

# PASS-Befragungsdaten verknüpft mit administrativen Daten des IAB (PASS-ADIAB) 1975-2015

Manfred Antoni (Institut für Arbeitsmarkt- und Berufsforschung (IAB))

Sandra Dummert (Institut für Arbeitsmarkt- und Berufsforschung (IAB))

Simon Trenkle (Institut für Arbeitsmarkt- und Berufsforschung (IAB))

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

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

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung und Kurzbeschreibung</b>	<b>8</b>
1.1	Einleitung	8
1.2	Datennutzung	10
1.2.1	Datenzugang und Datenhaltung	10
1.2.2	Dateinamen im PASS-ADIAB 7515	12
1.3	Veränderungen zu früheren Versionen	14
1.3.1	Einführung des Tätigkeitsschlüssels 2010	14
1.3.2	Aufnahme der Maßnahmeteilnehmergehistoriken (MTH)	15
1.3.3	Merkmalskanon	15
1.4	Steckbrief	16
1.5	Merkmalsliste	21
1.6	Mengengerüst und Zeiträume	25
<b>2</b>	<b>Datenquellen</b>	<b>26</b>
2.1	PASS-Befragungsdaten	26
2.2	Administrative Daten aus den IEB	28
2.2.1	Beschäftigtenhistorik (BeH)	29
2.2.2	Leistungsempfängerhistorik (LeH)	30
2.2.3	Leistungshistorik Grundsicherung (LHG)	30
2.2.4	Arbeitsuchendenhistorik (ASU/XASU)	31
2.2.5	Maßnahmeteilnehmergehistoriken (MTH)	32
<b>3</b>	<b>Datengenerierung und -verknüpfung</b>	<b>32</b>
3.1	PASS-Befragungsdaten	32
3.2	Administrative Daten aus den IEB	32
3.2.1	Bereinigungen und Aufbereitungen	32
3.2.1.1	Beschäftigtenhistorik (BeH)	33
3.2.1.2	Leistungsempfängerhistorik (LeH)	33
3.2.1.3	Leistungshistorik Grundsicherung (LHG)	34
3.2.1.4	Arbeitsuchendenhistorik ASU und XASU	34
3.2.1.5	Maßnahmeteilnehmergehistoriken (MTH)	35
3.2.1.6	SGB-II-Anonymisierung	36
3.2.2	Episodensplitting	36
3.2.3	Fehlende Werte	37
3.3	Datenverknüpfung	37
3.3.1	Bereinigung und Standardisierung personenbezogener Merkmale	40
3.3.2	Goldstandard-Linkage	40
3.3.3	Deterministisches Record Linkage	41
3.3.4	Fehlertolerantes Record Linkage	41
<b>4</b>	<b>Datenqualität und Probleme</b>	<b>42</b>
4.1	PASS-Befragungsdaten	42
4.2	Gesamte IEB	42
4.3	Beschäftigtenhistorik (BeH)	45
4.4	Leistungsempfängerhistorik (LeH)	47
4.5	Leistungshistorik Grundsicherung (LHG)	47
4.6	Arbeitsuchendenhistorik (ASU/XASU)	48
4.6.1	ASU	48

4.6.2	XASU.....	49
4.7	Maßnahmeteilnehmehistoriken (MTH).....	49
<b>5</b>	<b>Variablenbeschreibung.....</b>	<b>50</b>
5.1	Identifikatoren.....	50
5.1.1	Unveränderliche Personennummer (pnr).....	50
5.1.2	Systemfreie Betriebsnummer (betnr).....	51
5.2	Generierte technische Merkmale.....	52
5.2.1	Quelle des Satzes (quelle).....	52
5.2.2	Satzzähler pro Konto (spell).....	52
5.2.3	Jahr (jahr).....	52
5.2.4	Vorliegen administrativer Personendaten (match_admin).....	53
5.2.5	Typ des Matches (match_type).....	53
5.2.6	Wellenindikator erfolgreicher Match (match_wave).....	53
5.3	Gültigkeitszeitraum.....	54
5.3.1	Beginndatum Originalbeobachtung (begorig).....	54
5.3.2	Enddatum Originalbeobachtung (endorig).....	54
5.3.3	Beginndatum der Episode (begepi).....	55
5.3.4	Enddatum der Episode (endepe).....	55
5.4	Informationen zur Person.....	56
5.4.1	Geschlecht (frau).....	56
5.4.2	Geburtsjahr (gebjahr).....	56
5.4.3	Staatsangehörigkeit (nation).....	56
5.4.4	Staatsangehörigkeit vergrößert (nation_gr).....	56
5.4.5	Familienstand (famst).....	57
5.4.6	Kinderzahl (kind).....	57
5.4.7	Ausbildung (ausbildung).....	58
5.4.8	Schulausbildung (schule).....	59
5.5	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche.....	60
5.5.1	Tagesentgelt/ täglicher Leistungssatz (tentgelt).....	60
5.5.2	Beruf – ausgeübte/letzte Tätigkeit (KldB 1988) (beruf).....	62
5.5.3	Berufsgruppe – ausgeübte/letzte Tätigkeit (KldB 2010) (beruf2010_3).....	62
5.5.4	Berufsuntergruppe – ausgeübte/letzte Tätigkeit (KldB 2010) (beruf2010_4).....	63
5.5.5	Anforderungsniveau – ausgeübte/letzte Tätigkeit (KldB 2010) (niveau).....	64
5.5.6	Teilzeit (teilzeit).....	65
5.5.7	Erwerbsstatus (erwstat).....	65
5.5.8	Gleitzone (gleitz).....	67
5.5.9	Leiharbeit (leih).....	68
5.5.10	Befristung (befrist).....	68
5.5.11	Abmeldegrund / Abgabegrund / Beendigungsgrund (grund).....	68
5.5.12	Erwerbsstatus vor Arbeitsuche (estatvor).....	70
5.5.13	Status nach Arbeitsuche (estatnach).....	70
5.5.14	Profillage (profil).....	71
5.5.15	Art der Kündigung der letzten Tätigkeit (art_kuend).....	71
5.5.16	Arbeitszeit des Stellengesuchs (arbzeit).....	72
5.5.17	Restanspruch/geplante Dauer (restanspruch).....	72
5.5.18	Trägerart (traeger).....	72
5.5.19	Beginndatum der Arbeitslosigkeit (alo_beg).....	73
5.5.20	Dauer der Arbeitslosigkeit (alo_dau).....	73
5.6	Ortsangaben.....	74
5.6.1	Wohnort - Kreis (wo_kreis).....	74
5.6.2	Wohnort – Bundesland (wo_bula).....	75
5.6.3	Wohnort - Arbeitsagentur (wo_aa).....	75
5.6.4	Wohnort - Regionaldirektion (wo_rd).....	76
5.7	Betriebsmerkmale.....	76

5.7.1	WZ73 3-Steller (w73_3).....	76
5.7.2	WZ93 5-Steller (w93_5).....	77
5.7.3	WZ93 3-Steller (w93_3).....	77
5.7.4	WZ03 5-Steller (w03_5).....	78
5.7.5	WZ03 3-Steller (w03_3).....	78
5.7.6	WZ08 5-Steller (w08_5).....	79
5.7.7	WZ08 3-Steller (w08_3).....	79
5.7.8	W73_3 vervollständigt durch Extrapolation/Imputation (w73_3_gen) .....	80
5.7.9	Art der Vervollständigung w73_3 (group_w73_3) .....	80
5.7.10	W93_3 vervollständigt durch Extrapolation/Imputation (w93_3_gen) .....	80
5.7.11	Art der Vervollständigung w93_3 (group_w93_3) .....	81
5.7.12	Jahr des 1. Auftretens der Betriebsnummer (grd_jahr) .....	81
5.7.13	Erstes Auftreten Betriebsnummer (grd_dat).....	82
5.7.14	Jahr des letzten Auftretens der Betriebsnummer (lzt_jahr).....	82
5.7.15	Letztes Auftreten Betriebsnummer (lzt_dat).....	83
5.7.16	Anzahl Beschäftigte gesamt (az_ges).....	83
5.7.17	Anzahl in Vollzeit (Normalbeschäftigte + sonstige) (az_vz) .....	83
5.7.18	Anzahl geringfügig Beschäftigte (az_gf) .....	84
5.7.19	Mittelwert imputiertes Bruttotagesentgelt Vollzeitbeschäftigte (te_imp_mw) .....	84
5.7.20	Arbeitsort Kreis (ao_kreis).....	85
5.7.21	Arbeitsort Bundesland (ao_bula).....	85
<b>6</b>	<b>Literaturverzeichnis .....</b>	<b>86</b>
<b>7</b>	<b>Anhang .....</b>	<b>89</b>
7.1	<i>Auszählungen</i> .....	89
<b>8</b>	<b>Abkürzungsverzeichnis .....</b>	<b>90</b>

## Zusammenfassung

Dieser Datenreport beschreibt die verknüpften Daten des Panels Arbeitsmarkt und soziale Sicherung (PASS) mit administrativen Daten des Instituts für Arbeitsmarkt- und Berufsforschung (IAB).

## Abstract

This data report describes the linked data of the panel study Labour Market and Social Security (PASS) with administrative data of the Institute for Employment Research (IAB).

**Keywords:** German administrative micro data, labour market data, linked employer-employee data, record linkage, data manual

## Danksagung

Unser Dank gilt Tobias Bachteler, Arne Bethmann, Johannes Eggs, Martina Huber und Anja Gruhl, die für die Datenverknüpfung bis einschließlich Welle 8 mitverantwortlich waren. Darüber hinaus danken wir Cerstin Rauscher (DIM) und Stefan Seth für ihre Unterstützung im Rahmen der Datenziehung und -aufbereitung, sowie Kevin Ruf für seine Unterstützung bei der Erstellung des Datenreports.

Zur Beschreibung der administrativen Daten wurden Textteile aus Datenbeschreibungen von DIM und aus früheren FDZ-Datenreporten übernommen. Die Verantwortung für die Richtigkeit der Inhalte dieses Datenreportes obliegt allein den hier genannten Autoren.

## **Datenverfügbarkeit**

Der in diesem Beitrag beschriebene Datensatz ist für die Fachöffentlichkeit zugänglich. Nähere Informationen dazu auf der Internetseite <http://fdz.iab.de/>.

# 1 Einleitung und Kurzbeschreibung

## 1.1 Einleitung

Dieser Datenreport beschreibt die verknüpften Befragungsdaten des Panels Arbeitsmarkt- und soziale Sicherung (PASS) mit den administrativen Daten des Instituts für Arbeitsmarkt- und Berufsforschung (IAB). Diese verknüpften Prozess- und Befragungsdaten bestehen aus in Interviews erhobenen Informationen von Teilnehmerinnen und Teilnehmern an der IAB-Erhebung PASS und ihren Personendaten aus den administrativen Daten der BA, sofern sie ihre Zustimmung zur Verknüpfung der Daten erteilt haben und dort identifiziert werden konnten. Um PASS-ADIAB analysieren zu können, ist dieser FDZ-Datenreport alleine nicht ausreichend, sondern als Ergänzung zu den bestehenden Dokumentationen von PASS zu sehen. Bitte nutzen Sie zusätzlich die FDZ-Datenreporte zu den Wellen 1 bis 9 des PASS (siehe z.B. Berg et al., 2016; Bethmann et al., 2013).

Die Befragungsdaten stammen aus den ersten neun Wellen des Panels Arbeitsmarkt und soziale Sicherung. Der gleichnamige Forschungsbereich (FB E3) des IAB ist verantwortlich für die Planung und Durchführung der Studie, während das FDZ für Aufbereitung, Dokumentation und Verbreitung der Daten verantwortlich ist. Die Studie wird vom Bundesministerium für Arbeit und Soziales (BMAS) finanziert und seit Dezember 2006 jährlich durchgeführt. Die ursprüngliche Stichprobe umfasste Interviews von 18.954 Personen in 12.794 Haushalten und bestand zu ungefähr gleichen Teilen aus einer bevölkerungsrepräsentativen Teilstichprobe und einer Teilstichprobe aus dem Bestand der Bezieher von Arbeitslosengeld II (ALG II). In den folgenden Wellen wurden die Fälle jeweils um eine Stichprobe von ca. 1.000 Haushalten aus den Zugängen in den ALG-II-Bezug seit dem Zeitpunkt der letzten Stichprobenziehung ergänzt. In Welle 5 folgte zusätzlich je eine Bestandsauffrischung der Bevölkerungs- und der ALG-II-Bezieherstichprobe. Unter anderem erlaubt das Gewichtungskonzept eine kombinierte Hochrechnung beider Teilstichproben auf die deutsche Wohnbevölkerung. Die Feldarbeit der ersten drei Wellen wurde von TNS Infratest Sozialforschung durchgeführt. Seit der vierten Welle übernimmt das ifas Institut für angewandte Sozialwissenschaft diese Aufgaben. Als Haushaltspanel werden im Rahmen von PASS alle Haushaltsmitglieder ab 15 Jahren befragt. Die PASS-Daten enthalten u. a. detaillierte Informationen zur Dynamik des Bezugs von Arbeitslosengeld II sowie zur Versorgungslage und sozialen Teilhabe von Haushalten. Zu den befragten Personen wurden u. a. Details zur Arbeitsmarktintegration, arbeitsmarktrelevanten Einstellungen und Verhaltensweisen, physischer und psychischer Gesundheit, sowie zur sozialen Vernetzung erhoben. Für Welle 9 liegen Informationen von 13.271 Personen in 8.921 Haushalten vor.

Die verwendeten administrativen Personendaten stammen aus den Integrierten Erwerbsbiographien des IAB (IEB). Die IEB enthalten Daten zu allen Personen, die im Beobachtungszeitraum mindestens einmal einen der folgenden Zustände aufweisen:

- sozialversicherungspflichtige Beschäftigung in Deutschland (erfasst ab 1975),
- geringfügige Beschäftigung in Deutschland (erfasst ab 1999),
- Bezug von Leistungen nach dem Rechtskreis SGB III (erfasst ab 1975) oder SGB II (erfasst ab 2005),
- bei der Bundesagentur für Arbeit (BA) bzw. den Grundsicherungsträgern als arbeitsuchend gemeldet (erfasst ab 2000),
- Teilnahme an arbeitsmarktpolitischer Maßnahme (erfasst ab 2000).

Diese aus unterschiedlichen Datenquellen stammenden Informationen werden in den IEB zusammengeführt und die Zustände jeweils tagesgenau abgebildet. Während die Beschäftigtenhistorik (BeH) Ursprung der Informationen zu sozialversicherungspflichtiger und geringfügiger Beschäftigung ist, wird der Bezug von Leistungen nach den Rechtskreisen des SGB III und des SGB II in der Leistungsempfängerhistorik (LeH) bzw. in der Leistungshistorik Grundsicherung (LHG) erfasst. Die Arbeitsuchendenhistoriken (ASU und XASU) stellen die Datenquelle für die durch die BA bzw. die durch die kommunalen Grundsicherungsträger erfassten Zeiten der Arbeitsuche dar, während Teilnahmen an arbeitsmarktpolitischen Maßnahmen in der Maßnahmeteilnahmehistorik (MTH) erfasst werden. Anders als in der früheren Version PASS-ADIAB 7511 sind Informationen zur Teilnahme an arbeitsmarktpolitischen Maßnahmen aus der MTH in der aktuellen Version der PASS-ADIAB enthalten<sup>1</sup>.

Dieser Datenreport gliedert sich wie folgt: Neben der Einleitung enthält Kapitel 1 unter anderem Informationen zum Datenzugang sowie eine Kurzbeschreibung der Daten, das Mengengerüst und eine Merkmalsliste. Eine Beschreibung der einzelnen Datenquellen findet sich in Kapitel 2. Datenaufbereitung und -qualität sowie die Verknüpfung werden in den Kapiteln 3 und 4 dargestellt, während die einzelnen Merkmale in Kapitel 5 beschrieben werden.

---

<sup>1</sup> Die bereitgestellten administrativen Daten entsprechen was die Aufbereitung, die Auswahl und die Bezeichnung der Variablen angeht denen der Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) (Antoni et al., 2016). Für Bezieher von Leistungen des SGB II, die in den Zuständigkeitsbereich von zugelassenen kommunalen Trägern (optierende Kommunen) fallen, sind dabei keine Daten zur Teilnahme an Maßnahmen enthalten.



## 1.2 Datennutzung

### 1.2.1 Datenzugang und Datenhaltung

PASS-ADIAB kann ausschließlich im Rahmen von Gastaufenthalten am FDZ und anschließender Datenfernverarbeitung ausgewertet werden. Für die Datennutzung ist zunächst ein Antrag beim FDZ einzureichen. Über die Genehmigung des Forschungsprojekts entscheidet das Bundesministerium für Arbeit und Soziales (BMAS). Nach erfolgter Genehmigung wird ein Nutzungsvertrag mit der Institution des Forschenden geschlossen. Details zur Beantragung des Datensatzes und Datenverarbeitungsmöglichkeiten können der FDZ-Website entnommen werden.

Der PASS-ADIAB-Datensatz ist nach dem Ursprung der Variablen gegliedert und in mehreren Dateien abgelegt. Als Ausgangspunkt dienen die *PASS-Personendaten*. Sie enthalten die im Interview erhobenen persönlichen Merkmale der PASS-Befragten sowie deren Personen- und Haushaltsidentifikatoren. Über die Identifikatoren können andere Teildatensätze der PASS-Befragungsdaten – wie z. B. Haushaltsdaten oder Informationen zu Erwerbs- oder ALG-II-Biographien – zugespielt werden.

Die administrativen Daten in PASS-ADIAB bestehen zum einen aus der *administrativen Personendatei*. Diese enthält Identifikatoren (Personen-ID und systemfreie Betriebsnummer), persönliche Merkmale, Informationen zu Beschäftigung, Leistungsbezug, Arbeitsuche und Teilnahme an arbeitsmarktpolitischen Maßnahmen, Merkmale zum Wohnort sowie technische Merkmale. Zum anderen besteht der administrative Teil von PASS-ADIAB aus der *Betriebsdatei*. Diese enthält eine systemfreie Betriebsnummer, die eine Verknüpfung mit der administrativen Personendatei erlaubt, sowie jahresbezogene Betriebsmerkmale. Die in der Merkmalsliste (siehe S. 21ff.) mit einem „\*“ gekennzeichneten Variablen sind in der Betriebsdatei enthalten. Somit sind Befragungsdaten und administrative Betriebs- bzw. Personenmerkmale organisatorisch voneinander getrennt, was die Struktur der Daten deutlich macht und Speicherplatz in der Datenhaltung spart.

Ein weiteres Element von PASS-ADIAB ist die *Verknüpfungsdatei*, welche die Personennummer sowie technische Merkmale enthält, die angeben, ob bzw. wie die Verknüpfung mit den administrativen Daten gelungen ist. Die Datei enthält auch die PASS-Befragten, zu denen keine verknüpften administrativen Daten gefunden werden konnten. Damit können eigene Ausfallanalysen auf Basis aller PASS-Teilnehmer durchgeführt werden.

Bestimmte Merkmale, die eine Deanonymisierung von Personen oder Betrieben erleichtern, werden nur dann im Original weitergegeben, wenn es für das Analyseziel notwendig ist und dies im Antrag auf den Datenzugang explizit begründet wird. Diese aus datenschutzrechtlicher Sicht besonders sensiblen Merkmale sind:

### **Administrative Personendatei:**

- Staatsangehörigkeit (nation)
- Wohnort Arbeitsagentur (wo\_aa)
- Wohnort Kreis (wo\_kreis)
- Berufsuntergruppe (beruf2010\_4)

### **Administrative Basisbetriebsdatei:**

- Arbeitsort Kreis (ao\_kreis)
- Wirtschaftszweig - Wirtschaftsunterklasse (w93\_5)
- Wirtschaftszweig - Wirtschaftsunterklasse (w03\_5)
- Wirtschaftszweig - Wirtschaftsunterklasse (w08\_5)
- Erstes Auftreten Betriebsnummer (grd\_dat)
- Letztes Auftreten Betriebsnummer (lzt\_dat)

Beachten Sie, dass die sensiblen Merkmale in der Regel in vergrößerter Version bereits in den Daten enthalten sind. Die sensiblen Merkmale werden nur weitergegeben, wenn die Information im vergrößerten Merkmal zur Erreichung des Forschungsziels nicht ausreichend ist. Auf begründeten Antrag hin können aus dem Betriebs-Historik-Panel (BHP) (Schmucker et al., 2016) weitere Betriebsmerkmale bereitgestellt werden.

Folgende **Variablenblöcke** aus dem BHP 1975-2014 können für PASS-ADIAB 7515

beantragt werden:

- Struktur der Beschäftigten allgemein (az\_ges az\_f az\_reg az\_gf az\_azubi az\_atz az\_vz az\_tz az\_f\_vz az\_f\_tz az\_reg\_vz)
- Struktur der Beschäftigten nach Schul- und Berufsbildung (az\_gq az\_mq az\_hq az\_gq\_vz az\_mq\_vz az\_hq\_vz)
- Altersstruktur der Beschäftigten (az\_15\_19 az\_15\_19\_vz az\_20\_24 az\_20\_24\_vz az\_25\_29 az\_25\_29\_vz az\_30\_34 az\_30\_34\_vz az\_35\_39 az\_35\_39\_vz az\_40\_44 az\_40\_44\_vz az\_45\_49 az\_45\_49\_vz az\_50\_54 az\_50\_54\_vz az\_55\_59 az\_55\_59\_vz az\_60\_64 az\_60\_64\_vz az\_ab65 az\_ab65\_vz alter\_mw alter\_mw\_vz)
- Aktivitäten Forschung und Entwicklung (az\_ingnat)
- Anzahl atypisch Beschäftigter (az\_leih az\_bfr)
- Struktur der Beschäftigten nach Staatsangehörigkeit (az\_d az\_d\_vz az\_eu)
- Struktur der Beschäftigten nach Blossfeld Berufsqualifikation (az\_bf\_agr az\_bf\_emb az\_bf\_edi az\_bf\_evb az\_bf\_qmb az\_bf\_qdi az\_bf\_qvb az\_bf\_tec az\_bf\_semi az\_bf\_ing az\_bf\_prof az\_bf\_man)

- Struktur der Beschäftigten nach Tätigkeitsniveau (az\_niv1 az\_niv2 az\_niv3 az\_niv4)
- Gehaltsstruktur der Vollzeitbeschäftigten (az\_zens te\_med te\_p25 te\_p75 te\_imp\_mw te\_imp\_med te\_imp\_p25 te\_imp\_p75 te\_imp\_mw\_f te\_imp\_med\_f te\_imp\_med\_m te\_imp\_mw\_gq te\_imp\_med\_gq te\_imp\_mw\_mq te\_imp\_med\_mq te\_imp\_mw\_hq te\_imp\_med\_hq te\_imp\_med\_uq te\_imp\_mw\_d te\_imp\_med\_d te\_imp\_med\_a)

Darüber hinaus können auch folgende **Erweiterungsmodule** des BHP beantragt werden:

- Erweiterungsmodul – Beschäftigtenströme (ein\_ges ein\_f ein\_reg ein\_gf ein\_azubi ein\_vz ein\_tz ein\_vz\_f ein\_tz\_f ein\_reg\_vz ein\_wdr ein\_wdr\_f ein\_bw ein\_bw\_f ein\_bf\_agr ein\_bf\_emb ein\_bg\_edi ein\_bf\_evb ein\_bf\_qmb az\_bf\_qdi ein\_bf\_qvb ein\_bf\_tec ein\_bf\_semi ein\_bf\_ing ein\_bf\_prof ein\_bf\_man ein\_15\_19 ein\_20\_24 ein\_25\_29 ein\_30\_34 ein\_35\_39 ein\_40\_44 ein\_45\_49 ein\_50\_54 ein\_55\_59 ein\_60\_64 ein\_ab65 aus\_ges aus\_f aus\_reg aus\_gf aus\_azubi aus\_vz aus\_tz aus\_vz\_f aus\_tz\_f aus\_reg\_vz aus\_temp aus\_bw aus\_bw\_f aus\_bf\_agr aus\_bf\_emb aus\_bf\_edi aus\_bf\_evb aus\_bf\_qmb aus\_bf\_qdi aus\_bf\_qvb aus\_bf\_tec aus\_bf\_semi aus\_bf\_ing aus\_bf\_prof aus\_bf\_man aus\_15\_19 aus\_20\_24 aus\_25\_29 aus\_30\_34 aus\_35\_39 aus\_40\_44 aus\_45\_49 aus\_50\_54 aus\_55\_59 aus\_60\_64 aus\_ab65 aus\_senio\_1 aus\_senio\_2 aus\_senio\_3)
- Erweiterungsmodul – Betriebsdynamik (eintritt besch inflow betnr\_vor besch\_vor status\_vor austritt besch outflow betnr\_nach besch\_nach status\_nach)

**Die Erforderlichkeit der zusätzlichen Daten muss begründet werden: Bei den sensiblen Merkmalen für jedes einzelne Merkmal, bei den Variablenblöcken pro Block und für jedes einzelne Erweiterungsmodul.**

Ausführliche Informationen zu den einzelnen Variablen und der Erweiterungsmodule des BHP finden sich im [FDZ-Datenreport 03/2016](#).

## 1.2.2 Dateinamen im PASS-ADIAB 7515

### Dateinamen der administrativen Daten

#### Personendatei

PASS-ADIAB\_7515\_v1.dta

#### Basisbetriebsdatei

PASS-ADIAB\_7515\_v1\_bhp\_basis\_v1.dta

Verknüpfungsdatei

PASS\_ADIAB\_7515\_v1\_link.dta

Zusatzmodule BHP

Bestandsdaten mit Variablenblöcken:

PASS-ADIAB\_7515\_v1\_bhp\_v1\_YYYY.dta, YYYY = 1975 – 2014

Erweiterungsmodul – Beschäftigtenströme:

PASS-ADIAB\_7515\_v1\_bhp\_inflow\_v1.dta

PASS-ADIAB\_7515\_v1\_bhp\_outflow\_v1.dta

Erweiterungsmodul – Betriebsdynamik:

PASS-ADIAB\_7515\_v1\_bhp\_ein\_v1.dta

PASS-ADIAB\_7515\_v1\_bhp\_aus\_v1.dta

**Dateinamen der PASS-Befragungsdaten**

<b>Datensatz</b>	<b>Dateiname</b>	<b>Besonderheiten</b>
Haushaltsdatensatz	HHENDDAT.dta	
Personendatensatz	PENDDAT.dta	
Spelldaten ALG I	alg1_spells.dta	nur Welle 1
Spelldaten ALG II	alg2_spells.dta	
Spelldaten Arbeitslosigkeiten	al_spells.dta	Wellen 2 und 3
Spelldaten Erwerbstätigkeiten	et_spells.dta	Wellen 2 und 3
Spelldaten Lücken	lu_spells.dta	Wellen 2 und 3
Spelldaten Erwerbstätigkeiten, Arbeitslosigkeiten und Lücken integriert	bio_spells.dta	ab Welle 2
Spelldaten Maßnahmeteilnahme	massnahmespells.dta	nur Welle 1
Spelldaten Maßnahmen	mn_spells.dta	Wellen 2 und 3
Spelldaten Ein-Euro-Jobs	ee_spells.dta	ab Welle 4
Registerdaten Haushaltsebene	hh_register.dta	
Registerdaten Personenebene	p_register.dta	
Gewichtungsdaten Haushaltsebene	hweights.dta	
Gewichtungsdaten Personenebene	pweights.dta	
Altersvorsorge Haushaltsebene	HAVDAT.dta	nur Welle 3
Altersvorsorge Personenebene	PAVDAT.dta	nur Welle 3
Vignettendaten	VIGDAT.dat	nur Welle 5
Kinderdatensatz	KINDER.dta	ab Welle 6

## 1.3 Veränderungen zu früheren Versionen

### 1.3.1 Einführung des Tätigkeitsschlüssels 2010

Die Übermittlung von Informationen zur Ausbildung, ausgeübten Tätigkeit und Stellung im Beruf erfolgt durch Meldungen der Arbeitgeber nach DEÜV (siehe Abschnitt 2.2.1) über einen sogenannten Tätigkeitsschlüssel. Für Zeiträume mit einem Enddatum nach dem 30.11.2011 erfolgte der Wechsel zum neuen Tätigkeitsschlüssel 2010 (siehe Bertat et al., 2013 für mehr Details). Dieser Wechsel wurde von den Spitzenorganisationen der Sozialversicherungsträger beschlossen, da der Tätigkeitsschlüssel 2003 eine Reihe von Sachverhalten nicht mehr zeitgemäß und nicht mehr realitätsnah erfassen konnte.

