

Forschungsdatenzentrum

der Bundesagentur für Arbeit
im Institut für Arbeitsmarkt-
und Berufsforschung

FDZ

FDZ-Datenreport

Dokumentation zu Arbeitsmarktdaten

03/2014

DE

PASS-Befragungsdaten verknüpft mit administrativen Daten des IAB

(PASS-ADIAB) 1975-2011

Manfred Antoni,
Arne Bethmann



Bundesagentur für Arbeit

PASS-Befragungsdaten verknüpft mit administrativen Daten des IAB

(PASS-ADIAB) 1975–2011

Manfred Antoni, Arne Bethmann (Institut für Arbeitsmarkt- und Berufsforschung - IAB)

Datensatz Version: PASS-ADIAB 7511 v1

Dokumentation Version: PASS-ADIAB 7511 DE v1 dok1

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

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

Inhaltsverzeichnis

Zusammenfassung	9
Abstract	9
1 Einleitung und Kurzbeschreibung	10
1.1 Einleitung	10
1.2 Datennutzung	11
1.3 Veränderungen zu früheren Versionen	12
1.4 Kurzbeschreibung	12
1.5 Merkmalsliste	16
1.6 Mengengerüst und Zeiträume	19
2 Datenquellen	20
2.1 PASS-Befragungsdaten	20
2.2 Administrative Daten aus den IEB	21
2.2.1 Beschäftigten-Historik (BeH)	21
2.2.2 Leistungsempfänger-Historik (LeH)	22
2.2.3 Leistungs-Historiken Grundsicherung (LHG und XLHG)	23
2.2.4 Arbeitsuchenden-Historiken (ASU und XASU)	24
3 Datengenese und -verknüpfung	25
3.1 PASS-Befragungsdaten	25
3.2 Administrative Daten aus den IEB	25
3.2.1 Bereinigungen und Aufbereitungen	25
3.2.2 Episodensplitting	27
3.2.3 Fehlende Werte	28
3.3 Datenverknüpfung	28
3.3.1 Bereinigung und Standardisierung personenbezogener Merkmale	29
3.3.2 Goldstandard-Linkage	30
3.3.3 Exaktes Record Linkage	30
3.3.4 Probabilistisches Record Linkage	30
4 Datenqualität und Probleme	32
4.1 PASS-Befragungsdaten	32
4.2 Gesamte IEB	32
4.3 Beschäftigten-Historik (BeH)	33
4.4 Leistungsempfänger-Historik (LeH)	34
4.5 Leistungs-Historiken Grundsicherung (LHG und XLHG)	34
4.6 Arbeitsuchenden-Historiken (ASU und XASU)	35
5 Variablenbeschreibung	37
5.1 Identifikatoren	37
5.1.1 Unveraenderliche Personennummer (pnr)	37
5.1.2 Systemfreie Betriebsnummer (betnr)	37
5.2 Gültigkeitszeitraum	38
5.2.1 Beginndatum Originalsatz (begorig)	38

5.2.2	Enddatum Originalsatz (endorig)	39
5.2.3	Beginndatum der Episode (begepi)	39
5.2.4	Enddatum der Episode (endepe)	39
5.3	Informationen zur Person	40
5.3.1	Geschlecht (frau)	40
5.3.2	Geburtsjahr (gebjahr)	40
5.3.3	Staatsangehörigkeit (nation)	40
5.3.4	Staatsangehörigkeit vergrößert (nation_gr)	41
5.3.5	Familienstand (famst)	41
5.3.6	Kinderzahl (kind)	41
5.3.7	Ausbildung (bild)	42
5.3.8	Schulausbildung (schbild)	44
5.3.9	Qualifikation (quali)	44
5.3.10	Schwerbehindertenstatus (schweb)	45
5.3.11	Gesundheitliche Einschränkungen (gesund_ein)	45
5.4	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche	46
5.4.1	Abgabegrund/ Beendigungsgrund/ SGB-II-Einstellungsgrund /Abmeldegrund (grund)	46
5.4.2	Tagesentgelt, täglicher Leistungssatz (tentgelt)	47
5.4.3	Gleitzone (gleitz)	49
5.4.4	Beruf — ausgeübte Tätigkeit (beruf)	49
5.4.5	Stellung im Beruf und Arbeitszeit (stib)	50
5.4.6	Erwerbsstatus (erwstat)	50
5.4.7	Erwerbsstatus vor Arbeitsuche (estatvor)	52
5.4.8	Leistungsart (lart)	52
5.4.9	BA-Kundengruppe (kundengr)	52
5.4.10	Art der Kündigung (art_kuend)	52
5.4.11	Arbeitszeit des Stellengesuchs (arbzeit)	53
5.4.12	Restanspruchsdauer Arbeitslosengeld (restanspruch)	53
5.4.13	Trägerart (traeger)	53
5.5	Betriebsmerkmale	54
5.5.1	Wirtschaftszweig 73 (w73_3)	54
5.5.2	Wirtschaftszweig 93, 5-Steller (w93_5)	54
5.5.3	Wirtschaftszweig 93, 3-Steller (w93_3)	55
5.5.4	Wirtschaftszweig 03, 5-Steller (w03_5)	55
5.5.5	Wirtschaftszweig 03, 3-Steller (w03_3)	56
5.5.6	Wirtschaftszweig 08, 5-Steller (w08_5)	56
5.5.7	Wirtschaftszweig 08, 3-Steller (w08_3)	57
5.5.8	Wirtschaftszweig 73 generiert — vervollständigt durch Extrapolation/Imputation (w73_3_gen)	57
5.5.9	Wirtschaftszweig 73 generiert — Art der Vervollständigung (group_w73_3)	58
5.5.10	Wirtschaftszweig 93 generiert — vervollständigt durch Extrapolation/Imputation (w93_3_gen)	58
5.5.11	Wirtschaftszweig 93 generiert — Art der Vervollständigung (group_w93_3)	58
5.5.12	Jahr des ersten Auftretens der Betriebsnummer (grd_jahr)	59
5.5.13	Erstes Auftreten der Betriebsnummer (grd_dat)	59

5.5.14	Jahr des letzten Auftretens der Betriebsnummer (lzt_jahr)	60
5.5.15	Letztes Auftreten der Betriebsnummer (lzt_dat)	60
5.5.16	Anzahl der Beschäftigten gesamt (az_ges)	60
5.5.17	Anzahl Vollzeitbeschäftigte (az_ges_vz)	61
5.5.18	Anzahl geringfügig Beschäftigte (az_gf)	61
5.5.19	Median Bruttotagesentgelt Vollzeitbeschäftigte (te_med)	61
5.6	Ortsangaben	62
5.6.1	Wohnort Kreis (wo_kreis)	62
5.6.2	Wohnort Arbeitsagentur (wo_aa)	62
5.6.3	Wohnort Regionaldirektion (wo_rd)	63
5.6.4	Wohnort Bundesland (wo_bula)	63
5.6.5	Arbeitsort Kreis (ao_kreis)	64
5.6.6	Arbeitsort Arbeitsagentur (ao_aa)	64
5.6.7	Arbeitsort Regionaldirektion (ao_rd)	65
5.6.8	Arbeitsort Bundesland (ao_bula)	65
5.7	Generierte technische Merkmale	65
5.7.1	Quelle des Satzes (quelle)	65
5.7.2	Satzzähler Konto (spell)	65
5.7.3	Jahr (jahr)	66
5.7.4	Vorliegen administrativer Personendaten (match_admin)	66
5.7.5	Typ des Matches (match_typ)	66
5.7.6	Art des exakten Abgleichs (exakt_typ)	67
5.7.7	Probabilistisches Ähnlichkeitsmaß (quality)	67
5.7.8	Wellenindikator erfolgreicher Match (match_wave)	68
6	Auszählungen	69
	Literatur	70

Abbildungsverzeichnis

1	Quellen von PASS-ADIAB	22
---	----------------------------------	----

Tabellenverzeichnis

1	Kurzbeschreibung	12
2	Merkmalsliste	17
3	Mengengerüst	19
4	Biografische Lücken und Identifikationsmöglichkeiten	32

Abkürzungsverzeichnis

A2LL	Arbeitslosengeld II - Leistungen zum Lebensunterhalt
ALG	Arbeitslosengeld
ARGE	Arbeitsgemeinschaft
ASU	Arbeitsuchende-Historik
BA	Bundesagentur für Arbeit
BeH	Beschäftigten-Historik
BfA	Bundesversicherungsanstalt für Angestellte
BHP	Betriebs-Historik-Panel
BMAS	Bundesministerium für Arbeit und Soziales
CATI	Computer Assisted Telephone Interview
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
IAB	Institut für Arbeitsmarkt- und Berufsforschung
IABS	IAB-Beschäftigtenstichprobe
IEB	Integrierte Erwerbsbiographien
LeH	Leistungsempfänger-Historik
LHG	Leistungs-Historik Grundsicherung
LVA	Landesversicherungsanstalt
MTH	Maßnahmeteilnehmer-Historik
SchwBG	Gesetz zur Sicherung der Eingliederung Schwerbehinderter in Arbeit, Beruf und Gesellschaft – Schwerbehindertengesetz
SGB	Sozialgesetzbuch
SIAB	Stichprobe der Integrierten Arbeitsmarktbiographien
VerBIS	Vermittlungs- und Beratungsinformationssystem
XASU	Arbeitsuchenden-Historik aus XSozial-BA-SGB II
XLHG	Leistungsempfänger-Historik Grundsicherung aus XSozial-BA-SGB

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” with administrative data of the Institute for Employment Research (IAB).

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

Danksagung: Wir danken Tobias Bachteler, Johannes Eggs und Martina Huber, die für die Datenverknüpfung vor Welle 5 verantwortlich waren. Darüber hinaus danken wir Stephan Griebemer, Cerstin Rauscher und Stefan Seth für ihre Unterstützung im Rahmen der Datenziehung und -aufbereitung, sowie Johanna Eberle und Anja Wurdack für wertvolle Hinweis zu diesem Datenreport.

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

1 Einleitung und Kurzbeschreibung

1.1 Einleitung

Dieser Datenreport beschreibt die verknüpften Daten 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 5 des PASS (siehe z. B. Berg u. a. (2012) und Bethmann/Fuchs/Wurdack (2013)).

Die Befragungsdaten stammen aus den ersten fünf 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 infas 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 5 liegen Informationen von 15.607 Personen in 10.235 Haushalten vor.

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

- Sozialversicherungspflichtige Beschäftigung (erfasst ab 1975)
- Geringfügige Beschäftigung (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) als arbeitssuchend gemeldet (erfasst ab 2000)
- (Geplante) 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 Beschäftigten-Historik (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 Leistungsempfänger-Historik (LeH) bzw. in der Leistungs-Historiken-Grundsicherung (LHG und XLHG) erfasst. Die Arbeitsuchenden-Historiken (ASU und XASU) stellen die Datenquelle für die durch die BA erfassten Zeiten der Arbeitsuche dar, während die (geplante) Teilnahme an arbeitsmarktpolitischen Maßnahmen in der Maßnahme-Teilnahme-Historik (MTH) erfasst werden.

Mit Ausnahme der (geplanten) Teilnahme an arbeitsmarktpolitischen Maßnahmen aus der MTH sind alle Datenquellen der IEB auch im PASS-ADIAB enthalten.¹ Neben diesen personenbezogenen Daten können auch Angaben zu den beschäftigenden Betrieben aus dem Betriebs-Historik-Panel (BHP) des IAB (Gruhl/Schmucker/Seth, 2012) zugespielt werden.

Der Datenreport zu PASS-ADIAB 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 diskutiert, während die einzelnen Merkmale in Kapitel 5 dargestellt werden.

1.2 Datennutzung

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 BMAS. Nach erfolgter Genehmigung wird ein Nutzungsvertrag mit der Institution der Forscherin/des Forschers geschlossen. Details zur Beantragung des Datensatzes und Datenverarbeitungsmöglichkeiten können der FDZ-Homepage entnommen werden (siehe <http://fdz.iab.de>).

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 persönlichen Interview erhobenen 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.²

Die administrativen Daten in PASS-ADIAB bestehen zum einen aus der *administrativen Personendatei*. Diese enthält Identifikatoren (Personen-ID und systemfreie Betriebsnummer), die persönlichen Merkmale, die Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche, die Merkmale zum Wohnort sowie technische Merkmale. Zum anderen besteht der administrative Teil von PASS-ADIAB aus den *Betriebsdateien*. Diese enthalten Betriebsnummern, die eine Verknüpfung mit der administrativen Personendatei erlauben, sowie jahresbezogene Betriebsmerkmale.

