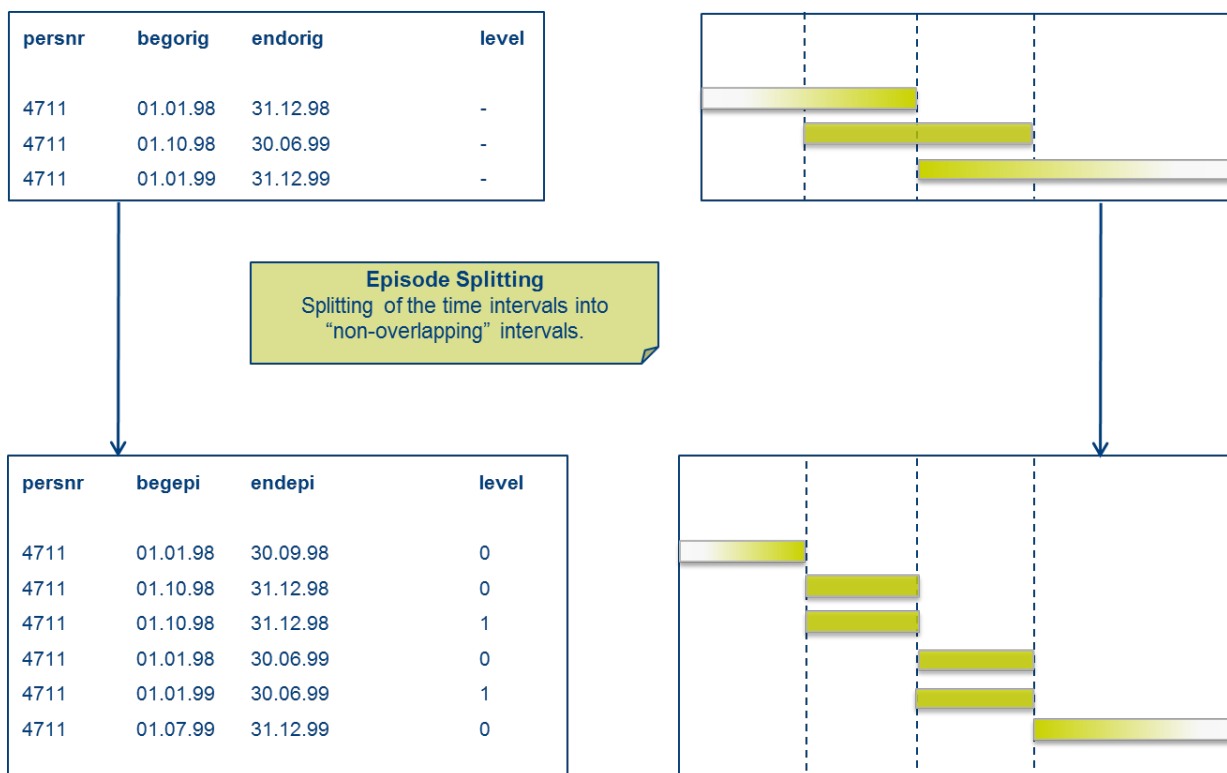


Figure 3: Episode splitting

The original date variables for the beginning and the end of the original observation ("begorig" and "endorig", or "BEGINORIG" and "ENDORIG") are retained. The variables for beginning and end of an episode ("begepi" and "endepi") mark the beginning and the end of the split episodes. It is possible to establish whether observations have been split by comparing the original period ("begorig" and "endorig") with the episode period ("begepi" and "endepi"). The BASiD data are sorted by person and episode counter (spell). The spell was generated using the following sorting:

1. Person
2. Start date of split episode
3. End date of split episode
4. State of split episode: employment, received benefit
5. Source of episode
6. Payment of remuneration: in descending order
7. Artificial establishment ID
8. East-West indicator

The episode counter assigns numbers to the episodes in one person's account, thus stating the existing number of episodes.

3.4 Sample Drawing and Projection

Since the VSKT serves as the basis of data merging, the structure of this sample also serves as a template for the BASiD data. The sample (VSKT) is a disproportionate stratified random sample according to agency, sex, nationality, and year of birth. The gross sample is drawn from RV master data. In another step, only those accounts are transferred to the VSKT which had an existing insurance account with the German Statutory Pension Insurance as of December 31st 2007 for which contributions were paid in 2007. The insured persons in the gross sample were identified in the IAB data, but only those who were integrated in the VSKT were integrated in the BASiD dataset. The net sample or basic population of BASiD data thus corresponds to all persons for whom pension insurance contributions were recorded in 2007 (see Fig. 4). They are 586,468 individuals.

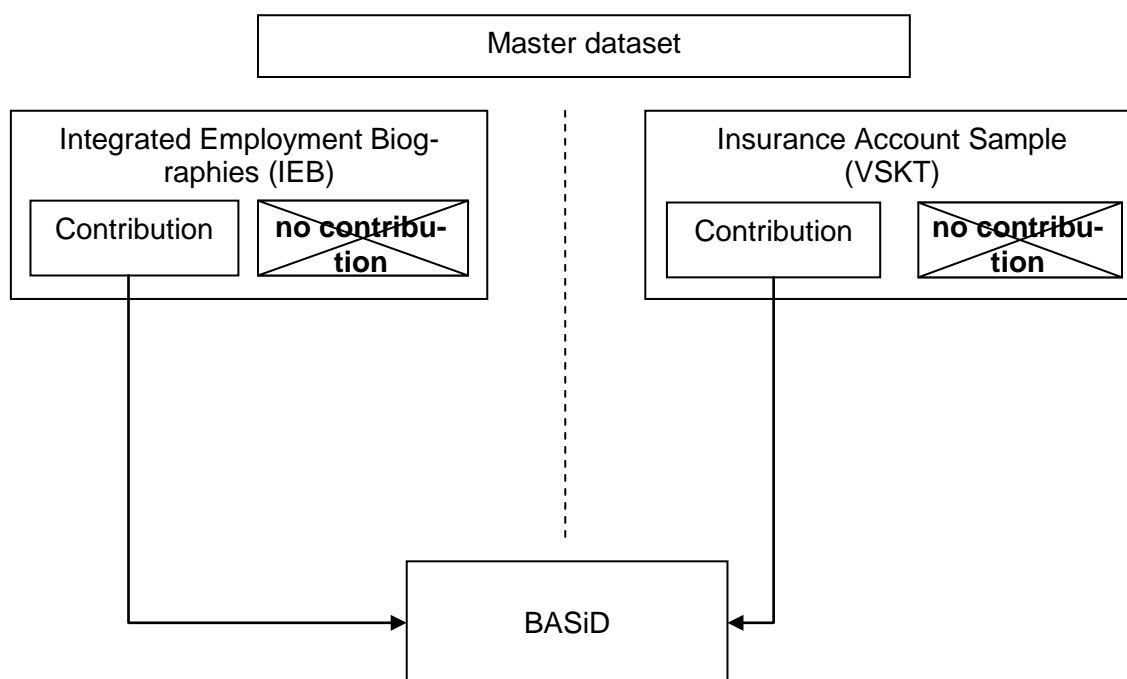


Figure 4: Sample drawing

It must be pointed out that not every person contained therein necessarily has an account with both institutions. An example would be a person who is self-employed but took out voluntary insurance. Due to these circumstances, this person can only be found in pension insurance data.

Due to the insurance numbers having been drawn from the pension insurance accounts in a disproportionate stratified manner, it is not possible to use the dataset is not representative of the German population in the longitudinal perspective. For example, it is more likely for women, foreigners, and individuals insured under social miners' and mine employees' insur-

ance to be drawn. However, there is an extrapolation factor which balances out this disproportionateness and extrapolates to the basic population. The weighting factor only exists for 2007. Therefore, the only representative evaluations are those for the 2007 reporting year. Due to the panel structure, people who are not alive at the current drawing date do not turn up in the data, either. They are replaced subsequently. For a more detailed description of the VSKT sample design, which also applies to the BASiD dataset, see Richter/Himmelreicher (2008).

3.5 Missing Values

In BASiD, missing values are coded as follows:

Label	Value	Description
No (valid) details available	.z	Values of a variable which are not missing systematically, i.e. the variable is, in principle, available for the data source, but no details are available for the value considered or cannot be interpreted reasonably
Not filled systematically	.n	A variable is not available for a data source in general (dark grey cells in the overview of variables) or is not available for a certain period
	0	Varies in reference to the respective variable (please refer to detailed description of variables)
	99	
	9999	No valid entry
	blank	System-missing value

4 Time-Sensitive Personal Attributes (variable part)

Frequency count and abstracts about individual values and labels of attributes can be found in separate files at: <http://fdz.iab.de>.

4.1 Identifiers

4.1.1 Artificial Individual ID (persnr)

Variable label	Artificial individual ID
Variable name	persnr
Category	Identifiers
Origin	Generated
Data type	Numerical
Hierarchy	None
Detailed description	The artificial individual ID is based on the social security number, because the social security number uniquely identifies an individual. However, to keep the individuals in the sample anonymous, an artificial number is developed.

4.1.2 Artificial Establishment ID (betnr)

Variable label	Artificial establishment ID
Variable name	betnr
Category	Identifiers
Origin	BeH
Data type	Numerical
Hierarchy	The artificial establishment ID is hierarchically sorted by establishment number zone, serial number, and error checking number.

Detailed description	<p>The artificial establishment ID enables an automatic reporting process to social security of employees liable to social security. The allocation of artificial establishment IDs for establishments with employees liable to social security has been centrally realised since January 1st 2008 by the Establishment Number Service (BNS) of the Federal Employment Agency (BA) in Saarbrücken; before that, it was assigned peripherally by employment offices. The artificial establishment ID is useful in two ways. On the one hand, employers can be identified at social security agencies; on the other hand, it is possible to draw conclusions about the employees of an establishment. Part of this is the assignment of an employee to a certain class of economic activity or region. Furthermore, the assignment of artificial establishment IDs follows some elementary rules. If the company has one office only in one municipality, this office is the establishment and receives an establishment number. If the company has several branch offices in one municipality, these offices are merged into a single establishment under one establishment ID if they belong to the same economic class. Vice versa, if they do not belong to the same economic class, each branch office is an establishment and receives an establishment ID. If the company has several branch offices in several municipalities, each of these branch offices is an establishment and receives an establishment ID.</p> <p>A new artificial establishment ID is assigned if an establishment is founded, the establishment owner changes, the establishment's economic class changes, or if the employer filed an application for amalgamation of several offices. No new artificial establishment ID can be assigned if the establishment changes local conditions, e.g. by relocating to another community. The old one remains valid until the health insurance or the employer demands a new artificial establishment ID. The artificial establishment ID is referred to as “non-active” if the establishment currently has no employees liable to social security. The artificial establishment ID can be reactivated if there are once more employees liable to social security and there are no objections by neither the health insurance company nor the employer. There is also the possibility to resume the artificial establishment ID when the employer founds a new establishment.</p> <p>The attribute “Artificial establishment ID” stems from IAB data and has been populated with employment notifications since 1975.</p>
----------------------	--

4.2 Period of Validity

4.2.1 Original Start Date of Observation (begorig)

Variable label	Original start date of observation from the Institute for Employment Research (IAB)
Variable name	begorig
Category	Period of validity
Origin	BeH, LeH, (X)LHG, (X)ASU, MTH
Data type	Date
Hierarchy	The original start date of IAB data is hierarchically sorted by day, month, and year.