Da die Meldungen der Arbeitgeber nach DEÜV in die BeH einfließen, hat die Umstellung des Tätigkeitsschlüssels fast ausschließlich Auswirkungen für Beobachtungen aus dieser Quelle. Die Messung folgender bisher enthaltener Angaben ist von dieser Umstellung betroffen: Arbeitszeit, Beruf, Stellung im Beruf sowie schulisches und berufliches Qualifikationsniveau. Darüber hinaus gibt es seit der Umstellung auch Angaben zur Befristung eines Beschäftigungsverhältnisses sowie dazu, ob eine Person als Verleihpersonal in einem Zeitarbeitsunternehmen angestellt ist. Die wichtigste Auswirkung ist der Umstieg auf eine neue Berufsklassifikation. Statt der bisherigen Klassifikation der Berufe 1988 (KIdB 1988) wird mit dem neuen Tätigkeitsschlüssel die differenziertere KIdB 2010 gemeldet.<sup>2</sup> Eine weitere wichtige Änderung ist der Wegfall der Variable „Stellung im Beruf und Arbeitszeit“ (stib), da die meisten der darin gegebenen Ausprägungen mit dem neuen Tätigkeitsschlüssel nicht mehr nachvollzogen werden können. Die Kategorien der früheren Variable stib, die über den gesamten Beobachtungszeitraum konsistent gefüllt werden können (Erwerbsstatus als Auszubildender, Unterscheidung von Teil- und Vollzeit), sind künftig in anderen Variablen nachzuvollziehen. Im Abschnitt 1.3.3 wird aufgezeigt, welche Merkmale durch die Umstellung hinzukommen, entfallen oder künftig verändert in den Daten enthalten sind.

Abschnitt 4.3 zeigt auf, welche Auswirkungen die Umstellung auf die Datenqualität der BeH-Beobachtungen hat. Eine der schwerwiegendsten Folgen des Umstellungsprozesses ist ein temporärer Anstieg von fehlenden Werten in den von Arbeitgebern über den Tätigkeitsschlüssel zu meldenden Merkmalen. Dieser Anstieg betraf auch das neue Merkmal „Arbeitszeit“, weshalb für PASS-ADIAB auf eine imputierte Version des Merkmals zurückgegriffen wurde. Die

<sup>2</sup> Siehe Paulus/Matthes (2013) für Details zur Klassifikation der Berufe 2010.

Probleme verstärkt auftretender fehlender Werte betreffen im Wesentlichen die Jahre 2011 und 2012.

### 1.3.2 Aufnahme der Maßnahmeteilnehmehistoriken (MTH)

Die PASS-ADIAB 7515 enthält, anders als die Vorgängerversion, auch Informationen aus den ebenfalls in den IEB enthaltenen Maßnahmeteilnehmehistoriken (MTH), wobei die Maßnahmekategorien in PASS-ADIAB jedoch nur in vergrößerter Form angeboten werden. Die in der XMTH gespeicherten Maßnahmeteilnahmen der zugelassenen kommunalen Träger (bzw. der optierenden Kommunen) sind hingegen nicht enthalten. Da die XMTH in anderen Standarddatensätzen des FDZ (wie der SIAB) ebenfalls nicht enthalten ist, wurde auf eine Aufnahme der XMTH verzichtet.

### 1.3.3 Merkmalskanon

Aus Gründen der Datensparsamkeit oder wegen mangelnder Nachfrage entfallen einige Merkmale im Vergleich zur PASS-ADIAB 7511. Darüber hinaus mussten Merkmale entfernt werden, die durch die Umstellung auf den neuen Tätigkeitsschlüssel nicht mehr für den ganzen Beobachtungszeitraum zu füllen gewesen wären. Aufgrund dieser Umstellung konnte eine Reihe von Merkmalen neu aufgenommen werden, teilweise durch Umschlüsselung für den ganzen Zeitraum, teilweise jedoch nur für Beobachtungen ab der Umstellung Ende 2011. Tabelle 1 gibt einen Überblick über die Änderungen im Merkmalskanon über alle Quellen hinweg.

**Tabelle 1: Änderungen des Merkmalskanons**

Merkmal	Art der Änderung	Erläuterung
Beginn der Arbeitslosigkeit (alo_beg)	N	
Dauer der Arbeitslosigkeit (alo_dau)	N	
Arbeitsort Arbeitsagentur (ao_aa)	E	Entfällt ersatzlos; relevante Information jedoch in wo_aa enthalten
Arbeitsort Regionaldirektion (ao_rd)	E	Entfällt ersatzlos; relevante Information jedoch in wo_rd enthalten
Befristung (befrist)	N	Mit neuem Tätigkeitsschlüssel eingeführt; ab 2011 gefüllt
Berufsgruppe – ausgeübte/letzte Tätigkeit (KIdB 2010), 3-Steller (be-ruf2010_3)/ 4-Steller (beruf2010_4)	N	Mit neuem Tätigkeitsschlüssel eingeführt; durch Umschlüsselung aus alter Variable für ganzen Zeitraum gefüllt.
Ausbildung (bild)	E	Ersetzt durch ausbildung und schule
Status nach Arbeitssuche (estatnach)	N	

Merkmal	Art der Änderung	Erläuterung
Vermittlungsrelevante gesundheitliche Einschränkungen (gesund_ein)	E	Entfällt ersatzlos
BA-Kundengruppe (kunden_gr)	E	Durch profil ersetzt
Leistungsart (lart)	E	Entfällt ersatzlos
Leiharbeit (leih)	N	Mit neuem Tätigkeitsschlüssel eingeführt; ab 2011 gefüllt
Anforderungsniveau - ausgeübte/letzte Tätigkeit (KldB 2010) (niveau)	N	Mit neuem Tätigkeitsschlüssel eingeführt; durch Umschlüsselung aus alter Variable für ganzen Zeitraum gefüllt.
Profillage (profil)	N	Ersetzt alte Variable kunden_gr
Qualifikation (quali)	E	Nun in ausbildung enthalten
Schulausbildung (schbild)	E	Nun in schule enthalten
Schwerbehindertenstatus (schweb)	E	Entfällt ersatzlos
Stellung im Beruf und Arbeitszeit (stib)	E	Entfällt weil im Meldeverfahren nicht mehr erfasst; Unterscheidung zwischen Voll- und Teilzeit in Variable teilzeit nachzuvollziehen; Erwerbsstatus „Auszubildender“ in Variable erwstat erfasst.

N = neu, E = entfällt

Zum Wegfall von Arbeitsort – Arbeitsagentur / Regionaldirektion (ao\_aa / ao\_rd) ist zu beachten, dass die zuständige Geschäftsstelle durch den Wohnort und nicht durch den Arbeitsort definiert wird. Eine Unterscheidung ist demnach ohnehin nicht sinnvoll, so dass der Wegfall der beiden Variablen keinen Informationsverlust bedeutet.

## 1.4 Steckbrief

**Tabelle 2: Steckbrief**

Inhaltliche Charakteristika	
Themen/ Merkmalsgruppen	<u>Befragungsdaten:</u> Soziodemographische Merkmale, Erwerbsbiographien mit Erwerbstätigkeiten/Arbeitslosigkeiten und Zeiten der Nichterwerbstätigkeit, Arbeitssuche, Bezugshistorien von Arbeitslosengeld I und II, Haushaltsstruktur, Kinderbetreuung, Lebensstandard, Wohnsituation, Gesundheit, soziale Netzwerke, Migration, soziale Herkunft, Zufriedenheit, arbeitsmarktrelevante Einstellungen, Selbstwirksamkeit, Kontakte zu ALG-II-Trägern, Maßnahmeteilnahme, Mindestlohn  <u>Administrative Personendaten:</u>

	<p><b>Beschäftigtenhistorik (BeH):</b> Jahres- und Abmeldungen zur Sozialversicherung für sozialversicherungspflichtig und geringfügig Beschäftigte</p> <p><b>Leistungsempfängerhistorik (LeH):</b> Informationen zu Leistungsbezug nach SGB III für Bezieher von Arbeitslosengeld, Arbeitslosenhilfe und Unterhaltsgeld</p> <p><b>Leistungshistorik Grundsicherung aus A2LL (LHG):</b> Daten zu Personen, die Leistungen zur Sicherung des Lebensunterhalts nach SGB II beziehen (Trägerarten: ARGE/gemeinsame Einrichtungen, getrennte Trägerschaft/getrennte Aufgabenwahrnehmung, zugelassene kommunale Träger)</p> <p><b>Arbeitsuchendenhistorik (ASU):</b> Informationen zur Arbeitsuche</p> <p><b>Arbeitsuchendenhistorik aus XSozial-BA-SGB II (XASU):</b> Informationen zur Arbeitsuche, die über den Übermittlungsstandard XSozial-BA-SGB II von zugelassenen kommunalen Trägern an die BA gemeldet werden.</p> <p><b>Maßnahmeteilnahmehistoriken (MTH):</b> Informationen zu Teilnahmen an Maßnahmen der aktiven Arbeitsmarktförderung (keine Maßnahmen der zugelassenen kommunalen Träger im SGB II)</p> <p><u>Administrative Betriebsdaten:</u></p> <p><b>Kerndatensatz:</b> Betriebsmerkmale (z.B. systemfreie Betriebsnummer, Wirtschaftszweig, Bundesland), Erwerbsstruktur der Beschäftigten (z.B. Anzahl der Beschäftigten nach Geschlecht, Art der Beschäftigung), Altersstruktur der Beschäftigten, Struktur der Beschäftigten nach Schul- und Berufsausbildung, Struktur der Beschäftigten nach Berufsklassifikation (Blossfeld-Klassifikation), Durchschnittslöhne der Vollzeitbeschäftigten</p> <p><b>Erweiterungsmodul- Beschäftigtenströme:</b> Anzahl Zugänge gesamt, Struktur der Zugänge nach Geschlecht, Beschäftigungsart, Berufsklassifikation (Blossfeld-Klassifikation), Wiedereinstellung, Alter, Anzahl Abgänge gesamt, Struktur der Abgänge nach Geschlecht, Beschäftigungsart, Berufsklassifikation (Blossfeld-Klassifikation), Betriebszugehörigkeit, Alter</p> <p><b>Erweiterungsmodul- Betriebsdynamik:</b> Gründungsjahr, Art der Gründung, Hilfsvariablen zur Klassifizierung der Gründung, Schließungsjahr, Art der Schließung, Hilfsvariablen zur Klassifizierung der Schließung</p>
Untersuchungseinheit	<p><u>Befragungsdaten:</u> Personen und Haushalte der deutschen Wohnbevölkerung bzw. mit ALG-II-Bezug.</p>



	<u>Administrative Personendaten:</u> Sozialversicherungspflichtig Beschäftigte (ab 1999 auch geringfügig Beschäftigte), Leistungsempfänger, Arbeitssuchende, Maßnahmeteilnehmer
Fallzahlen	<u>Befragungsdaten:</u> Welle 1: 18.954 Personen (in 12.794 Haushalten) Welle 2: 12.487 Personen (in 8.429 Haushalten) Welle 3: 13.439 Personen (in 9.535 Haushalten) Welle 4: 11.768 Personen (in 7.848 Haushalten) Welle 5: 15.607 Personen (in 10.235 Haushalten) Welle 6: 14.619 Personen (in 9.513 Haushalten) Welle 7: 14.449 Personen (in 9.509 Haushalten) Welle 8: 13.460 Personen (in 8.998 Haushalten) Welle 9: 13.271 Personen (in 8.921 Haushalten) <u>Administrative Personendaten:</u> 25.048 Personen mit 1,980,360 Beobachtungen
Zeitraum	<u>Befragungsdaten</u> Welle 1: 2006/2007 Welle 2: 2007/2008 Welle 3: 2008/2009 Welle 4: 2010 Welle 5: 2011 Welle 6: 2012 Welle 7: 2013 Welle 8: 2014 Welle 9: 2015  <u>Administrative Daten:</u> Der Zeitraum hängt von der Datenquelle ab. BeH: 01.01.1975 - 31.12.2014 (2012: 30-Monats-Datei, 2013: 18-Monats-Datei; 2014: 6-Monats-Datei) LeH: 01.01.1975 - 31.12.2014 ASU: 01.01.1997 - 31.12.2014 LHG: 01.01.2005 - 31.12.2014 XASU: 01.01.2005 - 31.12.2014 MTH: 01.01.2000 - 31.12.2014
Zeitbezug	<u>Befragungsdaten:</u> Befragungszeitpunkt, bei Spelldaten zwei Jahre retrospektiv bzw. seit letzter Befragung <u>Administrative Personendaten:</u> tagesgenaue Erwerbsbiografien <u>Administrative Betriebsdaten:</u>

	Stichtag 30. Juni
Gebietsstand	<u>Befragungsdaten:</u> Befragungszeitpunkt <u>Administrative Daten:</u> Gebietsstandbereinigt zum 31.12.2014
Regionale Gliederung	<u>Befragungsdaten:</u> Bundesland, Ost/West <u>Administrative Daten:</u> Bundesländer, Kreise
<b>Methodische Charakteristika</b>	
Erhebungsdesign	<u>Befragungsdaten:</u> <p>Welle 1: Zweistufige Zufallsstichprobe mit zwei Teilpopulationen. In der ersten Stufe wurden 300 Postleitzahlenbezirke (PLZ) nach dem Probability-Proportional-to-Size-Verfahren (PPS) anhand der Einwohnerzahl ausgewählt. In Stufe 2 wurde eine Teilstichprobe von ALG-II-Bedarfsgemeinschaften aus den administrativen Daten der BA zum Stichtag Juli 2006 in diesen PLZs gezogen. Eine weitere Teilstichprobe wurde in den PLZs mit Hilfe der kommerziellen Datenbank Microm mosaic und einer anschließenden Begehung gezogen.</p> <p>Wellen 2–9: Jeweils Ergänzung um eine Zugangsstichprobe von Bedarfsgemeinschaften aus den administrativen Daten der BA für die ursprünglichen PLZs. Es wurde nur aus den Fällen gezogen, die zum Juli-Stichtag vor Feldbeginn ALG-II-Leistungen bezogen haben, nicht jedoch zum Zeitpunkt der vorangegangenen Stichprobenziehung.</p> <p>Welle 5: Bestandsauffrischungen der Bevölkerungs- und BA-Stichprobe über die Auswahl von 100 zusätzlichen PLZs. Die Bevölkerungsstichprobe wurde aus den Registern der betroffenen Einwohnermeldeämter gezogen, die BA-Stichprobe aus den ALG-II-Bestandsfällen der jeweiligen PLZs.</p> <p>Weitere Informationen zum Erhebungsdesign sind in Bethmann/Gebhardt (2011) und Berg u. a. (2016) zu finden.</p> <u>Administrative Daten:</u> Identifikation der Befragten in den Integrierten Erwerbsbiografien des IAB
An der Erhebung beteiligte Institutionen	<u>Befragungsdaten:</u> Auftraggeber: Institut für Arbeitsmarkt- und Berufsforschung Durchführung: TNS Infratest Sozialforschung (Wellen 1-3), infas Institut für angewandte Sozialwissenschaft (ab Welle 4)

	<u>Administrative Daten:</u> Sozialversicherungsträger, Bundesagentur für Arbeit
Frequenz der Datensammlung	<u>Befragungsdaten:</u> Jährlich (Panel) <u>Administrative Daten:</u> Laufend
Dateiorganisation	<u>Befragungsdaten</u> Die Daten werden in mehrere Dateien abgelegt. <u>Administrative Daten</u> Die Daten werden in zwei Dateien abgelegt. Die eine enthält personenbezogene und die andere betriebsbezogene Informationen. Auf begründeten Antrag werden weitere Dateien mit zusätzlichen Betriebsinformationen bereitgestellt.
Dateiformat	Stata
<b>Datenzugang</b>	
Zugangswege	Gastaufenthalt, Datenfernverarbeitung (nach vorangegangenem Gastaufenthalt)
Anonymisierungsgrad	Schwach anonymisiert
Sensible Merkmale	<u>Befragungsdaten:</u> Keine <u>Administrative Personendaten:</u> Staatsangehörigkeit (nation), Berufsuntergruppe (beruf2010_4), Wohnort - Kreis (wo_kreis), Wohnort – Arbeitsagentur (wo_aa) <u>Administrative Betriebsdaten:</u> Arbeitsort - Kreis (ao_kreis), Wirtschaftszweigklassifikation WZ93 5-Steller (wz93_5), Wirtschaftszweigklassifikation WZ03 5-Steller (wz03_5), Wirtschaftszweigklassifikation WZ08 5-Steller (wz08_5), Erstes Auftreten der Betriebsnummer (grd_dat), Letztes Auftreten der Betriebsnummer (lzt_dat)
Zusatzmerkmale	Es können weitere Betriebsmerkmale aus dem BHP sowie die BHP-Erweiterungsmodule zur Zuspierung an Personendaten beantragt werden. Bitte benennen Sie im Antrag für Gastaufenthalte die sensiblen Merkmale, Variablenblöcke und Erweiterungsmodule entsprechend der Liste der BHP-Merkmale in Abschnitt 1.2.1.
Vorgaben zum Zitieren der Daten und Datendokumentation	<u>Daten:</u> „Die Datengrundlage dieses Beitrags bilden die Daten des Panels Arbeitsmarkt und soziale Sicherung (PASS) verknüpft mit administrativen Daten des IAB (1975-2015). Der Datenzugang erfolgte über Gastaufenthalte am Forschungsdatenzentrum (FDZ) der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung und/oder mittels kontrollierter Datenfernverarbeitung.“

	<p><u>Datendokumentation:</u></p> <ul style="list-style-type: none"> <li>▪ Antoni, Manfred; Dummert, Sandra; Trenkle, Simon (2017): PASS-Befragungsdaten verknüpft mit administrativen Daten des IAB (PASS-ADIAB) 1975-2015. FDZ-Datenreport, 06/2017, Nürnberg.</li> <li>▪ Berg, Marco; Cramer, Ralph; Dickmann, Christian; Gilberg, Reiner; Jesske, Birgit; Kleudgen, Martin; Bethmann, Arne; Fuchs, Benjamin; Huber, Martina; Schwarz, Stefan; Trappmann, Mark; Reindl, Alice (2016): Codebuch und Dokumentation des 'Panel Arbeitsmarkt und soziale Sicherung' (PASS). Datenreport Welle 9. FDZ-Datenreport, 07/2016, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.</li> <li>▪ Bethmann, Arne; Fuchs, Benjamin; Wurdack, Anja (2013): User Guide "Panel Study Labour Market and Social Security" (PASS). Wave 6. FDZ-Datenreport, 07/2013 (en), Nürnberg.</li> <li>▪ Trappmann, Mark; Beste, Jonas; Bethmann, Arne; Müller, Gerrit (2013): The PASS panel survey after six waves. In: Journal for Labour Market Research, Bd. 46, Nr. 4, S. 275-281.</li> </ul>
--	---

## 1.5 Merkmalsliste

Die Merkmalsübersicht in Tabelle 4 listet die Variablennamen und die längeren Merkmalsbezeichnungen der Prozessdaten auf. Weiterhin bietet sie eine Übersicht darüber, welche Merkmale für welche Datenquellen gefüllt sind. Bedingt durch die hohe Anzahl an Merkmalen in den Befragungsdaten werden die PASS-Merkmale nur in der zugehörigen Originaldokumentation aufgeführt. Eine Erläuterung der Füllgrade findet sich in Tabelle 3.

**Tabelle 3: Füllgrad der Merkmale**

	Merkmal ist für die Quelle gefüllt. Füllgrad immer > 0,85
	Merkmal ist für die Quelle gefüllt. Niedriger bzw. wechselnder Füllgrad, siehe Merkmalsbeschreibung und Häufigkeitsauszählung
	Merkmal ist für diese Quelle nicht gefüllt. Füllgrad immer < 0,05

Beispiel: Das Merkmal Tagesentgelt bzw. täglicher Leistungssatz ist nur für BeH- und LeH-Sätze gefüllt. Die Sätze der anderen Datenquellen enthalten in dieser Variablen den Missing-Wert „.n“. Eine weitere Besonderheit liegt darin, dass einige Merkmale je nach Datenquelle

unterschiedliche Inhalte haben. So enthält der Erwerbsstatus für BeH-Sätze die Personengruppe des Meldeverfahrens, für LeH-Sätze die Leistungsart, für LHG-Sätze den Status SGB II, für ASU- bzw. XASU-Sätze den Arbeitsuche-Status und für MTH die Maßnahmekategorie. Nicht bei jedem Merkmal sind diese Unterschiede direkt aus dem Variablennamen ersichtlich.

**Tabelle 4: Merkmalsliste mit Füllgrad**

Merkmalsliste	BHP Basis	Seite	BeH	LeH	LHG	ASU	XASU	MTH
<b>Identifikatoren</b>		<b>50</b>						
Unveränderliche Personennummer (pnr)		50						
Systemfreie Betriebsnummer (betnr)	*	51						
<b>Generierte technische Merkmale</b>		<b>52</b>						
Quelle des Satzes (quelle)								
Satzzähler pro Konto (spell)		52						
Jahr (jahr)	*	52						
Vorliegen administrativer Personendaten (match_admin)		53	(nur in Verknüpfungsdatei)					
Typ des Matches (match_type)		53	(nur in Verknüpfungsdatei)					
Wellenindikator erfolgreicher Match (match_wave)		53	(nur in Verknüpfungsdatei)					
<b>Gültigkeitszeitraum</b>								
Beginndatum Originalbeobachtung (begorig)		54						
Enddatum Originalbeobachtung (endorig)		54						
Beginndatum der Episode (begepi)		55						
Enddatum der Episode (endepe)		55						
<b>Informationen zur Person</b>		<b>56</b>						
Geschlecht (frau)		56						
Geburtsjahr (gebjahr)		56						
Staatsangehörigkeit (nation)		56						
Staatsangehörigkeit vergrößert (nation_gr)		56						

<b>Merkmalsliste</b>	<b>BHP Basis</b>	<b>Seite</b>	<b>BeH</b>	<b>LeH</b>	<b>LHG</b>	<b>ASU</b>	<b>XASU</b>	<b>MTH</b>
Familienstand (famst)		57						
Kinderzahl (kind)		57						
Ausbildung (ausbildung)		58						
Schulbildung (schule)		59						
<b>Informationen</b> zu Beschäftigung, Leistungsbezug und Arbeitsuche		<b>60</b>						
Tagesentgelt/ täglicher Leistungssatz (tentgelt)		60						
Beruf – ausgeübte/letzte Tätigkeit (KldB 1988) (beruf)		62						
Berufsgruppe – ausgeübte/letzte Tätigkeit (KldB 2010) (beruf2010_3)		62						
Berufsuntergruppe – ausgeübte/letzte Tätigkeit (KldB 2010) (beruf2010_4)		63						
Anforderungsniveau – ausgeübte/letzte Tätigkeit (KldB 2010) (niveau)		64						
Teilzeit (teilzeit)		65						
Erwerbsstatus (erwstat)		65						
Gleitzone (gleitz)		67						
Leiharbeit (leih)		68						
Befristung (befrist)		68						
Abmeldegrund / Abgabegrund / Beendigungsgrund (grund)		68						
Erwerbsstatus vor Arbeitsuche (estatvor)		70						
Status nach Arbeitsuche (estatach)		70						
Profillage (profil)		71						
Art der Kündigung der letzten Tätigkeit (art_kuend)		71						
Arbeitszeit des Stellengesuchs (arbeitszeit)		72						
Restanspruch/geplante Dauer (restanspruch)		72						
Trägerart (traeger)		72						
Beginndatum der Arbeitslosigkeit (alo_beg)		73						
Dauer der Arbeitslosigkeit (alo_dau)		73						
<b>Ortsangaben</b>		<b>74</b>						
Wohnort - Kreis (wo_kreis)		74						

<b>Merkmalsliste</b>	<b>BHP Basis</b>	<b>Seite</b>	<b>BeH</b>	<b>LeH</b>	<b>LHG</b>	<b>ASU</b>	<b>XASU</b>	<b>MTH</b>
Wohnort – Bundesland (wo_bula)		75						
Wohnort - Arbeitsagentur (wo_aa)		75						
Wohnort - Regionaldirektion (wo_rd)		76						
<b>Betriebsmerkmale</b>		<b>76</b>						
WZ73 3-Steller (w73_3)	*	76						
WZ93 5-Steller (w93_5)	*	77						
WZ93 3-Steller (w93_3)	*	77						
WZ03 5-Steller (w03_5)	*	78						
WZ03 3-Steller (w03_3)	*	78						
WZ08 5-Steller (w08_5)	*	79						
WZ08 3-Steller (w08_3)	*	79						
W73_3 vervollständigt durch Extrapolation/Imputation (w73_3_gen)	*	80						
Art der Vervollständigung w73_3 (group_w73_3)	*	80						
W93_3 vervollständigt durch Extrapolation/Imputation (w93_3_gen)	*	80						
Art der Vervollständigung w93_3 (group_w93_3)	*	81						
Jahr des 1. Auftretens der Betriebsnummer (grd_jahr)	*	81						
Erstes Auftreten Betriebsnummer (grd_dat)	*	82						
Jahr des letzten Auftretens der Betriebsnummer (lzt_jahr)	*	82						
Letztes Auftreten Betriebsnummer (lzt_dat)	*	83						
Anzahl Beschäftigte gesamt (az_ges)	*	83						
Anzahl in Vollzeit (Normalbeschäftigte + sonstige) (az_vz)	*	83						
Anzahl geringfügig Beschäftigte (az_gf)	*	84						
Mittelwert imputiertes Bruttotagesentgelt Vollzeitbeschäftigte (te_imp_mw)	*	84						
Arbeitsort Kreis (ao_kreis)	*	85						
Arbeitsort Bundesland (ao_bula)	*	85						

## 1.6 Mengengerüst und Zeiträume

Die Zahl der Datenzeilen in den administrativen Personendaten ist höher als die hier ausgewiesene Zahl der Personen, da es parallele Beschäftigung, Leistungsbezugszeiten, Arbeitsuche oder Maßnahmeteilnahmen gibt, die jeweils eine eigene Datenzeile erhalten. Dies gilt auch für die in den Befragungsdaten erfassten Spelldaten. Für einen Teil der PASS-Befragten konnten keine administrativen Daten in den IEB gefunden werden, obwohl sie erfolgreich verknüpft wurden. Tabelle 5 zeigt daher an, für wie viele PASS-Teilnehmer IEB-Daten vorliegen, und wie viele Zeilen diese in den administrativen Personendaten aufweisen.



**Tabelle 5: Mengengerüst**

PASS-Befragte	37.993
Erfolgreich verknüpfte Befragte	25.048
Zugehörige Zeilen in administrativen Personendaten	1.980.360

Quelle: PASS-ADIAB, eigene Berechnungen.

Da die einzelnen Datensätze aus unterschiedlichen Quellen stammen, ist es nicht möglich, dass alle Teildatensätze den gleichen Zeitraum umfassen. Die PASS-Befragungsdaten decken die Beobachtungszeiträume der Wellen 1 bis 9 ab, die in den Jahren 2006 bis 2015 erhoben wurden. Analog zu den Datenrändern der übrigen administrativen Datenprodukte des FDZ stehen die administrativen Personen- und Betriebsdaten in PASS-ADIAB für den Zeitraum 1975 - 2014 zur Verfügung. Bei den Betriebsdaten beziehen sich die Angaben auf den 30.6. des jeweiligen Jahres. Für einzelne Personen können die Zeiträume entsprechend der Teilnahme an Befragungen bzw. ihrer für die administrativen Daten relevanten Zustände deutlich kürzer sein.