Somit sind Befragungsdaten und administrative Merkmale organisatorisch voneinander getrennt, was Speicherplatz in der Datenhaltung spart. 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:

¹ Die bereitgestellten administrativen Daten entsprechen was die Aufbereitung, die Auswahl und die Bezeichnung der Variablen angeht denen der Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) (vom Berge/König/Seth, 2013)

² Bei der Nutzung der Daten muss berücksichtigt werden, dass die administrativen Daten, die in der vorliegenden Version von PASS-ADIAB zur Verfügung stehen, nur den Zeitraum bis Ende 2010 abdecken. Die aktuellsten Befragungsdaten reichen dagegen bis September 2011. Die administrativen Daten werden um zusätzliche Jahre erweitert, sobald die übrigen administrativen Datensätze des FDZ aktualisiert werden.

Administrative Personendaten:

- Staatsangehörigkeit (*nation*)
- Gesundheitliche Einschränkungen (*gesund_ein*)
- Schwerbehindertenstatus (*schweb*)
- Wohnort Arbeitsagentur (*wo_aa*)
- Wohnort Kreis (*wo_kreis*)
- Arbeitsort Arbeitsagentur (*ao_aa*)

Administrative Betriebsdaten:

- Arbeitsort Kreis (*ao_kreis*)
- Wirtschaftszweig – Wirtschaftsunterklasse (*w93_5*)
- Wirtschaftszweig – Wirtschaftsunterklasse (*w03_5*)
- Wirtschaftszweig – Wirtschaftsunterklasse (*w08_5*)

Beachten Sie, dass die sensiblen Merkmale meist in vergrößerter Version bereits in den Daten enthalten sind (Ausnahmen: *gesund_ein*, *schweb*). 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) weitere Betriebsmerkmale bereitgestellt werden. Nähere Informationen zum BHP finden sich unter <http://fdz.iab.de/>.

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.

1.3 Veränderungen zu früheren Versionen

entfällt

1.4 Kurzbeschreibung

Tabelle 1: Kurzbeschreibung

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

	<p><u>Administrative Personendaten:</u> Informationen zur Person: Geschlecht, Geburtsjahr, Staatsangehörigkeit (original/vergrößert), Familienstand, Kinderzahl, Ausbildung, Schulausbildung, Qualifikation, Schwerbehindertenstatus, gesundheitliche Einschränkungen Information zu Beschäftigung, Leistungsbezug und Arbeitssuche: Beginn und Ende von Beschäftigung, Leistungsbezug und Arbeitssuche, Tagesentgelt/täglicher Leistungssatz, Gleitzone, Beruf – ausgeübte Tätigkeit, Stellung im Beruf und Arbeitszeit, Erwerbsstatus, Erwerbsstatus vor Arbeitssuche, Leistungsart, BA-Kundengruppe, Art der Kündigung, Bereitschaft zur bundesweiten Stellensuche, Arbeitszeit des Stellengesuchs, Restanspruchsdauer Arbeitslosengeld, Trägerart Ortsangaben: Wohnort (Bundesland, Kreis, Regionaldirektion, Arbeitsagentur), Arbeitsort (Bundesland, Kreis, Regionaldirektion, Arbeitsagentur)</p> <p><u>Administrative Betriebsdaten:</u> Betriebsmerkmale, Anzahl der Beschäftigten gesamt, Anzahl Vollzeitbeschäftigte, Anzahl Beschäftigte nach Haupterwerbskonzept, Anzahl Beschäftigte mit Entgelt Null, Anzahl Frauen, Altersstruktur der Beschäftigten, Struktur der Beschäftigten nach Schul- und Berufsausbildung, Struktur der Beschäftigten nach Stellung im Beruf, Struktur der Beschäftigten nach Personengruppenschlüssel, Struktur der Beschäftigten nach Berufsqualifikation, Gehaltsstruktur der Vollzeit-Beschäftigten Erweiterungsmodul – Beschäftigungsströme: Anzahl Eintritte und Austritte gesamt, Anzahl Eintritte und Austritte Frauen, Anzahl Eintritte und Austritte Betriebswechsler, Altersstruktur der Eintritte und Austritte, Anzahl Eintritte und Austritte nach Stellung im Beruf, Anzahl Austritte nach Betriebszugehörigkeitsdauer</p>
Untersuchungseinheit	<p><u>Befragungsdaten:</u> Personen und Haushalte der deutschen Wohnbevölkerung bzw. mit ALG-II-Bezug</p> <p><u>Administrative Personendaten:</u> Sozialversicherungspflichtig Beschäftigte (ab 1999 auch geringfügig Beschäftigte), Leistungsempfänger, Arbeitssuchende</p>
Fallzahlen	<p><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)</p> <p><u>Administrative Personendaten:</u> 21.073 Personen mit 1.333.775 Beobachtungen</p>
Befragungszeitraum	<p>Welle 1: 2006/2007 Welle 2: 2007/2008 Welle 3: 2008/2009 Welle 4: 2010 Welle 5: 2011</p>
Zeitbezug	<p><u>Befragungsdaten:</u> Befragungszeitpunkt, bei Spelldaten zwei Jahre retrospektiv bzw. seit letzter Befragung</p> <p><u>Administrative Personendaten:</u> tagesgenaue Erwerbsbiografien</p> <p><u>Administrative Betriebsdaten:</u> Stichtag 30. Juni</p>

Regionale Gliederung	<u>Befragungsdaten:</u> Bundesland <u>Administrative Daten:</u> Arbeitsort und Wohnort, Bundesländer, Kreise, Regionaldirektionen, Arbeitsagenturbezirke (Einzelne dieser Merkmale sind nicht für alle Datenquellen und Zeiträume gefüllt.)
Gebietsstand	<u>Administrative Personendaten:</u> gebietsstandsbereinigt zum 31.12.2011 <u>Administrative Betriebsdaten:</u> gebietsstandsbereinigt zum 31.12.2011
Methodische Charakteristika	
Erhebungsdesign	<u>Befragungsdaten:</u> 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. Wellen 2–4: 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. 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. Zusätzlich wurde eine Zugangsstichprobe nach dem üblichen Schema gezogen. Weitere Informationen zum Erhebungsdesign sind in Bethmann/Gebhardt (2011) und Berg u. a. (2012) zu finden. <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), ifas 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> 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) ab Welle 4: Spelldaten Erwerbstätigkeiten, Arbeitslosigkeiten und Lücken integriert: bio_spells . dta

	<p>Spelldaten Maßnahmen: mn_spells.dta (ab Welle 2) Spelldaten Maßnahmeteiln.: massnahmespells.dta (nur Welle 1) 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)</p> <p><u>Administrative Personendaten:</u> In einer Datei abgelegt: PASS-ADIAB_7511_v1.dta</p> <p><u>Administrative Betriebsdaten:</u> Basisbetriebsdatei: PASS-ADIAB_7511_v1_bhp_basis_v1.dta Weitere Betriebsinformationen nach Jahren abgelegt (PASS-ADIAB_7511_v1_bhp_v1_1975.dta bis PASS-ADIAB_7511_v1_bhp_v1_2010.dta)</p> <p><u>Erweiterungsmodul – Beschäftigtenströme:</u> PASS-ADIAB_7511_v1_bhp_inflow_v1.dta PASS-ADIAB_7511_v1_bhp_outflow_v1.dta</p> <p><u>Erweiterungsmodulare – Betriebsdynamik:</u> PASS-ADIAB_7511_v1_bhp_ein_v1.dta PASS-ADIAB_7511_v1_bhp_aus_v1.dta</p>
Dateiformat	Stata
Datenzugang	
Zugangswege	Gastaufenthalt, Datenfernverarbeitung (nach vorangegangenem Gastaufenthalt)
Anonymisierungsgrad	Schwach anonymisiert
Sensible Merkmale	<p><u>Befragungsdaten:</u> Keine</p> <p><u>Administrative Personendaten:</u> Staatsangehörigkeit (original) Schwerbehindertenstatus Gesundheitliche Einschränkungen Wirtschaftszweig – Wirtschaftsunterklasse 93, 03, 08: 5-Steller Arbeitsort: Kreis, Arbeitsagentur; Wohnort: Kreis, Arbeitsagentur</p> <p><u>Administrative Betriebsdaten:</u> Kreis, Wirtschaftszweigklassifikation WZ93 5-Steller, Wirtschaftszweigklassifikation WZ03 5-Steller, Wirtschaftszweigklassifikation WZ08 5-Steller³</p>
Zusatzmerkmale	<p>Es können weitere Betriebsmerkmale aus dem BHP sowie die BHP-Erweiterungsmodulare zur Zuspierung an Personendaten beantragt werden. Bitte benennen Sie im Antrag für Gastaufenthalte die sensiblen Merkmale, Variablenblöcke und Erweiterungsmodulare entsprechend dieser Liste der BHP-Merkmale (http://doku.iab.de/fdz/access/BHP_Variablen.pdf).</p>

³ Die Variablen mit 3-stelligem Wirtschaftszweig sind also *nicht* sensibel.

Vorgaben zum Zitieren der Daten und Datendokumentation

Daten:

„Die Datengrundlage dieses Beitrags bilden die faktisch anonymisierten Daten des Panels „Arbeitsmarkt und soziale Sicherung“ (PASS verknüpft mit administrativen Daten des IAB (Version Jahre JJJJ - JJJJ)). Der Datenzugang erfolgte über einen Gastaufenthalt am Forschungsdatenzentrum der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung (FDZ) und/oder mittels kontrollierter Datenfernverarbeitung beim FDZ.“

Datendokumentation:

- Antoni, Manfred; Bethmann, Arne (2014): PASS-Befragungsdaten verknüpft mit administrativen Daten des IAB (PASS-ADIAB) 1975-2011. FDZ-Datenreport, 03/2014, Nürnberg
- Berg, Marco; Cramer, Ralph; Dickmann, Christian; Gilberg, Reiner; Jesske, Birgit; Kleudgen, Martin; Bethmann, Arne; Fuchs, Benjamin; Trappmann, Mark; Wurdack, Anja (2012): Codebuch und Dokumentation des „Panel Arbeitsmarkt und soziale Sicherung“ (PASS). Band I: Datenreport Welle 5. FDZ-Datenreport 06/2012, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.
- Bethmann, Arne (Hrsg.); Fuchs, Benjamin (Hrsg.); Wurdack, Anja (Hrsg.) (2013): User Guide "Panel Study Labour Market and Social Security"(PASS). Wave 6. FDZ-Datenreport, 07/2013 (en), 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.

1.5 Merkmalsliste

Die Merkmalsübersicht listet die Variablennamen und die längeren Merkmalsbezeichnungen der Prozessdaten auf. Weiterhin bietet sie eine Übersicht darüber, welche Merkmale für welche Datenquellen gefüllt sind. Die in der Spalte „BHP Basis“ mit einem „*“ gekennzeichneten Variablen sind in der Basisbetriebsdatei enthalten. Bedingt durch die hohe Anzahl an Merkmalen in den Befragungsdaten werden die PASS-Merkmale nur in der zugehörigen Originaldokumentation aufgeführt.

	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: Der Schulabschluss ist lediglich für ASU- und XASU-Sätze gefüllt, die Sätze der anderen Datenquellen enthalten in dieser Variablen „n“ (siehe auch Abschnitt „Fehlende Werte“ in Kapitel 3). Zu beachten ist, 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- bzw. XLHG-Sätze den Status SGB-II und für ASU- bzw. XASU-Sätze den Arbeitsuche-Status. Nicht bei jedem Merkmal sind diese Unterschiede direkt aus dem Variablennamen ersichtlich.