Detailed description	<p>This attribute has its origin in IAB data and has been recorded since 1975. In this case, it deals with the original start date of the specific notification, passed on by the social security notifications (employer notification) to the BA, or in line with BA working processes, such as claiming unemployment benefit, attending subsidised advanced training, or registering as job-seeking.</p> <p>1. BeH Due to the fact that the employer is obliged to file annual notifications, the BeH notifications always include an identical start and end year. Therefore, all employees liable to social security employed at December 31st are included. The annual notification has to be filed until April 15th of the following year. Annual notification can be skipped if at the turn of the year a disruption notification is filed, or if deregistration is effected by December 31st. One possible reason for deregistration may be a change in insurance conditions.</p> <p>2. LeH For notifications relating to LeH, the attribute marks the end of benefit receipt.</p> <p>3. (X)LHG and 4. (X)ASU Moreover, it is possible to be employed and registered as job-seeking at the same time. This is due to the fact that an employee registers as job-seeking at the Federal Employment Agency (BA) three months before the end of the contract of labour. This has been obligatory since July 1st 2003. If this deadline is missed, unemployment benefit may be reduced.</p> <p>5. MTH If the source of the start date is MTH, this includes the beginning of measure participation. The given date is not necessarily the day on which the person went to see the placement officer, but the date of allocation. This is used e. g. in Section 37.</p> <p>6.VSKT As the attribute has its origin in IAB data, there are systematic missing values if the only information is provided by pension insurance.</p> <p>7. Twins Twins are identical information found by the IAB and pension insurance. They always stem from two data sources at once and they cannot be assigned uniquely. The notification type for twins is determined by the variable "state." This variable implies whether the twin describes employment, for instance (BeH). BeH, LeH and (X)LHG sets can each become twins of one VSKT block. This allows for conclusions as to the content.</p>
----------------------	---

4.2.2 Original End Date of Observation (endorig)

Variable label	Original end date of observation from the Institute for Employment Research (IAB)
Variable name	endorig
Category	Period of validity
Origin	BeH, LeH, (X)LHG, (X)ASU, MTH
Data type	Date
Hierarchy	The original end date of IAB data is hierarchically sorted by day, month and year.
Detailed description	<p>This attribute has its origin in IAB data and has been recorded since 1975. In this case, it deals with the original end date of the specific notification, passed on by the social security notifications (employer notification) to the BA, or in line with BA working processes, such as claiming unemployment benefit, attending subsidised advanced training, or registering as job-seeking.</p> <p>1. BeH Due to the fact that the employer is obliged to file annual notifications, the BeH notifications always include an identical start and end year. Therefore, all employees liable to social security employed at December 31st are included. The annual notification has to be filed until April 15th of the following year. Annual notification can be skipped if at the turn of the year a disruption notification is filed, or if deregistration is effected by December 31st. One possible reason for deregistration may be a change in insurance conditions.</p> <p>2. LeH For notifications relating to LeH, the attribute marks the end of benefit receipt.</p> <p>3. (X)LHG and 4. (X)ASU Moreover, it is possible to be employed and registered as job-seeking at the same time. This is due to the fact that an employee registers as job-seeking at the Federal Employment Agency (BA) three months before the end of the contract of labour. This has been obligatory since July 1st 2003. If this deadline is missed, unemployment benefit may be reduced.</p> <p>5. MTH If the end date has its origin in an MTH source, the exit date of the participation rate of the last entry date is usually included (if an exit date is contained at all). If not, it includes the exit date of the participation rate with the first entry date.</p> <p>6. VSKT As the attribute has its origin in IAB data, there are systematic missing values if the only information is provided by pension insurance.</p>

	<p>7. Twins</p> <p>Twins are identical information found by the IAB and pension insurance. They always stem from two data sources at once and they cannot be assigned uniquely. The notification type for twins is determined by the variable “state.” This variable implies whether the twin describes employment, for instance (BeH). BeH, LeH and (X)LHG sets can each become twins of one VSKT block. This allows for conclusions as to the content.</p>
--	---

4.2.3 Original Start Date of Pension Insurance (BEGORIG)

Variable label	Original start date of pension insurance
Variable name	BEGORIG
Category	Period of validity
Origin	VSKT
Data type	Date
Hierarchy	The original start date of RV data is hierarchically sorted by day, month, and year.
Detailed description	<p>1. BeH and 2. LeH and 3. (X)LHG and 4. (X)ASU and 5.MTH</p> <p>As the attribute has its origin in RV data, there are systematic missing values if the only information is provided by the IAB.</p> <p>6. VSKT</p> <p>The start date of the original data equals the original start date of social security notifications or of the pension insurance working process. The start date has been available since before 1975 in pension insurance data.</p> <p>7. Twins</p> <p>Twins are identical information found by the IAB and pension insurance. They always stem from two data sources at once and they cannot be assigned uniquely. The notification type for twins is determined by the variable “state.” This variable implies whether the twin describes employment, for instance (BeH). BeH, LeH and (X)LHG sets can each become twins of one VSKT block. This allows for conclusions as to the content.</p>

4.2.4 Original End Date of Pension Insurance (ENDORIG)

Variable label	Original end date of pension insurance
Variable name	ENDORIG
Category	Period of validity
Origin	VSKT
Data type	Date
Hierarchy	The original end date of RV data is hierarchically sorted by day, month, and year.

Detailed description	<p>1. BeH and 2. LeH and 3. (X)LHG and 4. (X)ASU and 5. MTH As the attribute has its origin in RV data, there are systematic missing values if the only information is provided by the IAB.</p> <p>6. VSKT The end date of the original data equals the original end date of social security notifications or of the pension insurance working process. The end date has been available since before 1975 in pension insurance data.</p> <p>7. Twins Twins are identical information found by the IAB and pension insurance. They always stem from two data sources at once and they cannot be assigned uniquely. The notification type for twins is determined by the variable "state." This variable implies whether the twin describes employment, for instance (BeH). BeH, LeH and (X)LHG sets can each become twins of one VSKT block. This allows for conclusions as to the content.</p>
----------------------	---

4.2.5 Start Date of Split Episode (begepi)

Variable label	Start date of split episode
Variable name	begepi
Category	Generated period of validity
Origin	BeH, LeH, (X)LHG, (X)ASU, VSKT
Data type	Date
Hierarchy	The start date of split episodes is hierarchically sorted by day, month, and year.
Detailed description	This attribute contains the start dates of all episodes and is generated by episode splitting. This method is used if periods within one account overlap. The original start date (begorig) is kept. The original start date equals the start date of the split episode if it is the first of several periods for one person. If not, the two variables are different.

4.2.6 End Date of Split Episode (endepi)

Variable label	End date of split episode
Variable name	endepi
Category	Generated period of validity
Origin	BeH, LeH, (X)LHG, (X)ASU, VSKT
Data type	Date
Hierarchy	The end date of split episodes is hierarchically sorted by day, month, and year.
Detailed description	This attribute contains the end dates of all episodes and is generated by episode splitting. This method is used if periods within one account overlap. The original end date (begorig) is kept. The original end date equals the end date of the split episode if it is the first of several periods for one person. If not, the two variables are different.

4.3 Personal Attributes

4.3.1 Gender (frau)

Variable label	Gender
Variable name	frau
Category	Personal attributes
Origin	BeH, LeH, (X)LHG, (X)ASU, MTH, VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>The gender is generated out of the social security number. Digits ten and eleven out of the 12-digit social security number are between 00 and 49 for men and between 50 and 99 for women.</p> <p>The sample uses gender dummies (0 – man, 1 – woman). The gender information is constant within one individual account.</p>

4.3.2 Date of Birth (gebdat)

Variable label	Date of birth
Variable name	gebdat
Category	Personal attributes
Origin	BeH, LeH, (X)LHG, (X)ASU, MTH, VSKT
Data type	Date
Hierarchy	None
Detailed description	<p>Date of birth is generated out of the social security number, placed at digits three to eight. In the dataset, the date of birth includes the month and the year and is constantly shown as yyyy.mm.</p>

4.3.3 Nationality (nation)

Variable label	Nationality
Variable name	nation
Category	Personal attributes
Origin	BeH, LeH, (X)LHG, (X)ASU, MTH
Data type	Numerical
Hierarchy	None
Detailed description	<p>This variable includes the nationality of a person. The variable contains the nation codes used by the Federal Statistical Office (Statistisches Bundesamt). This three-digit code is uniquely defined for each state. The code is relevant for social security notifications and for foreigners employed in Germany who need the code to register at a health insurance company.</p> <p>Further information on the national code can be found on the following page:</p> <p>http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/C</p>

	<p>on-tent/Klassifikationen/Bevoelkerung/Staatsangehoerigkeitsgebietsschlüssel,property=file.pdf</p> <p>As of: September 7th 2011</p> <p>There is no data on nationality for episodes of pension insurance with no corresponding twins in the IAB data (source: VSKT).</p>
Notes	<p>Due 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>

4.3.4 Nationality, Aggregated (nation_gr)

Variable label	Nationality, grouped
Variable name	nation_gr
Category	Personal attributes
Origin	BeH, LeH, (X)LHG, (X)ASU, MTH
Data type	Numerical
Hierarchy	None
Detailed description	<p>The nation codes used by the Federal Statistical Office are divided into 48 categories.</p> <p>There is no data on aggregated nationality for episodes of pension insurance with no corresponding twins in the IAB data (source: VSKT).</p>

4.3.5 School Education and Vocational Training (bild)

Variable label	School education and vocational training
Variable name	bild
Category	Personal attributes
Origin	BeH, (X)ASU, MTH
Data type	Numerical
Hierarchy	None
Detailed description	<p>Depending on the source, this attribute has different meanings:</p> <p>1. BeH</p> <p>For BeH observations, this variable contains school education/vocational training as 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). The school qualifications are to be established first and then the vocational qualifications. Only for the values 5 (degree from a university of applied sciences) and 6 (university degree), no other vocational qualifications are determined (see BA 2005, p. IX.).</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>For the remaining individuals, the vocational qualifications at the beginning of the job search are included. The following options are possible:</p> <p>2. LeH and 3. (X)LHG The attribute is empty for those sources.</p> <p>4. (X)ASU and 5. MTH ASU and XASU observations contain the completed vocational training at the beginning of the job-search activity.</p> <p>6. VSKT There is no data for announcements of pension insurance with no corresponding twins in the IAB data (source: VSKT)</p> <p>7. Twin Twins are identical information found by the IAB and pension insurance. They always stem from two data sources at once and they cannot be assigned uniquely. The notification type for twins is determined by the variable “state.” This variable implies whether the twin describes employment, for instance (BeH). BeH, LeH and (X)LHG sets can each become twins of one VSKT block. This allows for conclusions as to the content.</p> <p>For more information on how to fill in or correct missing values or inconsistent chronologies of school education and vocational training in the precursor IAB sample, see Fitzenberger et al. (2006) and Drews (2006). For this, only BeH sources are available, because it they are the only IAB data with information on education.</p> <p>The attribute has different meanings in different sources and therefore cannot be evaluated for all sources at the same time.</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 to social security (this affects people in marginal part-time employment, for example).” (Meinken/Koch 2004, p. 63).</p> <p>Especially since the 1990s, the missing values for younger cohorts have been increasing. One reason may be that announcement is not mandatory.</p>

4.3.6 School-Leaving Qualification (schbild)

Variable label	School-leaving qualification
Variable name	schbild
Category	Personal attributes
Origin	(X)ASU, MTH
Data type	Numerical
Hierarchy	School-leaving qualification is sorted hierarchically. First of all, it is established whether school-leaving qualification exists at all. Then it further differentiated whether school-leaving qualification includes general qualification for university entrance.
Detailed description	<p>This attribute includes school-leaving qualification at the beginning of unemployment and/or job-search, or the attribute is transmitted from MTH one day before the measure starts. After June 1st 2006, this attribute has always been generated from two criteria. One is "Schulabschluss aus Profilangebot" (school qualification from profile offer) collected until August 2008, and the other is "Schulabschluss aus Werdegang" (school qualification from career), collected since December 2008. If there is no information on school-leaving qualification in either of these, the school-leaving qualification in MTH is set to "Not specified." If only one attribute is populated, this value is put in for school-leaving qualification. If both attributes are populated, the higher school-leaving qualification is used.</p> <p>The attribute is only available for (X)ASU and MTH sources.</p>
Notes on quality	The degree of completeness in the XASU is insufficient. In 2005, it was 33 %, with a gradual increase over time.

4.3.7 Severe Disability Status (schweb)

Variable label	Severe disability status
Variable name	schweb
Category	Personal attributes
Origin	(X)ASU, MTH
Data type	Numerical
Hierarchy	The hierarchy of this variable is determined by the disability level, which could also coincide with an equal status to severe disability. In that case, the level above that is the severe disability.
Detailed description	This attribute determines disability according to the Severely Disabled Persons Act (Schwerbehindertengesetz – SchwbG). Disability is classified according to the level of disability (GdB) and is graded in steps of ten between the values of 20 and 100. Generally, disability is defined as permanent damage to the physical, psychological or mental state. The damage may be limited to one of these or combine several types of damage. In this case, the GdB is determined based on the total impact of the damage. A severe disability status is recognised from a level of disability of over 50 % (Section 1 SchwbG). Equal status to severely disabled is granted on application in accordance with

	<p>Section 2 SchwbG when the level of disability is between 30 and 50 %. This would result in a severe disability status valid from receipt of the application. However, the equal severe disability status is temporary, and the limit is set by the employer.</p> <p>For further information on SchwbG, please see:</p> <p>http://behinderung.org/gesetze/Schwbg.htm</p> <p>This attribute is only available for (X)ASU and MTH sources.</p>
Notes	Due 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 XASU in 2005 and 2006.