## 2 Datenquellen

### 2.1 PASS-Befragungsdaten

Das PASS stellt eine Datengrundlage bereit, mit deren Hilfe die Dynamik des Grundsicherungsbezugs und die soziale Lage von Haushalten im Grundsicherungsbezug untersucht werden können. Die Notwendigkeit einer solchen Datengrundlage ergibt sich insbesondere aus zwei Feststellungen:

1. Da es sich bei der der Grundsicherung für Arbeitsuchende um eine haushalts- (beziehungsweise bedarfsgemeinschafts-) bezogene Sozialleistung handelt, müssen die Haushaltskontexte von Leistungsempfängern berücksichtigt werden, um die individuellen und sozialen Folgen der Reform adäquat untersuchen zu können. Gleiches gilt für eine adäquate Evaluation arbeitsmarktpolitischer Maßnahmen, da Strategien einer aktivierenden Arbeitsmarktpolitik immer auf haushaltstypische Lebensumstände treffen und abhängig von diesen kontextabhängige Wirkungen entfalten. Eine weitere Frage, die sich ohne Rückgriff auf die sozialen und ökonomischen Binnenstrukturen von Haushalten kaum hinreichend beantworten lässt, ist die nach den geschlechtsspezifischen Wirkungen einzelner Maßnahmen und der Bedeutung von Berufs- und Erwerbsorientierungen von Frauen und Männern für die Überwindung des Leistungsbezugs. Daher erschien es sinnvoll, das PASS als Haushaltsbefragung zu konzipieren.

2. Die Fokussierung der Untersuchungsperspektive auf Kontexte und Dynamiken von Armutshaushalten erfordert ein als Längsschnitt angelegtes Untersuchungsdesign, ein Sachverhalt, dem die Konzeption des PASS als Panelstudie Rechnung trägt. Ein solches Design gestattet nicht nur eine ereignisbezogene Untersuchung der Ein- und Austritte in den Leistungsbezug und deren Relation zu sozial- und arbeitsmarktpolitischen Maßnahmen des SGB II, sondern es ermöglicht darüber hinaus auch eine Ausweitung der Perspektive auf individuelle wie auch haushaltstypische Verfestigungen von Bedürftigkeit. Darüber hinaus können auch die in der Intention des SGB II liegenden Pfade aus der Hilfebedürftigkeit und in eine (Re-)Integration ins Erwerbsleben identifiziert werden. Dabei lässt sich auch untersuchen, inwieweit die im SGB II vorgesehenen Maßnahmepakete eine solche Integration im Kontext unterschiedlicher Haushaltskonstellationen erleichtern oder nicht.

Daneben sind noch drei weitere methodische Besonderheiten des PASS zu nennen: (a) die Kombination aus zwei ursprünglich in etwa gleich großen Teilstichproben zu je circa 6.000 Haushalten, (b) die Durchführung der Erhebung in einem Mixed Mode aus telefonischem (CATI) und persönlichem (CAPI) Interview, sowie (c) die mehrsprachige Durchführung der Erhebung.

Der erste Punkt (a) ergibt sich daraus, dass die forschungsleitenden Fragen des PASS zwar stark mit unterschiedlichen Aspekten des SGB-II-Leistungsbezugs verknüpft sind, sich in ihrer Gesamtheit aber nicht allein mit einer Stichprobe von SGB-II-Leistungsempfängern beantworten lassen. Beispiele für wichtige, mithilfe des PASS zu beantwortenden Fragen sind die nach

- den Wegen in die sowie aus der Abhängigkeit von staatlichen Transferleistungen,
- der Veränderung der sozialen Lage der vom Leistungsbezug betroffenen Haushalte und Personen,
- der subjektiven Bewältigung von Langzeitarbeitslosigkeit und Abhängigkeit von staatlichen Transferleistungen und der damit möglicherweise verknüpften Veränderung handlungsrelevanter Orientierungen der Befragten im Zeitverlauf,
- dem Kontakt zu den Trägern der Grundsicherung und der institutionellen Handlungspraxis zur Eingliederung in Arbeit.

Insbesondere bei der Analyse von Zugangsprozessen, für die Konstruktion von Kontrollgruppen und für die Einschätzung relativer Lebenslagen werden zur Beantwortung der genannten Fragestellungen auch Informationen über Bevölkerungsgruppen außerhalb des Leistungsbezugs benötigt. Aus diesem Grund umfasst das PASS neben der ersten Teilstichprobe, für die Bedarfsgemeinschaften aus

Prozessdaten der BA gezogen wurden, eine zweite Teilstichprobe, bei der es sich um eine Bevölkerungsstichprobe mit überproportionaler Berücksichtigung einkommensschwächerer Haushalte handelt.

Die zweite Besonderheit der Studie (b), die Verwendung sowohl telefonischer als auch persönlicher Interviews liegt in der teilweise geringen Erreichbarkeit einzelner Bevölkerungsgruppen über Festnetzanschlüsse begründet. Um die hieraus möglicherweise resultierenden systematischen Ausfälle zu vermeiden, werden telefonisch nicht erreichbare Personen über CAPI-Interviewer kontaktiert. Ab der vierten Befragungswelle wurde das CAPI-Feld ausgeweitet. Seitdem werden bei allen erstmalig befragten Haushalten persönlich-mündliche Interviews durchgeführt, unabhängig von ihrer telefonischen Erreichbarkeit.

Der dritte Punkt (c), die mehrsprachige Durchführung der Studie, ergibt sich vor allem daraus, dass sich Migranten überproportional häufig in prekären Lebenslagen befinden, wobei gerade auch eine schlechte Beherrschung der deutschen Sprache eine zusätzliche Hürde für eine Platzierung am Arbeitsmarkt darstellen dürfte. Damit auch bei dieser wichtigen Zielgruppe keine systematischen Ausfälle entstehen, ist eine Durchführung der Befragungen nicht nur auf Deutsch möglich, sondern zusätzlich auch in den Muttersprachen zweier wichtiger Migrantengruppen, in Türkisch und Russisch.

## **2.2 Administrative Daten aus den IEB**

Die administrativen Personendaten wurden aus den Integrierten Erwerbsbiografien (IEB) des IAB gezogen. Diese vereinen Daten aus fünf verschiedenen Datenquellen, die jeweils Informationen aus unterschiedlichen EDV-Verfahren enthalten können. Es werden Datensätze aus allen in Abbildung 1 dargestellten Quellen aus der IEB übernommen.

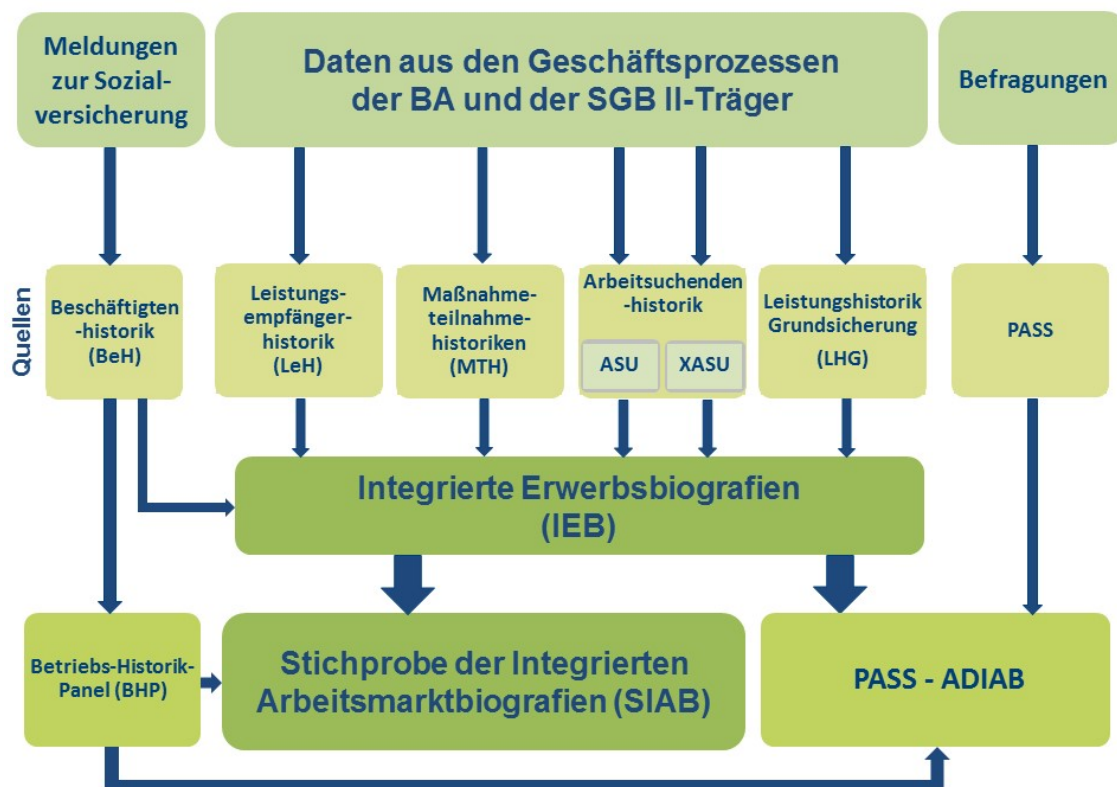


Abbildung 1: Quellen von PASS-ADIAB

### 2.2.1 Beschäftigtenhistorik (BeH)

Die Quelle für Informationen über Beschäftigung ist die Beschäftigtenhistorik (BeH) des IAB. Die Grundlage der Daten ist das mit Wirkung vom 1. Januar 1973 eingeführte (und seit dem 1. Januar 1991 auf Ostdeutschland ausgedehnte) integrierte Meldeverfahren zur Kranken-, Renten- und Arbeitslosenversicherung, das durch das Kürzel DEÜV (früher DEVO/DÜVO) bezeichnet ist (siehe für weitere Details: Bender et al., 1996, S. 4ff.; Wermter/Cramer, 1988). Es verpflichtet Arbeitgeber, mindestens einmal jährlich Meldungen über alle sozialversicherungspflichtig beschäftigten Arbeitnehmer an die zuständigen Sozialversicherungsträger abzugeben. Die BeH erfasst alle Arbeiter und Angestellten sowie alle Auszubildenden, soweit sie nicht von der Sozialversicherungspflicht befreit sind. Dies bedeutet, dass verbeamtete und selbstständige Tätigkeiten sowie Studierende<sup>3</sup> (siehe Cramer, 1985) grundsätzlich nicht in der BeH erfasst werden. Seit der Änderung des Meldeverfahrens zum 1. Januar 1999 werden auch geringfügig Beschäftigte und mithelfende Familienangehörige erfasst (enthalten erst ab 1. April 1999). Die Daten werden von den Krankenkassen aufgenommen, von

<sup>3</sup> Studierende können dennoch in der BeH erscheinen, wenn sie parallel zum Studium z. B. geringfügig beschäftigt waren.

der Bundesagentur für Arbeit gesammelt und aufbereitet und anschließend vom IAB in dessen Historikdatei integriert.

Die administrativen Personendaten werden durch Betriebsdaten (Basisbetriebsdaten und BHP-Erweiterungsmodule) ergänzt. Diese werden dem Betriebs-Historik-Panel (BHP) entnommen, das ebenso auf der BeH basiert.

Bei der Verknüpfung von Personen- mit Betriebsdaten muss berücksichtigt werden, dass die Merkmale in der Basisbetriebsdatei sowie den BHP-Erweiterungsmodulen zum 30.6. eines Jahres aggregiert werden. Die Betriebsmerkmale sind deshalb, im Unterschied zu den Personenangaben, keine zeitraumbezogenen Daten, sondern nur am 30.6. jedes Jahres exakt gültig.<sup>4</sup>

```
use PASS-ADIAB_7515_v1.dta
gen jahr = year(begepi)
sort betnr jahr
merge m:1 betnr jahr using PASS-ADIAB_7515_v1_bhp_basis_v1.dta
```

#### **Kasten 1: Beispielcode für Stata 14**

Die Verknüpfung der Betriebs- und Personendatei erfolgt über die programmspezifischen Kommandos der zur Datenaufbereitung und -analyse genutzten Softwarepakete. In Stata können die Daten beispielsweise unter Verwendung des „merge“ - Befehls und der Jahresangabe sowie der Betriebsnummer in Verbindung mit der entsprechenden Pfadangabe verknüpft werden (siehe Kasten 1).

#### **2.2.2 Leistungsempfängerhistorik (LeH)**

Die Leistungsempfängerhistorik (LeH) erfasst Zeiträume, in denen Personen Lohnersatzleistungen der Bundesagentur für Arbeit beziehen (Rechtskreis SGB III). Die Leistungen beinhalten Arbeitslosengeld, Arbeitslosenhilfe und Unterhaltsgeld, also keine Leistungen nach dem Rechtskreis SGB II (z.B. Arbeitslosengeld II). Da der Leistungsanspruch von der Erfüllung gesetzlicher Voraussetzungen abhängt, werden Arbeitslosigkeitsperioden, in denen die Voraussetzungen nicht vorliegen (z.B. fehlende Bedürftigkeit im Falle der Arbeitslosenhilfe oder Nichterfüllung der Anwartschaftszeit beim Arbeitslosengeld) nicht in der Leistungsempfängerhistorik berichtet. Der linke Datenrand der LeH ist der 1. Januar 1975.

#### **2.2.3 Leistungshistorik Grundsicherung (LHG)**

Die Leistungshistorik Grundsicherung (LHG) enthält Informationen zu erwerbsfähigen Leistungsberechtigten, ihren Bedarfsgemeinschaftsmitgliedern nach § 7 SGB II und besonderen

---

<sup>4</sup> Extrembeispiel: Es liegen Beschäftigungsmeldungen vom 1.1. bis zum 30.5.2006 vor, der Betrieb meldet jedoch Konkurs an, weshalb alle Beschäftigten entlassen werden. Dann gibt es im BHP zu diesem Betrieb für 2006 keine Informationen.

Personen im Umfeld einer Bedarfsgemeinschaft. Leistungsbezug nach SGB II umfasst sowohl Leistungen zur Sicherung des Lebensunterhalts (z.B. ALG II), wie auch Zuschläge zum ALG II oder auch zusätzliche Leistungen. Die LHG enthält jedoch keine Angaben zu den jeweiligen Leistungssätzen. Da die Ermittlung der Höhe des Leistungsbezugs in der Grundsicherung nicht auf individueller Ebene erfolgt, sondern auf der Ebene von Bedarfsgemeinschaften, ist die Zuweisung eines individuellen Leistungssatzes problematisch und erfolgt in der SIAB nicht. Im Gegensatz zu den Leistungen aus dem Rechtskreis des SGB III ist die Bundesagentur für Arbeit nicht alleiniger Träger der Leistungen. Es wird demnach unterschieden nach den drei möglichen Trägerarten:

- Arbeitsgemeinschaften (ARGEEn, bis 2010)/gemeinsame Einrichtungen (ab 2011) aus gemeinsamer Aufgabenwahrnehmung von BA und Kommune,
- Getrennte Trägerschaft/getrennte Aufgabenwahrnehmung (bis 2011) – hier besteht eine Aufgabenteilung zwischen BA und Kommune<sup>5</sup>,
- Zugelassene kommunale Träger, die nach der anfänglichen Experimentierklausel des § 6a auch als optierende Kommunen oder Optionskommunen bezeichnet werden – hier übernimmt die Kommune alle Aufgaben aus dem Bereich des SGB II.

Die Daten der LHG stammen aus verschiedenen Meldeverfahren. In der Regel findet das IT-Verfahren A2LL bis 2010 bei ARGEEn und ab 2011 bei gemeinsamen Einrichtungen Verwendung<sup>6</sup>. Zugelassene kommunale Träger verwenden verschiedene eigene IT-Verfahren und übermitteln ihre Daten mittels des Standards XSozial-BA-SGB II an die Bundesagentur für Arbeit. Bei Kommunen mit getrennter Trägerschaft kommen beide Verfahren zum Einsatz. Die unterschiedlichen Datenstandards haben Auswirkungen auf Umfang und Qualität der gelieferten Daten.

Der linke Datenrand der LHG ist der 1. Januar 2005. Die Datenquelle ist jedoch bis Anfang 2007 lückenhaft (siehe Kapitel 4).

#### **2.2.4 Arbeitssuchendenhistorik (ASU/XASU)**

Daten über Arbeitssuchende werden in der Arbeitssuchendenhistorik (ASU und XASU) gespeichert. Die Quelle ASU enthält Informationen zu Arbeitssuchenden, welche bei Arbeitsagenturen verzeichnet sind, ab 2005 auch solche, die bei Arbeitsgemeinschaften und getrennten Trägerschaften gemeldet sind. In der Quelle XASU sind ab 2005 die Daten von

---

<sup>5</sup> Die Kommune übernimmt die Kosten der Unterkunft und Heizung (§ 22 SGB II) sowie die abweichend zu erbringenden Leistungen (§ 23 (3) SGB II) und die flankierenden Leistungen nach §16 (2) Satz 2 Nr. 1 - 4 SGB II. Von der BA werden dagegen Regelleistungen, Sozialversicherungsbeiträge und Eingliederungsleistungen (SGB III und SGB II) sowie spezifische Leistungen ohne die oben genannten flankierenden Leistungen erbracht.

<sup>6</sup> 2014 wurde A2LL sukzessive durch ALLEGRO als neues IT-Verfahren für die Grundsicherung im SGB II in gemeinsamen Einrichtungen ersetzt.

arbeitsuchenden Personen im ALG-II-Bezug von zugelassenen kommunalen Trägern enthalten, welche nach dem XSozial-Standard konsolidiert gemeldet werden.

### **2.2.5 Maßnahmeteilnahmehistoriken (MTH)**

Die Maßnahmeteilnahmehistoriken (MTH) enthalten Informationen, die verschiedenen Rechtskreisen zugeordnet werden können. Zum einen sind Maßnahmen und Maßnahmeteilnahmen der aktiven Arbeitsmarktförderung nach dem SGB III enthalten. Zum anderen beinhalten die MTH Maßnahmen aus dem Rechtskreis SGB II, wenn diese in Verfahren der BA erfasst werden. Dies bedeutet insbesondere, dass in den MTH keine Maßnahmen der zugelassenen kommunalen Träger geführt werden, da diese über den abweichenden Standard XSozial gemeldet werden. Informationen von diesen Trägern werden wegen einer Reihe datentechnischer Probleme nicht in die IEB aufgenommen. Der linke Datenrand der MTH ist der 1. Januar 2000.

## **3 Datengenese und -verknüpfung**

### **3.1 PASS-Befragungsdaten**

Das Stichprobenkonzept der PASS-Studie wird bei Trappmann et al. (2009) vorgestellt. Im FDZ-Methodenreport 04/2016 (Jesske et al., 2016) finden sich detaillierte Informationen zur Datenerhebung. Datenstruktur und -aufbereitung werden in den FDZ-Datenreporten 07/2016 (Berg et al., 2016) und 04/2011 (Bethmann/Gebhardt, 2011) ausführlich beschrieben. Die verwendeten Haushalts- und Personenfragebögen, sowie die Dokumentation der für den verknüpften Datensatz verwendeten Befragungsdaten im Format des ebenfalls über das FDZ angebotenen Scientific Use File, stehen unter [http://fdz.iab.de/de/FDZ\\_Individual\\_Data/PASS.aspx](http://fdz.iab.de/de/FDZ_Individual_Data/PASS.aspx) zur Verfügung. Dort finden sich auch weitere Publikationen zu methodischen Aspekten der PASS-Studie.

### **3.2 Administrative Daten aus den IEB**

#### **3.2.1 Bereinigungen und Aufbereitungen**

Bevor die Daten aus den in Kapitel 2 genannten Quellen zur IEB zusammengefügt werden, durchlaufen sie quellenspezifische Bereinigungsverfahren (siehe Folgeabschnitte). Die Gesamt-IEB erfahren die folgenden Bereinigungen:

- Sätze, bei denen das Alter an ihrem Ende kleiner 13 oder zu Beginn größer 75 ist, werden gelöscht.
- Datensätze, bei denen das Enddatum kleiner als das Beginndatum ist, werden gelöscht.

- Widersprüchliche Angaben zu Geschlecht oder Geburtsdatum innerhalb eines Kontos werden bereinigt.
- Datensätze ohne Angabe von Geburtsdatum und Geschlecht nach Bereinigung werden gelöscht.

### **3.2.1.1 Beschäftigtenhistorik (BeH)**

- Um einen im Zeitablauf möglichst konstanten Personenkreis zu erfassen, werden einige Personengruppen, deren Daten nicht für den gesamten Beobachtungszeitraum zur Verfügung stehen, ausgeschlossen. Ab dem Meldejahr 2011 stammen die Daten der BeH aus neu konzipierten Quelldaten. Dadurch wurde eine Reihe von Personengruppen reaktiviert bzw. neu eingeführt, da sie von der BA-Statistik als sozialversicherungspflichtig eingestuft werden. Damit sind fortan die Personengruppen 101 - 107, 111 - 114, 118, 119, 120, 140, 141, 142, 143, 149, 201 und 203 - 205 enthalten sowie die beiden Gruppen 109 und 209, die geringfügig Beschäftigte kennzeichnen. Nicht enthalten sind unter anderem kurzfristig Beschäftigte, d. h. Personengruppen 110, 202 und 210.
- Die Personengruppen 123, 124 und 127 wurden neu eingeführt.
- Aus datenschutzrechtlichen Gründen werden die Personengruppen 107, 111, 113, 114, 127 und 204 zur Personengruppe "Sonstige Erwerbstätige" (599) zusammengefasst.
- Ab dem Meldejahr 2012 wurden Auszubildende mit den Personengruppen 121 und 122 neu aufgenommen.
- Sätze mit einem Entgelt in Höhe von Null bzw. mit fehlendem Entgelt sowie der Ausprägung 101 beim Merkmal Erwerbsstatus und dem Wert 50 beim Abgabegrund (Jahresmeldung) werden nicht in die IEB übernommen.
- Geschlecht und Geburtsdatum werden aus dem Data Warehouse (DWH) der Statistik der BA übernommen. Dort liegen diese Angaben quellenübergreifend harmonisiert vor.
- Der Gebietsstand für Arbeitsort und Wohnort entspricht dem 31.12.2014.

### **3.2.1.2 Leistungsempfängerhistorik (LeH)**

- Sätze ohne gültiges Beginndatum wurden ausgeschlossen.
- Sätze mit einem Enddatum, das vor dem Beginndatum liegt, wurden ausgeschlossen.
- Liegt das Enddatum bei Bezug von Arbeitslosenhilfe einen Tag vor dem Beginndatum, so wurde das Enddatum um ein Jahr hochgesetzt und der Spell nicht gelöscht.
- Zwischen 2004 und 2006 wurde das zugrunde liegende Meldeverfahren umgestellt. Auftretende Überschneidungen zwischen dem Alt- und Neuverfahren wurden bereinigt.
- Sätze mit leerem oder ungültigem Enddatum wurden ausgeschlossen, da hier nicht davon ausgegangen werden kann, dass überhaupt eine Zahlung geleistet worden ist.
- Es wird eine Gebietsstandbereinigung analog zur BeH durchgeführt.



### **3.2.1.3 Leistungshistorik Grundsicherung (LHG)**

- Sätze ohne BA-Kundennummer werden gelöscht.
- Konten ohne gültiges Geburtsdatum werden gelöscht.
- Stornierte Datensätze werden nicht genutzt.
- Es sind nur Sätze erwerbsfähiger Personen und von Personen über 65 Jahren enthalten.
- Es sind jeweils überschneidungsfreie Zeiträume der Leistungsberechtigung einer Person in einer bestimmten Bedarfsgemeinschaft abgebildet. Aus folgenden administrativen Gründen beginnt ein neuer Satz:
  - an gesetzlich festgelegten, für strukturelle Änderungen der Bedarfsgemeinschaft relevanten Geburtstagen (14, 15, 18 und 25) sowie dem individuell gültigen Renteneintrittsalter der Mitglieder einer Bedarfsgemeinschaft (siehe Abschnitt 3.2.1.6),
  - bei Veränderungen in der Zusammensetzung einer Bedarfsgemeinschaft (z.B. durch Zu-/Abgänge),
  - bei Merkmalsänderungen des Bedarfsgemeinschafts-Kunden sowie
  - bei Beginn und Ende eines Sanktionssachverhalts für Sätze ab dem 1. April 2006. Allerdings ist zu beachten, dass anhand der Daten weder die Dauer, die Art noch der Zeitpunkt der Verhängung bzw. des Beginns der Sanktion festgestellt werden kann. Grund hierfür ist das Fehlen eines entsprechenden Merkmals oder einer Ausprägung, welches den Beginn, die Art oder die Dauer der Sanktion anzeigt.
- Aus obigem Grund sind alle personenbezogenen Merkmale, die für die Quelle LHG gefüllt sind, für den gesamten Zeitraum des Satzes gültig.
- Doppelmeldungen aufgrund der Gebietsreformen 2009/2011 und Trägerneuorganisation 2012 wurden soweit möglich bereinigt.
- Es wird eine Gebietsstandbereinigung analog zur BeH durchgeführt.

### **3.2.1.4 Arbeitsuchendehistorik ASU und XASU**

- Sätze, deren Enddatum vor dem Beginndatum liegt, wurden nicht in die ASU übernommen.
- Es erfolgt keine Konsolidierung der ASU-Sätze einer einzelnen Person. Es kann daher vorkommen, dass Überschneidungen zwischen ASU-Sätzen existieren.
- Die personenbezogenen Merkmale, die für die Quellen (X)ASU gefüllt sind, beziehen sich immer auf den Beginn des Satzes.
- Ein neuer ASU-Satz wird angelegt, sobald ein Statuswechsel (z.B. von arbeitsuchend zu arbeitslos) erfolgt. Dies gilt auch, wenn sich die Trägerart (Agentur für Arbeit, ARGE, gemeinsame Einrichtung, zugelassener kommunaler Träger, getrennte Aufgabenwahrnehmung) ändert. In der Datenbasis der ASU gibt es nur die Unterscheidung zwischen Sätzen mit Status "arbeitslos", "arbeitsuchend" sowie seit 2006 „ratsuchend“ und

„ohne Status“. In den IEB wird jedoch ein zusätzlicher Status „krank“ unterschieden. Der Erwerbsstatus "krank" wird IEB-Sätzen zugeordnet, wenn in der ASU-Datenbasis

- ein Vorgängersatz mit Status „arbeitslos“ existiert, lückenlos anschließt und den Abgangsgrund "krank" hat und
  - ein Nachfolgesatz mit Status „arbeitslos“ existiert und ebenfalls lückenlos anschließt
  - und der Satz selbst nicht den Status "arbeitslos" sondern den Status "nicht arbeitslos arbeitsuchend" hat.
- Im Gegensatz zur Quelle ASU werden in der XASU nur die Zustände
- „nicht arbeitslos, aber arbeitsuchend“ und
  - „arbeitsuchend und gleichzeitig arbeitslos“
- unterschieden.
- Krankheitsepisoden werden in der XASU im Gegensatz zur ASU bei der Bildung des Erwerbsstatus nicht berücksichtigt, da in den XASU-Daten Informationen zu Krankheiten nicht erschlossen sind. Bei der Berechnung der Arbeitslosigkeitsdauer unter Einschluss von XASU-Sätzen können daher krankheitsbedingte Lücken nicht identifiziert werden.
- In der XASU werden überschneidungsfreie Zeiträume bezogen auf eine einzelne Person abgebildet. Bei Änderungen folgender Merkmale wird für die XASU jeweils ein neuer Datensatz gebildet:
- Wechsel des Arbeitsuche-Status
  - Wechsel der Verfügbarkeit
  - Wechsel des SGB-II-Trägers
  - Wechsel des Wohnortes
- Eine Gebietsstandbereinigung wird analog zur BeH durchgeführt.