Tabelle 2: Merkmalsliste

Merkmalsliste	BHP Basis	Seite	BeH	LeH	LHG	XLHG	ASU	XASU
Identifikatoren								
Unveraenderliche Personennummer (pnr)	*	37	■	■	■	■	■	■
Systemfreie Betriebsnummer (betnr)		37	■					
Gültigkeitszeitraum								
Beginndatum Originalsatz (begorig)		38	■	■	■	■	■	■
Enddatum Originalsatz (endorig)		39	■	■	■	■	■	■
Beginndatum der Episode (begepi)		39	■	■	■	■	■	■
Enddatum der Episode (endepe)		39	■	■	■	■	■	■
Informationen zur Person								
Geschlecht (frau)		40	■	■	■	■	■	■
Geburtsjahr (gebjahr)		40	■	■	■	■	■	■
Staatsangehörigkeit (nation)		40	■	■	■	■	■	■
Staatsangehörigkeit vergrößert (nation_gr)		41	■	■	■	■	■	■
Familienstand (famst)		41	■	■	■	■	■	■
Kinderzahl (kind)		41	■	■	■	■	■	■
Ausbildung (bild)		42	■	■	■	■	■	■
Schulbildung (schbild)		44	■	■	■	■	■	■
Qualifikation (quali)		44	■	■	■	■	■	■
Schwerbehindertenstatus (schweb)		45	■	■	■	■	■	■
Gesundheitliche Einschränkungen (gesund_ein)		45	■	■	■	■	■	■
Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche								
Abgabegrund/ Beendigungsgrund/ SGB-II-Einstellungsgrund/ Abmeldegrund (grund)		46	■	■	■	■	■	■
Tagesentgelt, täglicher Leistungssatz (tentgelt)		47	■	■	■	■	■	■
Gleitzone (gleitz)		49	■	■	■	■	■	■
Beruf -- ausgeübte bzw. letzte Tätigkeit (beruf)		49	■	■	■	■	■	■
Stellung im Beruf und Arbeitszeit (stib)		50	■	■	■	■	■	■
Erwerbsstatus (erwstat)		50	■	■	■	■	■	■
Erwerbsstatus vor Arbeitsuche (estatvor)		52	■	■	■	■	■	■
Leistungsart (lart)		52	■	■	■	■	■	■
BA-Kundengruppe (kundengr)		52	■	■	■	■	■	■
Art der Kündigung (art_kuend)		52	■	■	■	■	■	■
Arbeitszeit des Stellengesuchs (arbeitszeit)		53	■	■	■	■	■	■
Restanspruchsdauer Arbeitslosengeld (restanspruch)		53	■	■	■	■	■	■
Trägerart (traeger)		53	■	■	■	■	■	■

	BHP Basis	Seite	BeH	LeH	LHG	XLHG	ASU	XASU
Merkmalsliste								
Betriebsmerkmale								
Wirtschaftszweig 73 (w73_3)	*	54						
Wirtschaftszweig 93, 5-Steller (w93_5)	*	54						
Wirtschaftszweig 93, 3-Steller (w93_3)	*	55						
Wirtschaftszweig 03, 5-Steller (w03_5)	*	55						
Wirtschaftszweig 03, 3-Steller (w03_3)	*	56						
Wirtschaftszweig 08, 5-Steller (w08_5)	*	56						
Wirtschaftszweig 08, 3-Steller (w08_3)	*	57						
Wirtschaftszweig 73 generiert -- vervollständigt durch Extrapolation/ Imputation (w73_3_gen)	*	57						
Wirtschaftszweig 73 generiert -- Art der Vervollständigung (group_w73_3)	*	58						
Wirtschaftszweig 93 generiert -- vervollständigt durch Extrapolation/ Imputation (w93_3_gen)	*	58						
Wirtschaftszweig 93 generiert -- Art der Vervollständigung (group_w93_3)	*	58						
Jahr des ersten Auftretens der Betriebsnummer (grd_jahr)	*	59						
Erstes Auftreten der Betriebsnummer (grd_dat)	*	59						
Jahr des letzten Auftretens der Betriebsnummer (lzt_jahr)	*	60						
Letztes Auftreten der Betriebsnummer (lzt_dat)	*	60						
Anzahl der Beschäftigten gesamt (az_ges)	*	60						
Anzahl Vollzeitbeschäftigte (az_ges_vz)	*	61						
Anzahl geringfügig Beschäftigte (az_gf)	*	61						
Median Bruttotagesentgelt Vollzeitbeschäftigte (te_med)	*	61						
Ortsangaben								
Wohnort Kreis (wo_kreis)		62						
Wohnort Bundesland (wo_bula)		63						
Wohnort Arbeitsagentur (wo_aa)		62						
Wohnort Regionaldirektion (wo_rd)		63						
Arbeitsort Kreis (ao_kreis)	*	64						
Arbeitsort Bundesland (ao_bula)	*	65						
Arbeitsort Arbeitsagentur (ao_aa)		64						
Arbeitsort Regionaldirektion (ao_rd)		65						
Generierte technische Merkmale								
Quelle des Satzes (quelle)		65						
Satzzähler Konto (spell)		65						
Jahr (jahr)	*	66						
Vorliegen administrativer Personendaten (match_admin)		66						(nur in Verknüpfungsdatei)
Typ des Matches (match_typ)		66						(nur in Verknüpfungsdatei)
Art des exakten Abgleichs (exakt_typ)		67						(nur in Verknüpfungsdatei)
Probabilistisches Ähnlichkeitsmaß (quality)		67						(nur in Verknüpfungsdatei)
Wellenindikator erfolgreicher Match (match_wave)		68						(nur in Verknüpfungsdatei)

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, Leistungsbezüge und Arbeitsuche 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 3 zeigt daher an, für wie viele PASS-Teilnehmer IEB-Daten vorliegen, und wie viele Zeilen diese in den administrativen Personendaten aufweisen.

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 5 ab, die in den Jahren 2006 - 2011 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 - 2010 zur Verfügung. Bei den Betriebsdaten beziehen sich die Angaben auf den 30.6. des jeweiligen Jahres. Für einzelne Personen oder Betriebe können die Zeiträume entsprechend der Teilnahme an Befragungen bzw. ihres für die administrativen Daten relevanten Status deutlich kürzer sein.

Tabelle 3: Mengengerüst

PASS-Befragte	31.177
Erfolgreich verknüpfte Befragte mit administrativen Personendaten	21.073
Zugehörige Zeilen in administrativen Personendaten	1.333.775

Quelle: PASS-ADIAB, eigene Berechnungen.

2 Datenquellen

2.1 PASS-Befragungsdaten

Das PASS stellt eine neue 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 Maßnahmenpakete 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 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 Bevölkerungstichprobe 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 mit Ausnahme der Teilnahmen an arbeitsmarktpolitischen Maßnahmen (MTH) aus der IEB übernommen.

2.2.1 Beschäftigten-Historik (BeH)

Die Quelle für Informationen über Beschäftigung ist die Beschäftigten-Historik (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 (vgl. für weitere Details: Bender u. a. 1996: 4ff.; Wermter/Cramer 1988). Es verlangt von den Arbeitgebern, mindestens einmal jährlich Meldungen für 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 Beamte, Selbstständige und ordentliche Studierende (vgl. 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 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

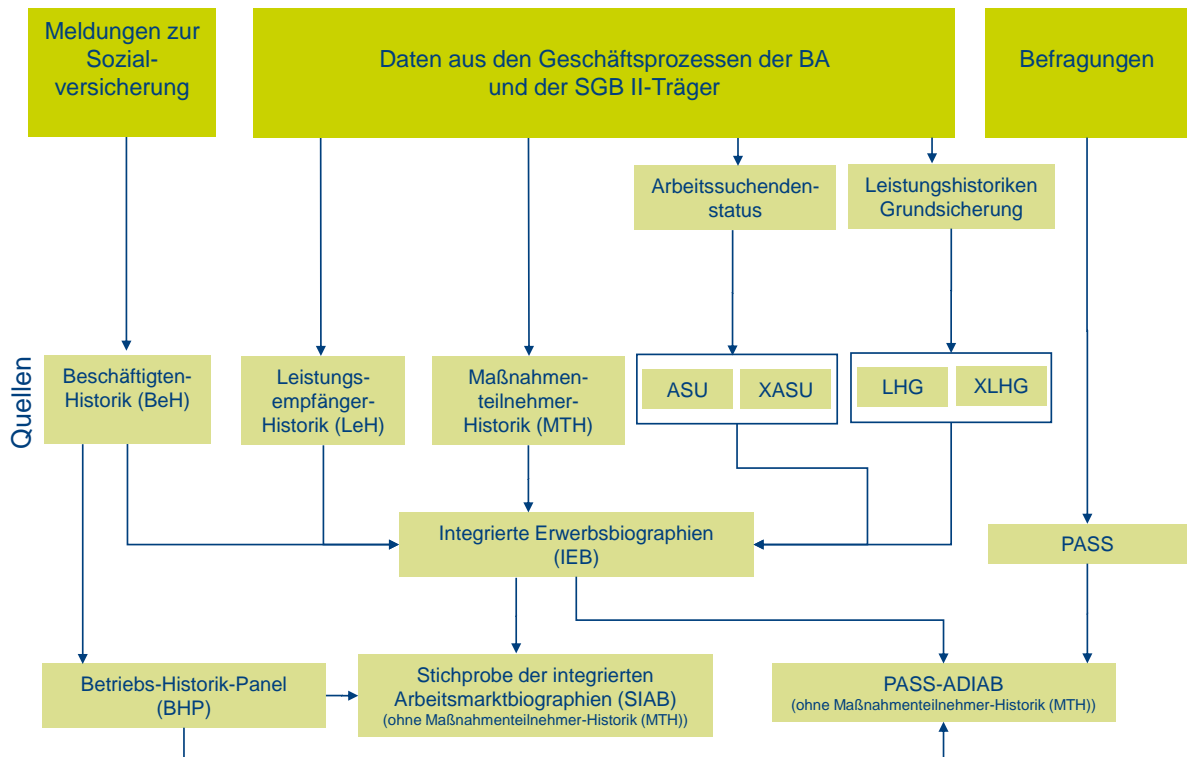


Abbildung 1: Quellen von PASS-ADIAB

Betriebsmerkmale sind deshalb, im Unterschied zu den Personenangaben, keine zeitraumbezogenen Daten, sondern nur am 30.6. jedes Jahres exakt gültig.⁴

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 Listing 1).

2.2.2 Leistungsempfänger-Historik (LeH)

Die Leistungsempfänger-Historik (LeH) des IAB erfasst Zeiträume, in denen Personen Lohnersatzleistungen von der Bundesagentur für Arbeit beziehen. 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,

Listing 1: Beispielcode für Stata 12

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

⁴ Extrembeispiel: Es liegt eine Beschäftigungsmeldung vom 1.1. bis zum 30.5.2006 vor, der Betrieb geht im Juni 2006 Konkurs. Dann gibt es im BHP zu diesem Betrieb für 2006 keine Informationen.

in denen die Voraussetzungen nicht vorliegen (z.B. fehlende Bedürftigkeit im Falle der Arbeitslosenhilfe oder Nichterfüllung der Anwartschaftszeit beim Arbeitslosengeld) nicht in der Leistungsempfänger-Historik berichtet.

2.2.3 Leistungs-Historiken Grundsicherung (LHG und XLHG)

Die Leistungs-Historiken Grundsicherung (LHG/XLHG) dokumentieren Leistungsbezug nach SGB II. Dies umfasst sowohl Leistungen zur Sicherung des Lebensunterhalts (z.B. ALG II), wie auch Zuschläge zum ALG oder zusätzliche Leistungen. Im Gegensatz zu den Leistungen aus dem Rechtskreis des SGB III ist die Bundesagentur für Arbeit nicht alleiniger Träger der Leistungen. Es wird demnach nach den drei möglichen Trägerarten unterschieden:

- Gemeinsame Einrichtungen/Arbeitsgemeinschaften (ARGEn, bis Ende 2010) aus gemeinsamer Aufgabewahrnehmung von BA und Kommune,
- Getrennte Trägerschaft (bis Ende 2011) – hier bleibt die Aufgabenteilung zwischen BA und Kommune bestehen,⁵
- 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 („Leistungs-Historik Grundsicherung aus A2LL“) stammen aus dem BA-Verfahren A2LL, während die XLHG („Leistungs-Historik Grundsicherung aus XSozial-BA-SGB II“) Daten erfasst, welche nach dem Übermittlungsstandard XSozial-BA-SGB II gemeldet wurden.

In der Regel findet A2LL bis 2010 bei ARGEn und ab 2011 bei gemeinsamen Einrichtungen Verwendung, während der Standard XSozial-BA-SGB II von den optierenden Kommunen genutzt wird. Bei Kommunen mit getrennter Trägerschaft kommen beide Verfahren zum Einsatz.

Ein wichtiger Unterschied gegenüber der LeH besteht darin, dass die Ermittlung der Höhe des Leistungsbezugs nicht auf individueller Ebene erfolgt, sondern auf der Ebene von Bedarfsgemeinschaften. Eine Verknüpfung von Individuen mit SGB-II-Leistungsbezug auf Ebene der Bedarfsgemeinschaften ist in der SIAB jedoch nicht möglich. In LHG und XLHG sind grundsätzlich alle gemeldeten Berechtigten nach § 7 SGB II, also erwerbsfähige Hilfebedürftige und ggf. ihre Bedarfsgemeinschaftsmitglieder enthalten. Jedoch werden nur Zeiträume übernommen, in denen eine Person erwerbsfähig oder älter als 64 Jahre ist. 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 65.) der Mitglieder einer Bedarfsgemeinschaft (siehe Abschnitt 3.2.1),
- bei Veränderungen in der Zusammensetzung einer Bedarfsgemeinschaft (z.B. durch Zu-/ Abgänge),
- bei Merkmalsänderungen des Bedarfsgemeinschafts-Kunden sowie

⁵ 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.

- 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.

Der linke Datenrand der LHG ist der 1. Januar 2005, der der XLHG der 1. November 2005. Beide Datenquellen sind jedoch bis Anfang 2007 lückenhaft (siehe Kapitel 4).

2.2.4 Arbeitssuchenden-Historiken (ASU und XASU)

Daten über Arbeitssuchende werden in den Arbeitssuchenden-Historiken (ASU/XASU) gespeichert. Die Quelle ASU enthält Informationen zu Arbeitssuchenden, welche bei Arbeitsagenturen verzeichnet sind und schließt ab 2005 auch Arbeitsgemeinschaften und getrennte Trägerschaften mit ein. In der Quelle XASU sind ab 2005 die Daten von arbeitssuchenden Personen im ALG-II-Bezug von zugelassenen kommunalen Trägern enthalten, welche nach dem XSozial-Standard gemeldet werden.