4.3.8 State in employment history (zustand)

Variable label	State in employment history
Variable name	zustand
Category	Generated technical variables
Origin	BeH, LeH, (X)LHG, (X)ASU, MTH, VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>This attribute indicates the state of each person in working life. It is a generated technical variable and consists of the employment situation in pension insurance and the employment status as recorded with the IAB. The following are possible states that can be achieved in the different sources.</p> <p>1. BeH</p> <p>BeH includes employees subject to social security contributions as well as efficiency and benefit characteristics of the employment. There is a difference between employees liable to social security and, since April 1st 1999, marginal part-time employees. Marginal part-time employees reported by household cheque procedure have been integrated in the reporting procedure since April 1st 2003. The household cheque procedure is a simplified reporting and commitment procedure with household cheques which serves for registration of a mini job in private households. Due to the fact that marginal part-time employees are included in the reporting procedure, the number of BeH datasets increases and more overlaps with other datasets occur, because it is possible that marginal part-time employees receive benefits.</p> <p>For further information on household cheque procedure, please see:</p> <p>http://www.minijobzentrale.de/sid_DAE15D5BA6DF6EB1C3742CE3C6A6FB4E/DE/haushaltsscheck/Node.html?__nnn=true</p>

	<p>2. LeH LeH indicates the benefit type defined in SGB II, summarised in the four main groups (ALG, ALHI, UHG).</p> <p>3. (X)LHG (X)LHG sets indicate receipt of benefit according to SGB II, i.e. the ALG II notifications.</p> <p>4. (X)ASU ASU datasets include the (generated) job-seeking status. This includes the calculated unemployment status while taking the illness status into account. XASU datasets contain the employment status. The value could either be “not unemployed, but job-seeking” or “job-seeking and unemployed.”</p> <p>5. MTH For MTH data sets, the employment status is developed from the attribute “Type of measure – differentiation” (“Maßnahmeart – Differenzierung“). It indicates the exact type of measure.</p> <p>6. VSKT VSKT information is based on the social employment situation. It includes the following information: school education and vocational training, caregiving on a non-commercial basis, parenting and household as well as inability to work and illness. Moreover, information on unemployment status, such as Unemployment Benefit II, unemployment benefits, and credited periods for pension insurance, is included. Other circumstances are military and civilian service, marginal part-time employment, self-employment, employment subject to social security, reduced earnings pension, and pension.</p> <p>7. Twins Twins are identical information found by the IAB and pension insurance. They always stem from two data sources at once and they cannot be assigned uniquely. The notification type for twins is determined by the variable “state.” This variable implies whether the twin describes employment, for instance (BeH). BeH, LeH and (X)LHG sets can each become twins of one VSKT block. This allows for conclusions as to the content.</p>
--	--

4.3.9 Type of Contribution (BYAT)

Variable label	Type of contribution
Variable name	BYAT
Category	Personal attributes
Origin	VSKT
Data type	Numerical

Hierarchy	None
Detailed description	This attribute includes detailed social security contributions to pension insurance. It is not included by default in the data, because it is contained in the variables "zustand" and "BYAT_GR."
Notes	Due to its particular sensitivity with regard to data protection legislation, this variable is only made available in non-aggregated form.

4.3.10 Type of Contribution, Aggregated (BYAT_GR)

Variable label	Type of contribution, grouped
Variable name	BYAT_GR
Category	Personal attributes
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	The aggregated form of this variable states if credited periods, compulsory contributions etc. exist. The attribute exists for notifications from pension insurance and twins (source: VSKT and twins).

4.3.11 Type of Insurance (VSGR)

Variable label	Type of insurance
Variable name	VSGR
Category	Personal attributes
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	The "Type of insurance" attribute is divided into subgroups: 1 "Pension insurance for wage-earners, AR" 2 "Pension insurance for salaried employees AV" 3 "Craftsman, AR" 4 "Craftsman, AV" 5 "German miners' pension insurance, KN, wage-earner" 6 "German miners' pension insurance, KN, salaried employee" If state = 20, 21, 26, "0" is acceptable if the type of insurance is unknown. The attribute exists for notifications from pension insurance and twins (source: VSKT and twins).

4.3.12 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	(X)ASU
Data type	Numerical
Hierarchy	None
Detailed description	<p>1. BeH and 2. LeH and 3. (X)LHG The attribute is not available for those sources.</p> <p>4. (X)ASU For ASU and XASU observations, this variable shows the employment status prior to the job-search activity. From December 2002 onwards, the values of the variable have been reduced substantially. Furthermore, the values of older observations were recoded to fit the currently valid values. They are therefore valid for the entire period. The specification is valid for the unsplit dataset and for the first split episode, because the attribute is assembled before the episode-splitting process.</p> <p>5. MTH and 6. VSKT and 7. Twins This attribute is not populated for those sources.</p>
Note on quality	A continuous decrease in the population degree can be observed.

4.4 Information on Employment, Benefit Receipt, and Job Search

4.4.1 Reason for Notification/ Reason for End of Benefit Receipt/ Reason for Discontinuation of Unemployment Benefit II / Reason for Deregistration (grund)

Variable label	Reason of notification
Variable name	grund
Category	Information on employment, benefit receipt and job search
Origin	BeH, LeH, (X)LHG, (X)ASU
Data type	Numerical
Hierarchy	None
Detailed description	<p>1. BeH In BeH observations, the “Reason of 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 IEB. For instance, 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</p>

	<p>notification. The reasons for submitting employment notifications are encoded according to the regulations of the notification procedure, which has been in effect since January 1st 1999 (in accordance with DEÜV).</p> <p>2. LeH In LeH observations, the variable indicates the reason for the end of receipt of unemployment benefit, unemployment assistance, or maintenance allowance. However, LeH contains no information about the reasons for the start of benefit receipt, as the information for LeH comes from the notifications submitted by the employment agencies to the health insurance institutions regarding finished periods of benefit receipt.</p> <p>3. (X)LHG The 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 refers to the individual, not to the benefit community. If receipt of Unemployment Benefit II by a different member of the benefit community than the one observed is discontinued, all observations of the members of the benefit community are split on this date, but the reason for discontinuation of Unemployment Benefit II is only available for the individual whose benefit is discontinued.</p> <p>4. (X)ASU In the case of ASU records, the variable contains the reason for de-registration or exit. It must be taken into account here that the number of values of the variable has been reduced from April 26th 2003 onwards. As far as possible, the old values were recoded into new values. The “Reason for notification” attribute has its origins in IAB data.</p> <p>5. MTH and 6. VSKT This variable is not available for those sources.</p> <p>7. Twins Twins are identical information found in the IAB and pension insurance data. They always stem from two data sources at once and they cannot be assigned uniquely to a specific source. The notification type for twins is determined by the variable “state.” This variable implies whether the twin describes employment or non employment, because BeH, LeH and (X)LHG sets can each be paired with one VSKT block.</p>
Note on quality	The degree of completeness of the reason for notification variable is poor in (X)LHG und (X)ASU data after 2005.

4.4.2 Daily Wage, Daily Benefit Rate (tengtelt)

Variable label	Daily wage
----------------	------------

Variable name	tentgelt
Category	Generated technical attributes
Origin	BeH, LeH, VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>1. BeH This attribute includes daily wage, generated from fixed period pay referring to the original duration of employment. All indications are consistently stated in euros, converted in reference to calendar days, and rounded. The wage specifications of pension data have already been evaluated pursuant to Appendix 10 SGB VI.</p> <p>2. LeH For LeH observations, the variable shows the daily benefit rate, converted into euros in each case. 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>
Note on quality	The daily wage for twins is only computed from the IAB data, even if it is available in both IAB and RV data sources. This is because IAB executed quality checks of the wage reportings, whereas the RV does not. If values in IAB data are missing, the RV values are used. This mechanism leads to more complete data.

4.4.3 Original Earnings points (EGPT)

Variable label	Original earnings points
Variable name	EGPT
Category	Personal attributes
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>This variable indicates the original earnings points for contribution periods. Increases due to payments in kind are already included. Not included are increases due to parenting, personal earnings points if the income is low, and additional earnings points for periods of reduced contributions.</p> <p>Consideration periods include earnings points from the assessment of standard benefits of pensions; where applicable, only the earnings points of the increase may be included.</p> <p>If there was no assessment of standard benefits, blanks are inserted.</p>

	<p>If state = 20, 21, the additional earnings points for consideration periods due to parenting or non-commercial caregiving to a care-dependent child (Section 70(3a) Letter a SGB VI, Section 83(1) Sentence 3 SGB VI) are included.</p> <p>If state = 26, the sum of AR/AV earnings points, which form the basis of the pension benefit, must be stated.</p> <p>The attribute is valid for notifications coming from pension accounts and twins only (source: VSKT and twins).</p>
--	---

4.4.4 Accredited Earnings points (EGPTAN)

Variable label	Accredited earnings points
Variable name	EGPTAN
Category	Personal attributes
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>This attribute indicates the credited contribution period of the original earnings points. In addition to EGPT, increases due to parenting are included, but earnings points if the income is low and additional earnings points for periods of reduced contribution are not included. Moreover, increases due to consideration periods are not included.</p> <p>Earnings points for employer's contribution according to Section 172(3) SGB VI for low-income employment without waiver of exemption from social security payments (state = 5) are not included in this attribute; if state = 5, the attribute is set to blank.</p> <p>If state = 20, 21, the additional earnings points for consideration periods due to parenting or non-commercial caregiving to a care-dependent child (Section 70(3a) Letter a SGB VI, Section 83(1) Sentence 3 SGB VI) are included.</p> <p>With non-contributory periods and consideration periods, the attribute contains blanks. This also applies to lower-ranking contribution periods coinciding with other contribution periods.</p> <p>If state = 26, the sum of credited AR/AV earnings points, which form the basis of the pension benefit receipt, must be stated.</p> <p>The attribute is valid for notifications coming from pension accounts and twins only (source: VSKT and twins).</p>

4.4.5 Transition Zone (gleitz)

Variable label	Transition zone
----------------	-----------------

Variable name	gleitz
Category	Information on employment, benefit receipt and job search
Origin	BeH
Data type	Numerical
Hierarchy	None
Detailed description	<p>This variable 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 (known as “midi jobs”), for which the employee has to pay only 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 April 1st 2003.</p> <p>The attribute contains information for notifications from BeH and for twins.</p>

4.4.6 Occupation – Activity Performed/Most Recent Activity (beruf)

Variable label	Occupation – activity performed/most recent activity
Variable name	beruf
Category	Employment/benefit receipt
Origin	BeH, (X)ASU, MTH
Data type	Numerical
Hierarchy	Occupation is hierarchically sorted by occupational field, chapter, group, category, and class.
Detailed description	<p>1. BeH</p> <p>The occupational title of the job performed by the employee during the notification period is a component of the employment details submitted by the employer.</p> <p>Information on activity is included in the activity code that the employer has to report. Digits one to three of the 5-digit code include information on the performed activity; digit four identifies the job position, and digit five, vocational training.</p> <p>To report the occupational title, the employer encodes the employee’s job in accordance with the reference work “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>2. LeH and 3. (X)LHG This attribute is not available for those sources.</p> <p>4. (X)ASU and 5. MTH For ASU observations, this variable contains the most recent occupation. For XASU observations, the variable contains the last successfully completed vocational training.</p> <p>7. Twins Twins are identical information found in the IAB and pension insurance data. They always stem from two data sources at once and they cannot be assigned uniquely to a specific source. The notification type for twins is determined by the variable “state.” This variable implies whether the twin describes employment or non employment, because BeH, LeH and (X)LHG sets can each be paired with one VSKT block.</p>
--	--

4.4.7 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.” Activity specifications are provided in the activity code. Digits one to three of the 5-digit code include information on the performed activity; digit four identifies the job position, and digit five, vocational training. The attribute at hand includes digit 4, i.e. the B1 code of the activity code, and makes the distinction between full-time and part-time employees possible.</p> <p>The decisive factor here is the ratio between the contracted hours and the regular 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 difference between blue-collar full-time employees (1, 2) and white-collar employees is established by the different types of pension insurance obligation (BfA for salaried employees and LVA for hourly 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 on-boarding training.</p> <p>If more than one code is possible for an employee, the employer is</p>