### **3.2.1.5 Maßnahmeteilnahmehistoriken (MTH)**

- Sätze mit einem Enddatum, das vor dem Beginndatum liegt, wurden ausgeschlossen.
- Sätze, die mehr als ein Jahr nach Maßnahmeende erstellt wurden, werden gelöscht, falls ein anderer Satz existiert, innerhalb eines Jahres nach Abschluss der Maßnahme erstellt wurde.
- Es wird lediglich der aktuellste Datensatz zu einer einzelnen Maßnahmeteilnahme verwendet.
- In die IEB werden nur Maßnahmeteilnahmen aufgenommen, die als „tatsächlich stattgefunden“ klassifiziert sind. Teilnahmen, die nicht oder noch nicht stattgefunden haben, werden gelöscht. Teilnahmen werden auch dann als nicht stattgefunden gewertet, wenn während der Teilnahme an einer Maßnahme ein Löschdatum gesetzt wird.
- Bestimmte Maßnahmearten werden nicht aufgenommen. Hierzu zählen Unterstützung der Beratung und Vermittlung, Mobilitätshilfen und reine Reha-Maßnahmen.

### **3.2.1.6 SGB-II-Anonymisierung**

Um das Deanonymisierungsrisiko zu verringern, wird in den administrativen Daten grundsätzlich nur das Geburtsjahr ausgewiesen. In den Quellen LHG und (X)ASU besteht allerdings die Gefahr, dass durch die zeitliche Struktur der Beobachtungen dennoch das exakte Geburtsdatum sichtbar werden könnte. Sätze könnten systematisch an bestimmten Geburtstagen oder dem Renteneintritt bzw. am Tag davor enden oder aber an diesem Tag neu beginnen.

Um eine indirekte Bestimmung des exakten Geburtstags zu verhindern, wird daher das folgende Verfahren angewendet. Sätze, die am 18., 25. oder 65. Geburtstag geteilt wurden, werden zu einem einzelnen Satz zusammengefügt. Hierdurch geht außer dem exakten Geburtstag keine weitere Information verloren. Bei Sätzen, die am 15. Geburtstag oder dem Renteneintritt beginnen bzw. am Renteneintrittstag enden wird das Beginndatum (Enddatum) des Satzes auf den Beginn (das Ende) des jeweiligen Quartals gesetzt. Bei dieser Korrektur kommt es zu einer Verzerrung der Satzdauer. Korrigierte Sätze sind in der Variable "Erwerbsstatus" gekennzeichnet.

### **3.2.2 Episodensplitting**

Die administrativen Personendaten liegen „gesplittet“ vor. Falls sich innerhalb eines Kontos Sätze überlappen, werden diese durch künstliche Sätze mit neuen Datumsangaben derart ersetzt, dass vollständig parallele oder überschneidungsfreie Zeiträume entstehen (Episodensplitting). Die Anzahl der Sätze erhöht sich dadurch (siehe Abbildung 2).

Die ursprünglichen Datumsvariablen Beginn und Ende des Originalsatzes (`begorig` und `endorig`) bleiben erhalten, die Variablen Beginn und Ende der Episode (`begepi` und `endepi`) kennzeichnen Anfang und Ende der gesplitteten Episoden. Ob Sätze gesplittet wurden, kann an einem Vergleich des ursprünglichen Zeitraums (`begorig` und `endorig`) mit dem Episodenzeitraum (`begepi` und `endepi`) festgestellt werden. Um die Originaldaten ohne Episodensplitting wieder herzustellen bzw. die durch das Episodensplitting künstlich erzeugten Datensätze zu löschen, selektiert man alle Sätze, bei denen der Beginn des Originalsatzes gleich dem Beginn der gesplitteten Episode ist (`begepi == begorig`).

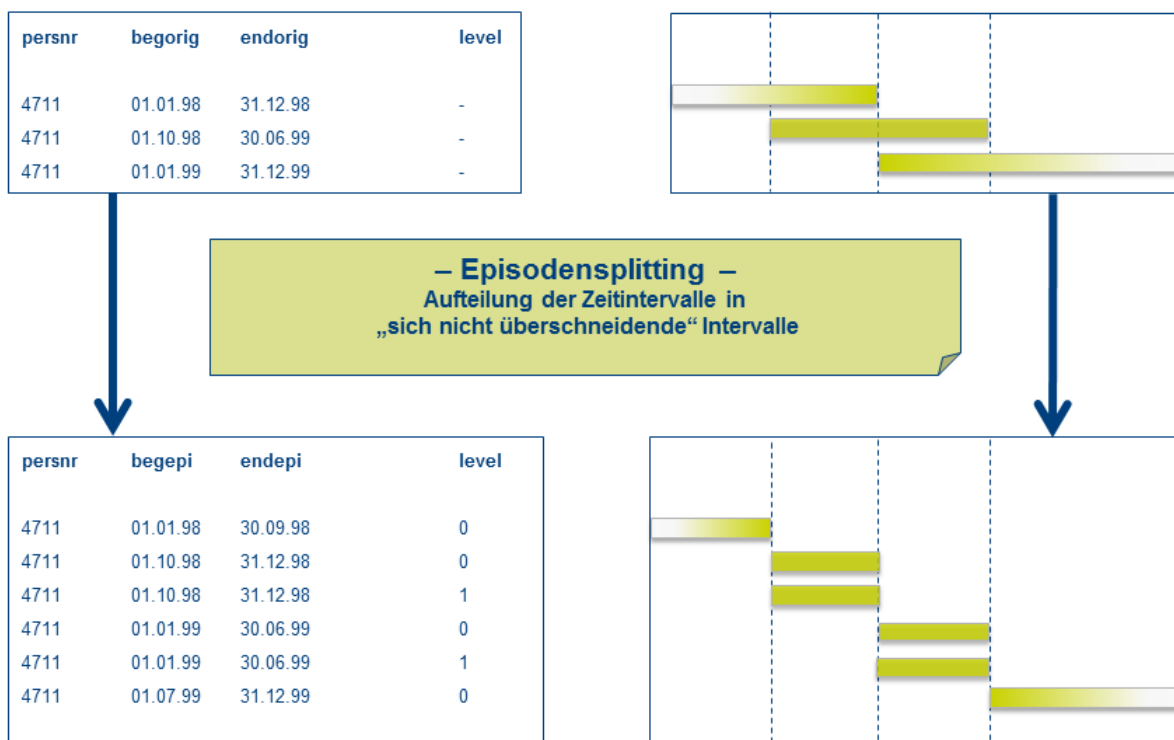


Abbildung 2: Episodensplitting

### 3.2.3 Fehlende Werte

In den administrativen Daten gilt die folgende Codierung von fehlenden Werten:

Tabelle 6: Codierung fehlender Werte

Bezeichnung	Wert	Beschreibung
Keine (valide) Angabe	.z	Nicht systematisch fehlende Ausprägungen eines Merkmals, d.h. das Merkmal ist prinzipiell für die Quelle gefüllt, jedoch ist für die betrachtete Ausprägung keine Angabe vorhanden bzw. keine Zuordnung möglich.
Systematisch nicht gefüllt	.n	Ein Merkmal ist grundsätzlich für eine Quelle (dunkelgraue Felder der Merkmalsübersicht im Abschnitt 1.5) oder für einen gewissen Zeitraum nicht gefüllt.

### 3.3 Datenverknüpfung

Ausgangspunkt der Verknüpfung der Befragungs- und administrativen Daten sind die PASS-Teilnehmerinnen und Teilnehmer bis einschließlich Welle 8.<sup>7</sup> Diese wurden im Rahmen des

<sup>7</sup> Die Befragungsdaten in PASS-ADIAB enthalten auch Daten zu Personen, die in PASS Welle 9 erstmals an der Befragung teilgenommen oder Verknüpfungsbereitschaft erklärt haben. Zu diesen Personen hat jedoch bisher kein Verknüpfungsversuch stattgefunden.

Interviews um Erlaubnis gebeten, ihre Befragungsdaten mit den über sie bei der BA vorhandenen administrativen Daten für Forschungszwecke verknüpfen zu dürfen. Der Wortlaut der Frage lautete wie folgt:

*„Um das Interview im Folgenden möglichst kurz zu halten, würden wir gerne bei der Auswertung der Befragung Auszüge aus Daten einbeziehen, die beim Institut für Arbeitsmarkt- und Berufsforschung in Nürnberg vorliegen.*

*Dabei handelt es sich zum Beispiel um zusätzliche Informationen zu vorausgegangenen Zeiten der Beschäftigung, der Arbeitslosigkeit und der Teilnahme an Maßnahmen während der Arbeitslosigkeit. Zum Zweck der Zuspiegelung dieser Daten an die Interviewdaten möchte ich Sie herzlich um Ihr Einverständnis bitten.*

*Werden diese Informationen ausgewertet, so ist absolut sichergestellt, dass alle datenschutzrechtlichen Bestimmungen strengstens eingehalten werden. Ihr Einverständnis ist selbstverständlich freiwillig. Sie können es auch jederzeit wieder zurückziehen. Sind Sie damit einverstanden?“*

Über die Wellen 1 bis 8 hinweg haben 81,07% der Befragten ihre Zustimmung erteilt. Anschließend wurden die zustimmenden Befragten in den administrativen Datenquellen der BA gesucht. Dazu wurden aus allen verfügbaren Adressdatenquellen folgende personenbezogenen Merkmale herangezogen:

- Vor- und Nachname
- Geschlecht
- Adresse
- Geburtsdatum
- Bedarfsgemeinschaftsidentifikator

Bis einschließlich Welle 5 wurde jeweils nach Abschluss einer jeden Befragungswelle versucht, zuspiegelbereite Befragungsteilnehmer in den administrativen Daten des IAB zu identifizieren, die entweder in der jeweiligen Welle erstmals teilgenommen hatten, bereits vorher teilgenommen aber noch keine Verknüpfungserlaubnis erteilt hatten oder bereits vorher verknüpfungsbereit waren, aber bei vorherigen Anläufen noch nicht erfolgreich verknüpft werden konnten. Da direkt nach den Wellen 6 und 7 kein Verknüpfungsversuch unternommen wurde, wurden bei dem bisher letzten Anlauf alle in den Wellen 6 bis 8 neu hinzugekommenen, verknüpfungsbereiten Befragten gemeinsam bearbeitet. Hinzu kamen alle Personen, die bereits bis einschließlich Welle 5 verknüpfungsbereit gewesen waren aber während der ersten fünf Anläufe nicht erfolgreich verknüpft werden konnten. Diese wurden im letzten Anlauf erneut gesucht.<sup>8</sup>

---

<sup>8</sup> Die Verknüpfungsanläufe ab Welle 5 wurden durch Mitarbeiterinnen und Mitarbeiter des German Record Linkage Center (GRLC, siehe Antoni/Schnell (2017) oder <http://www.record-linkage.de>) durchgeführt. Bei Fragen zu diesen Verknüpfungen wenden Sie sich bitte direkt an das GRLC.

Über die Zeit hinweg wurde das Vorgehen bei der Verknüpfung der beiden Datenquellen mehrfach leicht verändert. Es folgte jedoch stets im Wesentlichen dem in den Abschnitten 3.3.1 bis 3.3.4 beschriebenen Ablauf. Da nicht jeder der hier genannten Arbeitsschritte in jedem der Verknüpfungsanläufe angewendet wurde, wird auf Abweichungen bei der Beschreibung des jeweiligen Prozessschritts hingewiesen.

Generierte technische Merkmale im Datensatz kennzeichnen für jede Befragte und jeden Befragten, in welchem Verknüpfungsanlauf und in welchem Arbeitsschritt eine erfolgreiche Verknüpfung zustande gekommen ist. Damit kann jeder Nutzer nach eigenem Ermessen entscheiden, welche der Fälle in den eigenen Analysen Verwendung finden sollen. Unsere Empfehlung ist, alle in den Daten enthaltenen Matches zu verwenden, da nur Matches in den Datensatz eingeflossen sind, die mit sehr hoher Wahrscheinlichkeit eine richtige Zuordnung darstellen.<sup>9</sup>

Von den Befragten, für die erfolgreich entsprechende Datensätze in den BA-Prozessdaten gefunden werden konnten, weisen nicht alle tatsächlich Verlaufsdaten in den IEB auf. Die wenigen Ausfälle ergeben sich dadurch, dass möglicherweise personenidentifizierende Angaben in den Datenquellen der BA über eine Person vorliegen können, obwohl sie nie IEB-relevante Zustände durchlaufen haben. Beispiele hierfür sind Personen, die sich in einer Arbeitsagentur haben beraten lassen, anschließend aber nicht in den Arbeitsmarkt eingetreten sind oder sich selbständig gemacht haben. Die ratsuchende Person wäre damit nie als arbeitsuchend, Maßnahmeteilnehmer/in, Leistungsbezieher/in oder sozialversicherungspflichtig Beschäftigte/r in Erscheinung getreten. Ein solcher Ausfall bedeutet daher keinen Verlust für die verknüpften Daten, da zu dieser Person ohnehin keine administrativen Erwerbsverlaufsdaten erhoben wurden.

**Tabelle 7** stellt die Anzahl der Befragten der Wellen 1 – 8 dar, die in irgendeinem Anlauf erfolgreich verknüpft werden konnten.<sup>10</sup> Zusätzlich differenziert die Tabelle danach, über welche Verknüpfungsmethode eine Person in den administrativen Daten des IAB identifiziert werden konnte. Zudem stellt die Tabelle dar, wie viele der identifizierten Personen tatsächlich Beobachtungen in den IEB-Daten aufweisen.

---

<sup>9</sup> Nutzer, die mutmaßlich fehlerhafte Verknüpfungen in PASS-ADIAB entdecken, die den Prüfungen durch die Datenverantwortlichen entgangen sind, werden gebeten, diese Fälle an das FDZ zu melden. Für spätere Versionen von PASS-ADIAB werden diese Zuordnungen ggf. korrigiert.

<sup>10</sup> Bei der Verknüpfung wurden nur Personen berücksichtigt, die bis einschließlich Welle 8 Ihre Zustimmung gaben. Aus Gründen der Vergleichbarkeit wird daher in Tabelle 7 der Anteil der Verknüpften in Relation zu den Befragten bis einschließlich Welle 8 angegeben.

**Tabelle 7: Anzahl verknüpfter Befragter nach Typ des Matches (Wellen 1 - 8)**

	Anzahl	Anteil an Zustimmenden	Anteil
Befragte (bis Welle 8)	36.381		100,00%
Zustimmung zur Verknüpfung	29.488	100,00%	81,05%
Matches insgesamt	25.183	85,40%	69,22%
<i>darunter:</i>			
Goldstandard-Matches	14.878	50,45%	40,89%
Deterministische Matches	7.015	23,79%	19,28%
Fehlertolerante Matches	2.972	10,08%	8,17%
Valider Match, unb. Matchtyp	318	1,08%	0,87%
Matches mit IEB-Daten	25.048	84,94%	68,85%
<i>darunter:</i>			
Goldstandard-Matches	14.863	50,40%	40,85%
Deterministische Matches	6.936	23,52%	19,06%
Fehlertolerante Matches	2.933	9,95%	8,06%
Valider Match, unb. Matchtyp	316	1,08%	0,87%

Quelle: PASS-ADIAB, eigene Berechnungen

### 3.3.1 Bereinigung und Standardisierung personenbezogener Merkmale

Die oben genannten personenidentifizierenden Merkmale wurden Bereinigungsroutinen unterzogen, um eine möglichst hohe Übereinstimmung zwischen beiden Datenquellen zu erreichen. Die Routinen unterschieden sich zwar leicht zwischen den Verknüpfungsanläufen, wurden jedoch innerhalb einzelner Anläufe einheitlich auf beide Datenquellen angewendet.

Konkret wurden Umlaute und Sonderzeichen durch ihre jeweiligen ASCII-konformen Äquivalente ersetzt, alle Buchstaben in Großbuchstaben umgewandelt sowie überflüssige Leerzeichen und Piktationszeichen entfernt. Mehrere Vor- bzw. Nachnamen wurden in einem Feld ohne Leerzeichen zusammengefasst. Neben diesen grundlegenden Bereinigungen werden seit Welle 5 zusätzliche, variablen-spezifische Routinen zur Standardisierung der Daten durchgeführt. So wird die Hausnummer aus einem kombinierten Straße-/Hausnummernfeld extrahiert und alle übrigen Adresszusätze (z.B. Stockwerk, Appartement-Nr.) gelöscht. Bei Straßennamen werden häufig vorkommende Muster und Abkürzungen standardisiert. Bei Ortsnamen werden Anhänge auf eine einheitliche Form gebracht (z.B. bei Ortschaften mit zusätzlicher Information wie Fluß oder Region) bzw. entfernt (z.B. bei Berlin-Kreuzberg).

### 3.3.2 Goldstandard-Linkage

Nach Bereinigung und Standardisierung der Namens- und Adresskomponenten erfolgte zunächst ein von uns als Goldstandard-Linkage bezeichnetes Vorgehen. Bei Befragten aus ALG-II-Bezieherstichproben wurde die Tatsache genutzt, dass die Stichprobenziehung auf Ebene der Bedarfsgemeinschaft stattfand. Dadurch lag ein für alle Personen eindeutiger Bedarfsgemeinschaftsidentifikator vor. Für Abgleiche innerhalb der jeweiligen Bedarfsgemeinschaft konnte auf die Adressbestandteile als Identifikatoren verzichtet werden. Innerhalb der Bedarfsgemeinschaften war es daher ausreichend, Datensatzpaare auf eine exakte Übereinstimmung der Merkmale Geschlecht und Geburtsdatum zu überprüfen. Im Falle von Mehrfachzuordnungen wurde zusätzlich auf den Namen abgestellt. Durch den Vorteil der

eindeutigen Zuordnung von Personen zu Bedarfsgemeinschaften konnte in diesem Schritt eine hohe Verknüpfungsquote bei einem gleichzeitig hohen Maß an Sicherheit erreicht werden.

### **3.3.3 Deterministisches Record Linkage**

Anschließend wurde ein deterministischer Abgleich der Felder zwischen den Einträgen der beiden Datenquellen durchgeführt, falls eine Person über den vorherigen Schritt nicht erfolgreich verknüpft werden konnte. Dies war generell für alle Personen der Fall, die nicht Teil der ALG-II-Bezieher-Stichproben waren. In diesem Schritt wurde nicht mehr auf den Bedarfsgemeinschaftsidentifikator zurückgegriffen.

Im ersten deterministischen Linkage-Schritt war die exakte Übereinstimmung in allen sonst zur Verfügung stehenden Identifikatoren ausschlaggebend, also Vor- und Nachname, Geburtsdatum, Geschlecht, Postleitzahl, Ort, Straße und Hausnummer. In den nächsten Schritten wurde jeweils bei weiterhin nicht verknüpften Personen eine Abweichung auf bestimmten Feldern toleriert, während die jeweils übrigen Felder weiterhin exakt übereinstimmen mussten. Dieser iterative Prozess begann mit der strengsten, am wenigsten risikobehafteten Merkmalskombination und wurde nach und nach gelockert. Es wurden jedoch nur Merkmalskombinationen berücksichtigt, bei denen das Risiko, zwei unterschiedliche Personen zu verknüpfen (sog. falsch-positive Matches), sehr gering war.

### **3.3.4 Fehlertolerantes Record Linkage**

Für Personen, für die durch keinen der bisherigen Schritte eine Entsprechung in den administrativen Daten ermittelt werden konnte, wurde anschließend ein fehlertolerantes Matching-Verfahren verwendet. Dabei wurde für jedes abzugleichende Variablenpaar aus beiden Datenquellen ein Ähnlichkeitsmaß berechnet und diese in einem zusammenfassenden Qualitätsindex aufsummiert. Die Berechnungen werden mit der Software Merge ToolBox (MTB, Version 0.742, siehe Schnell/Bachteler/Reiher, 2005; Bachteler, 2012) durchgeführt.

Bis einschließlich des Verknüpfungsanlaufs nach Welle 5 wurde bei der Berechnung des Ähnlichkeitsmaßes ein probabilistisches Matching angewandt, bei dem rechnerisch berücksichtigt wird, dass die Wahrscheinlichkeit einer Übereinstimmung von Merkmalen von deren Varianz in der Grundgesamtheit abhängig ist (siehe Fellegi/Sunter, 1969; Christen, 2012, S. 133-137). Hierbei werden so genannte m- und u-Parameter berechnet, die angeben, mit welcher Wahrscheinlichkeit die Angaben zu einer Person zwischen zwei verschiedenen Quellen variieren bzw. wie wahrscheinlich das Auftreten eines identischen Merkmals bei unterschiedlichen Personen ist. Beispielsweise deutet eine Übereinstimmung bei der Variable Geschlecht weniger auf einen tatsächlichen Match hin als eine Übereinstimmung beim Nachnamen und wird deshalb niedriger gewichtet. Die variablenspezifischen m- und u-Parameter, die dem probabilistischen Matching zugrunde liegen, wurden im Rahmen des Verknüpfungsanlaufs nach Welle 1 ermittelt (siehe Bachteler, 2008). Es werden alle zur Verfügung stehenden Merkmale außer dem Bedarfsgemeinschaftsidentifikator zwischen den beiden Datenquellen verglichen und ein additives Ähnlichkeitsmaß berechnet. Je höher der Wert dieses Index, desto größer die Gesamtähnlichkeit der beiden Zeilen und desto



wahrscheinlicher liegt eine tatsächliche Übereinstimmung vor. Da für den fehlertoleranten Linkage-Schritt im Verknüpfungsanlauf nach den Wellen 6 - 8 nur sehr wenige noch nicht verknüpfte Fälle übrig waren, wurden diese Fälle einem distanzbasierten Abgleich ohne probabilistische Klassifizierung unterzogen.

Datenpaare, deren Ähnlichkeitsmaß über durch vorläufige Sichtung bestimmten Schwellenwerten lag, wurden als fehlertolerante Matches eingestuft. Die Adresspaare mit einem Ähnlichkeitswert unterhalb des Schwellenwerts wurden anschließend einem Sichtvergleich unterzogen. Die dadurch ermittelten Verknüpfungen wurden nachträglich als fehlertolerante Matches klassifiziert.

## **4 Datenqualität und Probleme**

### **4.1 PASS-Befragungsdaten**

Non-Response, Gewichtung und andere Aspekte der Datenqualität der PASS-Daten werden im FDZ-Datenreport 07/2016 (Berg et al., 2016) diskutiert. Im FDZ-Methodenreport 09/2011 (Beste, 2011) werden potentielle Verzerrungen durch Selektionsprozesse bei der Zustimmung zur Datenzuspielung untersucht.

Unter [http://fdz.iab.de/de/FDZ\\_Individual\\_Data/PASS.aspx](http://fdz.iab.de/de/FDZ_Individual_Data/PASS.aspx) finden sich zahlreiche weitere surveymethodologische Publikationen, die sich mit der Qualität der PASS-Daten beschäftigen.

### **4.2 Gesamte IEB**

Die IEB beinhalten Erwerbsbiografien. Allerdings ist in den administrativen Daten nicht jede Art von Erwerbstätigkeit erfasst. Bei entsprechenden Lebensverläufen sind bestimmte Personen überhaupt nicht in den IEB abgebildet.

Für Auswertungszwecke ist es oft relevant, Lücken in den abgebildeten Biografien zu kennen (z.B. für Kontrollgruppenbildung, Lebensverlaufsanalysen etc.). Die im Folgenden aufgeführten Lücken sind definiert als Zeiträume nach Ende der Schulausbildung, für die kein Satz in den IEB enthalten ist. Diese Lücken können unterschieden werden in

- Lücken, für die keinerlei Information vorliegt und
- Lücken, für die anhand des Merkmals Abgabegrund/Beendigungsgrund/SGB-II-Einstellungsgrund/Abmeldegrund des direkten Vorgängersatzes eine Information vorliegen kann (sofern ein entsprechender Satz vorhanden ist).

Diese Lücken wurden identifiziert anhand der Ausprägungen der Merkmale Abgabegrund/Beendigungsgrund/SGB-II-Einstellungsgrund/Abmeldegrund und Erwerbsstatus in den verschiedenen Quellen. Die Auflistung erhebt keinen Anspruch auf Vollständigkeit.



**Tabelle 8: Biographische Lücken und Identifikationsmöglichkeiten**

<b>Biografische Lücke</b>	<b>Information über Lücke, ggf. identifizierbar anhand der Ausprägung des Merkmals „grund“ im Vorgängersatz der Quelle</b>
Beamte, Berufssoldaten, Richter, Angestellte von Körperschaften oder Stiftungen des öffentlichen Rechts	XASU
Ungeförderte Selbstständige	LeH, ASU
Studierende, Personen in schulischer Weiterbildung	LeH, LHG, ASU, XASU
Personen, die länger als 6 Wochen krank / arbeitsunfähig sind (Krankheit während Arbeitslosigkeit ist aber unter bestimmten Bedingungen in der Quelle ASU abgebildet, siehe Abschnitt 3.2.1.4)	BeH, LeH, ASU
Personen in Altersrente ohne Beschäftigung wenn nicht Mitglied in einer BG	LeH, LHG, ASU
Personen in Mutterschutz / Elternzeit	XASU
Bezieher von Vorruhestandsgeld	LeH, ASU
Hausgewerbetreibende	
Kurzfristig Beschäftigte	ASU
Personen in Einrichtungen der Jugendhilfe, Berufsbildungswerken, anerkannten Werkstätten oder ähnlichen Einrichtungen für behinderte Menschen	ASU
Teilnehmende an Leistungen zur Teilhabe am Arbeitsleben (Rehabilitanden)	ASU
(Nebenerwerbs-)Landwirte	
Pflegepersonen i.S. von § 19 SGB XI	
Grundwehrdienstleistende	BeH, LeH, LHG, ASU, XASU
Wehrübungsleistende	BeH, LeH, LHG, ASU, XASU
Zivildienstleistende	BeH, LeH, LHG, ASU, XASU
Ableistende eines freiwilligen sozialen bzw. ökologischen Jahres anstelle des Zivildienstes	
Sonstige Personen, die nicht bei der Rentenversicherung oder der BA registriert sind (z.B. Sabbatjahr, Finanzierung aus privatem Vermögen oder Rentenbezügen, Emigration, Beschäftigung im Ausland, ehrenamtliche Tätigkeit, etc.)	BeH, LeH, ASU
Streikende bei Streikdauer über einem Monat	LeH
Sozialhilfe-Bezieher (vor Einführung des SGB II 2005), Sozialgeld-Bezieher (im Sinne des SGB II)	

### 4.3 Beschäftigtenhistorik (BeH)

- Die Einführung des neuen Tätigkeitsschlüssels in 2011 hat zu einer Reihe von Problemen geführt. So kam es während des Übergangszeitraums, der Arbeitgebern im Meldeverfahren zur Sozialversicherung eingeräumt wurde,<sup>11</sup> zu einer temporären Zunahme fehlender Angaben. Analysen der Statistik der BA (Bertat et al., 2013, S. 10) zeigen, dass die neuen bzw. umgestellten Merkmale „ausgeübte Tätigkeit“, „Arbeitszeit“ und „Ausbildungsabschluss“ nach der Umstellung zu etwa 20 bis 30 Prozent keine Angaben enthielten. Eine deutliche Verbesserung stellt sich im ersten Halbjahr 2013 ein. Um die Qualität der Arbeitszeitangabe im Übergangszeitrum zu verbessern, haben Ludsteck/Thomsen (2016) ein Imputationsverfahren entwickelt, das die fehlenden durch imputierte Werte ersetzt. Die imputierten Daten werden in der PASS-ADIAB 7515 bereits verwendet.
- Aufgrund der Einführung des Meldeverfahrens in den neuen Bundesländern kann erst ab 1993 von einer hinreichenden Vollständigkeit der Meldungen ostdeutscher Betriebe ausgegangen werden. Aus demselben Grund weisen sehr viele Sätze des Jahres 1991 für einige Merkmale (u.a. Erwerbsstatus) fehlende Werte auf.
- Der Anstieg der Zahl der BeH-Sätze ab 1999 folgt aus der Einführung der Meldepflicht für geringfügig Beschäftigte ab dem 1. April 1999.
- Es lässt sich insbesondere 1999 ein deutlicher Anstieg bei den Meldungen zu Teilzeitbeschäftigung feststellen. Verursacht wird dieser Befund sowohl vom tatsächlich zu beobachtenden Anstieg der Teilzeitbeschäftigung als auch von der Tatsache, dass seit 1999 die Beschäftigtenmeldungen generell besser ausgefüllt werden.
- Innerhalb des Meldeverfahrens ist eine gewisse Zeitverzögerung nicht vermeidbar. Zwar müssen Änderungen von Beschäftigungsverhältnissen sofort gemeldet und bestehende Beschäftigungsverhältnisse jährlich bis zum April (bis Jahresmeldung 2012) bzw. Mitte Februar (ab Jahresmeldung 2013) des folgenden Jahres bestätigt werden. Faktisch treffen Meldungen jedoch zum Teil um Jahre verspätet ein. Die Historikdatei des IAB wird aber nicht kontinuierlich fortgeschrieben, sondern in bestimmten Abständen aktualisiert. Dies geschieht unter Verwendung von Dateien mit den Beschäftigten-Meldungen jeweils eines Jahres, die 36, 18, 12 oder 6 Monate nach Ablauf des Berichtsjahres eingetroffen sind (frühestens im

<sup>11</sup>Die im Rahmen des Meldeverfahrens eingesetzten Prüfprogramme ließen bis einschließlich Mai 2012 leere Angaben beim Tätigkeitsschlüssel 2010 zu.