3 Datengenerierung und -verknüpfung

3.1 PASS-Befragungsdaten

Das Stichprobenkonzept der PASS-Studie wird bei Trappmann u. a. (2009) vorgestellt. Im FDZ-Methodenreport 11/2012 (Jesske/Schulz, 2012) finden sich detaillierte Informationen zur Datenerhebung. Datenstruktur und -aufbereitung werden in den FDZ-Datenreporten 06/2012 (Berg u. a., 2012) 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 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 quellspezifische Bereinigungsverfahren (siehe Folgeabschnitte). Die Gesamt-IEB erfahren die folgenden Bereinigungen:

- Sätze, bei denen das Alter kleiner 13 oder 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 werden gelöscht.

Weitergehende Bereinigungen, wie sie z. B. in der IAB Beschäftigtenstichprobe vorgenommen wurden (Ergänzen von vermutlich fehlenden Meldungen, Streikbereinigungen, etc.), erfolgen nicht.

Beschäftigten-Historik (BeH)

- Um einen im Zeitablauf möglichst konstanten Personenkreis zu erfassen, werden einige Personengruppen, deren Daten nicht für den gesamten Beobachtungszeitraum zur Verfügung stehen, ausgeschlossen. Konkret sind ab 1999 nur die Personengruppen 101, 102, 103, 105, 106, 112, 118, 119, 120, 140, 141, 142, 143, 149, 201, 203 und 205 enthalten sowie die beiden Gruppen 109 und 209, die geringfügig Beschäftigte kennzeichnen. Nicht enthalten sind ab 1999 unter anderem kurzfristig Beschäftigte, d. h. Personengruppen 110, 202 und 210.
- 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 der Versicherungsnummer übernommen. Zu beachten ist aber, dass diese Angaben quellenübergreifend harmonisiert werden.
- Der Gebietsstand für Arbeitsort und Wohnort entspricht dem 31.12.2011.

Leistungsempfänger-Historik (LeH)

- Zu beachten ist, dass bei noch laufendem Bezug das Enddatum dem Ende des Bewilligungszeitraums entspricht und nicht dem Ende des Bezugszeitraums. LeH-Sätze, die über den 14.07.2011 hinausreichen, bilden somit den Bewilligungszeitraum ab. Sätze mit leerem oder ungültigem Enddatum wurden ausgeschlossen, da hier nicht davon ausgegangen werden kann, dass überhaupt eine Zahlung geleistet worden ist.
- Sätze ohne gültiges Beginndatum wurden ausgeschlossen.
- Ein sehr geringer Anteil der LeH-Sätze beginnt bereits vor 1975. Dies erklärt sich dadurch, dass in den IEB alle LeH-Sätze enthalten sind, die 1975 oder später enden.
- 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.
- Eine Gebietsstandsberichtigung wird analog zur BeH durchgeführt.

Leistungs-Historiken Grundsicherung (LHG und XLHG)

- Es sind nur Sätze erwerbsfähiger Personen und von Personen über 65 Jahren enthalten.
- Die personenbezogenen Merkmale, die für die Quellen (X)LHG gefüllt sind, sind für den gesamten Zeitraum des Satzes gültig.
- Eine Gebietsstandsberichtigung wird analog zur BeH durchgeführt.

Arbeitsuchenden-Historiken (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, getrennter Träger) 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 arbeitssuchend“ oder
- „arbeitssuchend und gleichzeitig arbeitslos“

unterschieden.

- Krankheitsepisoden werden 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 Arbeitssuche-Status
 - Wechsel der Verfügbarkeit
 - Wechsel des SGB-II-Trägers (aufgrund der Meldelogik)
 - Wechsel des Wohnortes
- Eine Gebietsstandsberichtigung wird analog zur BeH durchgeführt.

SGB-II-Anonymisierung

Um das Deanonymisierungsrisiko zu verringern, wird in den administrativen Daten grundsätzlich nur das Geburtsjahr ausgewiesen. In den Quellen (X)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 bzw. am Tag davor enden oder aber am Geburtstag neu beginnen.

Um eine indirekte Bestimmung des exakten Geburtstags zu verhindern, wird daher das folgende Verfahren angewendet. Sätze, die am 18. 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. oder 65. Geburtstag beginnen (am 65. Geburtstag 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 bzw. überschneidungsfreie Zeiträume entstehen (Episodensplitting). Die Anzahl der Sätze erhöht sich dadurch.

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`).

3.2.3 Fehlende Werte

In den Daten gilt die folgende Codierung von fehlenden Werten:

Bezeichnung	Wert	Beschreibung
Keine (valide) Angabe	.z	Nicht systematisch fehlender Wert. Das Merkmal ist prinzipiell für die Quelle gefüllt, jedoch ist für den betrachteten Satz ein Wert nicht vorhanden bzw. nicht sinnvoll zu interpretieren.
Fehlt systematisch	.n	Ein Merkmal ist grundsätzlich für eine Quelle (dunkel-graue Felder der Merkmalsübersicht im Abschnitt 1.5) oder für einen gewissen Zeitraum nicht gefüllt.

3.3 Datenverknüpfung

Ausgangspunkt der Verknüpfung der Befragungs- und Prozessdaten sind die Teilnehmerinnen und Teilnehmer der PASS-Erhebung. Diese wurden im Rahmen des Interviews um Erlaubnis gebeten, ihre Befragungsdaten mit den über sie bei der BA vorhandenen administrativen Daten für Forschungszwecke zu verknüpfen. Diese Frage wurde Erstbefragten aber auch Panelbefragten, welche die Zuspieldaten in früheren Wellen noch nicht erteilt hatten, gestellt. Der Wortlaut der Frage blieb über die Wellen hinweg fast identisch, in Welle 5 lautete er 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 Zuspieldaten 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 5 hinweg haben 79% 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

Jeweils nach Abschluss einer jeden Befragungswelle wurde versucht, zuspieldatenbereite Befragungsteilnehmer zu verknüpfen, 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. Über die Zeit hinweg wurde das

Vorgehen bei der Verknüpfung der beiden Datenquellen 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.⁶

Generierte technische Merkmale im Datensatz kennzeichnen für jede Befragte und jeden Befragten, in welchem Verknüpfungsanlauf, in welchem Arbeitsschritt und mit welchem Maß an Sicherheit 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.⁷

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 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 arbeitssuchend, 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 Verlaufsdaten erhoben wurden.

3.3.1 Bereinigung und Standardisierung personenbezogener Merkmale

Die oben genannten personenbezogenen Merkmale wurden Bereinigungsverfahren 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 Punctuationszeichen 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-/Hausnummerfeld 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).

⁶ Eine deutlich detailliertere Darstellung des Vorgehens bei jedem der Anläufe wird im Jahr 2015 als FDZ-Methodenreport veröffentlicht. Der Verknüpfungsanlauf nach Welle 5 wurde durch die Mitarbeiter des German Record Linkage Center (GRLC, vgl. <http://www.record-linkage.de>) durchgeführt. Bei Fragen zu dieser Verknüpfung wenden Sie sich bitte direkt an das GRLC. Das GRLC wurde durch die Deutsche Forschungsgemeinschaft (DFG) gefördert.

⁷ Nutzer, die mutmaßlich fehlerhafte Verknüpfungen in den Daten entdecken, die den Prüfungen durch die Datenverantwortlichen entgangen sind, werden gebeten, diese unter Angabe des einschlägigen Identifikators an das FDZ zu melden. Für spätere Versionen von PASS-ADIAB werden diese Zuordnungen ggf. korrigiert.

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 Exaktes Record Linkage

Anschließend wurde ein exakter 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 exakten Linkage-Schritt war die vollständige Übereinstimmung in allen sonst zur Verfügung stehenden linkage-relevanten Feldern ausschlaggebend, also Vor- und Nachname, Geburtsdatum, Geschlecht, Straße, Hausnummer, Postleitzahl sowie Ort. 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. Wie im generierten technischen Merkmal `match_typ` ersichtlich, begann dieser iterative Prozess mit der strengsten, am wenigsten risikobehafteten Merkmalskombination und wurde nach und nach gelockert. Es wurden jedoch nur Merkmalskombinationen berücksichtigt, bei denen das Risiko, zwei unterschiedlichen Personen zu verknüpfen (sog. falsch-positive Matches), sehr gering war. Insgesamt wurden im Rahmen des exakten Record Linkage mit folgenden Merkmalskombinationen auf Übereinstimmung zwischen den beiden Datenquellen geprüft:

1. Name, Vorname, Geburtsdatum, Geschlecht, PLZ, Straße, Ort, Hausnummer
2. Name, Vorname, Geburtsdatum, PLZ, Straße
3. Name, Vorname, Geburtsdatum, (PLZ oder Ort)
4. Name, Vorname, Geburtsdatum
5. Name, Geburtsdatum, Geschlecht, (PLZ oder Ort)

3.3.4 Probabilistisches Record Linkage

Für Personen, für die durch keinen der bisherigen Schritte eine Entsprechung in den administrativen Daten ermittelt werden konnte, wurden anschließend fehlertolerante Matching-Verfahren verwendet (vgl. Herzog/Scheuren/Winkler, 2007). 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, vgl. Schnell/Bachteler/Reiher (2005)) durchgeführt.

Bei der Berechnung des Ähnlichkeitsmaßes wurde ein probabilistisches Matching angewandt, bei dem rechnerisch berücksichtigt wird, dass die Wahrscheinlichkeit einer Übereinstimmung von Merkmalen von deren Varianz in der Grundgesamtheit abhängig ist. 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 (vgl. 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.

Adresspaare, deren Ähnlichkeitsmaß über einem durch vorläufige Sichtung bestimmten Schwellenwert lag, wurden als probabilistische Matches eingestuft. Die Adresspaare mit einem Ähnlichkeitswert unterhalb des Schwellenwerts wurden anschließend einem Sichtvergleich unterzogen. Die dadurch ermittelten Verknüpfungen wurden nachträglich als probabilistische 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 06/2012 (Berg u. a., 2012) diskutiert. Im 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 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 unterteilt 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 4: Biografische Lücken und Identifikationsmöglichkeiten

Biografische Lücke	Information über Lücke, ggf. identifizierbar anhand der Ausprägung des Merkmals „grund“ im Vorgängersatz der Quelle
Beamte, Berufssoldaten, Richter, Angestellte von Körperschaften oder Stiftungen des öffentlichen Rechts, ungeforderte Selbstständige	LeH, ASU
Studierende, Personen in schulischer Weiterbildung	LeH, ASU
Personen, die länger als 6 Wochen krank/ arbeitsunfähig sind (Krankheit während Arbeitslosigkeit ist aber unter bestimmten Bedingungen in der Quelle ASU abgebildet, vgl. Abschnitt 3.2.1)	BeH, LeH, ASU
Personen in Altersrente ohne Beschäftigung wenn nicht Mitglied in einer BG	LeH, ASU
Bezieher von Vorruhestandsgeld	LeH, ASU
Hausgewerbetreibende	
Kurzfristig Beschäftigte	ASU
Personen in Einrichtungen der Jugendhilfe, Berufsbildungswerken, anerkannten Werkstätten oder ähnlichen Einrichtungen für behinderte Menschen	ASU

Biografische Lücke	Information über Lücke, ggf. identifizierbar anhand der Ausprägung des Merkmals „grund“ im Vorgängersatz der Quelle
Teilnehmende an Leistungen zur Teilhabe am Arbeitsleben (Rehabilitanden)	ASU
(Nebenerwerbs-) Landwirte	
Pflegepersonen i. S. von § 19 SGB XI	
Grundwehrdienstleistende	BeH, LeH, ASU
Wehrübungsleistende	BeH, LeH, ASU
Zivildienstleistende	BeH, LeH, ASU
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.)	LeH, ASU, BeH
Streikende bei Streikdauer über einem Monat	LeH
Sozialhilfe-Bezieher (vor Einführung des SGB II 2005), Sozialgeld-Bezieher (im Sinne des SGB II)	
Ausgleichsempfänger nach dem FELEG (Gesetz zur Förderung der Einstellung der landwirtschaftlichen Erwerbstätigkeit)	
Personen in Maßnahmen	ASU

4.3 Beschäftigten-Historik (BeH)

- 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 Jahrs 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 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 Juli 2009 z. B. kann die 18-Monats-Datei für das Jahr 2007 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 im Jahr 2007 letztmalig bei 100%. Für die Jahre 2008 und 2009 wurden 18-Monats-Dateien verwendet, für 2010 eine 6-Monats-Datei. Es ist daher davon auszugehen, dass die BeH-Meldungen für die Jahre 2008 und 2009 leicht und für 2010 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 (vgl. Bender u. a., 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 Taunus (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.