	<p>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). The attribute is only available for BeH sources and twins.</p>
--	--

4.4.8 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, MTH
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 changed to four years. If the end date of the benefit receipt is before January 1st 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>Both duration of unemployment and age of the person have certain effects on the duration of the covering and result in different maximum periods of time. If a person is younger than 50 years, the maximum period of time equals 12 months. Persons over 50 achieve, under certain circumstances, a maximum period of time of 24 months.</p> <p>The attribute is valid for notifications coming from pension accounts and twins only (source: VSKT and twins).</p>

4.4.9 Type of Provider (traeger)

Variable label	Type of provider
Variable name	traeger
Category	Information on employment, benefit receipt and job search
Origin	(X)LHG, (X)ASU, MTH
Data type	Numerical
Hierarchy	None
Detailed description	<p>This attribute includes not only the agencies of the BA (SGB III) but also the three municipal agencies (SGB II) which have existed since January 1st 2005. They include cooperation of employment agencies and municipalities (ARGE), authorised municipalities (zkT), also called "opting local authorities," and separate responsibilities (gT).</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), the relevant municipality takes on the tasks in the field of SGB II independently. In the case of separate responsibilities (gAw), also called</p>

	<p>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>(X)ASU observations contain the type of agency that holds the records of the applicant pool data. The information is added from the applicant pool data at the beginning of the unsplit sets. BA clients who on January 1st 2005 transformed into opting local authorities, therefore entering the legal sphere of SGB III, were deleted from the applicant pool data (the BeW sets) and from the table of applicant search status (ASU spells) on December 31st 2004. Afterwards, on January 1st 2005, a new set with the same status was applied. On the same date in 2006, these artificial ASU sets were closed again, but can still be identified by the end date (endorig = 1/16/2006) and the deregistration reason (grund = 81). They were transferred to IEB because they constitute relevant information about employment biographies for the time without reliable data from XSozial.</p> <p>The attribute is defined for data from sources (X)LHG, (X)ASU and MTH</p>
--	--

4.5 Location Data

4.5.1 Place of Residence: Employment Agency (wo_aa)

Variable label	Place of residence: employment agency
Variable name	wo_aa
Category	Location data
Origin	BeH, LeH, (X)LHG, (X)ASU, MTH
Data type	Numerical
Hierarchy	The hierarchical order is regional directorate (Regionaldirektion), employment agency, and local labour office (Geschäftsstelle). The variable "wo_aa" is on the employment agency level.
Detailed description	This variable contains the employment agency of the employee's/BA client's place of residence. In order to guarantee consistent regional allocations across the entire observation period, the information on the employment agency is recoded to reflect the most recent standard territory allocation. As the boundaries of the employment agency territories have changed over time, cases would occur in which a person's employment agency territory changed 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 territories have been changed repeatedly over the years, which could not be fully corrected even by recoding the territorial allocations. Moreover, until 1989, hardly any assignment to different employment agencies was carried out in Berlin. Instead, the default entry for most cases

	<p>was AA 901 (= Landesarbeitsamt Berlin).</p> <p>1. BeH The attribute states which employment agency region is in charge for administering unemployment insurance for the respective place of residence of the employee. In BeH datasets, the district of the employment agency has only been available since 1999.</p> <p>2. LeH In the unemployment benefit receipt notifications, the responsible local labour office at the place of residence of the unemployed is included. The attribute has been available from 1996 onwards.</p> <p>3. (X)LHG For (X)LHG sources, the place of residence of the recipient for the respective notification is contained in this attribute.. In case of a change of residence (even if only the local labour office or the community changed), the notification is split. It has been available for X(LHG) since 2005.</p> <p>4. (X)ASU For (X)ASU, the variable contains the place of residence of the job-seeker at the start of the period of unemployment or the job search. For (X)ASU, the attribute has been available since 2005; for ASU, since 1999.</p> <p>5. MTH For MTH data, the attribute includes the local labour office that is closest to the main place of residence of the benefit receiver.</p> <p>6. VSKT The attribute is not available for this source.</p> <p>7. Twins Twins are identical information found in the IAB and pension insurance data. They always stem from two data sources at once and they cannot be assigned uniquely to a specific source. The notification type for twins is determined by the variable "state." This variable implies whether the twin describes employment or non employment, because BeH, LeH and (X)LHG sets can each be paired with one VSKT block.</p>
Notes	<p>Due 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.</p>

4.5.2 Place of Residence: District (wo_kreis)

Variable label	Place of residence: district
Variable name	wo_kreis
Category	Location data
Origin	BeH, LeH, (X)LHG, (X)ASU, MTH
Data type	Numerical
Hierarchy	The hierarchical order is federal state, administrative region, district, and municipality. The variable “wo_kreis” is on the district level.
Detailed description	<p>This variable indicates the district (urban or rural) in which the social security contributor lives. The first two digits of the 5-digit district code show the code for the federal state, positions 1 – 3 indicate the regional authority, and positions 1 – 5 show the district authority. 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 employment agency was recoded to reflect the territorial allocation of December 31st 2009 for all data source observations, i.e. in all calendar years, a place of residence is assigned to an employment agency territory corresponding to the boundaries that the employment agency territory had on December 31st 2009. As the boundaries of the employment agency territories have changed over time, cases would occur in which a person’s employment agency territory changed without him/her having relocated if the territorial allocations were not updated.</p> <p>1. BeH and 2. LeH In BeH and LeH datasets, the district of residence has been available since 1999.</p> <p>3. (X)LHG This attribute includes the place of residence of the BG client (the person entitled to basic social care).</p> <p>4. (X)ASU This attribute includes the district of applicants' residences.</p> <p>5. MTH This attribute includes the district of applicants' residences.</p> <p>6. VSKT The attribute is not available for this source.</p> <p>7. Twin Twins are identical information found in the IAB and pension insurance data. They always stem from two data sources at once and they cannot be assigned uniquely to a specific source. The notification type for twins is determined by the variable “state.” This variable implies whether the twin describes employment (BeH). This variable implies whether the twin describes employment or non employment,</p>

	because BeH, LeH and (X)LHG sets can each be paired with one VSKT block.
Notes	Due 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	For (X)LHG and (X)ASU sets, the attribute has only been available since 2005.

4.5.3 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, (X)LHG, (X)ASU, MTH
Data type	Numerical
Hierarchy	The hierarchical order is federal state, administrative region, district, and municipality. The variable “wo_bula” is on the federal state level.
Detailed description	This variable assigns the “district” variable to the 16 German federal states. Further information on the district of residence may be found in wo_kreis. The attribute is not available for notifications coming from pension insurance accounts only (source: VSKT).

4.5.4 Place of Work: Employment Agency (ao_aa)

Variable label	Place of work: employment agency
Variable name	ao_aa
Category	Location Data
Origin	BeH, LeH, (X)ASU, MTH
Data type	Numerical
Hierarchy	The hierarchical order is regional directorate, employment agency, and local labour office. The variable “ao_aa” is on the employment agency level.
Detailed description	<p>1. BeH For employment notifications from the BeH, the attribute contains the employment agency of the job location. In contrast to the “Place of residence: employment agency” attribute, this attribute has been available for the BeH datasets since 1975.</p> <p>2. LeH For LeH, this attribute indicates the employment agency responsible for the payment. The allocation is in the original district.</p> <p>3. (X)LHG The attribute is not populated for this source.</p>

	<p>4. (X)ASU The field contains the local labour office from BewA and has been populated in ASU observations since 2000.</p> <p>5. MTH This block includes the local labour office number of the local labour office which handled participation in active labor market programs.</p> <p>6. VSKT The attribute is not populated for this source.</p> <p>7. Twin Twins are identical information found in the IAB and pension insurance data. They always stem from two data sources at once and they cannot be assigned uniquely to a specific source. The notification type for twins is determined by the variable "state." This variable implies whether the twin describes employment or non employment, because BeH, LeH and (X)LHG sets can each be paired with one VSKT block.</p>
Notes	Due 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 territory of the regional directorate in which the social security contributor's place of residence is located is shown.

4.5.5 Place of Work: District (ao_kreis)

Variable label	Place of work: district
Variable name	ao_kreis
Category	Location data
Origin	BeH
Data type	Numerical
Hierarchy	The hierarchical order is federal state, administrative region, district, and municipality. The variable "ao_kreis" is on the district level.
Detailed description	<p>This variable indicates the district (urban or rural) in which the employee's establishment is located. The first two digits of the 5-digit district code show the code for the federal state, positions 1 – 3 indicate the regional authority, and positions 1 – 5 show the district authority. Federal states without a regional authority have a 0 in the third position.</p> <p>In order to guarantee consistent territorial allocations across the entire observation period, the information on the district was recoded to the territorial allocation of December 31st 2009. As the district boundaries have changed over time, cases would occur in which the district code of the location of the establishment changes without the establishment concerned having relocated if the territorial allocations were not updated.</p> <p>The attribute is only populated for employment notifications (source:</p>

	BeH, twins).
Notes	Due 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.

4.5.6 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	The hierarchical order is federal state, administrative region, district, and municipality. The variable "ao_bula" is on the federal state level. The attribute is only populated for employment notifications (source: BeH, twins).
Detailed description	This attribute is generated from the variable "ao_kreis" and assigns the districts to the 16 German federal states. The attribute is only populated for employment notifications (source: BeH, twins).

4.5.7 East-West Attributes (ow_knz)

Variable label	East-West attributes
Variable name	ow_knz
Category	Location data
Origin	BeH
Data type	Numerical
Hierarchy	None
Detailed description	The official legal territory notification, provided by DEUV notification, has been reported by employers since January 1 st 1990. For the year 1999, i.e. for the DEVO/DÜVO notifications, the data was generated from the establishment number, activity code, and year of notification. Due to the depleting stock of establishment numbers for "East," establishment numbers were only divided into East and West until December 31 st 2002. From then on, it has no longer been possible to identify the legal territory (East or West) for which the contribution is made. Therefore, two additional multiple-choice boxes for the legal territory (East or West) were added to the social security notification form. The legal territory for which the contribution is made must be checked in the form. If an employer contributes for employees in the old federal states (including West Berlin) as well as the new federal states (including East Berlin), he has to file two different statements of contributions paid separately for "East" and "West." Before 1990, there are no datasets including the legal territory "East." It is possible however that municipalities and local labour offices may refer to an establishment in West Germany but the legal territory of

	<p>accession is stated as “East,” and vice versa. The reason for that is that the current establishment location is always stored, whereas the attribute calculated out of the establishment number before 1998 may be outdated. An example would be that an establishment relocated from East to West Germany, or the other way around, without receiving a new establishment number.</p> <p>The attribute is only populated for employment notifications (source: BeH, twins).</p>
--	--

4.5.8 Legal Basis to Determine Payment of Remuneration (RCEG)

Variable label	Legal basis to determine payment of remuneration
Variable name	RCEG
Category	Personal attributes
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>This variable describes the legal basis to determine payment of remuneration.</p> <p>0 “Code 1 – 7 not applicable”</p> <p>1 “Payment of remuneration from Appendix 1 – 16 German Foreign Pensions Act (Fremdrentengesetz, FRG), legal status as of June 30th 1990”</p> <p>2 “Payment of remuneration from Appendix 1 – 16 FRG, legal status as of July 1st 1990”</p> <p>4 “Payment of remuneration from Appendix 17 FRG”</p> <p>5 “Payment of remuneration from Appendix 13/14 SGB VI (for times starting from January 1st 1950)” – Differentiation between FRG and SGB VI possible by the “frgld” box</p> <p>6 “Payment of remuneration in territory of accession after conversion pursuant to Appendix 10 SGB VI”</p> <p>7 “In order to determine the payment of remuneration, the German Act on Devolution of Claims (Anspruchs- und Anwartschaftenüberführungsgesetz, AAÜG) was applied or deemed analogous”</p> <p>If rcat = 03 (only possible for current pension payment), the determination for payment of remuneration (digit 2 or 4) must be coded. The attribute is only populated for employment notifications (source: BeH, twins).</p>
Notes	<p>Further special cases where periods in the new federal states cannot be allocated to this area: ALG II generally leads to EGPT (never EGPT (East)), therefore no upgrading pursuant to Appendix 10 SGB VI is necessary.</p>