Juli 2015 z.B. kann die 18-Monats-Datei für das Jahr 2013 erstellt werden). Mehr als 3 Jahre verspätete Meldungen werden im IAB nicht berücksichtigt, so dass eine 36-Monats-Datei per definitionem einen Füllgrad von 100% aufweist.

- In der zugrundeliegenden Version der IEB liegt der Füllgrad der BeH-Meldungen zum Jahr 2012 bei knapp 100%, da für dieses Jahr nur das 30-Monatsmaterial verwendet werden konnte.<sup>12</sup> Für das Jahr 2013 wurde eine 18-Monats-Datei verwendet, für 2014 eine 6-Monats-Datei. Es ist daher davon auszugehen, dass die BeH-Meldungen für das Jahr 2013 leicht und für 2014 etwas stärker untererfasst sind. Die Auswertbarkeit der Daten auf Personenebene sollte dadurch aber nicht eingeschränkt sein. Allerdings häufen sich die fehlenden Meldungen in wenigen Betrieben. Dies bedeutet, dass in Einzelfällen die Betriebsangaben, z.B. die Betriebsgröße, grob falsch sind und sich in Folgeversionen stark verändern werden.
- Im Jahr 1984 erfolgte eine Änderung des Meldeverfahrens. Einmalig gezahltes Bruttoarbeitsentgelt wurde ab diesem Zeitpunkt als Teil des beitragspflichtigen Jahresentgelts gemeldet, was zu einer Erhöhung des durchschnittlichen Tagesentgelts führt. Insbesondere steigt ab diesem Jahr der Anteil der Entgelte oberhalb der Beitragsbemessungsgrenze stark an (siehe Bender et al., 1996).
- Für die Jahre 1992 bis 2000 wurden in den IEB auffällige Einbrüche und Anstiege in der Anzahl der Meldungen festgestellt. Es lassen sich v.a. Einbrüche bei folgenden zehn Kreisen beobachten: Braunschweig (03101), Wolfsburg (03103), Emden (03402), Kassel (06633), Essen (05113), Neuss (05162), Erftkreis (05362), Hersfeld-Rotenburg (06632), Miltenberg (09676) und Kempten (Allgäu) (09763). Grund hierfür sind Meldeprobleme bei einzelnen bzw. mehreren Betrieben in diesen Regionen.
- Nennenswerte Einbrüche verzeichnen auch die Kreise Salzgitter (03102) und Hoyerswerda (14264).
- Bei den Meldungen zu Vollzeitbeschäftigung in den IEB sind besonders die Kreise Main-Taunus-Kreis (06436) und Alzey-Worms (07331) erwähnenswert. Hier sind überdurchschnittliche Anstiege zu beobachten. Als Grund sind ebenfalls Meldeprobleme bei einzelnen bzw. mehreren Betrieben in diesen Regionen anzuführen.
- In den Jahren 1996 bis 1998 kommen die Werte 841-844 (Ärzte und Apotheker) in der Variable Beruf – ausgeübte Tätigkeit (beruf) im Vergleich zu benachbarten Jahren sehr selten vor. Die Gründe dafür sind nicht bekannt.

---

<sup>12</sup>Durch ein Redesign der Datengrundlage der BeH stand zum Zeitpunkt der Erstellung der BeH ausnahmsweise nur das 30- statt des 36-Monatsmaterials zur Verfügung. Analysen mit früheren Datenversionen haben jedoch gezeigt, dass i.d.R. bereits die 18-Monats-Datei einen Füllgrad von rund 99% aufweist. Demnach kommt es in den weiteren 18 Monaten nur noch zu extrem wenigen Nach- oder Korrekturmeldungen.

#### 4.4 Leistungsempfängerhistorik (LeH)

- Für die neuen Bundesländer werden die LeH-Sätze erst ab 1992 vollständig erfasst.
- Die Leistungsbezugsdaten wurden früher auf Magnetbändern gesichert. Aufgrund eines Schadens an einem Magnetband sind die Leistungsbezugsdaten bis einschließlich des Jahres 1980 nur teilweise enthalten. Daher ist für das vorliegende Datenprodukt auch davon auszugehen, dass die Informationen zum Leistungsbezug im genannten Zeitraum nur zum Teil enthalten sind.
- Durch einen internen Systemwechsel kommt es 2004 zu einem Bruch in der Erfassung von Sperr- und Säumniszeiten. Bis 1. Juli 2004 können Sperr und Säumniszeiten nur durch den Abmeldegrund im vorhergehenden LeH-Satz identifiziert werden. Danach existiert ein eigener Satz mit täglichem Leistungssatz = 0 für Versagens-, Entziehens- und Ruhezeiten.

#### 4.5 Leistungshistorik Grundsicherung (LHG)

- In den Jahren 2005 und 2006 ist die LHG nicht vollständig gefüllt. Es kommt zu erheblichen trägerspezifischen Ausfällen. Es wird daher abgeraten, für diesen Zeitraum Analysen allein auf Basis der LHG zu machen.
- Längsschnittuntersuchungen für Personen sind mit Unsicherheiten behaftet, da Umzüge in und aus Gebieten von Trägern mit Lieferproblemen nicht von Änderungen des Status der Leistungsberechtigung unterschieden werden können.
- Auch ab 2007 kommt es hauptsächlich bei zugelassenen kommunalen Trägern immer wieder zu meist einmonatiger trägerspezifischer Untererfassung.
- Im Zusammenhang mit Trägerartwechseln kommt es zu Unter-/Übererfassungen:
  - Im Rahmen der Trägergebietsreform vom 1. Januar 2011 kommt es zu Untererfassungen im Gebiet der Arbeitsagenturen Dessau-Roßlau, Halberstadt, Halle und Sangerhausen.
  - Doppelmeldungen aufgrund der Gebietsreformen 2009/2011 und der Trägerformwechsel zum 1. Januar 2012 werden in der IEB bereits soweit möglich bereinigt. Dennoch können Doppelmeldungen auftreten.
- Bei folgenden Jobcentern bestehen Unschärfen bei der Zuordnung von Leistungsfällen:
  - zwischen Emden und Norden zwischen September und Dezember 2009
  - zwischen Döbeln und Mittelsachsen von Oktober bis Dezember 2012
  - zwischen Tirschenreuth und Wunsiedel von November 2012 bis März 2013
- Manche Personen, für die eine LHG-Beobachtung vorliegt, sind vom Leistungsbezug nach SGB II ganz oder teilweise ausgeschlossen, da sie z. B. eine geförderte Ausbildung machen, Altersrente beziehen, in einer stationären Einrichtung untergebracht sind oder Versicherungsleistungen zur Vermeidung von Hilfebedürftigkeit bekommen. Dies betrifft pro

Jahr ca. 3 bis 5 Prozent der Fälle. In XSozial kommt es zu trägerspezifischen Untererfassungen dieser Personengruppe. In der SIAB ist der Ausschluss von Leistungen nicht identifizierbar.

## 4.6 Arbeitssuchendenhistorik (ASU/XASU)

### 4.6.1 ASU

- Die registrierten Zeiten der Arbeitssuche der Quelle ASU können ab 1997 als vollständig angesehen werden. Das Analysepotenzial der ASU-Sätze vor 1997 ist damit eingeschränkt.
- Für die Vermittlungskräfte ist es nicht möglich, die Zuordnung zum Rechtskreis immer unmittelbar zu erfassen, da vielfach erst zeitverzögert feststeht, welcher Träger aufgrund einer möglichen SGB-II-Leistungsberechtigung vorrangig zuständig ist. Es wird daher empfohlen, die Ausprägung des Merkmals „Trägerart“ in der ASU mit der Ausprägung in der LHG im gleichen Zeitraum abzugleichen. Aufgrund der Erfassungslücken bei LHG im Zeitraum 2005 bis 2006 ist dies nicht immer möglich.
- Für manche Personen, für die ab 2005 eine Optionskommune zuständig war, wurden von der BA parallel "künstliche" ASU-Sätze angelegt. Diese sind über estatvor (Wechsel zu einem zugelassenen kommunalen Träger) zu identifizieren.
- Von Mitte 2005 bis Mitte 2006 wurde in der BA das EDV-Verfahren coArb, aus dem die Arbeitssuchendendaten stammen, durch das Verfahren VerBIS abgelöst. Ab Juli 2005 wurde coArb erstmals in der Agentur Wiesbaden als Pilotprojekt, ab Dezember 2005 dann sukzessive in mehreren Tranchen in allen Arbeitsagenturen durch VerBIS ersetzt. Viele der erfassten Merkmale wurden in den beiden Systemen in unterschiedlicher Differenziertheit und unterschiedlicher qualitativer Gewichtung abgefragt. Die Integration dieser Merkmale in die IEB gestaltet sich daher schwierig und ist nur durch die Anwendung eines speziellen Verfahrens (Mapping) möglich. Leider kann durch das Mapping keine vollständige Umsetzung der betroffenen Merkmale von coArb auf VerBIS erreicht werden, sodass sich bei einigen Merkmalen die Ausprägungen „Keine Angabe“, „Sonstiges“ oder „Missing“ ungewöhnlich häufen. Außerdem können sich bei Häufigkeitsauszählungen auffällige Unterschiede ergeben, je nachdem ob die Ursprungsquelle der Daten coArb oder VerBIS war. Wichtige Einschränkungen des Analysepotentials werden in der jeweiligen Variablenbeschreibung angesprochen.
- Das bis Juni 2006 genutzte Verfahren coArb unterstützte ausschließlich die Vermittlung von Arbeitslosen und Arbeitssuchenden. Lückenhaft erfasst wurden auch Daten zu Personen, die sich lediglich beraten lassen wollten. Die Daten der Berufsberatung wurden in einem separaten System gehalten. Mit VerBIS erweiterten sich die Ausprägungen des Arbeitssuche-

Status um „Ratsuchende“ und Personen „ohne Status“. Zu Letzteren gehören ALG-II-Leistungsberechtigte, die eingeschränkt der Arbeitsvermittlung verfügbar sind. Ihre Erfassung in VerBIS gilt erst seit Januar 2008 als weitgehend vollständig.

- Ein Wechsel der Grundsicherungsträger oder des Wohnorts führen nicht zu einem neuen ASU-Satz, die Merkmalsausprägung jeweils zu Beginn einer Episode wird fortgeführt. Mit zunehmender Dauer des Spells wird also das Risiko größer, dass Träger oder Wohnort nicht mehr zutreffend sind.

#### **4.6.2 XASU**

- Im Gegensatz zu den Arbeitsuchendendaten aus den ARGEn und den getrennten Trägerschaften zeigen sich nach 1.1.2005 für die zugelassenen kommunalen Träger zwischenzeitlich systematische Untererfassungen. Deshalb sollten Daten aus der Quelle XASU erst ab 2007 ausgewertet werden.
- Verschiedene Merkmale weisen für die XASU zum Teil nur einen sehr geringen Füllgrad auf. Hiervon betroffen sind unter anderem die Merkmale Schulabschluss, Abgabegrund sowie Erwerbsstatus vor Arbeitsuche. Die Füllgrade dieser Merkmale verbessern sich zwar im Zeitablauf, bleiben aber dennoch zum Teil unbefriedigend. Das Merkmal Beruf ist nahezu für den gesamten zur Verfügung stehenden Zeitraum in der XASU nicht gefüllt.
- Bei etlichen Trägern (Kreisen) erscheint der Anteil der gemeldeten Arbeitsuchenden an den gemeldeten erwerbsfähigen Hilfebedürftigen in den IEB durchgängig oder zeitweise unplausibel hoch. Gründe dafür könnten eine möglicherweise fehlerhafte Bestimmung des Status „Nichtarbeitslos, arbeitssuchend“ oder eine technisch fehlerhafte Übermittlung mit hohem Anteil an Nachmeldungen bei diesen Trägern sein.
- Die träger- und zeitraumabhängige Plausibilität der XASU-Angaben sollte jeweils unter Berücksichtigung der Untersuchungsfrage vor der Verwendung geprüft werden.

#### **4.7 Maßnahmeteilnehmergehistoriken (MTH)**

- Die MTH sind für Maßnahmen mit einem Eintrittsdatum vor dem 1.1.2000 unvollständig.
- Zum 1.1.2005 kommt es zu einem Bruch, da mit Einführung des SGB II Maßnahmeteilnehmer auf verschiedene Träger aufgeteilt wurden (siehe Abschnitte 2.2.3 und 2.2.5).
- Die MTH enthält nur Meldungen, die in BA-Verfahren gebucht sind. Deren Nutzung nimmt in Arbeitsgemeinschaften/ getrennten Trägerschaften/ getrennten Aufgabenwahrnehmungen zwischen 2005 und 2007 kontinuierlich zu. Ab März 2007 sind die Meldungen für diese Träger komplett.



- Maßnahmen, die über den Standard XSozial gemeldet werden, sind nicht in der MTH oder PASS-ADIAB enthalten.
- Durch die Neuorganisation von SGB-II-Trägern kann die Dokumentation von Maßnahmeteilnahmen in der MTH abrechnen oder neu beginnen, wenn es zu einer Umstellung des Meldeverfahrens kommt.<sup>13</sup>
- Bei Meldungen zum Überbrückungsgeld wird zum Teil die maximal zulässige Dauer von einem halben Jahr überschritten. Dies kann in den meisten Fällen durch eine Voreinstellung der Eingabemaske im Erfassungssystem erklärt werden.
- Die MTH wird um Bewerbermerkmale (z. B. Ausbildung) aus anderen BA-Verfahren ergänzt. Für diese Merkmale erfolgte im Jahr 2006 die Umstellung des operativen Verfahrens von coArb auf VerBIS. Es gelten hier die gleichen Qualitätseinschränkungen wie für die ASU.

## 5 Variablenbeschreibung

Auszählungen sowie Übersichten über die einzelnen Ausprägungen und Labels der Merkmale finden sich unter [http://fdz.iab.de/de/FDZ\\_Individual\\_Data/PASS/Working\\_Tools.aspx](http://fdz.iab.de/de/FDZ_Individual_Data/PASS/Working_Tools.aspx) in gesonderten Dateien.

### 5.1 Identifikatoren

#### 5.1.1 Unveränderliche Personennummer (pnr)

Variablenlabel	Unveränderliche Personennummer
Variablenname	pnr
Kategorie	Identifikatoren
Herkunft	Generiert
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Die unveränderliche Personennummer zeigt an, welche Sätze zur gleichen Person gehören. Sie zeigt auch an, welche Personen in Befragungs- und administrativen Daten korrespondieren, d.h. die pnr ist sowohl in den Befragungsdaten als auch in den administrativen Daten enthalten. Von dieser Personennummer nicht auf Eigenschaften der Person oder Original-Identifikatoren geschlossen werden kann.

<sup>13</sup> Weitere Informationen zur Gebietsstruktur der SGB-II-Träger und relevanten Änderungen finden sich unter: <http://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen/Regionale-Gliederungen/Gebietsstruktur-Traeger-Grundsicherung-Nav.html>

	Da es in den verschiedenen Datenquellen keinen einheitlichen Personen-Identifikator gibt, ist die Zuordnung der Informationen aus den unterschiedlichen Datenquellen (z.B. Beschäftigung und Leistung) zu Personen(-Konten) nicht immer eindeutig. In solchen Fällen kann es zu unplausiblen Erwerbshistoriken kommen. Die Bildung des quellenübergreifenden Personen-Identifikators basiert auf einer von der BA entwickelten Heuristik.
--	---

### 5.1.2 Systemfreie Betriebsnummer (betnr)

Variablenlabel	Systemfreie Betriebsnummer
Variablenname	Betnr
Kategorie	Identifikatoren
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Die systemfreie Betriebsnummer zeigt an, welche Sätze zum gleichen Betrieb gehören. Sie basiert auf der von der BA vergebenen Betriebsnummer, die durch eine systemfreie Nummer ersetzt wurde. (Weitere Informationen zur Betriebsnummernvergabe der BA finden sich in Bender et al., 1996, S. 15f. und S. 27-30.) Die Betriebsnummer kann zusammen mit der Jahresangabe benutzt werden, um Personen- und Betriebsinformationen zusammenzuspielen.</p> <p>Für die Betriebsnummer sollte Folgendes grundsätzlich beachtet werden:</p> <ul style="list-style-type: none"> <li>a) Besteht das Unternehmen aus nur einer Niederlassung oder hat das Unternehmen in einer Gemeinde nur eine Niederlassung, so ist diese Niederlassung der Betrieb und erhält eine Betriebsnummer.</li> <li>b) Hat das Unternehmen <b>mehrere</b> Niederlassungen in <b>einer</b> Gemeinde, so sind diese Betriebs- und Arbeitsstätten zu einem Betrieb unter einer Betriebsnummer zusammenzufassen, wenn sie derselben Wirtschaftsklasse angehören bzw. ist jede Niederlassung ein Betrieb und erhält je eine Betriebsnummer, wenn sie <b>nicht</b> derselben Wirtschaftsklasse angehören.</li> <li>c) Hat das Unternehmen <b>mehrere</b> Niederlassungen in <b>mehreren</b> Gemeinden, so ist jede dieser Niederlassungen ein Betrieb und erhält je eine Betriebsnummer.</li> </ul> <p>In diesem Zusammenhang sind auch folgende Definitionen im Sinne der Betriebsnummernvergabe im Rahmen des Meldeverfahrens zur Sozialversicherung zu beachten:</p> <ul style="list-style-type: none"> <li>a) Ein <b>Betrieb</b> ist eine regional und wirtschaftsfachlich abgegrenzte Einheit, in der Beschäftigte tätig sind und für die nach o.a. Grundsätzen eine Betriebsnummer zu vergeben wird.</li> <li>b) Eine <b>Arbeitsstätte</b> ist eine Einheit, in der Beschäftigte tätig sind und für die nach o.a. Grundsätzen eine Betriebsnummer nicht zu vergeben ist.</li> </ul>

	<p>c) Ein <b>Unternehmen</b> als Begriff fasst Betriebs- und Arbeitsstätten desselben Arbeitgebers zusammen.</p> <p>d) Ein <b>Arbeitgeber</b> ist jede natürliche oder juristische Person, für die mindestens ein sozialversicherungspflichtig bzw. geringfügig Beschäftigter tätig ist.</p> <p>e) Betrieb und Betriebsstätten sind synonyme Begriffe; Niederlassung ist dann ein synonymer Begriff für Filialen, Zweigstellen, Auslagerungen, Dependancen, Arbeitsstätten u.ä., wenn sie kein Betrieb ist.</p>
Qualitätshinweis	Die Betriebsnummer fehlt in nur sehr wenigen BeH-Meldungen. Bei diesen Meldungen handelt es sich in den meisten Fällen um Meldungen der Personengruppe „205“ (Entgeltmeldungen für unständig Beschäftigte). Da Betriebsmerkmale (Arbeitsort, Wirtschaftszweig, Betriebsgröße usw.) über die Betriebsnummer zugespielt werden, fehlen sie in diesen Meldungen.

## 5.2 Generierte technische Merkmale

### 5.2.1 Quelle des Satzes (quelle)

Variablenlabel	Quelle des Satzes
Variablenname	Quelle
Kategorie	Generierte technische Merkmale
Herkunft	BeH, LeH, LHG, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Das Merkmal gibt die Datenquelle an.

### 5.2.2 Satzzähler pro Konto (spell)

Variablenlabel	Satzzähler pro Konto
Variablenname	Spell
Kategorie	Generierte technische Merkmale
Herkunft	BeH, LeH, LHG, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Der Satzzähler Konto nummeriert aufsteigend die Sätze einer Person. Das Merkmal wird beim Episodensplitting erzeugt und bezieht sich auf die gesplitteten Sätze. Mit dem Merkmal „Satzzähler Konto“ kann die ursprüngliche Sortierung auf einfache Art und Weise wiederhergestellt werden. Die Sortierung erfolgt grundsätzlich zunächst nach dem Beginndatum der Episode und dann nach der Quelle.

### 5.2.3 Jahr (jahr)

Variablenlabel	Jahr
Variablenname	jahr
Kategorie	Generierte technische Merkmale

Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Dieses Merkmal ist nur in der Betriebsdatei enthalten. Es gibt das Gültigkeitsjahr der Betriebsinformationen zum Stichtag 30. Juni an. Mit Hilfe dieses Merkmals und der Betriebsnummer können Personen- und Betriebsdatei verknüpft werden.

#### 5.2.4 Vorliegen administrativer Personendaten (match\_admin)

Variablenlabel	Vorliegen administrativer Personendaten
Variablenname	match_admin
Kategorie	Verknüpfungsbezogenes Merkmal
Herkunft	Generiert
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Dieses Merkmal ist nur in der Verknüpfungsdatei enthalten. Es zeigt für jeden PASS-Befragten an, ob zu dieser Person auch Beobachtungen in den administrativen Personendaten vorliegen. Diese Variable nimmt in wenigen Fällen den Wert 0 an, obwohl die Variable match_type einen erfolgreichen Match anzeigt. In diesen Fällen konnte zwar ein Match durch Abgleich der Identifikatoren beider Datenquellen erreicht werden, die IEB enthält jedoch zu dieser Person keine Episoden. Zu den möglichen Gründen für eine solche Diskrepanz siehe Abschnitt 3.3.

#### 5.2.5 Typ des Matches (match\_type)

Variablenlabel	Typ des Matches
Variablenname	match_type
Kategorie	Verknüpfungsbezogenes Merkmal
Herkunft	Generiert
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Dieses Merkmal ist nur in der Verknüpfungsdatei enthalten. Es beschreibt, auf welcher Stufe des Verknüpfungsprozesses eine Übereinstimmung zwischen der Befragungsadresse einer Person mit ihrer Adresse in den BA-Prozessdaten gefunden werden konnte. Die Ausprägungen 2-4 steigen mit der Abfolge der Stufen, d. h. Ausprägung 3 bedeutet, dass die Stufe des deterministischen Linkage nicht erfolgreich war usw. Ausprägung 0 steht für PASS-Befragte, die der Verknüpfung bis einschließlich Welle 8 nicht zugestimmt haben. Darunter sind auch Personen, die erst in Welle 9 Verknüpfungserlaubnis erteilt haben.

#### 5.2.6 Wellenindikator erfolgreicher Match (match\_wave)

Variablenlabel	Wellenindikator erfolgreicher Match
Variablenname	match_wave
Kategorie	Verknüpfungsbezogenes Merkmal

Herkunft	Generiert
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Dieses Merkmal ist nur in der Verknüpfungsdatei enthalten. Es zeigt an, nach welcher Befragungswelle der Match erfolgreich war. Hintergrund ist, dass nach jeder der ersten fünf Wellen und nach den Wellen 6 – 8 versucht wurde, neu hinzugekommene oder vorher nicht erfolgreich verknüpfte Befragte in den administrativen Daten zu finden. Da Personen nicht zwingend sofort im ersten Anlauf gefunden werden, kann die Ausprägung dieses Merkmal von der Welle abweichen, in der die Verknüpfungserlaubnis erstmals erteilt wurde.

## 5.3 Gültigkeitszeitraum

### 5.3.1 Beginndatum Originalbeobachtung (begorig)

Variablenlabel	Beginndatum Originalbeobachtung
Variablenname	Begorig
Kategorie	Gültigkeitszeitraum
Herkunft	BeH, LeH, LHG, ASU, XASU, MTH
Datentyp	Datum
Hierarchie	Keine
Detailbeschreibung	<p>Das Beginndatum des Originalsatzes entspricht dem ursprünglichen Beginndatum der Meldung. Dies kann vom Beginndatum der betreffenden Datenzeile, dem sogenannten Episoden-Beginndatum, abweichen (siehe auch die Erläuterungen zum Episodensplitting in Abschnitt 3.2.2). Da die Meldelogik unter bestimmten Umständen eine Reidentifikation des exakten Geburtsdatums ermöglichen würde, wurden die ursprünglichen Datumsangaben in diesen Fällen durch das in Abschnitt 3.2.1.6 beschriebene Anonymisierungsverfahren verändert.</p> <p><b>1) BeH</b> In BeH-Meldungen sind Beginn- und Endjahr wegen der Regeln des Meldeverfahrens immer identisch (Pflicht des Arbeitgebers zur Abgabe von Jahresmeldungen). Ein durchgehendes Beschäftigungsverhältnis kann also auf mehrere Meldungen verteilt sein.</p> <p><b>2) LHG, ASU, XASU</b> Bestimmte Änderungen führen zur Anlage eines neuen Datensatzes (siehe Abschnitte 2.2.3 und 2.2.4). begorig gibt das Beginndatum des neuen Zeitraums an.</p>

### 5.3.2 Enddatum Originalbeobachtung (endorig)

Variablenlabel	Enddatum Originalbeobachtung
Variablenname	endorig
Kategorie	Gültigkeitszeitraum

Herkunft	BeH, LeH, LHG, ASU, XASU, MTH
Datentyp	Datum
Hierarchie	Keine
Detailbeschreibung	<p>Das Enddatum des Originalsatzes entspricht dem ursprünglichen Enddatum der Meldung. Dies kann vom Enddatum der betreffenden Datenzeile, dem sogenannten Episoden-Enddatum, abweichen (siehe auch die Erläuterungen zum Episodensplitting in Abschnitt 3.2.2). Da die Meldelogik unter bestimmten Umständen eine Reidentifikation des exakten Geburtsdatums ermöglichen würde, wurden die ursprünglichen Datumsangaben in diesen Fällen durch das in Abschnitt 3.2.1.6 beschriebene Anonymisierungsverfahren verändert.</p> <p><b>1) BeH</b> In BeH-Meldungen sind Beginn- und Endjahr wegen der Regeln des Meldeverfahrens immer identisch (Pflicht des Arbeitgebers zur Abgabe von Jahresmeldungen). Ein durchgehendes Beschäftigungsverhältnis kann also auf mehrere Meldungen verteilt sein.</p> <p><b>2) LHG, ASU, XASU</b> Bestimmte Änderungen führen zur Anlage eines neuen Datensatzes (siehe Abschnitte 3.2.1.3 und 3.2.1.4). endorig gibt das Enddatum des neuen Zeitraums an.</p>

### 5.3.3 Beginndatum der Episode (begepi)

Variablenlabel	Beginndatum der Episode
Variablenname	begepi
Kategorie	Generierter Gültigkeitszeitraum
Herkunft	BeH, LeH, LHG, ASU, XASU, MTH
Datentyp	Datum
Hierarchie	Keine
Detailbeschreibung	Beginndatum der Episode. Es ist immer gleich dem oder größer als das Beginndatum der Originalbeobachtung (siehe auch die Erläuterungen zum Episodensplitting in Abschnitt 3.2.2).