4.4 Leistungsempfänger-Historik (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.

4.5 Leistungs-Historiken Grundsicherung (LHG und XLHG)

- In den Jahren 2005 und 2006 sind LHG und XLHG nicht vollständig gefüllt. Es wird daher abgeraten, für diesen Zeitraum Analysen allein auf Basis der (X)LHG zu machen.
- Auch ab 2007 kommt es hauptsächlich in der XLHG immer wieder zu meist ein-monatiger träger-spezifischer Untererfassung. Auch die LHG weist in Zusammenhang mit Trägerartwechseln Unter-/Übererfassungen auf.

- Sowohl in der LHG als auch in der XLHG ist zu beobachten, dass es überdurchschnittlich viele 15-jährige und im geringeren 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.
- Manche Personen, für die eine (X)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 ca. 3 bis 5 Prozent der Fälle.

4.6 Arbeitssuchenden-Historiken (ASU und XASU)

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 bzw. XLHG im gleichen Zeitraum zu abzugleichen. Aufgrund der Erfassungslücken bei LHG und XLHG 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.

XASU

- Im Gegensatz zu den Arbeitssuchendendaten aus den ARGEn und den getrennten Trägerschaften zeigen sich nach 1.1.2005 für die 69 zugelassenen kommunalen Träger zwischenzeitlich systematische Untererfassungen. Deshalb sollten Daten aus der Quelle XASU erst nach Juni 2006 ausgewertet werden.
- Verschiedene Merkmale weisen für die XASU zum Teil nur einen sehr geringen Füllgrad auf. Hiervon betroffen sind unter anderem die Merkmale Schulabschluss, Schwerbehindertenstatus, Abgabebegründung sowie Erwerbsstatus vor Arbeitssuche. 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. Ein Grund dafür könnte eine möglicherweise fehlerhafte Bestimmung des Status „Nichtarbeitslos, arbeitssuchend“ bei diesen Trägern sein.
- Die träger- und zeitraumabhängige Plausibilität der XASU-Angaben sollte jeweils unter Berücksichtigung der Untersuchungsfrage vor der Verwendung geprüft werden.

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 in gesonderten Dateien.

5.1 Identifikatoren

5.1.1 Unveraenderliche Personennummer (pnr)

Variablenlabel	Unveraenderliche Personennummer
Variablenname	pnr
Kategorie	Identifikatoren
Herkunft	Generiert
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Die systemfreie 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. Systemfrei bedeutet, dass von dieser Personennummer nicht auf Eigenschaften der Person oder Original-Identifikatoren geschlossen werden kann.</p> <p>Da es keinen einheitlichen BA-weiten Personen-Identifikator gibt, ist die Zuordnung von (BA-)Daten aus unterschiedlichen Datenquellen (z. B. Beschäftigung und Leistung) zu Personen nicht immer eindeutig. Die Bildung des quellenübergreifenden Personen-Identifikators basiert auf einer von der BA entwickelten Heuristik.</p>

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 u. a. 1996, S. 15 f. 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> <ol style="list-style-type: none"> 1. 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. 2. Hat das Unternehmen <i>mehrere</i> Niederlassungen in <i>einer</i> 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 <i>nicht</i> derselben Wirtschaftsklasse angehören. 3. Hat das Unternehmen <i>mehrere</i> Niederlassungen in <i>mehreren</i> Gemeinden, so ist jede dieser Niederlassungen ein Betrieb und erhält je eine Betriebsnummer. <p>In diesem Zusammenhang sind auch folgende Definitionen im Sinne der Betriebsnummernvergabe im Rahmen des Meldeverfahrens zur Sozialversicherung zu beachten:</p> <ol style="list-style-type: none"> 1. Ein <i>Betrieb</i> ist eine regional und wirtschaftsfachlich abgegrenzte Einheit, in der Beschäftigte tätig sind und für die nach o. a. Grundsätzen eine Betriebsnummer zu vergeben wird. 2. Eine <i>Arbeitsstätte</i> ist eine Einheit, in der Beschäftigte tätig sind und für die nach o. a. Grundsätzen eine Betriebsnummer nicht zu vergeben ist. 3. Ein <i>Unternehmen</i> als Begriff fasst Betriebs- und Arbeitsstätten desselben Arbeitgebers zusammen. 4. Ein <i>Arbeitgeber</i> ist jede natürliche oder juristische Person, für die mindestens ein sozialversicherungspflichtig bzw. geringfügig Beschäftigter tätig ist. 5. Betrieb und Betriebsstätten sind synonyme Begriffe; Niederlassung ist dann ein synonyme Begriff für Filialen, Zweigstellen, Auslagerungen, Dependancen, Arbeitsstätten u. ä., wenn sie kein Betrieb ist.
Qualitätshinweis	Die Betriebsnummer fehlt in wenigen BeH-Meldungen. Bei diesen Meldungen handelt es sich u. a. um Meldungen der Personengruppe „205“ (Entgeltmeldungen für unständig Beschäftigte). Da Betriebsmerkmale (Arbeitsort, Wirtschaftszweig, Betriebsgröße usw.) über die originale Betriebsnummer zugespielt werden, fehlen in diesen Meldungen die Betriebsinformationen.

5.2 Gültigkeitszeitraum

5.2.1 Beginndatum Originalsatz (*begorig*)

Variablenlabel	Beginndatum Originalsatz
Variablenname	<i>begorig</i>
Kategorie	Gültigkeitszeitraum
Herkunft	BeH, LeH, LHG, XLHG, ASU, XASU
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 (vgl. 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 beschriebene Anonymisierungsverfahren verändert.</p> <p>In BeH-Meldungen sind Beginn- und Endjahr wegen der Regeln des Meldeverfahrens immer identisch (Pflicht des Arbeitgebers zur Abgabe von Jahresmeldungen).</p>
--------------------	---

5.2.2 Enddatum Originalsatz (endorig)

Variablenlabel	Enddatum Originalsatz
Variablenname	endorig
Kategorie	Gültigkeitszeitraum
Herkunft	BeH, LeH, LHG, XLHG, ASU, XASU
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 (vgl. 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 beschriebene Anonymisierungsverfahren verändert.</p> <p>In BeH-Meldungen sind Beginn- und Endjahr wegen der Regeln des Meldeverfahrens immer identisch (Pflicht des Arbeitgebers zur Abgabe von Jahresmeldungen).</p>

5.2.3 Beginndatum der Episode (begepi)

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

5.2.4 Enddatum der Episode (endepi)

Variablenlabel	Enddatum der Episode
Variablenname	endepi
Kategorie	Generierter Gültigkeitszeitraum
Herkunft	BeH, LeH, LHG, XLHG, ASU, XASU
Datentyp	Datum

Hierarchie	Keine
Detailbeschreibung	Enddatum der Episode. Es ist immer gleich dem oder kleiner als das Enddatum der Originalbeobachtung (vgl. auch die Erläuterungen zum Episodensplitting in Abschnitt 3.2.2).

5.3 Informationen zur Person

5.3.1 Geschlecht (*frau*)

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

5.3.2 Geburtsjahr (*gebjahr*)

Variablenlabel	Geburtsjahr
Variablenname	<i>gebjahr</i>
Kategorie	Persönliche Merkmale
Herkunft	BeH, LeH, LHG, XLHG, ASU, XASU
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.

5.3.3 Staatsangehörigkeit (*nation*)

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

5.3.4 Staatsangehörigkeit vergrößert (*nation_gr*)

Variablenlabel	Staatsangehörigkeit vergrößert
Variablenname	<i>nation_gr</i>
Kategorie	Persönliche Merkmale
Herkunft	BeH, LeH, LHG, XLHG, ASU, XASU
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Das Merkmal enthält als Ausprägungen die zu Gruppen zusammengefassten Staatenschlüssel des Statistischen Bundesamtes.

5.3.5 Familienstand (*famst*)

Variablenlabel	Familienstand
Variablenname	<i>famst</i>
Kategorie	Persönliche Merkmale
Herkunft	LeH, ASU
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Das Merkmal beschreibt den Familienstand. In der Quelle LeH umfasst das Merkmal nur zwei Ausprägungen (0 – nicht verheiratet, 1 – verheiratet), während in den Quellen ASU/ XASU zwischen vier Ausprägungen und in den Quellen LHG/ XLHG zwischen sieben Ausprägungen unterschieden wird. Ein Abgleich der Informationen aus den Quellen wurde nicht vorgenommen.</p> <ol style="list-style-type: none"> ASU/XASU Der Familienstand schließt Kategorien der häuslichen Gemeinschaft mit Kindern oder Lebenspartnern ein. Allerdings ist nicht trennscharf, ob <ul style="list-style-type: none"> eine verheiratete/ in häuslicher Gemeinschaft lebende Person Kinder erzieht, eine verheiratete/ in häuslicher Gemeinschaft/ alleinstehende Person überhaupt Kinder hat, eine alleinerziehende/ in häuslicher Gemeinschaft/ alleinstehende Person verheiratet ist. XLHG Aufgrund eines Bruchs in der Meldelogik 2007 ist die Qualität der Variable eingeschränkt.

5.3.6 Kinderzahl (*kind*)

Variablenlabel	Kinderzahl
Variablenname	<i>kind</i>
Kategorie	Persönliche Merkmale
Herkunft	LeH, ASU
Datentyp	Numerisch
Hierarchie	Keine

Detailbeschreibung	<p>Je nach Datenquelle hat das Merkmal eine unterschiedliche Bedeutung. Außer in der Quelle LeH gibt es die Anzahl der Kinder unter 15 Jahren im Haushalt bzw. in der Bedarfsgemeinschaft zum Zeitpunkt der Antragstellung an.</p> <ol style="list-style-type: none"> 1. <i>LeH</i> In der LeH wird nur unterschieden, ob die Anzahl an Kindern gleich Null oder mindestens eins ist. Folgende Ausprägungen kommen daher vor: <ul style="list-style-type: none"> 0 Keine Kinder 100 Ein oder mehrere Kinder <p>Das Merkmal wird bei Änderungen der Leistungsart oder Bewilligungen nicht aktualisiert, sondern z.B. erst, wenn ein erneuter Leistungsfall auftritt. Kinder werden in der LeH grundsätzlich bis zur Vollendung des 18. Lebensjahrs berücksichtigt, unter bestimmten Voraussetzungen auch noch darüber hinaus, wenn diese für den Leistungsberechtigten einen für ihn erhöhten Leistungssatz begründen.</p> 2. <i>ASU</i> Hier entspricht die Ausprägung des Merkmals der Anzahl an Kindern. 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. 3. <i>LHG, XLHG, XASU</i> Hier entspricht die Ausprägung des Merkmals der Anzahl an Kindern in der Bedarfsgemeinschaft. Bei den Quellen LHG und XLHG gilt der Wert für den gesamten Originalzeitraum.
--------------------	---

5.3.7 Ausbildung (bild)

Variablenlabel	Ausbildung
Variablenname	bild
Kategorie	Persönliche Merkmale
Herkunft	BeH, ASU, XASU
Datentyp	Numerisch
Hierarchie	Keine

Zu beachten ist, dass das Merkmal je nach Quelle eine unterschiedliche Bedeutung hat.

1. *BeH* Das Merkmal enthält für BeH-Sätze die von den Arbeitgebern im Rahmen des Meldeverfahrens anzugebende Ausbildung. Bei dem Signierschlüssel „Ausbildung“ handelt es sich um einen kombinierten Schlüssel, durch den sowohl die erreichte Schulbildung als auch die abgeschlossene Berufsausbildung (jeweils höchster Abschluss) der/des Beschäftigten erhoben werden soll. Dabei ist grundsätzlich zunächst die Schulbildung festzustellen, danach die Berufsausbildung. Lediglich bei den Schlüsselpositionen 5 (Abschluss einer Fachhochschule) und 6 (Hochschul-/ Universitätsabschluss) wird auf die Feststellung einer sonstigen Berufsausbildung verzichtet (vgl. Bundesagentur für Arbeit, 2005: IX). Folgende Ausprägungen kommen vor:
 - 1 Volks-/ Hauptschule/ mittlere Reife oder gleichwertige Schulbildung, ohne abgeschlossene Berufsausbildung
 - 2 Volks-/ Hauptschule/ mittlere Reife oder gleichwertige Schulbildung, mit abgeschlossener Berufsausbildung
 - 3 Mit Abitur, ohne abgeschlossene Berufsausbildung
 - 4 Mit Abitur und abgeschlossener Berufsausbildung
 - 5 Fachhochschulabschluss
 - 6 Hochschulabschluss

Aus den Daten kann nicht geschlossen werden, ob eine abgeschlossene Schulausbildung vorliegt oder nicht, beide Fälle können in den Ausprägungen „1“ und „2“ vorkommen.