Notes on quality	<p>In general, employment periods and/or periods of earnings replacement benefit receipt (e.g. ALG I) in the employment area have to be upgraded pursuant to Appendix 10 SGB IV and afterwards awarded personal earnings points (East). This leads to the following combination: RCEG 6 and RTVS 5/6. A value of 6 in the RCEG box indicates which periods were spent in the territory of accession and must be upgraded pursuant to Appendix 10 SGB VI.</p> <p>There are some exceptions where “west evaluation” was applied to periods spent in the territory of accession: Moving from the GDR/the territory of accession before May 18th 1990: For the relocated citizens, periods in the territory of accession are rated pursuant to Appendix 10 SGB VI, but treated as EGPT (West). This also applies analogously to FRG (Law of Foreign Pensions) periods, i.e. when moving into federal territory from the area of origin, personal earnings points (West) are awarded. Moving into the territory of accession automatically led to personal earnings points (East) to be awarded. There are also some special cases stated in Art. 6 Section 4 Revised German Foreign Pensions Act (Fremdrenten- und Auslandsrenten-Neuregelungsgesetz, FANG).</p>
------------------	---

4.6 Generated Technical Variables

4.6.1 Source of the Observation, grouped (quelle_gr)

Variable label	Source of the observation, grouped
Variable name	quelle_gr
Category	Generated technical variables
Origin	BeH, LeH, (X)LHG, (X)ASU, MTH, VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>This attribute indicates the data source in grouped form. BeH, LeH, (X)LHG, (X)ASU, and MTH are datasets from the IAB. The individual data sources contain different notifications, e.g. BeH includes employment while LeH includes receipt of benefit pursuant to SGB III. VSKT is the only dataset from the pension insurance. Twins are identical information found in the IAB and pension insurance data. They always stem from two data sources at once and they cannot be assigned uniquely to a specific source. The notification type for twins is determined by the variable “state.” This variable implies whether the twin describes employment or non employment, because BeH, LeH and (X)LHG sets can each be paired with one VSKT block.</p>

4.6.2 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, MTH, VSKT
Data type	Numerical

Hierarchy	None
Detailed description	The observation counter per person counts a person's spells, beginning with 1, i.e. it indicates the number of the respective observation period in the employment biography.

4.6.3 Observation Counter per Episode (level2)

Variable label	Observation counter per episode
Variable name	level2
Category	Generated technical variables
Origin	BeH, LeH, LHG, XLHG, ASU, XASU, MTH, VSKT
Data type	Numerical
Hierarchy	None
Detailed description	This attribute contains the counter which counts parallel spells. This counter shows if any spells overlap within one insurance account. This may occur when several employments are held at the same time. The "level2" variable does not differentiate between sources, but merely counts all parallel spells within one episode, starting at 0.

4.6.4 Observation Counter per Episode and Source (level1)

Variable label	Observation counter per episode and source
Variable name	level1
Category	Generated technical variables
Origin	BeH, LeH, LHG, XLHG, ASU, XASU, MTH, VSKT
Data type	Numerical
Hierarchy	None
Detailed description	This attribute contains the counter which counts parallel spells of the same quality. This counter shows if any spells overlap within one insurance account. In contrast to "level2", "level1" also differentiates between the spells of different sources, numbering only parallel spells of the same source. The counter starts at 0 and is used to identify the main occupation in case a worker has more than one job.

4.6.5 Type of Reassessment (berrein)

Variable label	Type of reassessment
Variable name	berrein
Category	Identification number
Origin	Generated
Data type	Numerical
Hierarchy	None

Detailed description	The type of reassessment can be used to reconstruct the reassessment steps during data merging.
----------------------	---

4.7 Establishment Attributes

4.7.1 Economic Activity 73 (w73_3)

Variable label	Economic activity 73, 3-digit code
Variable name	w73_3
Category	Establishment attributes
Origin	BeH
Data type	Numerical
Hierarchy	The hierarchical assignment is done by division of economic activity, group of economic activity (3-digit code), and class 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, where the first digit of the code defines the division of economic activity out of a total of 10, and the first two digits together define the particular group of economic activity out 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 with due regard to the institutional orientation of the establishment.</p> <p>The attribute is only available for employment notifications (source: BeH, twins).</p>

4.7.2 Economic Activity 93, 5-Digit Code (w93_5)

Variable label	Economic activity 93, 5-digit code
Variable name	w93_5
Category	Establishment attributes
Origin	BeH
Data type	Numerical
Hierarchy	The hierarchical assignment is done by section of economic activity, sub-section of economic activity, division of economic activity, group of economic activity, and sub-class of economic activity (5-digit code).
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</p>

	<p>der Bundesanstalt für Arbeit, Ausgabe 1993”).</p> <p>To ensure a longer time series, there is the possibility to recode the WZ93 to WZ03 and the other way round. Therefore, comparability over a long time is guaranteed. In total, there are four ways of recoding:</p> <ul style="list-style-type: none"> • One class of WZ93 equals exactly one sub-class of WZ03. • One class of WZ93 equals more than one sub-class of WZ03. • More than one class of WZ93 can be merged into one sub-class of WZ03. • More than one class of WZ93 equals more than one sub-class of WZ03. <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> <p>The attribute is only available for employment notifications (source: BeH, twins).</p>
Notes	<p>Due 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).</p>

4.7.3 Economic Activity 93, 3-Digit Code (w93_3)

Variable label	Economic activity 93, 3-digit code
Variable name	w93_3
Category	Establishment attributes
Origin	BeH
Data type	Numerical
Hierarchy	The hierarchical assignment is done by division of economic activity, group of economic activity, and class of economic activity (3-digit code).
Detailed description	<p>This variable indicates the economic activity as a 3-digit code in accordance with the WZ93 classification and is available from 1999 up to and including 2003. WS93 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”).</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> <p>The attribute is only available for employment notifications (source: BeH, twins).</p>

4.7.4 Economic Activity 03, 5-Digit Code (w03_5)

Variable label	Economic activity 03, 5-digit code
Variable name	w03_5
Category	Establishment attributes
Origin	BeH
Data type	Numerical
Hierarchy	The hierarchical assignment is done by section of economic activity, sub-section of economic activity, division of economic activity, group of economic activity, class of economic activity, and sub-class of economic activity (5-digit code).
Detailed description	This variable indicates the economic activity as a 5-digit code in accordance with the WZ03 classification and is available from 2003 up to and including 2008. WZ03 stands for the "Classification of Economic Activities, Edition 2003" ("Klassifikation der Wirtschaftszweige Ausgabe 2003") of the Federal Statistical Office (eds.). Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected. The attribute is only available for employment notifications (source: BeH, twins).
Notes	Due 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).

4.7.5 Economic Activity 03, 3-Digit Code (w03_3)

Variable label	Economic activity 03, 3-digit code
Variable name	w03_3
Category	Establishment attributes
Origin	BeH
Data type	Numerical
Hierarchy	The hierarchical assignment is done by division of economic activity, group of economic activity, and class of economic activity (3-digit code).
Detailed description	This variable indicates the economic activity as a 3-digit code in accordance with the WZ03 classification and is available from 2003 up to and including 2008. WZ03 stands for the "Classification of Economic Activities, Edition 2003" ("Klassifikation der Wirtschaftszweige Ausgabe 2003") of the Federal Statistical Office (eds.). Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected. The attribute is only available for employment notifications (source: BeH, twins).

4.7.6 Economic Activity 08, 5-Digit Code (w08_5)

Variable label	Economic activity 08, 5-digit code
Variable name	w08_5
Category	Establishment attributes
Origin	BeH
Data type	Numerical
Hierarchy	The hierarchical assignment is done by section of economic activity, division of economic activity, group of economic activity, class of economic activity, and sub-class of economic activity (5-digit code).
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.). Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected.</p> <p>The attribute is only available for employment notifications (source: BeH, twins).</p>
Notes	Due 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).

4.7.7 Economic Activity 08, 3-Digit Code 08 (w08_3)

Variable label	Economic activity 08, 3-digit code 08 (w08_3)
Variable name	w08_3
Category	Establishment attributes
Origin	BeH
Data type	Numerical
Hierarchy	The hierarchical assignment is done by division of economic activity, group of economic activity, and class of economic activity (3-digit code).
Detailed description	<p>This variable indicates the economic activity as a 3-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.). Each establishment is only assigned one code. If an establishment is active in different economic sectors, the main economic activity should be reflected.</p> <p>The attribute is only available for employment notifications (source: BeH, twins).</p>

4.7.8 Pension Receipt From Own Insurance (RTVS)

Variable label	Pension receipt from own insurance/personal earnings points attributes in West and East Germany
Variable name	RTVS
Category	Personal attributes
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>This attribute states if during the period in question, (parts of) pension was drawn from own insurance, and during periods of contribution or consideration, if personal earnings points (East) should be awarded. Periods of contribution and/or consideration without awarding of personal earnings points (East) or non-contributory period:</p> <p>0 = No period of pension receipt from own insurance 1 = Period of pension receipt of own insurance</p> <p>Periods of contribution and/or consideration with earnings points (East) awarded:</p> <p>5 = No period of pension receipt from own insurance 6 = Period of pension receipt from own insurance</p> <p>The attribute is populated with the value 0 or 1 if in one calendar month both earnings points and earnings points East have to be considered (Section 254d(3) SGB VI).</p> <p>In general, periods of employment and/or periods obtaining earnings replacement benefit (ALG I and similar) in the territory of accession have to be upgraded pursuant to Appendix 10 SGB VI and awarded with earnings points (East). This leads to the combination RCEG 6 and RTVS 5/6. A value of 6 in the RCEG box indicates which periods were spent in the territory of accession and must be upgraded pursuant to Appendix 10 SGB VI.</p> <p>Exceptions where “west evaluation” was applied to periods spent in the territory of accession:</p> <p>Moving in from the GDR/the territory of accession before May 18th 1990: For the relocated citizens, periods in the territory of accession are rated pursuant to Appendix 10 SGB VI, but treated as EGPT (West) (Sk79: RCEG = 6 and RTVS = 0/1).</p> <p>This also applies analogously to FRG periods, i.e. when moving into federal territory from the area of origin, personal earnings points (west) are awarded; when moving into the territory of accession, personal earnings points (East) are awarded. There are also some special cases stated in Art. 6 Section 4 FANG.</p>

	<p>Dual job holding in old and new federal states in the same month: With pension starting in 2010, it is possible for earnings points (East) and (West) to occur during the same month, due to the cancellation of Section 254d(3) Sentence 1. Until 2009, east_EGPT were treated as EGPT periods. Therefore the combination of variables is: Sk79: RCEG=6 and RTVS=0/1.</p> <p>Further special cases where periods in the new federal states cannot be allocated to this area: ALG II (unemployment assistance) generally leads to EGPT (never EGPT (East)), therefore no upgrading due to Appendix 10 is carried out (RCEG=0 and RTVS=0/1).</p> <p>The attribute is only available for pension insurance data and twins (source: VSKT, twins).</p>
--	---

4.7.9 Month in the reference Period (ZRMO)

Variable label	Month in the reference period
Variable name	ZRMO
Category	Personal attributes
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>This attribute states the chargeable months for the mapped period. With dual job holding and/or overlap with other periods, only the superior time receives the monthly value, the other blocks are labelled "0." In the case of periods of reduced contributions, periods of contribution as well as the highest non-contributory period (if applicable) receive the periods of months. Otherwise, the sum of ZRMO may not exceed the duration of from-to periods.</p> <p>If state = 26, receipt duration can be stated in months.</p> <p>The attribute is only available for pension insurance data and twins (source: VSKT, twins).</p>

4.7.10 Parenting Period/Creditable Period (KI)

Variable label	Parenting period/creditable period (KI)
Variable name	KI
Category	Personal attributes
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	For acquisition periods of contribution, substitution, consideration and pension, it must be specified whether parenting or consideration

	<p>periods apply at the same time:</p> <p>0 “No parenting periods and consideration period at the same time” 1 “Parenting periods (and consideration period) at the same time” 2 “Parenting periods at the same time” 3 “Consideration periods for caregiving at the same time (January 1st 1992 – March 31st 1995)” 4 “Caregiving period for at least one care-dependent child up to the age of 18 at the same time”</p> <p>The classification code digits 2/3 must be increased by 5 if the person was self-employed parallel to the consideration period (only for pensions beginning before January 1st 2002).</p> <p>The lower value must be given priority in the statement. If state = 26, the first digit of the benefit type must be specified.</p>
--	--