### 5.3.4 Enddatum der Episode (endepe)

Variablenlabel	Enddatum der Episode
Variablenname	endepe
Kategorie	Generierter Gültigkeitszeitraum
Herkunft	BeH, LeH, LHG, ASU, XASU, MTH
Datentyp	Datum
Hierarchie	Keine
Detailbeschreibung	Enddatum der Episode. Es ist immer gleich dem oder kleiner als das Enddatum der Originalbeobachtung (siehe auch die Erläuterungen zum Episodensplitting in Abschnitt 3.2.2).

## 5.4 Informationen zur Person

### 5.4.1 Geschlecht (frau)

Variablenlabel	Geschlecht
Variablenname	frau
Kategorie	Persönliche Merkmale
Herkunft	BeH, LeH, LHG, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Geschlechter-Dummy (0 - Mann, 1 - Frau). Die Geschlechtsangabe ist innerhalb eines Personenkontos konstant.

### 5.4.2 Geburtsjahr (gebjahr)

Variablenlabel	Geburtsjahr
Variablenname	gebjahr
Kategorie	Persönliche Merkmale
Herkunft	BeH, LeH, LHG, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Das Geburtsjahr ist innerhalb eines Personenkontos konstant.
Qualitätshinweis	In den Originaldaten kann es vorkommen, dass das Geburtsjahr zwischen den Quellen wechselt. Dies wird bei der Datenaufbereitung bereinigt. Hierbei kommt Angaben aus dem Leistungsbezug die höchste Priorität zu.

### 5.4.3 Staatsangehörigkeit (nation)

Variablenlabel	Staatsangehörigkeit
Variablenname	nation
Kategorie	Persönliche Merkmale
Herkunft	BeH, LeH, LHG, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Das Merkmal enthält als Ausprägungen die Staatenschlüssel des Statistischen Bundesamtes.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt.

### 5.4.4 Staatsangehörigkeit vergrößert (nation\_gr)

Variablenlabel	Staatsangehörigkeit vergrößert
Variablenname	nation_gr
Kategorie	Persönliche Merkmale
Herkunft	BeH, LeH, LHG, ASU, XASU, MTH

Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Das Merkmal enthält als Ausprägungen die zu Gruppen zusammengefassten Staatenschlüssel des Statistischen Bundesamtes.

#### 5.4.5 Familienstand (famst)

Variablenlabel	Familienstand
Variablenname	famst
Kategorie	Persönliche Merkmale
Herkunft	LeH, LHG, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Das Merkmal beschreibt den Familienstand. In der Quelle LeH umfasst das Merkmal nur zwei Ausprägungen (0 - nicht verheiratet, 1 - verheiratet), während in den Quellen LHG/ASU/XASU/MTH zwischen sechs Ausprägungen unterschieden wird. Ein Abgleich der Informationen aus den Quellen wurde nicht vorgenommen.

#### 5.4.6 Kinderzahl (kind)

Variablenlabel	Kinderzahl
Variablenname	kind
Kategorie	Persönliche Merkmale
Herkunft	LeH, LHG, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Je nach Datenquelle hat das Merkmal eine unterschiedliche Bedeutung.</p> <p><b>1) LeH</b>  In der LeH gibt die Variable die Anzahl der Kinder unter 16 Jahren zum Zeitpunkt der Antragstellung an. Es wird nur unterschieden, ob die Anzahl an Kindern gleich Null oder mindestens eins ist. Folgende Ausprägungen kommen daher vor:  0       Keine Kinder  100     Ein oder mehrere Kinder  Das Merkmal wird bei Änderungen der Leistungsart oder neuen Bewilligungen nicht aktualisiert, sondern erst, wenn ein erneuter Leistungsfall auftritt. Dadurch ist die Datenqualität eingeschränkt.</p> <p><b>2) ASU, MTH</b>  Hier entspricht die Ausprägung des Merkmals der Anzahl an Kindern unter 15. Bis zum 30.6.2006 konnten nur bis zu neun Kinder erfasst werden. Der Wert 0 kommt nicht vor. Vor dem 30.6.2006 wurde der Wert Null auf missing umgeschlüsselt, da nicht klar ist, ob Null inhaltlich als "Keine Kinder" oder als "Feld nicht bearbeitet" zu interpretieren ist. Nach dem 30.6.2006 ist das Merkmal nur gefüllt, wenn Kinder vorhanden sind.</p>



	<p><b>3) XASU, LHG</b></p> <p>Hier entspricht die Ausprägung des Merkmals der Anzahl an Kindern unter 15 in der Bedarfsgemeinschaft. Bei der Quelle LHG gilt der Wert für den gesamten Originalzeitraum.</p>
--	--

#### 5.4.7 Ausbildung (ausbildung)

Variablenlabel	Ausbildung
Variablenname	ausbildung
Kategorie	Persönliche Merkmale
Herkunft	BeH, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Zu beachten ist, dass das Merkmal je nach Quelle eine unterschiedliche Bedeutung hat.</p> <p><b>1) BeH</b></p> <p>Das Merkmal enthält für BeH-Sätze die von den Arbeitgebern im Rahmen des Meldeverfahrens anzugebende Berufsausbildung. Folgende Ausprägungen kommen vor:</p> <ul style="list-style-type: none"> <li>1 ohne abgeschlossene Berufsausbildung</li> <li>2 betriebliche/außerbetriebliche Ausbildung</li> <li>11 Fachhochschulabschluss</li> <li>12 Hochschulabschluss</li> </ul> <p>Bei Meldungen nach dem neuen Tätigkeitsschlüssel (siehe Abschnitt 1.3.1) ist es nicht mehr möglich, FH-Absolventen wie bisher eindeutig zu identifizieren, da der neue Tätigkeitsschlüssel die separate Erfassung dieses Bildungsabschlusses nicht länger erlaubt. Sie werden der Kategorie 12 zugeordnet.</p> <p>„Zusammen mit einem Betriebswechsel treten häufig auch Änderungen im Ausbildungsstatus auf. Das liegt daran, dass beim neuen Betrieb die Meldedaten neu zusammengestellt werden. Wenn ein Beschäftigter z.B. durch berufsbegleitende Weiterbildung einen höheren Abschluss erlangt hat, dann wird dieser Statuswechsel wahrscheinlich erst von einem neuen Beschäftigungsbetrieb erfasst. Es ist in der Regel davon auszugehen, dass bei längeren Beschäftigungszeiten innerhalb eines Betriebs die einmal erhobenen personenbezogenen Daten fortgeschrieben werden.“ (Meinken/Koch, 2004, S. 63)</p> <p>Eine Methode, fehlende Werte oder inkonsistente zeitliche Verläufe der Bildungsangaben in der Vorgängerstichprobe IABS zu bereinigen, findet sich in Fitzenberger et al. (2006), sowie in Drews (2006). Hierfür wird aber nur die Quelle BeH herangezogen, die in der IABS die einzige Datenquelle mit Informationen zur Bildung war.</p> <p><b>2) ASU, XASU, MTH</b></p>

	<p>Für diese Sätze ist die letzte abgeschlossene Berufsausbildung enthalten. Folgende Ausprägungen kommen vor:</p> <ol style="list-style-type: none"> <li>1 ohne abgeschlossene Berufsausbildung</li> <li>2 betriebliche/außerbetriebliche Ausbildung</li> <li>3 Berufsfachschule</li> <li>4 Fachschule</li> <li>5 Fachhochschule</li> <li>6 Universität</li> <li>7 in Deutschland nicht anerkannte Berufsausbildung</li> <li>8 in Deutschland nicht anerkannter Hochschulabschluss</li> </ol> <p>Wobei die Ausprägungen 7 und 8 nur für die Quelle MTH gelten.</p>
Qualitätshinweis	<p>Die Zahl der fehlenden Werte steigt im Zeitverlauf kontinuierlich an und erreicht am rechten Datenrand über 40%. Fehlende Werte treten besonders häufig bei folgenden Personengruppen auf: geringfügig beschäftigte Personen, Personen in Teilzeit, ausländische Beschäftigte und ostdeutsche Arbeitnehmer. Hintergrund ist, dass das Merkmal versicherungsrechtlich keine besondere Bedeutung hat (siehe Meinken/Koch, 2004, S. 63).</p> <p>Durch die Umstellung von coArb auf VerBIS kann in den Quellen ASU und MTH zwischen 2006 und 2008 nicht korrekt zwischen „ohne abgeschlossene Berufsausbildung“ und „Keine Angabe“ unterschieden werden. Ein fehlender Wert in diesem Zeitraum besagt also nicht unbedingt, dass keine (Information zur) Ausbildung vorliegt, sondern kann auch bedeuten, dass das einschlägige Generierungsverfahren nicht angewandt werden konnte. Der Füllgrad in der Quelle XASU ist generell niedrig.</p>

#### 5.4.8 Schulausbildung (schule)

Variablenlabel	Schulausbildung
Variablenname	schule
Kategorie	Persönliche Merkmale
Herkunft	BeH, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Das Merkmal enthält den Schulabschluss. Je nach Quelle sind unterschiedliche Ausprägungen möglich.</p> <p><b>1) BeH</b> Mit der Umstellung des Tätigkeitsschlüssels ändern sich die möglichen Ausprägungen (siehe Abschnitt 1.3.1). Davor gelten:</p> <ol style="list-style-type: none"> <li>5 Volks-/Hauptschule o./m. Abschluss, Mittlere Reife oder gleichwertige Schulbildung</li> <li>8 Fachhochschulreife oder Abitur/Hochschulreife</li> <li>9 Abitur/Hochschulreife</li> </ol>

	<p>Danach gelten:</p> <ul style="list-style-type: none"> <li>1 Kein Hauptschulabschluss</li> <li>4 Hauptschulabschluss/Volksschulabschluss</li> <li>6 Mittlere Reife</li> <li>8 Fachhochschulreife oder Abitur/Hochschulreife</li> </ul> <p><b>2) ASU, XASU, MTH</b> Für diese Quellen sind die folgenden Ausprägungen möglich:</p> <ul style="list-style-type: none"> <li>1 Kein Hauptschulabschluss</li> <li>4 Hauptschulabschluss/Volksschulabschluss</li> <li>6 Mittlere Reife</li> <li>7 Fachhochschulreife</li> <li>9 Abitur/Hochschulreife</li> </ul> <p>Sie sind zu Beginn der Arbeitssuche/Maßnahmeteilnahme gültig. In der Quelle XASU kann das Merkmal bei Ausbildungsplatzsuchenden auch den angestrebten Abschluss enthalten.</p>
Qualitätshinweis	Der Füllgrad nimmt in der BeH über die Jahre kontinuierlich ab und liegt in jüngeren Jahren unter 2/3. In der XASU nimmt er kontinuierlich zu und liegt seit 2012 über 2/3. In ASU und MTH ist der Füllgrad insgesamt gut.

## 5.5 Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche

### 5.5.1 Tagesentgelt/ täglicher Leistungssatz (tentgelt)

Variablenlabel	Tagesentgelt/täglicher Leistungssatz
Variablenname	tentgelt
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH, LeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p><b>1) BeH</b> In BeH-Sätzen weist das Merkmal jeweils das Bruttotagesentgelt eines Beschäftigten aus. Es wird berechnet aus dem vom Arbeitgeber gemeldeten Zeitraumentgelt sowie der Dauer des Zeitraums der (ungesplitteten) Originalmeldung in Kalendertagen. Das Tagesentgelt ist in Euro angegeben.</p>

Bis 1998 meldeten Arbeitgeber grundsätzlich nur das sozialversicherungspflichtige Arbeitsentgelt. Entgelte, die unterhalb der Geringfügigkeitsgrenze lagen, wurden nicht gemeldet. Entgelte, welche die Beitragsbemessungsgrenze in der gesetzlichen Rentenversicherung überschritten, werden nur bis zu dieser Grenze gemeldet. In der gesetzlichen Rentenversicherung gibt es zwei Beitragsbemessungsgrenzen. Die Grenze der Knappschaftlichen Rentenversicherung liegt meist höher als die Grenze der Arbeiter- und Angestelltenrentenversicherung. Jedoch ist eine Unterscheidung der beiden Versicherungsträger in den Daten nicht möglich.

Seit der Einbeziehung der geringfügig Beschäftigten in das Meldeverfahren zum 1.4.1999 werden auch Entgelte unterhalb der Geringfügigkeitsgrenze erfasst. Die Beitragsbemessungsgrenze als obere Kappungsgrenze gilt jedoch weiterhin. In einigen Fällen übersteigen die gemeldeten Entgelte die Beitragsbemessungsgrenze allerdings trotzdem. Dies dürfte in der Regel auf Jahressonderzahlungen zurückzuführen sein, die der Arbeitgeber bei den Jahres-, Unterbrechungs- oder Abmeldungen zum normalen Entgelt hinzurechnen kann. Dabei ist es dann unerheblich, ob durch diese Addition die für den Meldezeitraum maßgebliche Beitragsbemessungsgrenze in der Rentenversicherung überschritten wird. Es könnte sich allerdings auch um fehlerhafte Angaben bei der Beschäftigungszeit handeln. (Die Entgelt-Informationen hingegen können auf Grund ihrer Versicherungsrelevanz als weniger fehleranfällig gesehen werden.)

Die Geringfügigkeits- und Beitragsbemessungsgrenzen unterscheiden sich von Jahr zu Jahr und in West- und Ostdeutschland (maßgeblich ist jeweils der Standort des Betriebs). Eine Übersicht über diese Grenzen findet sich unter <http://fdz.iab.de>.

Eine Angabe von 0 Euro ist auf "Unterbrechungsmeldungen" zurückzuführen. In diesen Zeiten besteht das Beschäftigungsverhältnis zwar rechtlich fort, es wird aber kein Entgelt gezahlt. Dies ist der Fall bei Krankheit nach Ende der Lohnfortzahlung, Mutterschaftszeiten und Sabbaticals.

Das Tagesentgelt wird mit zwei Nachkommastellen ausgewiesen. Alle Werte größer als 0 und kleiner als 0,01 wurden auf 0,01 aufgerundet. Somit können die o.a. Unterbrechungsmeldungen mit der Bedingung Tagesentgelt = 0 identifiziert werden.

## 2) LeH

Bei LeH-Datensätzen enthält das Merkmal den täglichen Leistungssatz, jeweils umgerechnet in Euro. Es ist zu berücksichtigen, dass der tägliche Leistungssatz bei Datensätzen mit einem Original-Beginndatum, das vor dem 1.1.1998 liegt, für Werktage und bei Datensätzen mit einem Original-Beginndatum ab 1.1.1998 für Kalendertage gilt.

Die Angabe von 0 Euro ist seit 1.1.2005 auf Sperrzeiten oder Leistungsunterbrechungen zurückzuführen. Ist für einen Satz mit Leistungssatz gleich 0 ein Beendigungsgrund angegeben, so handelt es sich um eine Unterbrechungsmeldung. Bei Sätzen, die eine Sperrzeit widerspiegeln, ist der Anspruch derselbe wie vor Beginn der Sperrzeit.

### 5.5.2 Beruf – ausgeübte/letzte Tätigkeit (KIdB 1988) (beruf)

Variablenlabel	Beruf – ausgeübte/letzte Tätigkeit (KIdB 1988)
Variablenname	beruf
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p><b>1) BeH</b> Die Berufsbezeichnung der im Meldezeitraum vom Beschäftigten ausgeübten Tätigkeit ist Bestandteil der vom Arbeitgeber übermittelten „Angaben zur Tätigkeit“. Treffen mehrere Berufsbezeichnungen mit verschiedenen Schlüsselzahlen für einen Beschäftigten zu, ist der Arbeitgeber gehalten, diejenige Berufsbezeichnung zu wählen, welche die überwiegend ausgeübte Tätigkeit kennzeichnet (siehe Bundesagentur für Arbeit, 2005, S. V). Der Arbeitgeber verschlüsselt die Tätigkeit der/des Beschäftigten hierbei gemäß der „Klassifizierung der Berufe 1988“ (Bundesanstalt für Arbeit, 1988), in der ca. 25.000 Berufsbezeichnungen zu finden sind. Die Berufsordnung besteht aus einer 3-stelligen Kennziffer und umfasst rund 330 Ausprägungen. Beschäftigungsmeldungen mit einem Enddatum nach dem 30.11.2011 werden nach dem neuen Tätigkeitsschlüssel 2010 (KIdB2010) gemeldet (siehe Abschnitt 1.3.1). Diese Werte werden durch einen Schwerpunktumstieg auf die KIdB1988 zurückgeschlüsselt. Dabei kommt es zu Unschärfen.</p> <p><b>2) ASU, XASU, MTH</b> Das Merkmal enthält den Beruf der letzten Tätigkeit. Zum Schlüssel siehe 1).</p>
Qualitätshinweis	<p>In 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben. In der XASU ist das Merkmal Beruf nahezu für den gesamten zur Verfügung stehenden Zeitraum nicht gefüllt.</p>

### 5.5.3 Berufsgruppe – ausgeübte/letzte Tätigkeit (KIdB 2010) (beruf2010\_3)

Variablenlabel	Berufsgruppe – ausgeübte/letzte Tätigkeit (KIdB 2010), 3-Steller
Variablenname	beruf2010_3
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH, LeH, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p><b>1) BeH</b></p>

	<p>Die Berufsbezeichnung der im Meldezeitraum vom Beschäftigten ausgeübten Tätigkeit ist Bestandteil der vom Arbeitgeber übermittelten „Angaben zur Tätigkeit“. Treffen mehrere Berufsbezeichnungen mit verschiedenen Schlüsselzahlen für einen Beschäftigten zu, ist der Arbeitgeber gehalten, diejenige Berufsbezeichnung zu wählen, welche die überwiegend ausgeübte Tätigkeit kennzeichnet (siehe Bundesagentur für Arbeit, 2005, S. V).</p> <p>Der Arbeitgeber verschlüsselt die Tätigkeit der/des Beschäftigten hierbei gemäß der „Klassifikation der Berufe 2010“ (Bundesagentur für Arbeit, 2011). Die Berufsgattung besteht aus einer 5-stelligen Kennziffer und umfasst rund 1.300 Ausprägungen. Die weniger detaillierte Berufsgruppe wird durch die ersten drei Stellen der Kennziffer erfasst. Das in der KIdB2010 an der fünften Stelle erfasste Anforderungsniveau einer Tätigkeit wird gesondert in der Variable Anforderungsniveau (niveau) zur Verfügung gestellt.</p> <p>Beschäftigungsmeldungen mit einem Enddatum vor dem 30.11.2011 werden nach dem alten Tätigkeitsschlüssel 1988 (KIdB 1988) gemeldet (siehe Abschnitt 1.3.1). Diese Werte werden durch einen Schwerpunktumstieg auf die KIdB2010 umgeschlüsselt. Da der neue Tätigkeitsschlüssel wesentlich detaillierter als der alte ist, kommt es dabei zu erheblichen Unschärfen. Dies ist bei Analysen zu beachten.</p> <p><b>2) LeH, ASU, XASU, MTH</b> Das Merkmal enthält den Beruf der letzten Tätigkeit. Zum Schlüssel siehe 1).</p>
Qualitätshinweis	In 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben.

#### 5.5.4 Berufsuntergruppe – ausgeübte/letzte Tätigkeit (KIdB 2010) (beruf2010\_4)

Variablenlabel	Berufsuntergruppe – ausgeübte/letzte Tätigkeit (KIdB 2010), 4-Steller
Variablenname	beruf2010_4
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH, LeH, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p><b>1) BeH</b> Die Berufsbezeichnung der im Meldezeitraum vom Beschäftigten ausgeübten Tätigkeit ist Bestandteil der vom Arbeitgeber übermittelten „Angaben zur Tätigkeit“. Treffen mehrere Berufsbezeichnungen mit verschiedenen Schlüsselzahlen für einen Beschäftigten zu, ist der Arbeitgeber gehalten, diejenige Berufsbezeichnung zu wählen, welche die überwiegend ausgeübte Tätigkeit kennzeichnet (siehe Bundesagentur für Arbeit, 2005, S. V).</p>

	<p>Der Arbeitgeber verschlüsselt die Tätigkeit der/des Beschäftigten hierbei gemäß der „Klassifikation der Berufe 2010“ (Bundesagentur für Arbeit, 2011). Die Berufsgattung besteht aus einer 5-stelligen Kennziffer und umfasst rund 1.300 Ausprägungen. Die weniger detaillierte Berufsuntergruppe wird durch die ersten vier Stellen der Kennziffer erfasst. Das in der KIdB2010 an der fünften Stelle erfasste Anforderungsniveau einer Tätigkeit wird gesondert in der Variable Anforderungsniveau (niveau) zur Verfügung gestellt.</p> <p>Beschäftigungsmeldungen mit einem Enddatum vor dem 30.11.2011 werden nach dem alten Tätigkeitsschlüssel 1988 (KIdB 1988) gemeldet (siehe Abschnitt 1.3.1). Diese Werte werden durch einen Schwerpunktumstieg auf die KIdB2010 umgeschlüsselt. Da der neue Tätigkeitsschlüssel wesentlich detaillierter als der alte ist, kommt es dabei zu erheblichen Unschärfen. Dies ist bei Analysen zu beachten.</p> <p><b>2) LeH, ASU, XASU, MTH</b></p> <p>Das Merkmal enthält den Beruf der letzten Tätigkeit. Zum Schlüssel siehe 1).</p>
Qualitätshinweis	In 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt. Ansonsten wird als Berufsangabe gemäß KIdB2010 lediglich die Berufsgruppe (beruf2010_3) bereitgestellt.

### 5.5.5 Anforderungsniveau – ausgeübte/letzte Tätigkeit (KIdB 2010) (niveau)

Variablenlabel	Anforderungsniveau - ausgeübte/letzte Tätigkeit (KIdB 2010)
Variablenname	niveau
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH, LeH, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Keine

Detailbeschreibung	<p><b>1) BeH</b> Die Berufsbezeichnung der im Meldezeitraum vom Beschäftigten ausgeübten Tätigkeit ist Bestandteil der vom Arbeitgeber übermittelten „Angaben zur Tätigkeit“. Treffen mehrere Berufsbezeichnungen mit verschiedenen Schlüsselzahlen für einen Beschäftigten zu, ist der Arbeitgeber gehalten, diejenige Berufsbezeichnung zu wählen, welche die überwiegend ausgeübte Tätigkeit kennzeichnet (siehe Bundesagentur für Arbeit, 2005, S. V).</p> <p>Der Arbeitgeber verschlüsselt die Tätigkeit der/des Beschäftigten hierbei gemäß der „Klassifikation der Berufe 2010“ (Bundesagentur für Arbeit, 2011). Die Berufsgattung besteht aus einer 5-stelligen Kennziffer und umfasst rund 1.300 Ausprägungen. Das an der fünften Stelle erfasste Anforderungsniveau einer Tätigkeit wird gesondert in der Variable Anforderungsniveau (niveau) zur Verfügung gestellt.</p> <p>Beschäftigungsmeldungen mit einem Enddatum vor dem 30.11.2011 werden nach dem alten Tätigkeitsschlüssel 1988 (KldB 1988) gemeldet (siehe Abschnitt 1.3.1). Diese Werte werden durch einen Schwerpunktumstieg auf die KldB2010 umgeschlüsselt. Da der neue Tätigkeitsschlüssel wesentlich detaillierter als der alte ist, kommt es dabei zu erheblichen Unschärfen. Dies ist bei Analysen zu beachten.</p> <p><b>2) LeH, ASU, XASU, MTH</b> Das Merkmal enthält den Beruf der letzten Tätigkeit. Zum Schlüssel siehe 1).</p>
Qualitätshinweis	In 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben.

### 5.5.6 Teilzeit (teilzeit)

Variablenlabel	Teilzeit
Variablenname	teilzeit
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Die berufliche Stellung eines Beschäftigten im Meldezeitraum wird vom Arbeitgeber im Rahmen der „Angaben zur Tätigkeit“ übermittelt. Das Merkmal „Berufsstellung“ unterscheidet zwischen Vollzeit- und Teilzeitbeschäftigten. Maßgeblich ist dabei das Verhältnis zwischen der arbeitsvertraglich vereinbarten und der betriebsüblichen Arbeitszeit.
Qualitätshinweis	In 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben. Um dieses Problem abzumildern, wurde am IAB eine Imputation der Arbeitszeit für den fraglichen Zeitraum vorgenommen. Nähere Informationen zum Verfahren können Ludsteck/Thomsen (2016) entnommen werden.

### 5.5.7 Erwerbsstatus (erwstat)

Variablenlabel	Erwerbsstatus
Variablenname	Erwstat



Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH, LeH, LHG, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Das Merkmal nimmt für jede Datenquelle unterschiedliche Werte mit unterschiedlichen Bedeutungen an.</p> <p><b>1) BeH</b>  Für BeH-Sätze entspricht das Merkmal Erwerbsstatus der im neuen Meldeverfahren (DEÜV) ab dem 1.1.1999 erfassten Personengruppe. Sie weist beitrags- oder leistungsrechtliche Besonderheiten des Beschäftigungsverhältnisses aus.  Treffen für eine Beschäftigungsmeldung mehrere Schlüssel zu, ist vom meldenden Arbeitgeber der kleinste anzugeben. Mehrheitlich handelt sich dabei um sozialversicherungspflichtige Beschäftigungsverhältnisse ohne Besonderheiten, die unter Schlüsselnummer 101 geführt werden. Entsprechend ist es möglich, dass diese Beschäftigungen leicht überschätzt werden.  Im Meldeverfahren ist geregelt, dass Wechsel im Erwerbsstatus - z.B. die Übernahme eines Auszubildenden nach dem Ende der Berufsausbildung im Ausbildungsbetrieb - durch eine erneute Meldung erfolgen müssen (siehe Deutsche BKK, 2016, S. 35).  Die Personengruppe kann in Beschäftigungsmeldungen, die sich auf die Jahre vor 1999 beziehen, aber erst 1999 oder später eingegangen sind, enthalten sein. Für Meldungen, die vor 1999 eingegangen sind, wird versucht, anhand bestimmter Regeln mit Hilfe der Merkmale „Ausbildung“, „Stellung im Beruf und Arbeitszeit“ und „Beruf“ sowie weiterer Informationen eine Zuordnung zu den Personengruppen durchzuführen. In vielen Fällen ist das jedoch nicht eindeutig möglich.  Seit dem 1.4.1999 werden im Meldeverfahren DEÜV auch geringfügig Beschäftigte erfasst. Diese Personengruppe lässt sich über die Ausprägungen 109 und 209 abgrenzen. Für geringfügig Beschäftigte konnten keine Daten vor dem Zeitpunkt der Einführung der Meldepflicht im Jahr 1999 zurückgeschrieben werden.</p> <p><b>2) LeH</b>  Für LeH-Sätze ist im Merkmal Erwerbsstatus die gruppierte Leistungsart angegeben. Es kann also unterschieden werden, ob eine Person Arbeitslosengeld, Arbeitslosenhilfe oder Unterhaltsgeld bezieht oder ob Beiträge zur privaten Pflegeversicherung von der BA übernommen werden.</p> <p><b>3) ASU/XASU</b>  Für ASU-Sätze zeigt das Merkmal Erwerbsstatus den Status der Arbeitsuche an. Als Ratsuchende werden Empfänger von Arbeitslosengeld (bzw. Arbeitslosengeld II), die Leistungen unter den erleichterten Voraussetzungen des § 428 SGB III (bzw. des § 65 Abs. 4 SGB II) beziehen, sowie Nichtleistungsempfänger ohne Vermittlungsbereitschaft i. S. des § 252 Abs. 8 SGB VI, wenn das 58. Lebensjahr vollendet wurde, geführt.</p>

	<p>Der Status "ohne Status" (Statistik: „nicht gesetzt“) impliziert hauptsächlich Personen, denen nach §10 SGB II eine Aktivierung bzw. Vermittlung nicht zuzumuten ist. Ebenso werden Personen mit einer Arbeitsunfähigkeit von mehr als 42 Tagen, die weiterhin ALG II beziehen, unter diesem Status im System geführt.</p> <p>In XASU-Sätzen enthält das Merkmal Erwerbsstatus bislang die Ausprägungen „nicht arbeitslos, aber arbeitsuchend“ sowie „arbeitsuchend und gleichzeitig arbeitslos“.</p> <p>Da die Meldelogik in vielen Fällen eine Reidentifikation des exakten Geburtsdatums ermöglichen würde, wurden die ursprünglichen Datumsangaben durch das in Abschnitt 3.2.1.6 beschriebene Anonymisierungsverfahren verändert.</p> <p><b>4) LHG</b></p> <p>Für LHG-Sätze zeigt das Merkmal Erwerbsstatus an, ob die Person minderjährig erwerbsfähig, volljährig erwerbsfähig oder nicht erwerbsfähig ab Altersrentengrenze geführt wird.</p> <p>Da die Meldelogik in vielen Fällen eine Reidentifikation des exakten Geburtsdatums ermöglichen würde, wurden die ursprünglichen Datumsangaben durch das in Abschnitt 3.2.1.6 beschriebene Anonymisierungsverfahren verändert.</p> <p><b>5) MTH</b></p> <p>Für MTH-Sätze zeigt das Merkmal Erwerbsstatus die Maßnahmeart-Kategorie an. Dies ist die höchste Hierarchiestufe der Maßnahmeart-Systematiken der Bundesagentur für Arbeit.</p>
Qualitätshinweis	<p>In der LHG ist zu beobachten, dass es überdurchschnittlich viele 15-jährige und in geringerem Maß 16- und 17-jährige mit Hinweis auf fehlende Erwerbsfähigkeit gibt. 15- bzw. 16-jährige Hilfebedürftige der zugelassenen kommunalen Träger sind deshalb möglicherweise unterrepräsentiert, da „Nichterwerbsfähige“ nicht in die IEB übernommen werden.</p>

### 5.5.8 Gleitzone (gleitz)

Variablenlabel	Gleitzone
Variablenname	gleitz
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Keine

Detailbeschreibung	Das Merkmal ist erst ab 2003 und nur für BeH-Datensätze gefüllt. Es zeigt an, ob es sich bei der Beschäftigungsmeldung um eine Beschäftigung im Niedriglohnbereich innerhalb der sogenannten Gleitzone mit einem monatlichen Bruttoarbeitsentgelt von 400,01 bis 800,00 € handelt (sog. Midi-Job), für die der Arbeitnehmer nur einen reduzierten Beitrag zum Gesamtsozialversicherungsbeitrag leisten muss. Da Arbeitnehmer mit einem Entgelt in der Gleitzone freiwillig den „normalen“ Sozialversicherungsbeitrag zahlen können, sind nicht automatisch alle Beschäftigungen mit einem entsprechen Entgelt als Gleitzone gekennzeichnet. Die entsprechende gesetzliche Regelung besteht seit dem 1.4.2003.
--------------------	---

### 5.5.9 Leiharbeit (leih)

Variablenlabel	Leiharbeit
Variablenname	leih
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Das Merkmal gibt an, ob es sich bei der Beschäftigung der Person um ein Leiharbeitsverhältnis handelt. Das Merkmal entstammt dem Tätigkeitsschlüssel 2010 und ist nur für Meldungen mit einem Enddatum nach dem 30.11.2011 gefüllt.
Qualitätshinweis	In 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben.