„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: 63)

Eine Methode, fehlende Werte oder inkonsistente zeitliche Verläufe der Bildungsangaben in der Vorgängerstichprobe IABS zu bereinigen, findet sich in Fitzenberger/Osikominu/Völter (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>2. <i>ASU/XASU</i></p> <p>Für ASU- und XASU-Sätze ist die abgeschlossene Berufsausbildung zum Beginn der Arbeitsuche enthalten. Folgende Ausprägungen kommen vor:</p> <p>21 ohne abgeschlossene Berufsausbildung 22 betriebliche Ausbildung 23 außerbetriebliche Ausbildung 24 Berufsfachschule 25 Fachschule 26 Fachhochschule 27 Universität</p> <p>In der Quelle ASU kann seit dem 1.7.2006 aufgrund einer Umstellung des Verfahrens nicht zwischen betrieblicher und außerbetrieblicher Ausbildung unterschieden werden. In der Quelle XASU ist diese Unterscheidung im gesamten Zeitraum nicht möglich. In beiden Fällen wird die Mischkategorie auf „betriebliche Ausbildung“ geschlüsselt. In der Quelle ASU kann darüber hinaus ab 1.7.2006 nicht zwischen „ohne abgeschlossene Berufsausbildung“ und „Keine Angabe“ unterschieden werden.</p>
Qualitätshinweis	<p>„Für bestimmte Teilgruppen gibt es einen hohen Anteil fehlender Angaben, weil das Merkmal versicherungsrechtlich keine besondere Bedeutung hat (das betrifft z.B. die geringfügig Beschäftigten).“ (Meinken/Koch, 2004: 63)</p> <p>Von einer quantitativen Auswertung des Merkmals für die Quelle ASU wird ab 2006 abgeraten.</p>

5.3.8 Schulausbildung (*schbild*)

Variablenlabel	Schulausbildung
Variablenname	<i>schbild</i>
Kategorie	Persönliche Merkmale
Herkunft	ASU, XASU
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Das Merkmal enthält den Schulabschluss zu Beginn der Arbeitslosigkeit bzw. Arbeitsuche.
Qualitätshinweis	Der Füllgrad in der XASU steigert sich zwar über die Jahre, bleibt aber relativ gering (zum Teil unter 50%).

5.3.9 Qualifikation (*quali*)

Variablenlabel	Qualifikation
Variablenname	<i>quali</i>
Kategorie	Persönliche Merkmale
Herkunft	ASU
Datentyp	Numerisch
Hierarchie	Keine

Detailbeschreibung	Bei diesem Merkmal wird nicht nur die Ausbildung berücksichtigt, sondern auch die Berufserfahrung. Bei gleichwertiger Befähigung aufgrund von Vorbildung, Leistung oder Erfahrung kann auch ohne einen entsprechenden Abschluss eine Qualifikationsstufe eingetragen werden.
Qualitätshinweis	Aufgrund einer Umstellung des Meldeverfahrens ist dieses Merkmal nur bis 2006 gefüllt.

5.3.10 Schwerbehindertenstatus (schweb)

Variablenlabel	Schwerbehindertenstatus
Variablenname	schweb
Kategorie	Persönliche Merkmale
Herkunft	ASU, XASU
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Die Anerkennung des Schwerbehindertenstatus erfolgt bei einem Grad der Behinderung von über 50%. Die Gleichstellung erfolgt nach §2 SchwbG auf Antrag bei einem Grad der Behinderung von 30 oder 40%. Das Merkmal bezieht sich auf das Beginndatum der Arbeitslosigkeit bzw. des Arbeitsuchendenstatus.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt.
Qualitätshinweis	Der Füllgrad des Merkmals ist in der XASU in den Jahren 2005 und 2006 sehr schlecht.

5.3.11 Gesundheitliche Einschränkungen (gesund_ein)

Variablenlabel	Vermittlungsrelevante gesundheitliche Einschränkungen
Variablenname	gesund_ein
Kategorie	Persönliche Merkmale
Herkunft	ASU
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Die Variable informiert darüber, ob für eine Person vermittlungsrelevante anerkannte gesundheitliche Einschränkungen vorliegen.
Besonderheiten	Das neue Merkmal <i>Gesundheitliche Einschränkungen – Auswirkungen auf Vermittlung</i> ersetzt das alte Merkmal <i>Gesundheitliche Einschränkungen</i> , das in der SIAB7508 enthalten war und aufgrund einer Meldeumstellung nur bis 2006 gefüllt ist. Für dieses Merkmal ist nicht bekannt, ob ein Kunde keine Angabe macht, oder Informationen durch den Arbeitsvermittler nicht erfasst wurden. Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt.

5.4 Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche

5.4.1 Abgabegrund/ Beendigungsgrund/ SGB-II-Einstellungsgrund /Abmeldegrund (*grund*)

Variablenlabel	Grund Abgabe Beschäftigungsmeldung/ Ende Leistungsbezug
Variablenname	<i>grund</i>
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH, LeH, LHG, XLHG, ASU, XASU
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>1. <i>BeH</i></p> <p>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>2. <i>LeH</i> 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 und damit auch hingegen keine Informationen, da die LeH mit den Meldungen der Arbeitsagenturen an die Krankenkassen über abgeschlossene Leistungsbezugsdauern gefüllt wird.</p> <p>3. <i>LHG/XLHG</i> Bei LHG- und XLHG-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>4. <i>ASU/XASU</i> Bei ASU-Sätzen ist der Abmelde- bzw. Abgangsgrund enthalten. Hier-bei ist zu beachten, dass ab dem 26.4.2003 eine Reduktion der Merkmalsausprägungen erfolgte. Für Auswertungen über längere Zeiträume können die alten Werte auf die aktuell gültigen anhand folgender Tabelle umgeschlüsselt werden:</p> <table border="1" data-bbox="614 1025 1209 1350"> <thead> <tr> <th>alt</th><th>→</th><th>neu</th><th>alt</th><th>→</th><th>neu</th><th>alt</th><th>→</th><th>neu</th></tr> </thead> <tbody> <tr><td>29</td><td>→</td><td>60</td><td>38</td><td>→</td><td>66</td><td>48</td><td>→</td><td>78</td></tr> <tr><td>30</td><td>→</td><td>60</td><td>39</td><td>→</td><td>71</td><td>49</td><td>→</td><td>69</td></tr> <tr><td>31</td><td>→</td><td>61</td><td>40</td><td>→</td><td>69</td><td>50</td><td>→</td><td>75</td></tr> <tr><td>32</td><td>→</td><td>60</td><td>42</td><td>→</td><td>65</td><td>51</td><td>→</td><td>74</td></tr> <tr><td>33</td><td>→</td><td>60</td><td>43</td><td>→</td><td>70</td><td>52</td><td>→</td><td>76</td></tr> <tr><td>34</td><td>→</td><td>60</td><td>44</td><td>→</td><td>74</td><td>53</td><td>→</td><td>68</td></tr> <tr><td>35</td><td>→</td><td>60</td><td>45</td><td>→</td><td>77</td><td>54</td><td>→</td><td>78</td></tr> <tr><td>36</td><td>→</td><td>61</td><td>46</td><td>→</td><td>67</td><td></td><td></td><td></td></tr> <tr><td>37</td><td>→</td><td>66</td><td>47</td><td>→</td><td>67</td><td></td><td></td><td></td></tr> </tbody> </table>	alt	→	neu	alt	→	neu	alt	→	neu	29	→	60	38	→	66	48	→	78	30	→	60	39	→	71	49	→	69	31	→	61	40	→	69	50	→	75	32	→	60	42	→	65	51	→	74	33	→	60	43	→	70	52	→	76	34	→	60	44	→	74	53	→	68	35	→	60	45	→	77	54	→	78	36	→	61	46	→	67				37	→	66	47	→	67			
alt	→	neu	alt	→	neu	alt	→	neu																																																																																			
29	→	60	38	→	66	48	→	78																																																																																			
30	→	60	39	→	71	49	→	69																																																																																			
31	→	61	40	→	69	50	→	75																																																																																			
32	→	60	42	→	65	51	→	74																																																																																			
33	→	60	43	→	70	52	→	76																																																																																			
34	→	60	44	→	74	53	→	68																																																																																			
35	→	60	45	→	77	54	→	78																																																																																			
36	→	61	46	→	67																																																																																						
37	→	66	47	→	67																																																																																						
Qualitätshinweis	<p>Anteil valider Werte (Füllgrad) des Abgabegrundes in den Quellen LHG, XLHG, ASU und XASU:</p> <table border="1" data-bbox="539 1429 1264 1585"> <thead> <tr> <th></th><th>2005</th><th>2006</th><th>2007</th><th>2008</th><th>2009</th><th>2010</th><th>2011</th></tr> </thead> <tbody> <tr><td>LHG</td><td>0,17</td><td>0,11</td><td>0,19</td><td>0,18</td><td>0,15</td><td>0,12</td><td>0,05</td></tr> <tr><td>XLHG</td><td>0,02</td><td>0,06</td><td>0,09</td><td>0,12</td><td>0,17</td><td>0,20</td><td>0,11</td></tr> <tr><td>ASU</td><td>0,83</td><td>0,95</td><td>0,98</td><td>0,97</td><td>0,96</td><td>0,89</td><td>0,48</td></tr> <tr><td>XASU</td><td>0,25</td><td>0,53</td><td>0,52</td><td>0,53</td><td>0,55</td><td>0,49</td><td>0,27</td></tr> </tbody> </table>		2005	2006	2007	2008	2009	2010	2011	LHG	0,17	0,11	0,19	0,18	0,15	0,12	0,05	XLHG	0,02	0,06	0,09	0,12	0,17	0,20	0,11	ASU	0,83	0,95	0,98	0,97	0,96	0,89	0,48	XASU	0,25	0,53	0,52	0,53	0,55	0,49	0,27																																																		
	2005	2006	2007	2008	2009	2010	2011																																																																																				
LHG	0,17	0,11	0,19	0,18	0,15	0,12	0,05																																																																																				
XLHG	0,02	0,06	0,09	0,12	0,17	0,20	0,11																																																																																				
ASU	0,83	0,95	0,98	0,97	0,96	0,89	0,48																																																																																				
XASU	0,25	0,53	0,52	0,53	0,55	0,49	0,27																																																																																				

5.4.2 Tagesentgelt, täglicher Leistungssatz (*tentgelt*)

Variablenlabel	Tagesentgelt, täglicher Leistungssatz
Variablenname	<i>tentgelt</i>
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH, LeH
Datentyp	Numerisch
Hierarchie	Keine

1. *BeH*

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.

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 Entgeltinformationen 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.

5.4.3 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 Euro 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 01.04.2003.

5.4.4 Beruf – ausgeübte Tätigkeit (beruf)

Variablenlabel	Beruf – ausgeübte bzw. letzte Tätigkeit
Variablenname	beruf
Kategorie	Beschäftigung/ Leistungsbezug
Herkunft	BeH, ASU
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<ol style="list-style-type: none"> 1. <i>BeH</i> Die Berufsbezeichnung der im Meldezeitraum vom Beschäftigten ausgeübten Tätigkeit ist Bestandteil der vom Arbeitgeber übermittelten „Angaben zur Tätigkeit“. Der Arbeitgeber verschlüsselt die Tätigkeit der/des Beschäftigten hierbei gemäß der „Klassifizierung der Berufe. Systematisches und alphabetisches Verzeichnis der Berufsbenennungen“ (Hrsg. von der Bundesanstalt für Arbeit, Nürnberg, 1988), in dem ca. 25.000 Berufsbezeichnungen zu finden sind. Die Berufsordnung besteht aus einer 3-stelligen Kennziffer und umfasst rund 330 Ausprägungen. Treffen mehrere Berufsbezeichnungen mit verschiedenen Schlüsselzahlen für einen Beschäftigten zu, ist der Arbeitgeber gehalten, diejenige Berufsbezeichnung zu wählen, welche die überwiegend ausgeübte Tätigkeit kennzeichnet (vgl. Bundesagentur für Arbeit, 2005: V). 2. <i>AUS/XASU</i> Bei ASU-Datensätzen bezeichnet das Merkmal den Beruf der letzten Tätigkeit. Bei XASU-Datensätzen bezeichnet das Merkmal die letzte erfolgreich abgeschlossene Berufsausbildung.