4.7.11 Application of special Regulations (KZSO)

Variable label	Application of special regulations
Variable name	KZSO
Category	Personal attributes
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>This block marks contribution periods and employment subject to application due to FRG, Section 256 a/b SGB VI, or FANG. If the state is 2, 3, 4, 6, 20, 22:</p> <p>0 “No application of FRG / Section 256 a/b SGB VI / FANG” 1 “Section 15 or Section 17 without Section 22(3) FRG (contribution periods, total EGPT)” 2 “Section 16 without Section 22(3) FRG (periods of employment, total EGPT)” 3 “Art. 6 Section 23 FANG without Section 22(3) FRG (total EGPT)” 4 “Section 256 a SGB VI (total EGPT)” 6 “Section 15 or Section 17 with Section 22(3) FRG (contribution periods, reduced EGPT)” 7 “Section 16 with Section 22(3) FRG (periods of employment, reduced EGPT)” 8 “Art. 6 Section 23 FANG with Section 22(3) FRG (reduced EGPT)” 9 “Section 256 a(3) Sentence 2 / Section 256 b SGB VI (reduced EGPT)”</p> <p>The attribute is only available for pension insurance data and twins (source: VSKT, twins).</p>

4.7.12 Group of Employment for Application of FRG (BHBR)

Variable label	Group of employment for Application of FRG
Variable name	BHBR
Category	Personal attributes
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>For periods based on FRG and/or Section 256 a/b SGB VI, the group of employment (with a leading 0) /branch /sector of the economy used by the corresponding block in RCEG must be encoded.</p> <p>Group of employment (with <i>RCEG</i> = 1, 2)</p> <p>(a) For AR/AV</p> <p>0 “Educational holiday in Poland” 1 “Workers outside of agriculture and forestry” 2 “Agriculture workers” 3 “Forestry workers” 4 “Employees” 7 “Voluntarily insured persons” 8 “Seamen”</p> <p>(b) For KN</p> <p>0 “Educational holiday in Poland” 1 “Workers in underground mining” 2 “Workers in surface mining” 3 “Voluntarily insured persons AR” 4 “Laboratory assistants in underground mining” 5 “Laboratory assistants in surface mining” 6 “Office clerks” 7 “Voluntarily insured persons AV”</p> <p>Sector of the economy (with <i>RCEG</i> = 4)</p> <p>00 “No sector of the economy stored and not code 50 – 99” 01 “Agriculture and forestry” ‘</p> <p>24 “Handicrafts without building trade”</p> <p>Other codes:</p> <p>50 “Lowest area” 79 “Voluntary contributions” 90 “German state railway” 95 “Basic military service and apprentice” 99 “Assignment to industry not relevant”</p> <p>Branch (with <i>RCEG</i> = 5)</p> <p>01 “Power generation industry and oil and fuel industry ”</p>

	<p>23 "Production cooperative of craftsmen"</p> <p>Other codes:</p> <p>50 "Lowest area"</p> <p>77 "Apprenticeship without contribution payment"</p> <p>79 "Voluntary contribution"</p> <p>90 "German state railway"</p> <p>95 "Basic military service and apprentice"</p> <p>The attribute is only available for pension insurance data and twins (source: VSKT, twins).</p>
--	--

4.7.13 FRG Group of Benefit/Group of Qualification (QLGR)

Variable label	FRG group of benefit/group of qualification
Variable name	QLGR
Category	Personal attributes
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>For periods based on FRG and/or Section 256 a/b SGB VI, the FRG group of benefit/group of qualification used by the corresponding block in RCEG must be encoded.</p> <p>FRG group of benefit, old (with RCEG = 1,2)</p> <p>0 "Apprenticeship"</p> <p>1 "Group of benefit 1"</p> <p>2 "Group of benefit 2"</p> <p>3 "Group of benefit 3"</p> <p>4 "Group of benefit 4"</p> <p>5 "Group of benefit 5"</p> <p>8 "Basic military service pursuant to Section 15 FRG"</p> <p>9 "No apprenticeship or group of benefit"</p> <p>FRG group of benefit (with RCEG = 4)</p> <p>0 "Apprenticeship"</p> <p>1 "Group of benefit 1"</p> <p>2 "Group of benefit 2"</p> <p>3 "Group of benefit 3"</p> <p>4 "Group of benefit 4"</p> <p>5 "Group of benefit 5"</p> <p>9 "No apprenticeship or group of benefit"</p> <p>Group of qualification (with RCEG = 5)</p> <p>0 "Time of apprenticeship or not assigned to any group of qualification"</p>

	<p>1 “Group of qualification 1” 2 “Group of qualification 2” 3 “Group of qualification 3” 4 “Group of qualification 4” 5 “Group of qualification 5” 7 “Apprenticeship without contribution payment” ‘</p> <p>Appendix 1 of FRG concerning groups of benefit regulates in detail the affiliation of persons in groups of benefits according to occupation, type of employment, and duration of employment.</p> <p>The attribute is only available for pension insurance data and twins (source: VSKT, twins).</p>
--	---

5 Personal Attributes, Due Date December 31st 2007 (fixed part) Artificial Individual ID (persnr)

Variable label	Artificial individual ID
Variable name	persnr
Category	Identifiers
Origin	Generated
Data type	Numerical
Hierarchy	None
Detailed description	The artificial individual ID is generated through the social security number, because the social security number uniquely identifies an individual. However, to maintain the anonymity of the individuals in the sample, an artificial number is developed.

5.2 Category of Persons (PSGR)

Variable label	Category of persons
Variable name	PSGR
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>This attribute indicates whether any pensions are drawn at the survey date (12/31). The following degrees exist:</p> <p>99 “No pension receipt” 10 “Compensatory payment by German miners' insurance”</p>

	<p>11 “Miners' pension due to reduced capacity for work in mining with pension starting before December 31st 2000”</p> <p>12 “Miners' pension due to many years of underground work and being 50 years of age at the start of pension before December 31st 2000”</p> <p>13 “Pension due to incapacity to work, with employment covered miners' insurance before December 31st 2000”</p> <p>14 “Pension due to incapacity to work, pension due to incapacity to work after terminating employment covered by miners' insurance”</p> <p>15 “Pension due to unemployability”</p> <p>16 “Standard old-age pension”</p> <p>17 “Old-age pension due to unemployment or following partial retirement”</p> <p>18 “Women's old-age pension”</p> <p>19 “Old-age pension for miners after many years of underground work”</p> <p>29 “Miners' pension with AR/AV shares due to incapacity to work”</p> <p>31 “Old-age pension (Art. 2 Section 4 German Pension Transference Act (Rentenüberleitungsgesetz, RÜG))”</p> <p>32 “Disability pension (Art. 2 Section 7 RÜG)”</p> <p>33 “Disability pension for disabled persons (Art. 2 Section 10 RÜG)”</p> <p>37 “Maintenance pension (Art. 2 Section 14 RÜG)”</p> <p>39 “Miners' pension with AR/AV shares due to unemployability”</p> <p>43 “Enhanced incapacity pension”</p> <p>45 “Child-raising benefits”</p> <p>51 “Pension based on higher insurance contributions for insured persons”</p> <p>62 “Old-age pension for severely disabled individuals”</p> <p>63 “Old-age pension for long-term insured persons”</p> <p>65 “Pension due to unemployability (converted pension Section 308 SGB VI)”</p> <p>71 “Miners' pension due to reduced capacity for work in mining with pension starting before January 1st 2001”</p> <p>72 “Miners' pension due to many years of underground work and being 50 years of age at the start of pension after January 1st 2001”</p> <p>73 “Pension due to partial reduction in earning capacity, with employment covered by miners' insurance”</p> <p>74 “Pension due to partial reduction in earning capacity, pension due to partial reduction in earning capacity after terminating employment covered by miners' insurance”</p> <p>75 “Pension due to full reduction in earning capacity (Section 43(2) SGB VI)”</p> <p>76 “Pension due to full reduction in earning capacity (Section 43(6) SGB VI)”</p> <p>91 “Miners' old-age pension (Art. 2 Section 5 RÜG)”</p> <p>92 “Miners' disability pension (Art. 2 Section 8 RÜG)”</p> <p>93 “Miners' full pension (Art. 2 Section 6 RÜG)”</p>
--	---

	94 "Miners' pension (Art. 2 Section 9 RÜG)"
--	---

5.3 Category of Persons, Aggregated (PSGR_GR)

Variable label	Category of persons, grouped
Variable name	PSGR_GR
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	Specification as to whether any pensions are drawn at the survey date (12/31).

5.4 Type of Insurance (VSAT)

Variable label	Type of insurance
Variable name	VSAT
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>This attribute specifies the type of insurance. The most recent legal insurance status has priority.</p> <p>1 = Marginal part-time employees waiving exemption from insurance 2 = Caregivers 3 = Other individual subject to compulsory insurance 4 = Compulsory insurance due to unemployment 5 = Compulsory insurance due to inability to work etc. 6 = Employees subject to insurance contributions 7 = Recipients of early retirement pension benefits 8 = Self-employed individuals 9 = Craftsmen 10 = Individuals performing military or civilian service 11 = Voluntarily insured individuals 12 = Individuals insured during credited periods 13 = Marginal part-time employees not waiving exemption from insurance 14 = Latently insured individuals 15 = Transition cases 18 = Compulsorily insured pensioner 19 = Voluntarily insured pensioner 20 = Insured pensioner without payment of contribution 25 = Compensatory payment by German miners' insurance</p>

5.5 Account Balance Indicator (KTSD)

Variable label	Account balance indicator – year, month
Variable name	KTSD
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>It can be assumed for Germans over 30 years of age that they have been contacted regarding adjustment of account, unless the insurance account has already been adjusted within the last 6 years. The year of the most recent adjustment of account is indicated in YYYY form.</p> <p>0000 = No adjustment of account so far, and no granting of pensions, or year of adjustment of account 2003 = 2003 and earlier 9999 = Application for adjustment of account or pension still pending</p>

5.6 Year of Most Recent Notification on Adjustment of Account, with the Collaboration of the Insured Person (BXKLVS)

Variable label	Year of most recent notification on adjustment of account, with the collaboration of the insured person
Variable name	BXKLVS
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>The year of the notification must be specified which announced the most recent adjustment of insurance account with the collaboration of the insured person. Unless the insurance account has not been adjusted at least once with the collaboration of the insured person, the value "0000" shall be entered.</p>

5.7 Time of Current Commencement of Retirement (ZTPTRTBE)

Variable label	Time of current commencement of retirement – year, month
Variable name	ZTPTRTBE
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None

Detailed description	If no pensions are drawn at the survey date (12/31), the value "000000" shall be entered. In all other cases, the current commencement of retirement must be indicated in YYYYMM format. If the person moved to Germany, the start of pension payment for the current pension by a federal German pension agency must be taken into account. If the month of current commencement of retirement is unknown, the code "YYYY00" shall be used.
----------------------	--

5.8 Partial Pension Indicator (TLRT)

Variable label	Partial pension indicator
Variable name	TLRT
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>For pensioners, this attribute must indicate whether the pension drawn is partial pension or a pension based on shares.</p> <p>0 = no pension or no partial pension</p> <p>For old-age pensions: 1 = 1/3 partial pension 2 = 1/2 partial pension 3 = 2/3 partial pension</p> <p>For pensions due to reduced earning capacity: 0 = Full pension (no additional income) 1 = Pension corresponding to 1/3 of disablement pension or 1/3 miners' pension 2 = Pension due to full/partial reduced earning capacity of half the amount 3 = Pension corresponding to 2/3 of disablement pension or 2/3 miners' pension 4 = EU pension in the full amount of disablement pension 5 = Pension is not fully paid due to coincidence with additional income 6 = Pension due to full reduced earning capacity of one-quarter the amount 7 = Pension due to full reduced earning capacity of three-quarters the amount</p> <p>The partial pension indicator shows if there is a current partial pension drawn on December 31st of the reporting year. It is not suited to record gradual phasing-out into retirement:</p> <p>If an insured individual gradually phased out into retirement but already received full pension benefits on December 31st of the reporting year, this gradual phasing-out is not recorded in the partial pension indicator. Therefore, the variable "Pension receipt from own insurance/personal earnings points attributes" should be consulted. It</p>

	must be noted, however, that the partial pension option has only been available since 1992.
--	---

5.9 Partial Pension Indicator, Aggregated (TLRT_GR)

Variable label	Partial pension indicator, aggregated
Variable name	TLRT_GR
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>This attribute indicates whether a partial pension and/or a pension based on shares is drawn. The following degrees exist:</p> <p>0 = No pension or no partial pension 1 = Partial pension or pension based on shares is drawn</p>

5.10 Information on Earnings points

5.10.1 Sum of Earnings points – West (SUEGPT_WEST)

Variable label	Sum of earnings points – West
Variable name	SUEGPT_WEST
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>The sum of earnings points from all of the following must be indicated:</p> <ul style="list-style-type: none"> • Contribution periods • Non-contributory periods • Extra pay for periods of reduced contribution • Supplementary benefit • Extra earnings points for remuneration from marginal insurance-free employment pursuant to Section 76b SGB VI • Extra pay or deduction from statutory equalisation of pensions • Compensation payment due to pension reduction • Earnings points from termination pay from a company pension scheme • Earnings points for remuneration from credit balance not used according to an agreement on flexible working time (already included in contribution periods) • Extra pay or deduction from pension splitting <p>The sum of all earnings points must be indicated before applying Section 256 d SGB VI.</p>

5.10.2 Sum of Earnings points – East (SUEGPT_OST)

Variable label	Sum of earnings points – East
Variable name	SUEGPT_OST
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	Earnings points registered as “East” are multiplied with a different, up to 2013 lower, pension value. Contributions registered as “East” are multiplied with a factor higher than 1. The same gross income leads to more earnings points if registered as “East”. These are later multiplied with the pension value east. The procedure leads in the end to a slightly higher pension than the same gross income registered as “West”.