### 5.5.10 Befristung (befrist)

Variablenlabel	Befristung
Variablenname	befrist
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	keine
Detailbeschreibung	Das Merkmal gibt an, ob es sich bei der Beschäftigung der Person um ein befristetes oder unbefristetes Arbeitsverhältnis handelt. Das Merkmal entstammt dem Tätigkeitsschlüssel 2010 und ist nur für Meldungen mit einem Enddatum nach dem 30.11.2011 gefüllt.
Qualitätshinweis	In 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben.

### 5.5.11 Abmeldegrund / Abgabegrund / Beendigungsgrund (grund)

Variablenlabel	Abmeldegrund / Abgabegrund / Beendigungsgrund
Variablenname	Grund
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH, LeH, LHG, ASU, XASU
Datentyp	Numerisch

Hierarchie	Keine																								
Detailbeschreibung	<p><b>1) BeH</b>  In BeH-Meldungen kennzeichnet der Abgabegrund den Anlass, aus dem der Arbeitgeber den Sozialversicherungsträgern die betreffende Beschäftigungsmeldung übermittelt hat. Allerdings werden in die IEB nicht alle im Rahmen des Meldeverfahrens mögliche Abgabegründe integriert. So beinhalten die IEB nur Meldungen mit Entgeltangaben (also Jahres-, Unterbrechungs- und Abmeldungen), während Anmeldungen wegen fehlender Entgeltinformationen nicht enthalten sind. Ein Informationsverlust ist damit jedoch nicht verbunden, da die Angaben aus einer Anmeldung mit der darauf folgenden Jahres-, Unterbrechungs- oder Abmeldung erneut übermittelt werden.  Die Abgabegründe für Beschäftigungsmeldungen sind nach den Regeln des seit 1.1.1999 geltenden Meldeverfahrens (gemäß DEÜV) kodiert.</p> <p><b>2) LeH</b>  In LeH-Meldungen gibt das Merkmal den Grund für das Ende des Bezugs von Arbeitslosengeld, -hilfe oder Unterhaltsgeld an. Über die Gründe für den Beginn des Leistungsbezugs finden sich in der LeH keine Informationen, da die LeH mit den Meldungen der Arbeitsagenturen an die Krankenkassen über abgeschlossene Leistungsbezugsdauern gefüllt wird.</p> <p><b>3) LHG</b>  Bei LHG-Datensätzen ist der SGB-II-Einstellungsgrund enthalten und gibt den Grund für die Beendigung von laufenden Leistungen an. Der SGB-II-Beendigungsgrund bezieht sich auf die Person, nicht aber auf die Bedarfsgemeinschaft. Wird bei einem anderen als dem betrachteten Mitglied der Bedarfsgemeinschaft der SGB-II-Leistungsbezug beendet, werden neue Sätze für alle Mitglieder in einer Bedarfsgemeinschaft zu diesem Datum begonnen, der SGB-II-Beendigungs-/Einstellungsgrund ist jedoch nur für die betroffene Person gefüllt. Das Merkmal ist genau zum Ende des Originalsatzes gültig.</p> <p><b>4) ASU</b>  Bei ASU-Sätzen ist der Abmelde- bzw. Abgangsgrund enthalten. Im Fall eines Rechtskreiswechsels wird der Satz künstlich gesplittet und als Abmeldegrund "DS durch Splittung erzeugt" eingetragen. Zur korrekten Abbildung der Abgangsgründe ist außerdem das Merkmal „Status nach Arbeitssuche“ zu beachten.  Ab dem 26.4.2003 erfolgte eine Reduktion der Merkmalsausprägungen. Für Auswertungen über längere Zeiträume können die alten Werte auf die aktuell gültigen anhand folgender Tabelle umgeschlüsselt werden:</p> <table border="0" data-bbox="496 1787 1171 1973"> <thead> <tr> <th><b>alt-&gt;neu</b></th> <th><b>alt-&gt;neu</b></th> <th><b>alt-&gt;neu</b></th> <th><b>alt-&gt;neu</b></th> </tr> </thead> <tbody> <tr> <td>29-&gt;60</td> <td>36-&gt;61</td> <td>44-&gt;74</td> <td>51-&gt;74</td> </tr> <tr> <td>30-&gt;60</td> <td>37-&gt;66</td> <td>45-&gt;77</td> <td>52-&gt;76</td> </tr> <tr> <td>31-&gt;61</td> <td>38-&gt;66</td> <td>46-&gt;67</td> <td>53-&gt;68</td> </tr> <tr> <td>32-&gt;60</td> <td>39-&gt;71</td> <td>47-&gt;67</td> <td>54-&gt;78</td> </tr> <tr> <td>33-&gt;60</td> <td>40-&gt;69</td> <td>48-&gt;78</td> <td></td> </tr> </tbody> </table>	<b>alt-&gt;neu</b>	<b>alt-&gt;neu</b>	<b>alt-&gt;neu</b>	<b>alt-&gt;neu</b>	29->60	36->61	44->74	51->74	30->60	37->66	45->77	52->76	31->61	38->66	46->67	53->68	32->60	39->71	47->67	54->78	33->60	40->69	48->78	
<b>alt-&gt;neu</b>	<b>alt-&gt;neu</b>	<b>alt-&gt;neu</b>	<b>alt-&gt;neu</b>																						
29->60	36->61	44->74	51->74																						
30->60	37->66	45->77	52->76																						
31->61	38->66	46->67	53->68																						
32->60	39->71	47->67	54->78																						
33->60	40->69	48->78																							

	34->60      42->65      49->69 35->60      43->70      50->75  <b>5) XASU</b> Bei XASU-Sätzen ist der Abmelde- bzw. Abgangsgrund enthalten. Im Fall eines Rechtskreiswechsels wird der Satz künstlich gesplittet und als Abmeldegrund "DS durch Splitting erzeugt" eingetragen.
Qualitätshinweis	Der Anteil valider Werte (Füllgrad) des Abgabegrundes ist in der Quelle LHG über alle Jahre sehr niedrig (< 20%). Auch in der XASU ist der Füllgrad gering, jedoch verbessert er sich im Zeitablauf.

### 5.5.12 Erwerbsstatus vor Arbeitsuche (estatvor)

Variablenlabel	Erwerbsstatus vor Arbeitsuche
Variablenname	estatvor
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	ASU, XASU
Datentyp	Numerisch
Hierarchie	keine
Detailbeschreibung	Die Variable zeigt für ASU- und XASU-Sätze den Erwerbsstatus vor der Arbeitsuche an. Ab 12/2002 erfolgte eine deutliche Reduzierung der Ausprägungen. Die Werte älterer Sätze wurden auf die aktuell gültigen Werte umgeschlüsselt, die somit für den gesamten Zeitraum gültig sind. Wird ein Datensatz innerhalb der ASU/XASU künstlich gesplittet (etwa bei einem Rechtskreiswechsel zwischen SGB III und SGB II), wird der Zugangsgrund „DS durch Splitting erzeugt“ vergeben. Dies gilt jedoch nicht für das Episodensplitting innerhalb der PASS-ADIAB. Die Angabe bezieht sich daher nicht notwendigerweise auf das Beginndatum der Episode, sondern auf den Beginn des Originalzeitraums.
Qualitätshinweis	Der Anteil valider Werte (Füllgrad) steigt für XASU-Sätze zwischen 2005 und 2013 von ca. 8 % auf ca. 78 %.

### 5.5.13 Status nach Arbeitsuche (estatnach)

Variablenlabel	Status nach Arbeitsuche
Variablenname	estatnach
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	ASU
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Das Merkmal enthält den Status der Person nach Abgang aus der Arbeitslosigkeit. Längere Krankheitsepisoden können über dieses Merkmal identifiziert werden. Die Ausprägungen sind folgendermaßen unterteilt: <ul style="list-style-type: none"> <li>• 1000er: Maßnahme (geförderte Beschäftigung)</li> <li>• 2000er: ungeförderte Beschäftigung</li> <li>• 3000er: in Ausbildung etc.</li> </ul>

	<ul style="list-style-type: none"> <li>• 4000er: in Selbständigkeit</li> <li>• 5000er: in Sperre</li> <li>• 6000er: in Sonstige</li> </ul>
--	--

#### 5.5.14 Profillage (profil)

Variablenlabel	Profillage
Variablenname	profil
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	ASU, MTH
Datentyp	numerisch
Hierarchie	keine
Detailbeschreibung	<p>Das Merkmal gibt die dem Kunden beim Profiling zugewiesene Profillage wieder. Das Profiling dient der Erstellung eines Kundenprofils, d. h. einer Auflistung der arbeitsmarktrelevanten Fähigkeiten, Erfahrungen und Interessen, so dass die Verortung des Kunden auf dem Arbeitsmarkt erleichtert wird. Gegen Ende des Profilings wird zusammenfassend eine Profillage festgestellt. Hierzu wird zunächst die allgemeine Integrationsprognose des Kunden bestimmt. Zur Auswahl stehen:</p> <ul style="list-style-type: none"> <li>• integrationsnah (Integration in den 1. Arbeitsmarkt innerhalb von 12 Monaten realistisch)</li> <li>• komplex (Integration in den 1. Arbeitsmarkt innerhalb von 12 Monaten nicht realistisch)</li> <li>• sonstiges</li> </ul> <p>Von der Bestimmung der Integrationsprognose ist die Zuweisung der Profillage abhängig. Den Kunden mit integrationsnaher Prognose können die Profillagen "Marktprofil", "Aktivierungsprofil" und "Förderprofil" zugewiesen werden, während die Kunden mit komplexer Prognose die Profillagen "Entwicklungsprofil", "Stabilisierungsprofil" oder "Unterstützungsprofil" erhalten. Die Auswahl der konkreten Profillage richtet sich nach dem vom Vermittler angenommenen Handlungsbedarf. Wenn eine "sonstige" Prognose vorliegt, kann bei der Profillage "Zuordnung nicht erforderlich" oder - nur für SGB-II-Kunden - "Integriert, aber hilfebedürftig" gewählt werden.</p>
Qualitätshinweis	Das Merkmal wurde 2009 eingeführt, aber bis 2006 aus anderen Merkmalen zurückgemappt. Die Qualität wird von der BA-Statistik seit 2010 als verlässlich eingestuft.

#### 5.5.15 Art der Kündigung der letzten Tätigkeit (art\_kuend)

Variablenlabel	Art der Kündigung der letzten Tätigkeit
Variablenname	art_kuend
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	ASU, MTH
Datentyp	Numerisch
Hierarchie	keine

Detailbeschreibung	Dieses Merkmal beschreibt die Art der Beendigung des letzten Arbeits- oder Ausbildungsverhältnisses vor einer erneuten Arbeitsuche. Es kann daher zur Identifikation von Job-to-Job-Vermittlungen verwendet werden.
--------------------	---

#### 5.5.16 Arbeitszeit des Stellengesuchs (arbeitszeit)

Variablenlabel	Arbeitszeit des Stellengesuchs
Variablenname	arbeitszeit
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	ASU, MTH
Datentyp	Numerisch
Hierarchie	keine
Detailbeschreibung	Arbeitsuchende geben beim Vermittlungsvorgang an, welchem zeitlichen Umfang die gesuchte Stelle entsprechen sollte.

#### 5.5.17 Restanspruch/geplante Dauer (restanspruch)

Variablenlabel	Restanspruch/geplante Dauer
Variablenname	restanspruch
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	LeH, MTH
Datentyp	Numerisch
Hierarchie	keine
Detailbeschreibung	<p>Je nach Datenquelle hat das Merkmal eine unterschiedliche Bedeutung.</p> <p><b>1) LeH</b>  Das Merkmal enthält den Anspruch, der nach Ende des aktuellen Leistungsbezugs übrig bleibt. Endet der Leistungsbezug vor dem Ende der maximalen Anspruchsdauer (z. B. durch Arbeitsaufnahme), bleibt ein Restanspruch mit der Dauer des nicht ausgeschöpften Anspruchs bestehen. Entsteht innerhalb von fünf Jahren ein neuer Anspruch, wird die Dauer des Restanspruchs zur neuen Anspruchsdauer addiert. Allerdings ist hierbei die Höchstdauer für das jeweilige Lebensalter die Obergrenze. Entsteht kein neuer Anspruch, so kann der Restanspruch auf Antrag innerhalb von vier Jahren für Leistungen verwendet werden. Liegt das Enddatum des Leistungsempfangs vor dem 1.1.1998, wird der Restanspruch in Werktagen angegeben. Danach erfolgt die Angabe in Kalendertagen.  Die Angabe bezieht sich nicht auf das Enddatum der Episode, sondern auf das Ende des Originalzeitraums.</p> <p><b>2) MTH</b>  Das Merkmal enthält die geplante Dauer der Maßnahme.</p>

#### 5.5.18 Trägerart (traeger)

Variablenlabel	Trägerart
Variablenname	traeger

Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	LHG, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Das Merkmal enthält die Trägerschaftsart des Trägers der Grundsicherung (LHG), des Trägers der Maßnahme (MTH) oder des Trägers der das Bewerberprofil verwaltet (ASU/XASU). Die Dimension enthält neben der SGB-III-Trägerschaftsart "Bundesagentur für Arbeit (BA)" drei Trägerschaftsarten des SGB II. Für weitere Informationen siehe Abschnitt 2.2.3. Die Trägerschaftsart von Grundsicherungsträgern kann im Zeitverlauf wechseln (z.B. von ARGE zu gAw oder gE zu zKT). Zu beachten ist außerdem, dass die Grenzen der Träger in Bezug auf Kreise nicht immer trennscharf sind und somit nicht mit den Grenzen der Kreise übereinstimmen müssen. Gleiches gilt für die Arbeitsagenturen.</p> <p>Bei ASU- und XASU-Sätzen ist die Art des Trägers enthalten, der das Bewerberangebot führt.</p>

#### 5.5.19 Beginndatum der Arbeitslosigkeit (alo\_beg)

Variablenlabel	Beginndatum der Arbeitslosigkeit
Variablenname	alo_beg
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	LeH, LHG, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Das Merkmal gibt das Beginndatum einer ununterbrochenen Folge von Arbeitslosigkeitsperioden an und ist zu Beginn des Satzes gültig.</p> <p>Folgende Lücken führen nicht zu einer Unterbrechung der Arbeitslosigkeit:</p> <ul style="list-style-type: none"> <li>• jede Lücke kleiner oder gleich 7 Tage</li> <li>• Krankheitslücken bis zu 42 Tagen (ASU)</li> </ul> <p>Bei XASU-Sätzen sind Informationen zu Krankheit nicht enthalten, weshalb Sie bei der Berechnung nicht berücksichtigt werden können.</p>

#### 5.5.20 Dauer der Arbeitslosigkeit (alo\_dau)

Variablenlabel	Dauer der Arbeitslosigkeit
Variablenname	alo_dau
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	LeH, LHG, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Keine



Detailbeschreibung	<p>Das Merkmal gibt die Dauer (in Tagen) einer ununterbrochenen Folge von Arbeitslosigkeitsperioden an und ist zu Beginn des Satzes gültig.</p> <p>Folgende Lücken führen nicht zu einer Unterbrechung der Arbeitslosigkeit:</p> <ul style="list-style-type: none"> <li>• jede Lücke kleiner oder gleich 7 Tage</li> <li>• Krankheitslücken bis zu 42 Tagen (ASU)</li> </ul> <p>Bei der Berechnung der Dauer werden diese Lücken jedoch nicht addiert. Bei XASU-Sätzen sind Informationen zu Krankheit nicht enthalten, weshalb Sie bei der Berechnung nicht berücksichtigt werden können.</p> <p>Vor 1997 bedeutet der Wert „0“ nicht, dass keine Arbeitslosigkeit vorliegt, da hier die Quellen ASU/XASU nicht gefüllt sind.</p>
--------------------	--

## 5.6 Ortsangaben

### 5.6.1 Wohnort - Kreis (wo\_kreis)

Variablenlabel	Wohnort - Kreis
Variablenname	wo_kreis
Kategorie	Ortsangaben
Herkunft	BeH, LeH, LHG, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Bundesland   Kreis
Detailbeschreibung	<p>In BeH- und LeH-Sätzen liegt der Kreis des Wohnortes erst für die Jahre ab 1999 vor. Darin ist jeweils der Kreis (kreisfreie Stadt oder Landkreis) angegeben, in dem der Versicherte wohnt. Der 5-stellige Kreisschlüssel enthält in den ersten beiden Stellen den Schlüssel des Bundeslandes, in der ersten bis dritten Stelle den Regierungsbezirk und in der ersten bis fünften Stelle den Kreis. Bei Ländern ohne Regierungsbezirk steht an der dritten Stelle eine Null.</p> <p>In der BeH wird der Wohnort jeweils zum Jahresende ermittelt und dann einheitlich für alle Sätze eines Jahres zugespielt. Für die Quellen LHG und XASU gilt der Wohnort für den Zeitraum der Originalbeobachtung. Bei ASU, LeH und MTH enthält das Merkmal den Wohnort zum Beginn des Originalzeitraums.</p> <p>Um konsistente regionale Zuordnungen über den gesamten Beobachtungszeitraum sicherzustellen, wurden die Kreisangaben für alle Quellen auf den Gebietsstand vom 31.12.2014 umgeschlüsselt, d.h. in allen Kalenderjahren richtet sich die Zuordnung eines Wohnortes zu einem Kreis nach den Grenzen, welche die Kreise zum 31.12.2014 hatten. Da sich die Grenzen der Kreise im Zeitverlauf geändert haben, würden ohne Gebietsstandsaktualisierung Fälle auftreten, bei denen die Kreiskennziffer des Wohnortes wechselt, ohne dass der betreffende Versicherte umgezogen ist.</p>

Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt. Ansonsten wird als Regionalangabe lediglich das Bundesland (wo_bula) bereitgestellt.
----------------	--

### 5.6.2 Wohnort – Bundesland (wo\_bula)

Variablenlabel	Wohnort - Bundesland
Variablenname	wo_bula
Kategorie	Ortsangaben
Herkunft	BeH, LeH, LHG, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Bundesland   Kreis
Detailbeschreibung	<p>Dieses Merkmal ist eine Aggregation aus dem Merkmal Wohnort Kreis auf die 16 deutschen Bundesländer.</p> <p>In BeH- und LeH-Sätzen liegt das Bundesland des Wohnortes erst für die Jahre ab 1999 vor. Weitere Informationen zum Kreis des Wohnortes finden sich unter Wohnort - Kreis (wo_kreis).</p>

### 5.6.3 Wohnort - Arbeitsagentur (wo\_aa)

Variablenlabel	Wohnort - Arbeitsagentur
Variablenname	wo_aa
Kategorie	Ortsangaben
Herkunft	BeH, LeH, LHG, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Regionaldirektion   Arbeitsagentur
Detailbeschreibung	<p>Das Merkmal enthält ab 1999 die Arbeitsagentur, die am Wohnort des Beschäftigten/Kunden zuständig ist. Sie wird aus der Wohnadresse ermittelt. Für die Quellen LHG und XASU gilt der Wohnort für den Zeitraum der Originalbeobachtung. Bei LeH, ASU und MTH enthält das Merkmal den Wohnort zum Beginn der Meldung. Es ist für die BeH und LeH erst ab 1999 vorhanden.</p> <p>Um konsistente regionale Zuordnungen über den gesamten Beobachtungszeitraum sicherzustellen, wurden die Arbeitsagenturangaben für alle Quellen auf den Gebietsstand vom 31.12.2014 umgeschlüsselt, d.h. in allen Kalenderjahren richtet sich die Zuordnung eines Wohnortes zu einem Arbeitsamtsbezirk nach den Grenzen, welche die Arbeitsagenturbezirke zum 31.12.2014 hatten. Da sich die Grenzen der Arbeitsagenturbezirke im Zeitverlauf geändert haben, würden ohne diese Gebietsstandsaktualisierung Fälle auftreten, bei denen der Arbeitsagenturbezirk einer Person wechselt, ohne dass diese Person umgezogen ist. Einen Problemfall stellt hinsichtlich der Gebietsstandsaktualisierung allerdings Berlin dar: Die Berliner Arbeitsagenturbezirke wurden im Laufe der Jahre immer wieder neu zugeschnitten, was auch durch die Gebietsstandsumschlüsselung nicht vollständig bereinigt werden konnte.</p>

	Vor 1999 enthält das Merkmal die Arbeitsagentur, die den Leistungsfall zuletzt bearbeitet hat. Sie wird nicht aus der Wohnortadresse ermittelt und ist nur für die Quelle LeH gefüllt. Außerdem ist eine Gebietsstandsaktualisierung nicht möglich. Dies ist bei Berechnungen im Zeitverlauf (z. B. regionale Arbeitslosenzahlen) zu beachten.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt. Ansonsten wird lediglich das Gebiet der Regionaldirektion, in dem der Wohnort des Versicherten liegt, ausgewiesen.

#### 5.6.4 Wohnort - Regionaldirektion (wo\_rd)

Variablenlabel	Wohnort - Regionaldirektion
Variablenname	wo_rd
Kategorie	Ortsangaben
Herkunft	BeH, LeH, LHG, ASU, XASU, MTH
Datentyp	Numerisch
Hierarchie	Regionaldirektion   Arbeitsagentur
Detailbeschreibung	Dieses Merkmal ist eine Aggregation aus dem Merkmal Wohnort - Arbeitsagentur auf die Ebene der Regionaldirektionen. Weitere Informationen finden sich dort.

### 5.7 Betriebsmerkmale

#### 5.7.1 WZ73 3-Steller (w73\_3)

Variablenlabel	WZ73 3-Steller
Variablenname	w73_3
Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Wirtschaftsabteilung (1-Steller)   Wirtschaftsgruppe (2-Steller)   Wirtschaftsklasse (3-Steller)
Detailbeschreibung	<p>Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WS73 als 3-Steller aus und ist für die Jahre 1975 bis einschließlich 2002 gefüllt.</p> <p>WS73 steht für die "Klassifikation der Wirtschaftszweige 1973" (Bundesagentur für Arbeit, 2009). Mittels eines dreistelligen Zahlencodes werden 269 Wirtschaftsklassen unterschieden, wobei die erste Stelle dieses Codes die Wirtschaftsabteilung, insgesamt 10, definiert und die beiden ersten Stellen zusammen die jeweilige Gruppe, insgesamt 95, festlegen.</p> <p>Jedem Betrieb ist nur eine Kennziffer zugeordnet. Die Zuordnung zu der entsprechenden Wirtschaftsklasse erfolgt unter Beachtung der institutionellen Ausrichtung des Betriebes.</p>

### 5.7.2 WZ93 5-Steller (w93\_5)

Variablenlabel	WZ93 5-Steller
Variablenname	w93_5
Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Wirtschaftsabschnitt (1-Steller)   Wirtschaftsabteilung (2-Steller)   Wirtschaftsgruppe (3-Steller)   Wirtschaftsklasse (4-Steller)   Wirtschaftsunterklasse (5-Steller)
Detailbeschreibung	<p>Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WZ93 als 5-Steller aus und ist für die Jahre 1999 bis inklusive 2003 gefüllt. WZ93 steht für die „Klassifikation der Wirtschaftszweige, Ausgabe 1993“ (Statistisches Bundesamt, 2002). Die WZ93 baut auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.1 („Nomenclature générale des activités économiques dans les communautés européennes“) auf, die vier Gliederungsebenen vorsieht und bezüglich der ersten beiden wiederum auf dem internationalen Standard ISIC Rev.3 („International Standard Industrial Classification of All Economic Activities“) basiert.</p> <p>Jedem Betrieb ist nur eine Kennziffer zugeordnet. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, soll die wirtschaftliche Haupttätigkeit widerspiegelt werden.</p>
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt. Ansonsten wird der Wirtschaftszweig lediglich als 3-Steller (w93_3) bereitgestellt.

### 5.7.3 WZ93 3-Steller (w93\_3)

Variablenlabel	WZ93 3-Steller
Variablenname	w93_3
Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Wirtschaftsabschnitt (1-Steller)   Wirtschaftsabteilung (2-Steller)   Wirtschaftsgruppe (3-Steller)   Wirtschaftsklasse (4-Steller)   Wirtschaftsunterklasse (5-Steller)
Detailbeschreibung	<p>Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WZ93 als 3-Steller aus und ist für die Jahre 1999 bis inklusive 2003 gefüllt. WZ93 steht für die „Klassifikation der Wirtschaftszweige, Ausgabe 1993“ (Statistisches Bundesamt, 2002). Die WZ93 baut auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.1 („Nomenclature générale des activités économiques dans les communautés européennes“) auf, die vier Gliederungsebenen vorsieht und bezüglich der ersten beiden wiederum auf dem internationalen Standard ISIC Rev.3 („International Standard Industrial Classification of All Economic Activities“) basiert.</p>

	Jedem Betrieb ist nur eine Kennziffer zugeordnet. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, soll die wirtschaftliche Haupttätigkeit widergespiegelt werden.
--	--

#### 5.7.4 WZ03 5-Steller (w03\_5)

Variablenlabel	WZ03 5-Steller
Variablenname	w03_5
Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Wirtschaftsabschnitt (1-Steller)   Wirtschaftsabteilung (2-Steller)   Wirtschaftsgruppe (3-Steller)   Wirtschaftsklasse (4-Steller)   Wirtschaftsunterklasse (5-Steller)
Detailbeschreibung	<p>Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WZ03 als 5-Steller aus und ist für die Jahre 2003 bis 2008 gefüllt. WZ03 steht für die "Klassifikation der Wirtschaftszweige, Ausgabe 2003" (Statistisches Bundesamt, 2003). Die WZ03 baut ebenso wie die WZ93 auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.1.1 auf (siehe Variablenbeschreibung w93_3, w93_5). Die Klassifikationen der Wirtschaftszweige wurden aktualisiert, wobei die Struktur der WZ93 weitgehend beibehalten wurde.</p> <p>Jedem Betrieb ist nur eine Kennziffer zugeordnet. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, soll die wirtschaftliche Haupttätigkeit widergespiegelt werden.</p>
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt. Ansonsten wird der Wirtschaftszweig lediglich als 3-Steller (w03_3) bereitgestellt.