5.4.5 Stellung im Beruf und Arbeitszeit (stib)

Variablenlabel	Stellung im Beruf und Arbeitszeit
Variablenname	stib
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Die berufliche Stellung eines Beschäftigten im Meldezeitraum wird vom Arbeitgeber im Rahmen der „Angaben zur Tätigkeit“ übermittelt.</p> <p>Das Merkmal „Berufsstellung“ unterscheidet zunächst zwischen Vollzeit- und Teilzeitbeschäftigten. Maßgeblich ist dabei das Verhältnis zwischen der arbeitsvertraglich vereinbarten und der betriebsüblichen Arbeitszeit. Nur für Vollzeitbeschäftigte liefert die Variable tatsächlich Angaben zur beruflichen Stellung, während für Teilzeitbeschäftigte nur erfasst wird, ob ihre Stundenzahl eine bestimmte Grenze übersteigt oder nicht. Diese Grenze lag bis 1978 bei 20 Wochenstunden, zwischen 1979 und 1987 bei 15 Wochenstunden und seit 1988 bei 18 Wochenstunden. Die Unterscheidung zwischen vollzeitbeschäftigten Arbeitern und Angestellten richtet sich ausschließlich nach der Art der Rentenversicherungspflicht (BfA für Angestellte und LVA für Arbeiter). Neben den Auszubildenden, Praktikanten und Volontären zählen auch Anlernlinge, Schüler an Schulen des Gesundheitswesens und Teilnehmer an geförderten Maßnahmen zur beruflichen Fortbildung, Umschulung und betrieblichen Einarbeitung zur Kategorie Beschäftigte in beruflicher Ausbildung.</p> <p>Falls mehrere Schlüsselzahlen in Betracht kommen, ist der Arbeitgeber dazu angehalten, die Einstufung nach der überwiegend ausgeübten Tätigkeit vorzunehmen. Lässt sich diese nicht eindeutig feststellen, ist die Schlüsselzahl der höherwertigen Stellung im Beruf einzutragen (vgl. Bundesagentur für Arbeit, 2005: VI).</p>

5.4.6 Erwerbsstatus (erwstat)

Variablenlabel	Erwerbsstatus
Variablenname	erwstat
Kategorie	Persönliche Merkmale
Herkunft	BeH, LeH, LHG, XLHG, ASU, XASU
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Das Merkmal nimmt für jede Datenquelle unterschiedliche Werte mit unterschiedlichen Bedeutungen an.</p> <ol style="list-style-type: none"> 1. <i>BeH</i> 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 (vgl. Deutsche BKK, 2012: 31).

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.

2. *LeH*

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.

3. *ASU/XASU*

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.

In XASU-Sätzen enthält das Merkmal Erwerbsstatus bislang die Ausprägungen „nicht arbeitslos, aber arbeitsuchend“ sowie „arbeitsuchend und gleichzeitig arbeitslos“.

Da die Meldelogik in vielen Fällen eine Reidentifikation des exakten Geburtsdatums ermöglichen würde, wurden die ursprünglichen Datumsangaben durch das in Abschnitt 3.2.1 beschriebene Anonymisierungsverfahren verändert.

4. *LHG/XLHG*

Für LHG- und XLHG-Sätze zeigt das Merkmal Erwerbsstatus an, ob die Person minderjährig erwerbsfähig, volljährig erwerbsfähig oder nicht erwerbsfähig ab Altersrentengrenze geführt wird.

Da die Meldelogik in vielen Fällen eine Reidentifikation des exakten Geburtsdatums ermöglichen würde, wurden die ursprünglichen Datumsangaben durch das in Abschnitt 3.2.1 beschriebene Anonymisierungsverfahren verändert.

5.4.7 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. Die Angabe bezieht sich nicht auf das Beginndatum der Episode, sondern auf den Beginn des Originalzeitraums.
Qualitätshinweis	Der Anteil valider Werte (Füllgrad) ist für XASU-Sätze sehr gering.

5.4.8 Leistungsart (lart)

Variablenlabel	Leistungsart
Variablenname	lart
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	LeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Art der von der BA gezahlten Entgeltersatzleistung. Im Merkmal „Erwerbsstatus“ wird dieses Merkmal für die Quelle LeH zusammengefasst.

5.4.9 BA-Kundengruppe (kundengr)

Variablenlabel	BA-Kundengruppe
Variablenname	kundengr
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	ASU
Datentyp	numerisch
Hierarchie	Keine
Detailbeschreibung	Das Merkmal enthält die Eingruppierung der BA-Kunden in verschiedene Typen wie Marktkunde, Beratungskunde (aktivieren), Beratungskunde (fördern) sowie Betreuungskunde. Diese Unterscheidung berücksichtigt auch den Handlungsbedarf hinsichtlich potenziell vorhandener individueller Integrationshemmnisse.

5.4.10 Art der Kündigung (art_kuend)

Variablenlabel	Art der Kündigung der letzten Tätigkeit
Variablenname	art_kuend
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche

Herkunft	ASU
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.

5.4.11 Arbeitszeit des Stellengesuchs (**arbzeit**)

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

5.4.12 Restanspruchsdauer Arbeitslosengeld (**restanspruch**)

Variablenlabel	Restanspruch Arbeitslosengeld
Variablenname	restanspruch
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	LeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Der Restanspruch ist der Anspruch, der nach Ende des aktuellen Leistungsbezugs übrig bleibt. Der Restanspruch blieb früher sieben Jahre bestehen, seit Februar 2006 sind es vier Jahre. Liegt das Enddatum des Leistungsempfangs vor dem 1.1.1998, wird der Restanspruch in Werktagen angegeben. Danach erfolgt die Angabe in Kalendertagen. Dieser Anspruch zählt maximal bis auf die Höchstdauer für das jeweilige Lebensalter. Die Angabe bezieht sich nicht auf das Enddatum der Episode, sondern auf das Ende des Originalzeitraums.

5.4.13 Trägerart (**traeger**)

Variablenlabel	Trägerart
Variablenname	traeger
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	LHG, XLHG, ASU, XASU
Datentyp	Numerisch
Hierarchie	Keine

Detailbeschreibung	<p>Das Merkmal zeigt die drei möglichen Trägerschaftsarten des Rechtskreises SGB II (Arbeitsgemeinschaften (ARGEn), optierende Kommunen, getrennte Träger) und die Trägerschaftsart des Rechtskreises SGB III (Bundesagentur für Arbeit). Arbeitsgemeinschaften (ARGEn) bilden sich aus einer gemeinsamen Aufgabewahrnehmung von BA und Kommunen. Bei den optierenden Kommunen übernimmt die jeweilige Kommune die Aufgaben des SGB II in Eigenregie. Bei der geteilten Trägerschaft bleibt die Aufgabenteilung zwischen BA und Kreis bestehen. Zu beachten ist, dass die Grenzen der Träger in Bezug auf Kreise nicht trennscharf sind und somit nicht mit den Grenzen der Kreise übereinstimmen müssen. Bei XASU-Sätzen ist die Art des Trägers enthalten, der das Bewerberangebot führt. Die Information wird aus dem Bewerberangebot jeweils zu Beginn des ungesplitteten Satzes zugespielt.</p>
--------------------	--

5.5 Betriebsmerkmale

5.5.1 Wirtschaftszweig 73 (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. Die Wirtschaftsklasse W73 ist von 1975 bis einschließlich 2002 gefüllt.</p> <p>WS73 steht für das „Verzeichnis der Wirtschaftszweige für die Statistik der Bundesanstalt für Arbeit, Ausgabe 1973“. 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 darf nur eine Kennziffer zugeordnet werden. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, muss die wirtschaftliche Haupttätigkeit ermittelt werden.</p>

5.5.2 Wirtschaftszweig 93, 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 1998 bis inklusive 2003 gefüllt. WZ93 steht für die „Klassifikation der Wirtschaftszweige für die Statistik der Bundesanstalt für Arbeit, Ausgabe 1993“. 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	<p>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.</p>

5.5.3 Wirtschaftszweig 93, 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 1998 bis inklusive 2003 gefüllt. WZ93 steht für die „Klassifikation der Wirtschaftszweige für die Statistik der Bundesanstalt für Arbeit, Ausgabe 1993“. 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>

5.5.4 Wirtschaftszweig 03, 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	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“ des Statistischen Bundesamtes (Hrsg.) 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. Jedem Betrieb ist nur eine Kennziffer zugeordnet. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, soll die wirtschaftliche Haupttätigkeit widerspiegelt werden.
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.

5.5.5 Wirtschaftszweig 03, 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	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“ des Statistischen Bundesamtes (Hrsg.) 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. Jedem Betrieb ist nur eine Kennziffer zugeordnet. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, soll die wirtschaftliche Haupttätigkeit widerspiegelt werden.

5.5.6 Wirtschaftszweig 08, 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" des Statistischen Bundesamtes (Hrsg.) 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 widerspiegelt werden.</p>
Besonderheiten	<p>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.</p>

5.5.7 Wirtschaftszweig 08, 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“ des Statistischen Bundesamtes (Hrsg.). 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 widerspiegelt werden.</p>

5.5.8 Wirtschaftszweig 73 generiert — 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 u. a. (2011). Weitere Informationen zur Klassifikation WS73 finden sich bei der Variablenbeschreibung w73_3.</p>

5.5.9 Wirtschaftszweig 73 generiert – Art der Vervollständigung (group_w73_3)

Variablenlabel	Art der Vervollständigung w73_3
Variablenname	group_w73_3
Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	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. Eine ausführliche Beschreibung zur Vorgehensweise findet sich bei Eberle u. a. (2011).

5.5.10 Wirtschaftszweig 93 generiert — 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	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 WS93 enthält. Eine ausführliche Beschreibung dazu findet sich bei Eberle u. a. (2011). Weitere Informationen zur Klassifikation WZ93 finden sich bei der Variablenbeschreibung w93_3.

5.5.11 Wirtschaftszweig 93 generiert — Art der Vervollständigung (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 <code>w93_3_gen</code> aus. Es ist angegeben, ob der jeweilige Wert in <code>w93_3_gen</code> dem Originalwert von <code>w93_3</code> entspricht, weiterhin fehlend ist bzw. extrapoliert oder auf Basis von Umschlüsselungstabellen imputiert wurde.</p> <p>Eine ausführliche Beschreibung zur Vorgehensweise findet sich bei Eberle u. a. (2011).</p>
--------------------	---

5.5.12 Jahr des ersten Auftretens der Betriebsnummer (`grd_jahr`)

Variablenlabel	Jahr des 1. Auftretens der Betriebsnummer
Variablenname	<code>grd_jahr</code>
Kategorie	Betriebsmerkmale
Herkunft	BeH
Hierarchie	Keine
Detailbeschreibung	<p>Bei dieser Variablen handelt es sich um das erste Auftreten der Nummer des Betriebes im Datensatz.</p> <p>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: 9–11). Ebenfalls könnte es sein, dass der Betrieb bereits vorher existierte, aber keine sozialversicherungspflichtig bzw. ab 1999 auch keine geringfügig Beschäftigten hatte.</p>

5.5.13 Erstes Auftreten der Betriebsnummer (`grd_dat`)

Variablenlabel	Erstes Auftreten Betriebsnummer
Variablenname	<code>grd_dat</code>
Kategorie	Betriebsmerkmale
Herkunft	BeH
Hierarchie	Keine
Detailbeschreibung	<p>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: 9–11). Ebenfalls könnte es sein, dass der Betrieb bereits vorher existierte, aber keine sozialversicherungspflichtig bzw. ab 1999 auch keine geringfügig Beschäftigten hatte.</p>
Besonderheiten	<p>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 (<code>grd_jahr</code>) bereitgestellt.</p>

5.5.14 Jahr des letzten Auftretens der Betriebsnummer (lzt_jahr)

Variablenlabel	Jahr des letzten Auftretens der Betriebsnummer
Variablenname	lzt_jahr
Kategorie	Betriebsmerkmale
Herkunft	BeH
Hierarchie	Keine
Detailbeschreibung	Bei dieser Variablen handelt es sich um das letztmalige Auftreten der Nummer des Betriebes im Datensatz (vgl. Bender u. a., 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 (vgl. Bender u. a. 1996 oder Bundesagentur für Arbeit 2007: 9–11).

5.5.15 Letztes Auftreten der Betriebsnummer (lzt_dat)

Variablenlabel	Letztes Auftreten Betriebsnummer
Variablenname	lzt_dat
Kategorie	Betriebsmerkmale
Herkunft	BeH
Hierarchie	Keine
Detailbeschreibung	Bei dieser Variablen handelt es sich um das tagesgenaue letztmalige Auftreten der Nummer des Betriebes im Datensatz (vgl. Bender u. a., 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 (vgl. Bender u. a. 1996 oder Bundesagentur für Arbeit 2007: 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.

5.5.16 Anzahl der Beschäftigten gesamt (az_ges)

Variablenlabel	Anz. 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.

5.5.17 Anzahl Vollzeitbeschäftigte (az_ges_vz)

Variablenlabel	Anzahl Vollzeitbeschäftigte gesamt
Variablenname	az_ges_vz
Kategorie	Generierte Betriebsmerkmale
Herkunft	BeH
Hierarchie	Keine
Detailbeschreibung	Die Variable enthält die Anzahl der Vollzeitbeschäftigten eines Betriebes zum 30. Juni eines Jahres. Nicht enthalten sind hier die geringfügig Beschäftigten sowie ruhende Beschäftigungsverhältnisse (Tagesentgelt von Null).

5.5.18 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.