5.10.3 Personal earnings points – West (PSEGPT_WEST)

Variable label	Personal earnings points – West
Variable name	PSEGPT_WEST
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	This attribute indicates the sum of personal earnings points that results from the SUEGPT variable while taking into account the respective access factor(s), the share of partial pension, and the reduced consideration of earnings points for parenting periods (Sections 256d, 307d Sentence 5 SGB VI). For pensions due to reduced earning capacity which are not paid in full due to coincidence with additional income, the full sum of personal earnings points is indicated.

5.10.4 Personal Earnings points – East (PSEGPT_OST)

Variable label	Personal earnings points – East
Variable name	PSEGPT_OST
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	This attribute indicates the sum of personal earnings points that results from the SUEGPT variable while taking into account the respective access factor(s), the share of partial pension, and the reduced

	consideration of earnings points for parenting periods (Sections 256d, 307d Sentence 5 SGB VI). For pensions due to reduced earning capacity which are not paid in full due to coincidence with additional income, the full sum of personal earnings points is indicated.
--	---

5.10.5 Sum of Earnings points for All Contribution Periods – West (BZEGPT_WEST)

Variable label	Sum of earnings points for all contribution periods – West
Variable name	BZEGPT_WEST
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	Without earnings points from non-contributory periods, additional earnings points for periods of reduced contribution, extra pay from statutory equalisation of pensions, deduction from statutory equalisation of pensions, extra pay from pension splitting, deduction from pension splitting, earnings points from supplementary benefit and/or extra earnings points pursuant to Section 76b SGB VI, earnings points from compensation payment due to pension reduction, earnings points from termination pay from a company pension scheme.

5.10.6 Sum of Earnings points for All Contribution Periods – East (BZEGPT_OST)

Variable label	Sum of earnings points for all contribution periods – East
Variable name	BZEGPT_OST
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	Without earnings points from non-contributory periods, additional earnings points for periods of reduced contribution, extra pay from statutory equalisation of pensions, deduction from statutory equalisation of pensions, extra pay from pension splitting, deduction from pension splitting, earnings points from supplementary benefit and/or extra earnings points pursuant to Section 76b SGB VI, earnings points from compensation payment due to pension reduction, earnings points from termination pay from a company pension scheme.

5.10.7 Sum of Earnings points for Non-Contributory Periods – West (BYFHEGPT_WEST)

Variable label	Sum of earnings points for non-contributory periods – West
Variable name	BYFHEGPT_WEST
Category	Personal attributes, due date December 31 st 2007

Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	According to Section 54(4) SGB VI, creditable non-contributory periods include only such calendar months that are registered only as creditable period, supplementary periods or substitute periods No contributions ought to be paid during these months.

5.10.8 Sum of Earnings points for Non-Contributory Periods – East (BYFHEGPT_OST)

Variable label	Sum of earnings points for non-contributory periods – East
Variable name	BYFHEGPT_OST
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None

5.10.9 Sum of Earnings points from Full-Value Contribution Periods – West (BYVLEGPT_WEST)

Variable label	Sum of earnings points from full-value contribution periods – West
Variable name	BYVLEGPT_WEST
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	The sum of earnings points for full-value contribution periods from the BYVL box must be indicated.

5.10.10 Sum of Earnings points from Full-Value Contribution Periods – East (BYVLEGPT_OST)

Variable label	Sum of earnings points from full-value contribution periods – East
Variable name	BYVLEGPT_OST
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	The sum of earnings points for full-value contribution periods from the BYVL box must be indicated.

5.10.11 Sum of Additional Earnings points for Periods of Reduced Contribution – West (BYGMEGPTZQ_WEST)

Variable label	Sum of additional earnings points for periods of reduced contribution- West
Variable name	BYGMEGPTZQ_WEST
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	Pursuant to Section 54 SGB VI, periods of reduced contribution include all months that show both contribution periods and credit periods, creditable additional periods of coverage, or substitute periods. Months of reduced contribution include mainly months of vocational training during which compulsory contributions were made.

5.10.12 Sum of Additional Earnings points for Periods of Reduced Contribution – East (BYGMEGPTZQ_OST)

Variable label	Sum of additional earnings points for periods of reduced contribution – East
Variable name	BYGMEGPTZQ_OST
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	Pursuant to Section 54 SGB VI, periods of reduced contribution include all months that show both contribution periods and credit periods, creditable additional periods of coverage, or substitute periods. Months of reduced contribution include mainly months of vocational training during which compulsory contributions were made.

5.10.13 Additional/Accredited Earnings points for Consideration Periods Due to Parenting and/or Caregiving – West (ZQEGPTKIPE_WEST)

Variable label	Earnings points for consideration periods due to parenting/caregiving – West
Variable name	ZQEGPTKIPE_WEST
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	Any additional/accredited earnings points for consideration periods due to parenting or non-commercial caregiving to a care-dependent child (Section 70(3a) SGB VI, Section 83(1) Sentence 3 SGB VI)

	must be entered here.
--	-----------------------

5.10.14 Additional/Accredited Earnings points for Consideration Periods Due to Parenting and/or Caregiving – East (ZQEGPTKIPE_OST)

Variable label	Earnings points for consideration periods due to parenting/caregiving – East
Variable name	ZQEGPTKIPE_OST
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	Any additional/accredited earnings points for consideration periods due to parenting or non-commercial caregiving to a care-dependent child (Section 70(3a) SGB VI, Section 83(1) Sentence 3 SGB VI) must be entered here.

5.10.15 Months with Additional/Accredited Earnings points for Consideration Periods Due to Parenting and/or Caregiving – West (ZQMOKIPE_WEST)

Variable label	Months with earnings points for consideration periods due to parenting/caregiving – West
Variable name	ZQMOKIPE_WEST
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	The number of months for which any additional/accredited earnings points for consideration periods due to parenting or non-commercial caregiving to a care-dependent child are taken into account (Section 70(3a) SGB VI, Section 83(1) Sentence 3 SGB VI) must be entered here.

5.10.16 Months with Additional/Accredited Earnings points for Consideration Periods Due to Parenting and/or Caregiving – East (ZQMOKIPE_OST)

Variable label	Months with earnings points for consideration periods due to parenting/caregiving – East
Variable name	ZQMOKIPE_OST
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	The number of months for which any additional/accredited earnings points for consideration periods due to parenting or non-commercial caregiving to a care-dependent child are taken into account (Section

	70(3a) SGB VI, Section 83(1) Sentence 3 SGB VI) must be entered here.
--	---

5.11 Full-Value Contribution Periods – West (BYVL_WEST)

Variable label	Full-value contribution periods – West
Variable name	BYVL_WEST
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	This attribute includes the months of full-value contribution, including contribution periods during pension receipt from own insurance. Full-value contribution periods include compulsory contribution periods and voluntary contribution periods.

5.12 Full-value Contribution Periods – East (BYVL_OST)

Variable label	Full-value contribution periods – East
Variable name	BYVL_OST
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	This attribute includes the months of full-value contribution, including contribution periods during pension receipt from own insurance. Full-value contribution periods include compulsory contribution periods and voluntary contribution periods.

5.13 Share of Contribution in Earnings points East (OPXAZ)

Variable label	Share of EGPT East
Variable name	OPXAZ
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	The factor which determines the relation of earnings points (East) and all earnings points of the (decisive) assessment of standard benefits of pensions (Section 263a SGB VI).

5.14 Pension-Relevant Periods (AR/AV and KN) (RTZTMO)

Variable label	Pension-relevant periods
Variable name	RTZTMO
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	The months in AR/AV, AR/AV (East), KN, and KN (East) which show pension-relevant periods. They result from the sum of the following periods: full-value contribution periods, periods of reduced contribution, credited periods, creditable additional periods of coverage, substitution periods, and consideration periods.

5.15 Consideration Periods (BUEZT)

Variable label	Consideration periods
Variable name	BUEZT
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	All absolute consideration periods, consideration periods during pension receipt from own insurance, and consideration periods during self-employed work which do not coincide with any other pension-relevant periods must be indicated here in the form of months.

5.16 Additional EGPT Taken Into Consideration for Consideration Periods (BUEZTEGPT)

Variable label	Additional EGPT taken into consideration for consideration periods
Variable name	BUEZTEGPT
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	All earnings points from consideration periods that were considered in the main assessment must be indicated here. In the event of an overlap with contribution periods, if applicable, only the difference of 0.0625 earnings points shall be considered per month.

5.17 Extra Pay From Settlement of Pension Entitlements (Bonus) (VAZU)

Variable label	Extra pay from settlement of pension entitlements in case of divorce – West and East
----------------	--

Variable name	VAZU
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	The most common case of statutory equalisation of pensions is that the party entitled to equalisation upon dissolution of marriage after July 1 st 1977 in West Germany and 1990 in East Germany receives an additional, independent claim, generally in statutory pension insurance. Entitlements of both partners are added up and divided by two. The difference between this sum and the own entitlement is transferred from the partner with more to the partner with less entitlements.

5.18 Deduction From Settlements of Pension Entitlements (Malus) (VAAB)

Variable label	Deduction from settlement of pension entitlements in case of divorce – West and East
Variable name	VAAB
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	The earnings points that are transferred in case settlement in case of divorce from statutory pension account.

5.19 Credited Periods

5.19.1 Total Credited Periods (AZ)

Variable label	Total credited periods
Variable name	AZ
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	Credited periods are non-contributory periods during which the insured person was unable to pay contributions (e.g. illness). Within the scope of social equilibrium, however, earnings points are awarded to them during the assessment of standard benefits of pensions. The total number of months of credited periods must be indicated, which are not periods of reduced contributions and not subject to Section 71(4) SGB VI. To the extent that ascertained credited periods before January 1 st 1957 cannot be included because the fixed credited period is at least of the same duration, any such ascertained credited periods are not to be included here; instead, the fixed credited period

	shall be included. No distinction is made between West and East.
--	--

5.19.2 Sum of Credited Periods Due to School-Based Training (SCHULAZ)

Variable label	Credited periods due to (technical) school-based or university-based training
Variable name	SCHULAZ
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	This attribute should include, in the form of months, all credited periods due to school-based, technical school-based or university-based training (Section 58(1) No. 4 SGB VI) contained in the AZ variable, including credited periods due to vocational training pursuant to the transitional regulation Section 252(4) SGB VI which are not periods of reduced contribution, and also including unassessed credited periods. In the blocks "AR/AV" and "KN," the corresponding credited periods must be coded. The blocks "AR/AV (East)" and "KN (East)" are backups and each digit should be "0."