#### 5.7.5 WZ03 3-Steller (w03\_3)

Variablenlabel	WZ03 3-Steller
Variablenname	w03_3
Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Wirtschaftsabschnitt (1-Steller)   Wirtschaftsabteilung (2-Steller)   Wirtschaftsgruppe (3-Steller)   Wirtschaftsklasse (4-Steller)  Wirtschaftsunterklasse (5-Steller)

Detailbeschreibung	<p>Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WZ03 als 3-Steller aus und ist für die Jahre 2003 bis 2008 gefüllt. WZ03 steht für die "Klassifikation der Wirtschaftszweige Ausgabe 2003" (Statistisches Bundesamt, 2003). Die WZ03 baut ebenso wie die WZ93 auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.1.1 auf (siehe Variablenbeschreibung w93_3, w93_5). Die Klassifikationen der Wirtschaftszweige wurden aktualisiert, wobei die Struktur der WZ93 weitgehend beibehalten wurde.</p> <p>Jedem Betrieb ist nur eine Kennziffer zugeordnet. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, soll die wirtschaftliche Haupttätigkeit widergespiegelt werden.</p>
--------------------	--

### 5.7.6 WZ08 5-Steller (w08\_5)

Variablenlabel	WZ08 5-Steller
Variablenname	w08_5
Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Wirtschaftsabschnitt (1-Steller)   Wirtschaftsabteilung (2-Steller)   Wirtschaftsgruppe (3-Steller)   Wirtschaftsklasse (4-Steller)   Wirtschaftsunterklasse (5-Steller)
Detailbeschreibung	<p>Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WZ08 als 5-Steller aus und ist für die Jahre ab 2008 gefüllt. WZ08 steht für die "Klassifikation der Wirtschaftszweige, Ausgabe 2008" (Statistisches Bundesamt, 2008). Die WZ08 baut auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.2 auf.</p> <p>Jedem Betrieb ist nur eine Kennziffer zugeordnet. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, soll die wirtschaftliche Haupttätigkeit widergespiegelt werden.</p>
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt. Ansonsten wird der Wirtschaftszweig lediglich als 3-Steller (w08_3) bereitgestellt.

### 5.7.7 WZ08 3-Steller (w08\_3)

Variablenlabel	WZ08 3-Steller
Variablenname	w08_3
Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Wirtschaftsabschnitt (1-Steller)   Wirtschaftsabteilung (2-Steller)   Wirtschaftsgruppe (3-Steller)   Wirtschaftsklasse (4-Steller)   Wirtschaftsunterklasse (5-Steller)

Detailbeschreibung	<p>Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WZ08 als 3-Steller aus und ist für die Jahre ab 2008 gefüllt. WZ08 steht für die "Klassifikation der Wirtschaftszweige, Ausgabe 2008" (Statistisches Bundesamt, 2008). Die WZ08 baut auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.2 auf.</p> <p>Jedem Betrieb ist nur eine Kennziffer zugeordnet. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, soll die wirtschaftliche Haupttätigkeit wiedergespiegelt werden.</p>
--------------------	---

### 5.7.8 W73\_3 vervollständigt durch Extrapolation/Imputation (w73\_3\_gen)

Variablenlabel	w73_3 vervollständigt durch Extrapolation/Imputation
Variablenname	w73_3_gen
Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Wirtschaftsabteilung (1-Steller)   Wirtschaftsgruppe (2-Steller)   Wirtschaftsklasse (3-Steller)
Detailbeschreibung	<p>Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WS73 als 3-Steller aus. Von 1975 bis einschließlich 2002 enthält das Merkmal die Originalwerte aus w73_3. Ab 2003 wird die Information entweder fortgeschrieben oder mit Hilfe von Umschlüsselungstabellen ersetzt. Das Merkmal bietet somit eine zeitkonsistente Wirtschaftszweigangabe auf Basis der Wirtschaftszweigklassifikation WS73. Eine ausführliche Beschreibung dazu findet sich bei Eberle et al. (2011).</p> <p>Weitere Informationen zur Klassifikation WS73 finden sich bei der Variablenbeschreibung w73_3.</p>

### 5.7.9 Art der Vervollständigung w73\_3 (group\_w73\_3)

Variablenlabel	Art der Vervollständigung w73_3
Variablenname	group_w73_3
Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Dieses Merkmal weist die Art der Vervollständigung für das Merkmal w73_3_gen aus. Es ist angegeben, ob der jeweilige Wert in w73_3_gen dem Originalwert von w73_3 entspricht, weiterhin fehlend ist bzw. extrapoliert oder auf Basis von Umschlüsselungstabellen imputiert wurde.</p> <p>Eine ausführliche Beschreibung zur Vorgehensweise findet sich bei Eberle et al. (2011).</p>

### 5.7.10 W93\_3 vervollständigt durch Extrapolation/Imputation (w93\_3\_gen)

Variablenlabel	W93_3 vervollständigt durch Extrapolation/Imputation
Variablenname	w93_3_gen

Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Wirtschaftsabschnitt (1-Steller)   Wirtschaftsabteilung (2-Steller)   Wirtschaftsgruppe (3-Steller)   Wirtschaftsklasse (4-Steller)   Wirtschaftsunterklasse (5-Steller)
Detailbeschreibung	<p>Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WZ93 als 3-Steller aus. Von 1998 bis einschließlich 2003 enthält das Merkmal die Originalwerte aus w93_3. Vor 1998 und nach 2003 wird die Information entweder rück- bzw. fortgeschrieben oder mit Hilfe von Umschlüsselungstabellen ersetzt, sodass das Merkmal eine zeitkonsistente Wirtschaftszweigangabe auf Basis der Wirtschaftszweigklassifikation WZ93 enthält. Eine ausführliche Beschreibung dazu findet sich bei Eberle et al. (2011).</p> <p>Weitere Informationen zur Klassifikation WZ93 finden sich bei der Variablenbeschreibung w93_3.</p>

#### 5.7.11 Art der Vervollständigung w93\_3 (group\_w93\_3)

Variablenlabel	Art der Vervollständigung w93_3
Variablenname	group_w93_3
Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Dieses Merkmal weist die Art der Vervollständigung für das Merkmal w93_3_gen aus. Es ist angegeben, ob der jeweilige Wert in w93_3_gen dem Originalwert von w93_3 entspricht, weiterhin fehlend ist bzw. extrapoliert oder auf Basis von Umschlüsselungstabellen imputiert wurde.</p> <p>Eine ausführliche Beschreibung zur Vorgehensweise findet sich bei Eberle et al. (2011).</p>

#### 5.7.12 Jahr des 1. Auftretens der Betriebsnummer (grd\_jahr)

Variablenlabel	Jahr des 1. Auftretens der Betriebsnummer
Variablenname	grd_jahr
Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Keine



Detailbeschreibung	Bei dieser Variablen handelt es sich um das erste Auftreten der Nummer des Betriebes im Datensatz. Wird eine Betriebsnummer in Westdeutschland erst nach 1975 bzw. in Ostdeutschland nach 1992 erstmalig nachgewiesen, könnte es sich um das Gründungsdatum des betreffenden Betriebes handeln. Allerdings könnte es sich auch um einen bereits länger existierenden Betrieb handeln, dem nach einem Wechsel des Inhabers oder einer Änderung der Rechtsform eine neue Betriebsnummer zugeordnet wurde (zur Betriebsnummernvergabe siehe Bundesagentur für Arbeit 2007, S. 9-11). Ebenfalls könnte es sein, dass der Betrieb bereits vorher existierte, aber keine sozialversicherungspflichtig bzw. ab 1999 auch keine geringfügig Beschäftigten hatte.
--------------------	--

### 5.7.13 Erstes Auftreten Betriebsnummer (grd\_dat)

Variablenlabel	Erstes Auftreten Betriebsnummer
Variablenname	grd_dat
Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Datum
Hierarchie	Keine
Detailbeschreibung	Bei dieser Variablen handelt es sich um das tagesgenaue erste Auftreten der Nummer des Betriebes in der BeH. Wird eine Betriebsnummer in Westdeutschland erst nach 1975 bzw. in Ostdeutschland nach 1992 erstmalig nachgewiesen, könnte es sich um das Gründungsdatum des betreffenden Betriebes handeln. Allerdings könnte es sich auch um einen bereits länger existierenden Betrieb handeln, dem nach einem Wechsel des Inhabers oder einer Änderung der Rechtsform eine neue Betriebsnummer zugeordnet wurde (zur Betriebsnummernvergabe siehe Bundesagentur für Arbeit, 2007, S. 9-11). Ebenfalls könnte es sein, dass der Betrieb bereits vorher existierte, aber keine sozialversicherungspflichtig bzw. ab 1999 auch keine geringfügig Beschäftigten hatte.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen tagesgenau bereitgestellt. Ansonsten wird lediglich das Jahr des ersten Auftretens der Betriebsnummer (grd_jahr) bereitgestellt.

### 5.7.14 Jahr des letzten Auftretens der Betriebsnummer (lzt\_jahr)

Variablenlabel	Jahr des letzten Auftretens der Betriebsnummer
Variablenname	lzt_jahr
Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Keine

Detailbeschreibung	Bei dieser Variablen handelt es sich um das letztmalige Auftreten der Nummer des Betriebes im Datensatz (siehe Bender et al., 1996). Endet die Existenz einer Betriebsnummer im BHP bereits vor dem Jahr 2008, könnte es sich um eine Betriebsschließung handeln. Allerdings kommen als Ursache auch ein „willkürlicher Wechsel der Betriebsnummer bei Wechsel des Inhabers oder Änderung der Rechtsform“, eine „Ausgliederung von Unternehmensteilen unter einer neue Nummer“ oder andere verwaltungstechnische Änderungen in Betracht (siehe Bender et al., 1996 oder Bundesagentur für Arbeit, 2007, S. 9-11).
--------------------	---

#### 5.7.15 Letztes Auftreten Betriebsnummer (lzt\_dat)

Variablenlabel	Letztes Auftreten Betriebsnummer
Variablenname	lzt_dat
Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Datum
Hierarchie	Keine
Detailbeschreibung	Bei dieser Variablen handelt es sich um das tagesgenaue letztmalige Auftreten der Nummer des Betriebes im Datensatz (siehe Bender et al., 1996). Endet die Existenz einer Betriebsnummer im BHP bereits vor dem Jahr 2008, könnte es sich um eine Betriebsschließung handeln. Allerdings kommen als Ursache auch ein „willkürlicher Wechsel der Betriebsnummer bei Wechsel des Inhabers oder Änderung der Rechtsform“, eine „Ausgliederung von Unternehmensteilen unter einer neue Nummer“ oder andere verwaltungstechnische Änderungen in Betracht (siehe Bender et al., 1996 oder Bundesagentur für Arbeit, 2007, S. 9-11).
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen tagesgenau bereitgestellt. Ansonsten wird lediglich das Jahr des letzten Auftretens der Betriebsnummer (lzt_jahr) bereitgestellt.

#### 5.7.16 Anzahl Beschäftigte gesamt (az\_ges)

Variablenlabel	Anzahl Beschäftigte gesamt
Variablenname	az_ges
Kategorie	Generierte Betriebsmerkmale
Herkunft	BeH
Detailbeschreibung	Die Variable enthält die Anzahl aller zum 30. Juni eines Jahres zur Sozialversicherung gemeldeten Mitarbeiter eines Betriebes. Seit Einführung der neuen Meldevorschriften im Jahr 1999 sind auch die geringfügig Beschäftigten erfasst. Ebenso sind Teilzeitbeschäftigte und ruhende Beschäftigungsverhältnisse (Tagesentgelt von Null) enthalten.

#### 5.7.17 Anzahl in Vollzeit (Normalbeschäftigte + sonstige) (az\_vz)

Variablenlabel	Anzahl Vollzeit (Normalbeschäftigte + sonstige)
Variablenname	az_vz
Kategorie	Generierte Betriebsmerkmale

Herkunft	BeH
Hierarchie	Keine
Detailbeschreibung	Die Variable enthält die Anzahl der Personen im Betrieb, die am 30. Juni eines Jahres mit den Personengruppenschlüsseln 101, 140, 143, 105, 106, 112, 118, 119, 120, 149, 201, 203, 205, 999 und YYY und als Vollzeitbeschäftigte gemeldet sind. Azubis, geringfüge Beschäftigte und Personen in Altersteilzeit werden nicht berücksichtigt.

#### 5.7.18 Anzahl geringfügig Beschäftigte (az\_gf)

Variablenlabel	Anzahl geringfügig Beschäftigte
Variablenname	az_gf
Kategorie	Generierte Betriebsmerkmale
Herkunft	BeH
Hierarchie	Keine
Detailbeschreibung	Die Anzahl der geringfügig Beschäftigten wird anhand des Personengruppenschlüssels – Ausprägungen 109 und 209 – gebildet. Dieses Merkmal ist erst seit 1999 im Datensatz enthalten, da die geringfügig Beschäftigten erst seit diesem Jahr in das Meldeverfahren zur Sozialversicherung integriert wurden.

#### 5.7.19 Mittelwert imputiertes Bruttotagesentgelt Vollzeitbeschäftigte (te\_imp\_mw)

Variablenlabel	Mittelwert imp. Bruttotagesentgelt Vollzeitbeschäftigte
Variablenname	te_imp_mw
Kategorie	Generierte Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Diese Variable enthält den Mittelwert des imputierten Bruttotagesentgeltes der Vollzeitbeschäftigten eines Betriebes. Nicht enthalten sind die Löhne der geringfügig Beschäftigten, der Azubis sowie der Personen in Altersteilzeit.</p> <p>Die Werte sind für alle Jahre in Euro angegeben.</p> <p>Gemäß den Meldevorschriften zur Sozialversicherung müssen Arbeitgeber angeben, wie hoch das beitragspflichtige Bruttoarbeitsentgelt des Beschäftigten für einen bestimmten Zeitraum war (Zeitraumentgelt). Bis Ende 1998 hatten Arbeitgeber nur den sozialversicherungspflichtigen Bruttolohn zu melden. Damit wurden nur Arbeitsentgelte erfasst, die oberhalb der Geringfügigkeits- und unterhalb der Beitragsbemessungsgrenze lagen. Seit 1999 werden im Rahmen des neuen Meldeverfahrens auch Entgelte gemeldet, die unterhalb der Geringfügigkeitsgrenze liegen. Bruttolöhne, die oberhalb der Beitragsbemessungsgrenze liegen, werden jedoch weiterhin gekappt.</p>

	Zur Ermittlung des Bruttotageslohnes wird das Zeitraumergeld durch die Anzahl der Kalendertage des Zeitraums geteilt. Zur Berechnung der Mittelwerte wurden diese zensierten Löhne imputiert (siehe Abschnitt 8.2 in Schmucker et al., 2016). Diese Daten wurden dann auf Betriebsebene aggregiert.
--	---

#### 5.7.20 Arbeitsort Kreis (ao\_kreis)

Variablenlabel	Arbeitsort Kreis
Variablenname	ao_kreis
Kategorie	Ortsangaben
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Bundesland   Kreis
Detailbeschreibung	<p>Angegeben ist jeweils der Kreis (kreisfreie Stadt oder Landkreis), in dem der Betrieb des Beschäftigten angesiedelt ist. Der 5-stellige Kreisschlüssel enthält in den ersten beiden Stellen den Schlüssel des Bundeslands, in der ersten bis dritten Stelle den Regierungsbezirk und in der ersten bis fünften Stelle den Kreis. Bei Ländern ohne Regierungsbezirk steht an der dritten Stelle eine Null.</p> <p>Um konsistente regionale Zuordnungen über den gesamten Beobachtungszeitraum sicherzustellen, wurden die Kreisangaben auf den Gebietsstand vom 31.12.2014 umgeschlüsselt, d.h. in allen Kalenderjahren richtet sich die Zuordnung eines Betriebsortes zu einem Kreis nach den Grenzen, welche die Kreise zum 31.12.2014 hatten. Da sich die Grenzen der Kreise im Zeitverlauf geändert haben, würden ohne Gebietsstandsaktualisierung Fälle auftreten, bei denen die Kreiskennziffer des Betriebsortes wechselt, ohne dass dieser Betrieb seinen Standort verlagert hat.</p>
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt. Ansonsten wird als Regionalangabe lediglich das Bundesland (ao_bula) bereitgestellt.

#### 5.7.21 Arbeitsort Bundesland (ao\_bula)

Variablenlabel	Arbeitsort Bundesland
Variablenname	ao_bula
Kategorie	Ortsangaben
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Bundesland   Kreis
Detailbeschreibung	Angegeben ist jeweils das Bundesland, in dem der Betrieb angesiedelt ist. Dieses Merkmal wird aus der Kreiskennziffer (ao_kreis) generiert. Die ersten beiden Stellen der Kreiskennziffer bezeichnen das Bundesland.

## 6 Literaturverzeichnis

Antoni, Manfred; Ganzer, Andreas; vom Berge, Philipp (2016): Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) 1975-2014. FDZ-Datenreport, 04/2016 (de), Nürnberg.

Antoni, Manfred; Bethmann, Arne (2014): PASS-Befragungsdaten verknüpft mit administrativen Daten des IAB (PASS-ADIAB) 1975-2011. FDZ-Datenreport, 03/2014, Nürnberg.

Antoni, Manfred; Schnell, Rainer (2017): The past, present and future of the German Record Linkage Center (GRLC). In: Jahrbücher für Nationalökonomie und Statistik, im Erscheinen.

Bachteler, Tobias (2008): Dokumentation Record Linkage IEB-PASS. Unveröffentlichtes Manuskript.

Bachteler, Tobias (2012): Merge ToolBox – MTB: Record Linkage Software, Version 0.742, Getting Started. Universität Duisburg-Essen. URL: [http://www.uni-due.de/~hq0215/documents/mtb\\_gettingstarted.pdf](http://www.uni-due.de/~hq0215/documents/mtb_gettingstarted.pdf) (Zugriff: 12.07.2017).

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

Berg, Marco; Cramer, Ralph; Dickmann, Christian; Gilberg, Reiner; Jesske, Birgit; Kleudgen, Martin; Bethmann, Arne; Fuchs, Benjamin; Huber, Martina; Schwarz, Stefan; Trappmann, Mark; Reindl, Alice (2016): Codebuch und Dokumentation des 'Panel Arbeitsmarkt und soziale Sicherung' (PASS). Datenreport Welle 9. FDZ-Datenreport, 07/2016, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

Bertat, Thomas; Dundler, Agnes; Grimm, Christopher; Kiewitt, Jochen; Schomaker, Christine; Schridde, Dr. Henning; Zemann, Dr. Christian (2013): Neue Erhebungsinhalte 'Arbeitszeit', 'ausgeübte Tätigkeit' sowie 'Schul- und Berufsabschluss' in der Beschäftigungsstatistik. Methodenbericht, Bundesagentur für Arbeit – Statistik. URL: <http://statistik.arbeitsagentur.de/Statischer-Content/Grundlagen/Methodenberichte/Beschaeftigungsstatistik/Generische-Publikationen/Methodenbericht-Neue-Erhebungsinhalte-Arbeitszeit-ausgeuebte-Taetigkeit-sowie-Schul-und-Berufsabschluss-in-der-Beschaeftigungsstatistik.pdf> (Zugriff: 31.03.2016).

Beste, Jonas (2011): Selektivitätsprozesse bei der Verknüpfung von Befragungs- mit Prozessdaten: Record Linkage mit Daten des Panels Arbeitsmarkt und soziale Sicherung und administrativen Daten der Bundesagentur für Arbeit. FDZ-Methodenreport, 09/2011, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

Bethmann, Arne; Fuchs, Benjamin; Wurdack, Anja (2013): User Guide „Panel Study Labour Market and Social Security“ (PASS). Wave 6. FDZ-Datenreport 07/2013 (en), Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

Bethmann, Arne; Gebhardt, Daniel (2011): User Guide „Panel Study Labour Market and Social Security“ (PASS). Wave 3. FDZ-Datenreport, 04/2011, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

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

Bundesagentur für Arbeit (Hrsg.) (2007): Handbuch für die Betriebsnummernvergabe und – pflege im Rahmen des Meldeverfahren zur Sozialversicherung. Ausgabe Dezember 2007, Nürnberg.

Bundesagentur für Arbeit (2009): Klassifikation der Wirtschaftszweige 1973, Nürnberg. URL: <http://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen/Klassifikation-der-Wirtschaftszweige/Klassifikation-der-Wirtschaftszweige-1973-2003/Klassifikationen-der-Wirtschaftszweige-1973-2003-Nav.html> (Zugriff: 21.04.2016)

Bundesagentur für Arbeit (2011): Klassifikation der Berufe 2010. Band 1: Systematischer und alphabetischer Teil mit Erläuterungen, Nürnberg. URL: <http://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen/Klassifikation-der-Berufe/KldB2010/Printausgabe-KldB2010/Printausgabe-KldB-2010-Nav.html> (Zugriff: 21.04.2016)

Bundesanstalt für Arbeit (1988): Klassifikation der Berufe – Systematisches und Alphabetisches Verzeichnis der Berufsbenennung, Nürnberg. URL: <http://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen/Klassifikation-der-Berufe/KldB1975-1992/KldB1975-1992-Nav.html> (Zugriff: 21.04.2016)

Christen, Peter (2012): Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection. Berlin, Springer.

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

Deutsche BKK (2016): Ratgeber Sozialversicherung 2016, Wolfsburg. URL: [https://www.deutschebkk.de/fileadmin/user\\_upload/microsites/arbeitgeber/medien/pdf/ratgeber-sozialversicherung-2016.pdf](https://www.deutschebkk.de/fileadmin/user_upload/microsites/arbeitgeber/medien/pdf/ratgeber-sozialversicherung-2016.pdf) (Zugriff: 31.03.2016).

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

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

Fellegi, Ivan; Sunter, Alan (1969): A Theory for Record Linkage. In: Journal of the American Statistical Association, 64(328): S. 1183-1210.

Fitzenberger, Bernd; Osikominu, Aderonke; Völter, Robert (2006): Imputation rules to improve the education variable in the IAB employment subsample. In: Schmollers Jahrbuch, Jg. 126, Heft 3, S. 405-436.

Jesse, Birgit; Knerr, Petra; Schulz, Sabine (2016): Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 9. Erhebungswelle - 2015. FDZ-Methodenreport, 04/2016, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

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

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

Paulus, Wiebke; Matthes, Britta (2013): Klassifikation der Berufe \* Struktur, Codierung und Umsteigeschlüssel. FDZ-Methodenreport, 08/2013 (de), Nürnberg.

Schnell, Rainer; Bachteler, Tobias; Reiher, Jörg (2005): MTB: Ein Record-Linkage-Programm für die empirische Sozialforschung. In: ZA-Information/Zentralarchiv für Empirische Sozialforschung, Nr. 56, S. 93-103.

Schmucker, Alexandra; Seth, Stefan; Ludsteck, Johannes; Eberle, Johanna; Ganzer, Andreas (2016): Betriebs-Historik-Panel 1975-2014. FDZ-Datenreport, 03/2016 (de), Nürnberg.

Statistisches Bundesamt (2002): Klassifikation der Wirtschaftszweige, Ausgabe 1993 (WZ 93), Wiesbaden. URL:

<https://www.destatis.de/DE/Methoden/Klassifikationen/GueterWirtschaftsklassifikationen/Content75/KlassifikationWZ93.html> (Zugriff: 21.04.2016)

Statistisches Bundesamt (2003): Klassifikation der Wirtschaftszweige, Ausgabe 2003 (WZ 2003), Wiesbaden. URL:

<https://www.destatis.de/DE/Methoden/Klassifikationen/GueterWirtschaftsklassifikationen/Content75/KlassifikationWZ2003.html> (Zugriff: 21.04.2016)

Statistisches Bundesamt (2008): Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008), Wiesbaden. URL:

<https://www.destatis.de/DE/Methoden/Klassifikationen/GueterWirtschaftsklassifikationen/Content75/KlassifikationWZ08.html> (Zugriff: 21.04.2016)

Trappmann, Mark; Christoph, Bernhard; Achatz, Juliane; Wenzig, Claudia; Müller, Gerrit; Gebhardt, Daniel (2009): Design and stratification of PASS. A new panel study for research on long term unemployment. IAB Discussion Paper 05/2009, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

Trappmann, Mark; Beste, Jonas; Bethmann, Arne; Müller, Gerrit (2013): The PASS panel survey after six waves. In: Journal for Labour Market Research, Bd. 46, Nr. 4, S. 275-281.

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

## 7 Anhang

### 7.1 Auszählungen

Auszählungen sowie Übersichten über die einzelnen Ausprägungen und Labels der Merkmale finden sich in gesonderten Dateien unter <http://fdz.iab.de>.



## 8 Abkürzungsverzeichnis

AA	Agentur für Arbeit / Arbeitsamt
ALG	Arbeitslosengeld
ARGE	Arbeitsgemeinschaft
ASU	Arbeitsuchende-Historik
A2LL	Arbeitslosengeld II – Leistungen zum Lebensunterhalt
BA	Bundesagentur für Arbeit
BeH	Beschäftigtenhistorik
BHP	Betriebs-Historik-Panel
BMAS	Bundesministerium für Arbeit und Soziales
coArb	Computerunterstützte Arbeitsvermittlung (operatives Verfahren zur Verwaltung der Vermittlung (Altverfahren))
DEÜV	Verordnung über die Erfassung und Übermittlung von Daten für die Träger der Sozialversicherung – Datenerfassungs- und –übermittlungs-Verordnung –
DEVO	Zweite VO über die Erfassung von Daten für die Träger der Sozialversicherung und für die BA – Datenerfassungs-Verordnung –
DÜVO	Zweite VO über die Datenübermittlung auf maschinell verwertbaren Datenträgern im Bereich der Sozialversicherung und der BA – Datenübermittlungs-Verordnung
FDZ	Forschungsdatenzentrum der Bundesagentur für Arbeit am IAB
gAw	Träger mit getrennter Aufgabenwahrnehmung
gE	Gemeinsame Einrichtungen
IAB	Institut für Arbeitsmarkt- und Berufsforschung
IABS	IAB-Beschäftigtenstichprobe
IEB	Integrierte Erwerbsbiographien
ISIC	International Standard Industrial Classification of All Economic Activities
infas	Institut für angewandte Sozialwissenschaft, Bonn
ITM	IT- und Informationsmanagement des Instituts für Arbeitsmarkt- und Berufsforschung
LeH	Leistungsempfängerhistorik
LHG	Leistungs-Historik Grundsicherung
MTH	Maßnahmeteilnehmer-Historik
NACE	Nomenclature générale des activités économiques dans les communautés européennes
PASS	Panel „Arbeitsmarkt und soziale Sicherung“
SGB	Sozialgesetzbuch
SIAB	Stichprobe der Integrierten Arbeitsmarktbiographien

VerBIS	Vermittlungs- und Beratungsinformationssystem
XASU	Arbeitsuchenden-Historik aus XSozial-BA-SGB II
zkT	Zugelassene kommunale Träger