5.5.19 Median Bruttotagesentgelt Vollzeitbeschäftigte (te_med)

Variablenlabel	Median Bruttotagesentgelt Vollzeitbeschäftigte
Variablenname	te_med
Kategorie	Generierte Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Diese Variable enthält den Median des Bruttotagesentgeltes der Vollzeitbeschäftigten eines Betriebes. Nicht enthalten sind die geringfügig Beschäftigten sowie ruhende Beschäftigungsverhältnisse (Tagesentgelt Null). 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. Zur Ermittlung des Bruttotageslohnes wird das Zeitraumentgelt durch die Anzahl der Kalendertage des Zeitraums geteilt und der Wert auf zwei Nachkommastellen gerundet. Diese Daten wurden dann auf Betriebsebene aggregiert.</p>

5.6 Ortsangaben

5.6.1 Wohnort Kreis (wo_kreis)

Variablenlabel	Wohnort Kreis
Variablenname	wo_kreis
Kategorie	Ortsangaben
Herkunft	BeH, LeH, LHG, XLHG, ASU, XASU
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, XLHG und XASU gilt der Wohnort für den Zeitraum der Originalbeobachtung. Bei ASU und LeH 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 31.12.2011 umgeschlüsselt, d. h. in allen Kalenderjahren richtet sich die Zuordnung eines Wohnortes zu einem Kreis nach den Grenzen, welche die Kreise zum 31.12.2011 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> <p>Für Sätze aus den Quellen LHG, XLHG und XASU ist das Merkmal ab 2005 gefüllt.</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 Arbeitsagentur (wo_aa)

Variablenlabel	Wohnort Arbeitsagentur
Variablenname	wo_aa
Kategorie	Ortsangaben
Herkunft	BeH, LeH, LHG, XLHG, ASU, XASU
Datentyp	Numerisch
Hierarchie	Regionaldirektion Arbeitsagentur

Detailbeschreibung	<p>Das Merkmal enthält die Arbeitsagentur, die am Wohnort des Beschäftigten/ Kunden zuständig ist. Sie wird aus der Wohnadresse ermittelt. Für die Quellen LHG, XLHG und XASU gilt der Wohnort für den Zeitraum der Originalbeobachtung. Bei LeH und ASU enthält das Merkmal den Wohnort zum Beginn der Arbeitslosigkeit bzw. der Arbeitsuche. Es ist für die BeH und LeH ab 1999 vorhanden, für LHG, XLHG und XASU ab 2005 und für ASU-Sätze bereits vor 1999.</p> <p>Um konsistente regionale Zuordnungen über den gesamten Beobachtungszeitraum sicherzustellen, wurden die Arbeitsagenturangaben für alle Quellen auf den Gebietsstand vom 31.12.2011 umgeschlüsselt, d.h. in allen Kalenderjahren richtet sich die Zuordnung eines Wohnortes zu einem Arbeitsamtsbezirk nach den Grenzen, welche die Arbeitsagenturbezirke zum 31.12.2011 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>
Besonderheiten	<p>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.</p>

5.6.3 Wohnort Regionaldirektion (wo_rd)

Variablenlabel	Wohnort Regionaldirektion
Variablenname	wo_rd
Kategorie	Ortsangaben
Herkunft	BeH, LeH, LHG, XLHG, ASU, XASU
Datentyp	Numerisch
Hierarchie	Regionaldirektion Arbeitsagentur
Detailbeschreibung	<p>Das Merkmal beschreibt die Regionaldirektion des Wohnorts. Für die Quellen LHG, XLHG und XASU gilt der Wohnort für den Zeitraum des Satzes. Ein Wechsel des Wohnorts (bereits auf Ebene der Geschäftsstellen bzw. Gemeinden) löst einen neuen Satz aus. Bei LeH und ASU enthält das Merkmal den Wohnort zum Beginn der Arbeitslosigkeit bzw. des Arbeitsuchendenstatus. Das Merkmal ist für die BeH und LeH ab 1999 vorhanden, für LHG, XLHG und XASU ab 2005. Für ASU Sätze ist das Merkmal bereits vor 1999 gefüllt.</p> <p>Das Merkmal ist für alle Quellen gebietsstandsbereinigt zum 31.12.2011.</p>

5.6.4 Wohnort Bundesland (wo_bula)

Variablenlabel	Wohnort Bundesland
Variablenname	wo_bula
Kategorie	Ortsangaben
Herkunft	BeH, LeH, LHG, XLHG, ASU, XASU
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-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>
--------------------	---

5.6.5 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 31.12.2011 umgeschlüsselt, d. h. in allen Kalenderjahren richtet sich die Zuordnung eines Betriebsortes zu einem Kreis nach den Grenzen, welche die Kreise zum 31.12.2011 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.

5.6.6 Arbeitsort Arbeitsagentur (ao_aa)

Variablenlabel	Arbeitsort Arbeitsagentur
Variablenname	ao_aa
Kategorie	Ortsangaben
Herkunft	BeH, LeH, ASU
Datentyp	Numerisch
Hierarchie	Regionaldirektion Arbeitsagentur
Detailbeschreibung	<p>Für Beschäftigungsmeldungen aus der BeH enthält das Merkmal die Arbeitsagentur des Arbeitsortes. Für die LeH weist dieses Merkmal die zahlungsleistende Arbeitsagentur aus. Die den Kunden betreuende Arbeitsagentur ist in der ASU enthalten. Im Gegensatz zum Merkmal „Wohnort Arbeitsagentur“ ist dieses Merkmal für BeH und LeH-Sätze ab 1975 gefüllt.</p> <p>Das Merkmal ist gebietsstandsbereinigt zum 31.12.2011. Dies gilt allerdings nicht für LeH und ASU.</p>
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 (ao_rd), in dem der Wohnort des Versicherten liegt, ausgewiesen.

5.6.7 Arbeitsort Regionaldirektion (ao_rd)

Variablenlabel	Arbeitsort Regionaldirektion
Variablenname	ao_rd
Kategorie	Ortsangaben
Herkunft	BeH, LeH, ASU
Datentyp	Numerisch
Hierarchie	Regionaldirektion Arbeitsagentur
Detailbeschreibung	Das Merkmal beschreibt die Regionaldirektion des Arbeitsortes des BA-Kunden. Im Gegensatz zum Merkmal „Wohnort Regionaldirektion“ ist dieses Merkmal für BeH und LeH-Sätze ab 1975, für ASU-Sätze ab dem Jahr 2000 gefüllt. Das Merkmal ist gebietsstandsbereinigt zum 31.12.2011.

5.6.8 Arbeitsort Bundesland (ao_bula)

Variablenlabel	Arbeitsort Bula
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.

5.7 Generierte technische Merkmale

5.7.1 Quelle des Satzes (quelle)

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

5.7.2 Satzzähler Konto (spell)

Variablenlabel	Satzzähler Konto
Variablenname	spell
Kategorie	Generierte technische Merkmale
Herkunft	BeH, LeH, LHG, XLHG, ASU, XASU
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 wieder hergestellt werden. Die Sortierung erfolgt grundsätzlich zunächst nach dem Beginndatum der Episode und dann nach der Quelle.

5.7.3 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.

5.7.4 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 aus den administrativen Personendaten vorliegen. Diese Variable nimmt in wenigen Fällen den Wert 0 annehmen, obwohl die Variable <code>match_typ</code> 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. Dies kann verschiedene Auslöser haben, z. B. wenn bei einer Interaktion der Person mit einer Dienststelle der BA (z.B. bei einer Berufsberatung) zwar deren Name und Adresse hinterlegt wurden, es jedoch nicht zu einem Ereignis gekommen ist, das für die IEB relevant gewesen wäre. Aus inhaltlicher Sicht stellen diese Fälle keinen Verlust dar, da hierzu keine relevanten Verlaufsdaten vorliegen, so dass diese Personen ohnehin nicht in Analysen hätten einbezogen werden können.

5.7.5 Typ des Matches (match_typ)

Variablenlabel	Typ des Matches
Variablenname	match_typ
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 ausreichend hohe Übereinstimmung zwischen den Identifikatoren beider Datenquellen gefunden werden konnte. Die Ausprägungen steigen mit der Abfolge der Stufen, d. h. Ausprägung 3 bedeutet, dass die Stufe des Goldstandard-Matchens nicht erfolgreich war usw. Ausprägung 0 steht für PASS-Befragte, die der Verknüpfung nicht zugestimmt haben.

5.7.6 Art des exakten Abgleichs (`exakt_typ`)

Variablenlabel	Art des exakten Abgleichs
Variablenname	<code>exakt_typ</code>
Kategorie	Verknüpfungsbezogenes Merkmal
Herkunft	Generiert
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Dieses Merkmal ist nur in der Verknüpfungsdatei enthalten. Es beschreibt, auf welche Identifikatoren ein exakter Abgleich durchgeführt wurde. Je größer der Wert des Merkmals, desto geringer die Sicherheit einer richtigen Zuordnung zwischen Befragungs- und administrativen Daten. Insgesamt enthält die Datei jedoch nur Fälle, die erfahrungsgemäß mit sehr hoher Sicherheit einen richtigen Match darstellen.

5.7.7 Probabilistisches Ähnlichkeitsmaß (`quality`)

Variablenlabel	Probabilistisches Ähnlichkeitsmaß
Variablenname	<code>quality</code>
Kategorie	Verknüpfungsbezogenes Merkmal
Herkunft	Generiert
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Dieses Merkmal ist nur in der Verknüpfungsdatei enthalten. Es zeigt an, welches Ähnlichkeitsmaß beim probabilistischen Abgleich der Identifikatoren beider Datenquellen errechnet wurde. Da Goldstandard- oder exakte Matches nicht mehr in den probabilistischen Vergleich einbezogen wurden, weisen diese Fälle kein gültiges Ähnlichkeitsmaß auf. Zu beachten ist auch, dass man die Ausprägungen der Variable nur innerhalb einzelner Matching-Anläufe vergleichen kann, die sich mithilfe des Merkmals <code>match_wave</code> unterscheiden lassen. Ein übergreifender Vergleich ist nicht möglich, da die Verteilung von des Merkmals <code>quality</code> von Art und Anzahl der jeweils verwendeten Identifikatoren, den Vergleichsalgorithmen und den zu Grunde gelegten m- und u-Parametern abhängt.

5.7.8 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 Welle wurde versucht, neu hinzugekommene oder vorher nicht erfolgreich verknüpfte Befragte in den administrativen Daten zu finden. Da Person nicht zwingend sofort im ersten Anlauf gefunden werden, kann die Ausprägung dieses Merkmal von der Welle abweichen, in der die Verknüpfungserlaubnis erstmals erteilt wurde.

6 Auszählungen

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

Literatur

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

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

Berg, Marco; Cramer, Ralph; Dickmann, Christian; Gilberg, Reiner; Jesske, Birgit; Kleudgen, Martin; Bethmann, Arne; Fuchs, Benjamin; Trappmann, Mark; Wurdack, Anja (2012): Codebuch und Dokumentation des „Panel Arbeitsmarkt und soziale Sicherung“ (PASS). Band I: Datenreport Welle 5. FDZ-Datenreport 06/2012, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

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.) (2007): Handbuch für die Betriebsnummernvergabe und -pflege im Rahmen des Meldeverfahren zur Sozialversicherung. Ausgabe Januar 2007, Nürnberg.

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

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

Deutsche BKK (2012): Ratgeber Sozialversicherung 2012. URL http://www.deutschebkk.de/fileadmin/user_upload/Service/Download-Center/Ratgeber_Sozialversicherung_2012_01.pdf, zugriff: 28.03.2012.

Drews, Nils (2006): Qualitätsverbesserung der Bildungsvariable in der IAB-Beschäftigtenstichprobe 1975–2001. FDZ Methodenreport 05/2006, Institut für Arbeitsmarkt- und Berufsforschung, 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, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

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

Gruhl, Anja; Schmucker, Alexandra; Seth, Stefan (2012): Das Betriebs-Historik-Panel 1975-2010. FDZ-Datenreport 04/2012, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

Herzog, Thomas N.; Scheuren, Fritz J.; Winkler, William E. (2007): Data Quality and Record Linkage Techniques. New York: Springer.

Jesske, Birgit; Schulz, Sabine (2012): Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 5. Erhebungswelle – 2011. FDZ-Methodenreport 11/2012, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

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

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

Trappmann, Mark; Christoph, Bernhard; Achatz, Juliane; Wenzig, Claudia; Müller, Daniel, Gerrit and Gebhardt (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.

vom Berge, Philipp; König, Marion; Seth, Stefan (2013): Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) 1975-2010. FDZ-Datenreport 01/2013, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

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, Bd. 21, Nr. 4, S. 468–482.

Impressum

FDZ-Datenreport 3/2014

Herausgeber

Forschungsdatenzentrum (FDZ)
der Bundesagentur für Arbeit
im Institut für Arbeitsmarkt- und Berufsforschung
Regensburger Str. 104
90478 Nürnberg

Redaktion

Stefan Bender, Dagmar Theune

Technische Herstellung

Dagmar Theune

Rechte

Nachdruck - auch auszugsweise - nur mit
Genehmigung des FDZ gestattet

Bezugsmöglichkeit

http://doku.iab.de/fdz/reporte/2014/DR_03-14.pdf

Internet

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

Rückfragen zum Inhalt an:

Manfred Antoni
Institut für Arbeitsmarkt- und
Berufsforschung (IAB)
Forschungsdatenzentrum
Regensburger Str. 104
90478 Nürnberg
E-Mail: Manfred.Antoni@iab.de