5.19.3 Sum of Assessed Credited Periods of Technical Training or Participation in a Work Preparation Scheme (FASCHULAZ)

Variable label	Credited periods due to technical training
Variable name	FASCHULAZ
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	Assuming a current commencement of retirement after 2005, this attribute is reserved exclusively for any assessed credited periods due to technical training or participation in a work preparation scheme (Section 58(1) No. 4 SGB VI in conjunction with Section 74 SGB VI). Periods should be specified in the form of months. In the blocks "AR/AV" and "KN," the corresponding credited periods must be coded. The blocks "AR/AV (East)" and "KN (East)" are backups and each digit should be "0."

5.19.4 Non-contributory supplementary period(ZZ)

Variable label	Non-contributory supplementary periods (in case of reduced earnings capacity)
Variable name	ZZ
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT

Data type	Numerical
Hierarchy	None
Detailed description	The creditable additional period of coverage considered notionally in annuity computation pursuant to Section 59 SGB VI, without periods of reduced contribution, must be entered in the form of months. No distinction is made between East and West.

5.20 Start Date of Refund (VNZRES)

Variable label	"From" date of refund, year, month
Variable name	VNZRES
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	"From" date (in YYYYMM format) of refund ("000000" also allowed for refund)

5.21 End Date of Refund (BSZRES)

Variable label	"To" date of refund, year, month
Variable name	BSZRES
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	"To" date (in YYYYMM format) of refund

5.22 Information on Children

5.22.1 Total Number of Children (KIND)

Variable label	Number of children
Variable name	KIND
Category	Personal attributes, due date December 31 st 2007
Origin	Generated from VSKT
Data type	Numerical
Hierarchy	None
Detailed description	This attribute contains the number of children per person. The following sub-division is used: 1 = If birth year of first child is on record (GBKI_JJ_1>=1) 2 = If birth year of first/second child is on record

	<p>3 = If birth year of first/second/third child is on record</p> <p>4 = If birth year of first to fourth child is on record</p> <p>5 = If birth year of first to fifth child is on record</p> <p>6 = If birth year of first to sixth child is on record</p> <p>7 = If birth year of first to seventh child is on record</p> <p>8 = If birth year of first to eighth child is on record</p> <p>9 = If birth year of first to ninth child is on record</p> <p>10 = If birth year of first to tenth child is on record</p>
--	--

5.22.2 Children's Birth Dates (GBKI1 – GBKI10)

Variable label	Birth dates of the individual children
Variable name	GBKI1, GBKI2 to GBKI10
Category	Personal attributes, due date December 31 st 2007
Origin	Generated from VSKT
Data type	Date
Hierarchy	None
Detailed description	This attribute indicates the dates of birth (month and year) of the children in YYYYMM format. The attribute is populated individually for every child. The birth dates are given in chronological order, left-justified, for the first 10 children at the most. Multiple births must be entered the respective number of times. If fewer than 10 children exist, the remainder must be populated with missings (.n).

5.22.3 Consideration in Variable Part – (GBKIZ1 – GBKIZ10)

Variable label	Children's birth dates
Variable name	GBKIZ1, GBKIZ2 to GBKIZ10
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>This attribute contains the children's birth dates.</p> <p>The values 1 and 2 are indicated even when parenting periods and/or consideration periods due to parenting cannot be credited in whole or in part. If, for example, pension benefits commence before the consideration period due to parenting is finished, the latter will be limited to the month before pension commencement.</p> <p>Values 0 and/or 5 imply that the parenting period and the corresponding consideration period have not yet had any impact on the variable part (with the exception of German Statutory Pension Insurance).</p> <p>0 = No child, or parenting periods and corresponding consideration periods for this birth have not yet been clarified, but a birth is registered in the insurance account.</p> <p>1 = Parenting periods for this birth are clarified</p> <p>2 = Parenting periods and corresponding consideration periods for</p>

	<p>this birth are clarified</p> <p>5 = Same as 0, but the child's birthday is the first day of the month (parenting period begins at the month of birth; parenting consideration period only lasts 120 calendar months).</p>
--	--

5.23 Country for Foreign Pensions (FRGLD)

Variable label	Country for foreign pensions
Variable name	FRGLD
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	Specification of FRG periods which should be considered pursuant to FRG. Content of the variable is the country of origin of the person that is eligible for entitlements according to the Act on Foreign Pensions (FRG).

5.24 Extrapolation Factor (HRF)

Variable label	Extrapolation factor
Variable name	HRF
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	The extrapolation factor is stored in this attribute by the German Statutory Pension Insurance. At the insurance agency, the attribute has a default value of 0.

5.25 Persons Originally Insured Under Miners' Insurance (VSKN)

Variable label	Persons originally insured under miners' insurance (VSKN)
Variable name	VSKN
Category	Personal attributes, due date December 31 st 2007
Origin	VSKT
Data type	Numerical
Hierarchy	None
Detailed description	<p>This attribute indicates whether the account of the insured person includes contribution periods for miners' pension insurance. The following options are possible:</p> <p>1 "General pension insurance (without any contribution periods for miners' pension insurance)"</p> <p>3. "Insured individuals with a contribution period for miners' pension</p>

	insurance of at least one month” 9 “No statement possible”
--	---

6 BHP Establishment Data

6.1 Artificial Establishment ID (betnr)

Variable label	Original artificial establishment ID
Variable name	betnr
Category	Identification number
Origin	BeH
Data type	Numerical
Hierarchy	The artificial establishment ID is hierarchically sorted by establishment number zone, serial number, and error checking number.
Detailed description	<p>The artificial establishment ID enables an automatic reporting process to social security of employees liable to social security. The allocation of artificial establishment IDs for establishments with employees liable to social security has been centrally realised since January 1st 2008 by the Establishment Number Service (BNS) of the Federal Employment Agency (BA) in Saarbrücken; before that, it was assigned peripherally by employment offices. The artificial establishment ID is useful in two ways. On the one hand, employers can be identified at social security agencies; on the other hand, it is possible to draw conclusions about the employees of an establishment. One example would be the assignment of an employee to a certain class of economic activity or a region. Furthermore, the assignment of artificial establishment IDs follows some elementary rules. If the company has one office only, this office is the establishment and receives an establishment ID. If a company has several offices in different municipalities, each office is assigned its own establishment ID. If the company has several branch offices in one municipality, these offices are merged into a single establishment under one establishment ID if they belong to the same economic class. Vice versa, if they do not belong to the same economic class, each branch office is an establishment and receives an establishment ID.</p> <p>A new artificial establishment ID is assigned if an establishment is founded, the establishment owner changes, the establishment's economic class changes, or if the employer filed an application for amalgamation of several offices. No new artificial establishment ID can be assigned if the establishment changes local conditions, e.g. by relocating to another community. The old one remains valid until the health insurance or the employer demands a new artificial establishment ID. The artificial establishment ID is referred to as “non-active” if the establishment currently has no employees liable to social security. The artificial establishment ID can be reactivated if there are once more employees liable to social security and there are no objections by neither the health insurance company nor the employer. There is also the possibility to take up the artificial establishment ID once again when the employer founds a new establishment.</p>

6.2 Year (jahr)

Variable label	Year
Variable name	jahr
Category	Generated technical variables
Origin	BeH
Data type	Numerical
Hierarchy	None
Detailed description	<p>This variable is only included in the establishment file. It indicates the year of validity of the establishment data as of June 30th.</p> <p>This variable can be used together with the establishment ID to link the individual file and the establishment file.</p>

6.3 Total Number of Employees (az_ges)

Variable label	Total number of full-time employees
Variable name	az_ges
Category	Generated establishment variables
Origin	BeH
Detailed description	<p>This variable contains the total number of an establishment's employees reported to the social security agencies as of Jun 30th of a given 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.</p>

6.4 Proportion of Women in an Establishment (at_f)

Variable label	Proportion of women in an establishment
Variable name	at_f
Category	Establishment attributes
Origin	Generated from BHP
Data type	Numerical
Hierarchy	None
Detailed description	<p>This attribute stores the share of women among all employees of an establishment. The proportion is calculated using the variables "Number of women (az_f)" and "Total number of employees (az_ges)."</p>

6.5 Proportion of Low-Qualified Employees in an Establishment (at_gq)

Variable label	Proportion of low-qualified employees in an establishment
Variable name	at_gq

Category	Establishment attributes
Origin	Generated from BHP
Data type	Numerical
Hierarchy	None
Detailed description	This attribute stores the share of low-qualified employees among all employees of an establishment. The proportion is calculated using the variables "Number of low-qualified employees (az_gq)" and "Total number of employees (az_ges)."

6.6 Proportion of Employees Under 30 in an Establishment (at_u30)

Variable label	Proportion of employees under 30 in an establishment
Variable name	at_u30
Category	Establishment attributes
Origin	Generated from BHP
Data type	Numerical
Hierarchy	None
Detailed description	This attribute stores the share of employees under the age of 30 among all employees of an establishment. The proportion is calculated using the variables "Number of employees between 15 and 19 (az_15_19)," "Number of employees between 20 and 24 (az_20_24)," "Number of employees between 25 and 29 (az_25_29)," and "Total number of employees (az_ges)."

6.7 Proportion of Employees Over 50 in an Establishment (at_u30)

Variable label	Proportion of employees over 50 in an establishment (at_u30)
Variable name	at_ü50
Category	Establishment attributes
Origin	Generated from BHP
Data type	Numerical
Hierarchy	None
Detailed description	This attribute stores the share of employees over the age of 50 among all employees of an establishment. The proportion is calculated using the variables "Number of employees between 50 and 54 (az_50_54)," "Number of employees between 55 and 59 (az_55_59)," "Number of employees between 60 and 64 (az_60_64)," and "Total number of employees (az_ges)."

6.8 Gross Daily Rate of Pay of All Full-Time Employees (te_p25, te_med, te_p75)

Variable label	Gross daily rate of pay of all full-time employees
Variable name	te_p25, te_med, te_p75

Category	Establishment attributes
Origin	Generated from BHP
Data type	Numerical
Hierarchy	None
Detailed description	This attribute includes the quartiles of the gross daily rate of pay of all full-time employees.

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 (Ed.) (2005): Schlüsselverzeichnis für die Angaben zur Tätigkeit in den Meldungen zur Sozialversicherung. January 2005 edition, Nuremberg.
- Cramer, Ulrich (1985): Probleme der Genauigkeit der Beschäftigtenstatistik. In: Allgemeines Statistisches Archiv 69: p. 56-68.
- Dorner, Matthias; Heining, Jörg; Jacobebbinghaus, Peter; Seth, Stefan (2010): Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) 1975-2008. FDZ data report, 01/2010 (de). Nuremberg, 63 p.
- Dorner, Matthias; Heining, Jörg; Jacobebbinghaus, Peter; Seth, Stefan (2010): The Sample of Integrated Labour Market Biographies. In: Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften, vol. 130(4), p. 599-608.
- Drews, Nils (2006): Qualitätsverbesserung der Bildungsvariable in der IAB-Beschäftigtenstichprobe 1975-2001. (FDZ method report, 05/2006 (de)), Nuremberg, 16 p.
- Fitzenberger, Bernd; Osikominu, Aderonke; Völter, Robert (2006): Imputation rules to improve the education variable in the IAB employment subsample. In: Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften, vol. 126(3), p. 405-436.
- Himmelreicher, Ralf K.; Stegmann, Michael (2008): New Possibilities for Socio-Economic Research through Longitudinal Data from the Research Data Centre of the German Federal Pension Insurance (FDZ-RV), Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften, 4, 647-660.
- Hochfellner, Daniela; Voigt, Axel; Budzak, Ute; Steppich, Birgit (2010): Das Projekt BASiD: Biographiedaten ausgewählter Sozialversicherungsträger, in: Deutschland. Projektinhalte, aktueller Stand der Arbeiten und Analysemöglichkeiten, in DRV Schriften.
- Jacobebbinghaus, Peter; Seth, Stefan (2007): The German integrated employment biographies sample IEBS, Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften, 2, 335-342.
- Meinken, Holger; Koch, Iris (2004): BA-Beschäftigtenpanel 1998-2002. Codebuch, Nuremberg.

Richter, Maria; Himmelreicher, Ralf K. (2008): Die Versicherungskontenstichprobe als Datengrundlage für Analysen von Versicherungsbiografien unterschiedlicher Altersjahrgänge, DRV Schriften 79, 34-61.

Spengler, Anja (2009): The Establishment History Panel, Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften, 3, 501-509.

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.